



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

ROBERTA CRISTINA DAL'EVEDOVE TARTAROTTI

**AVALIAÇÃO DO PROCESSO DE INDEXAÇÃO DE ASSUNTOS EM
REPOSITÓRIOS INSTITUCIONAIS PELA ABORDAGEM DA
RECUPERAÇÃO DA INFORMAÇÃO**

**MARÍLIA
2019**

ROBERTA CRISTINA DAL'EVEDOVE TARTAROTTI

**AVALIAÇÃO DO PROCESSO DE INDEXAÇÃO DE ASSUNTOS EM
REPOSITÓRIOS INSTITUCIONAIS PELA ABORDAGEM DA
RECUPERAÇÃO DA INFORMAÇÃO**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, como requisito parcial para a obtenção do título de Doutora em Ciência da Informação.

Área de concentração: Informação, Tecnologia e Conhecimento.

Linha de Pesquisa: Produção e Organização da Informação.

Orientadora: Profa. Dra. Mariângela Spotti Lopes Fujita.

MARÍLIA
2019

T194a Tartarotti, Roberta Cristina Dal'Evedove
Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação / Roberta Cristina Dal'Evedove Tartarotti. - Marília, 2019.
370 p.: il. color.; 30 cm.

Tese (Doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp). Faculdade de Filosofia e Ciências. Programa de Pós-graduação em Ciência da Informação.

Orientadora: Mariângela Spotti Lopes Fujita.

1. Indexação de assuntos. 2. Avaliação da indexação de assuntos. 3. Repositórios institucionais. 4. Recuperação por assuntos. 5. Bibliotecas universitárias. I. Fujita, Mariângela Spotti Lopes. II. Título.

CDD 025.524

CERTIFICADO DE APROVAÇÃO

Roberta Cristina Dal'Evedove Tartarotti

Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação

Área de concentração: Informação, Tecnologia e Conhecimento

Linha de pesquisa: Organização da Informação

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, como parte das exigências para a obtenção do título de Doutora em Ciência da Informação.

Profa. Dra. Mariângela Spotti Lopes Fujita (Orientadora) - Presidente da Banca Examinadora
Departamento de Ciência da Informação
Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP/Marília - SP

Prof. Dr. Edberto Ferneda (Membro interno)
Departamento de Ciência da Informação
Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP/Marília - SP

Profa. Dra. Brígida Maria Nogueira Cervantes (Membro externo)
Departamento de Ciência da Informação
Universidade Estadual de Londrina - UEL/Londrina - SP

Profa. Dra. Giovana Deliberali Maimone (Membro externo)
Departamento de Ciência da Informação
Escola de Comunicações e Artes - ECA
Universidade de São Paulo - USP/São Paulo - SP

Prof. Dr. Isidoro Gil Leiva (Membro externo)
Departamento de Información y Documentación
Universidad de Murcia - UM/Murcia - Espanha

Marília, 26 de setembro de 2019.

*Ao meu esposo **Renato**,
companheiro nesta vida e nas eternidades;
Aos meus filhos **Enzo e Pietro**,
meus “pacotinhos” de amor;
À minha filha **Laura**,
meu pequeno raio de sol,
que desde a barriga
me acompanha nesta trajetória ♥*

AGRADECIMENTOS

À minha **família e amigos**, por todo o incentivo e amor sempre presentes.

À minha orientadora Profa. Dra. **Mariângela Spotti Lopes Fujita**, pela amizade e sábia orientação que teve início ainda na iniciação científica.

À Profa. Dra. **Brígida Maria Nogueira Cervantes**, Profa. Dra. **Giovana Deliberali**, Prof. Dr. **Edberto Ferneda** e Prof. Dr. **Isidoro Gil Leiva**, pelas ricas contribuições nas Bancas de Qualificação e Defesa.

À Profa. Dra. **Vera Regina Casari Boccato** (in memoriam), pela amizade e orientação no mestrado, você estará sempre em nossos corações.

Aos **colegas e professores** do Programa de Pós-Graduação em Ciência da Informação da UNESP de Marília.

Ao **Sistema de Bibliotecas da UNICAMP (SBU)**, pelas oportunidades de atuação profissional nestes 15 anos.

Aos **diretores** das Bibliotecas e aos **participantes** da pesquisa, que possibilitaram a realização das coletas de dados.

E a **todos** que, de alguma forma, contribuíram para a realização desta pesquisa.

*“A university is just a group of buildings
gathered around a library.”*

(Shelby Foote)

RESUMO

Na contemporaneidade, os repositórios institucionais possibilitam a reunião, armazenamento, tratamento, preservação e visibilidade dos recursos informacionais produzidos no âmbito acadêmico. Considerando-se que o campo da Organização do Conhecimento possui os fundamentos teórico-práticos necessários no que tange aos problemas de representação e recuperação em sistemas de recuperação da informação, a proposição da tese foi investigar/avaliar o processo de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias e sua influência na recuperação da informação pelos usuários, visando o aprimoramento do tratamento/da representação temática da informação documental e da recuperação por assuntos. O objetivo geral da pesquisa foi contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação por assuntos nestes sistemas de recuperação da informação. Para tanto, foram delineados os seguintes objetivos específicos: a) realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos em perspectiva dicotômica: do tratamento/da representação e da recuperação por assuntos no contexto de bibliotecas universitárias; b) investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como metodologia qualitativa com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP, com coleta de documentação; c) verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada no Repositório Institucional da UNICAMP por meio da aplicação da avaliação da indexação de assuntos como metodologia quantitativa, delineada em: avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*; d) sistematizar a análise comparada e delinear recomendações para a indexação de assuntos para repositórios institucionais, visando propor um modelo pertinente de avaliação da indexação de assuntos em repositórios institucionais. De forma geral, os resultados revelaram a necessidade de adequação dos repositórios institucionais tanto no tratamento/na representação temática realizada pelo bibliotecário catalogador-indexador quanto na busca e recuperação por assuntos realizada pelos usuários. Como considerações finais, verifica-se a tendência e cada vez maior visibilidade dos repositórios institucionais não apenas no contexto de bibliotecas universitárias e, em uma perspectiva mais ampla, no contexto acadêmico no qual estão inseridas, tornando-se instrumentos de gestão universitária por meio de indicadores científicos. Neste cenário, reitera-se a relevância da recuperação da informação em repositórios institucionais condizente com as necessidades informacionais dos usuários, pautada em princípios/diretrizes de indexação de assuntos da Organização do Conhecimento e os atuais recursos tecnológicos de representação e recuperação da informação.

Palavras-chave: Indexação de assuntos. Recuperação por assuntos. Avaliação da indexação de assuntos. Repositórios institucionais. Bibliotecas universitárias.

ABSTRACT

In present times, the institutional repositories enable the gathering, storage, treatment, preservation and visibility of information resources produced in the academic field. Considering that the Knowledge Organization field has the necessary theoretical-practical foundations regarding the issues of representation and retrieval in information retrieval systems, the proposition of the thesis was to investigate/evaluate the subject indexing process in institutional repositories in the context of academic libraries and its influence on information retrieval by users, aiming at improving the treatment/thematic representation of documentary information and subject retrieval. The general objective of the research was to contribute to the implementation / improvement of subject index guidelines in institutional repositories in the context of academic libraries, enabling the adequate representation and subject retrieval in information retrieval systems. To this end, the following specific objectives were outlined: a) conducting a theoretical study on institutional repositories as information retrieval systems and the subject index process within in a dichotomous perspective: of the treatment/of the representation and of the subject retrieval in the context of academic libraries; b) investigating the treatment/the thematic representation of documentary information in institutional repositories in the context of academic libraries, by means of the application of organizational diagnostic interviews as a qualitative methodology with the managers of the institutional repositories and catalogers-indexers librarians from USP, UNESP and UNICAMP, with documentation collection; c) verifying the documentary retrieval by subjects, by comparison between natural language and controlled language at UNICAMP's Institutional Repository, by applying subject indexing evaluation as a quantitative methodology, outlined in: subject indexing evaluation in institutional repositories from the system's perspective - *Quantitative intrinsic evaluation through interconsistency* and from the user's perspective - *Extrinsic evaluation through retrieval*; d) systematize the comparative analysis and outline recommendations for subject indexing for institutional repositories, aiming to propose a pertinent model of the subject indexing evaluation in institutional repositories. In general, the results revealed the need for institutional repositories adequacy both in the treatment/in the thematic representation performed by the cataloguer-indexer librarian and in the search and subject retrieval done by users. As closing remarks, there is the tendency and an increasing visibility of institutional repositories not only in the context of academic libraries and, in a broader perspective, in the academic context, in which they are inserted, becoming university management instruments through scientific indicators. In this scenario, the relevance of information retrieval in institutional repositories in accordance with user's information needs, based on principals/subject indexing guidelines from Knowledge Organization, and current technological resources of representation and information retrieval.

Keywords: Subject indexing. Subject retrieval. Subject indexing evaluation. Institutional repositories. University libraries.

RESUMEN

En los tiempos contemporáneos, los repositorios institucionales permiten la recolección, almacenamiento, tratamiento, preservación y visibilidad de los recursos informacionales producidos en el campo académico. Considerando que el campo de la Organización del Conocimiento presenta los fundamentos teórico-prácticos necesarios en cuanto a los problemas de representación y recuperación en los sistemas de recuperación de información, la propuesta de la tesis fue investigar/evaluar el proceso de indización por materias en repositorios institucionales y su influencia en la recuperación de información por los usuarios, pretendiendo el perfeccionamiento del tratamiento temático de la información documental y de la recuperación por materias en estos sistemas de recuperación tanto en el contexto de bibliotecas universitarias como en el ámbito académico. El objetivo general de la investigación fue contribuir a la implementación/mejoras de directrices de indización por materias en repositorios institucionales en el contexto de bibliotecas universitarias, posibilitando la adecuada representación y recuperación por materias en estos sistemas de recuperación de información. Para ello, fueron delineados los siguientes objetivos específicos: a) realizar un estudio teórico sobre el proceso de indización por materias desde la perspectiva de la Organización del Conocimiento en perspectiva dicotómica: del tratamiento/de la representación y de la recuperación por materias en el contexto de bibliotecas universitarias y sobre los repositorios institucionales como sistemas de recuperación de información; b) investigar el tratamiento/la representación temática de la información documental en repositorios institucionales en el contexto de bibliotecas universitarias, a través de la aplicación de entrevistas de diagnóstico organizacional como metodología cualitativa a los responsables de los repositorios institucionales y bibliotecarios catalogadores-indizadores de la USP, UNESP y UNICAMP, con recolección de documentación; c) verificar la recuperación documental por materias de forma comparada entre lenguaje natural y lenguaje controlado en el repositorio institucional de la UNICAMP, por medio de la aplicación de la evaluación de la indización por materias como metodología cuantitativa, delineada en: evaluación de la indización por materias en repositorios institucionales por la perspectiva del sistema - *Evaluación intrínseca cuantitativa mediante la interconsistencia* y por la perspectiva del usuario - *Evaluación extrínseca mediante la recuperación*; d) sistematizar el análisis comparativo y delinear recomendaciones para la indización por materias para repositorios institucionales, buscando proponer un modelo pertinente de Evaluación de la indización por materias en repositorios institucionales. En general, los resultados revelaron la necesidad de que los repositorios institucionales se adapten tanto al tratamiento temático de la información realizada por el bibliotecario catalogador-indexador como a la búsqueda y recuperación de información por parte de los usuarios. Como consideraciones finales, se verifica la tendencia y visibilidad cada vez mayor de los repositorios institucionales como sistemas de recuperación de información, no sólo en el contexto de bibliotecas universitarias sino además, desde, desde una perspectiva más amplia, del contexto académico en el que se insertan, instrumentos de gestión universitaria mediante indicadores científicos. En este escenario, se reitera la relevancia de la recuperación de información en repositorios institucionales acorde con las necesidades informacionales de los usuarios, pautada en principios/directrices de indización por materias de la Organización del Conocimiento y los actuales recursos tecnológicos de representación y recuperación de información.

Palabras clave: Indización por materias. Recuperación por materias. Evaluación de la indización por materias. Repositorios institucionales. Bibliotecas universitarias.

LISTA DE QUADROS

Quadro 1. Sistematização da pesquisa	31
Quadro 2. O padrão de metadados Dublin Core.....	61
Quadro 3. A tematicidade textual na análise de assunto do processo de indexação	81
Quadro 4. Diferentes abordagens de avaliação da indexação de assuntos em repositórios institucionais.....	105
Quadro 5. Fórmulas matemáticas para obtenção do índice de consistência entre duas indexações no repositório institucional	123
Quadro 6. Seleção das Bibliotecas participantes da pesquisa para a Avaliação extrínseca mediante a recuperação	125
Quadro 7. Participantes da pesquisa e respectivas siglas para identificação e análise das entrevistas	135
Quadro 8. Exemplos de aplicação das fórmulas de consistência no Repositório Institucional da UNICAMP	186
Quadro 9. Participantes da pesquisa e respectivas siglas para identificação e análise da Avaliação extrínseca mediante a recuperação.....	204
Quadro 10. Relação entre as áreas do conhecimento das Bibliotecas e as linguagens de indexação utilizadas	242
Quadro 11. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Artes e Humanidades	248
Quadro 12. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Biomédicas	250
Quadro 13. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Ciências Aplicadas	251
Quadro 14. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Exatas.....	252
Quadro 15. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Tecnológicas.....	253
Quadro 16. Síntese dos principais resultados das entrevistas com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais	303

LISTA DE TABELAS

Tabela 1. Distribuição do número de assuntos atribuídos por autores e bibliotecários catalogadores-indexadores no Repositório Institucional da UNICAMP por área do conhecimento.....	177
Tabela 2. Índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores no Repositório Institucional da UNICAMP	187
Tabela 3. Índices de precisão em linguagem natural no Repositório Institucional da UNICAMP.....	266
Tabela 4. Índices de precisão em linguagem controlada no Repositório Institucional da UNICAMP.....	267

LISTA DE FIGURAS

Figura 1. O processo de indexação de assuntos em repositório institucional.....	74
Figura 2. Página inicial do Repositório Institucional da USP	111
Figura 3. Página inicial do Repositório Institucional da UNESP	111
Figura 4. Página inicial do Repositório Institucional da UNICAMP.....	112
Figura 5. Página inicial do Repositório Institucional do CRUESP.....	113
Figura 6. Avaliação da indexação de assuntos comparada entre autores e bibliotecários catalogadores-indexadores	124
Figura 7. Tela inicial do software Free Screen to Video 2.0.....	128
Figura 8. Total de recursos informacionais analisados no Repositório Institucional da UNICAMP por área do conhecimento.....	174
Figura 9. Número de assuntos atribuídos na indexação por autores e bibliotecários catalogadores-indexadores por recurso informacional.....	175
Figura 10. Número de assuntos em linguagem natural e linguagem controlada - Área de Artes e Humanidades.....	178
Figura 11. Número de assuntos em linguagem natural e linguagem controlada - Área de Biomédicas	179
Figura 12. Número de assuntos em linguagem natural e linguagem controlada - Área de Ciências Aplicadas	180
Figura 13. Número de assuntos em linguagem natural e linguagem controlada - Área de Exatas	181
Figura 14. Número de assuntos em linguagem natural e linguagem controlada - Área de Tecnológicas	182
Figura 15. Número de assuntos atribuídos em linguagem natural por área do conhecimento.....	183
Figura 16. Número de assuntos em linguagem controlada por área do conhecimento.....	184
Figura 17. Média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por tipo de índice.....	188
Figura 18. Índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores por área do conhecimento.....	188
Figura 19. Média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento.....	190
Figura 20. Escala dos índices de consistência	192
Figura 21. Total de recursos informacionais no Repositório Institucional da UNICAMP	195

Figura 22. Total de recursos informacionais no Repositório Institucional da UNICAMP por data de publicação	195
Figura 23. Opções de busca no Repositório Institucional da UNICAMP na página inicial.....	196
Figura 24. Visualização geral dos registros recuperados no Repositório Institucional da UNICAMP	197
Figura 25. Opções gerais de busca no Repositório Institucional da UNICAMP por tipo de produção	198
Figura 26. Opções de busca no Repositório Institucional da UNICAMP por tipo de recurso informacional.....	199
Figura 27. Opções de busca no Repositório Institucional da UNICAMP por assuntos.....	199
Figura 28. Opções de busca no Repositório Institucional da UNICAMP por assunto com o recurso autocompletar	200
Figura 29. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Assunto.....	200
Figura 30. Opções de busca no Repositório Institucional da UNICAMP por ordenação de registros - A.....	201
Figura 31. Opções de busca no Repositório Institucional da UNICAMP por ordenação de registros - B.....	202
Figura 32. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Agência de fomento.....	202
Figura 33. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Tipo de recurso informacional.....	203
Figura 34. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Data de publicação e Tipo de acesso	203
Figura 35. Tempo de duração das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP por área do conhecimento.....	207
Figura 36. Exemplo de estratégia de busca utilizada pelo usuário no Repositório Institucional da UNICAMP - A.....	222
Figura 37. Exemplo de estratégia de busca utilizada pelo usuário no Repositório Institucional da UNICAMP - B	222
Figura 38. Página inicial da Library of Congress Subject Headings (LCSH)	244
Figura 39. Página inicial da Terminologia de Assuntos da Biblioteca Nacional	245
Figura 40. Página inicial do Descritores em Ciências da Saúde (DeCS)	246
Figura 41. Página inicial do Ei Thesaurus	247
Figura 42. Página inicial do FSTA Thesaurus	248

Figura 43. Fórmula para a determinação do índice de precisão no Repositório Institucional da UNICAMP	266
Figura 44. Índices de precisão em linguagem natural por área do conhecimento	267
Figura 45. Índices de precisão em linguagem controlada por área do conhecimento	268
Figura 46. Índices de precisão em linguagem natural e linguagem controlada por área do conhecimento	269
Figura 47. Média dos índices de precisão entre linguagem natural e linguagem controlada por área do conhecimento no Repositório Institucional da UNICAMP	273
Figura 48. Média dos índices de precisão no Repositório Institucional da UNICAMP por tipo de índice	274
Figura 49. Índice total de precisão das linguagens de indexação na recuperação por assuntos no Repositório Institucional da UNICAMP	275
Figura 50. Exemplo de pré-visualização de registros no Repositório Institucional da UNICAMP	276
Figura 51. Exemplo de duplicação de registros entre artigo e artigo de periódico no Repositório Institucional da UNICAMP	277
Figura 52. Exemplo de duplicação de registros entre artigos de periódico no Repositório Institucional da UNICAMP	278
Figura 53. Exemplo de duplicação de registros em artigo de periódico com títulos em inglês e português no Repositório Institucional da UNICAMP	279
Figura 54. Exemplo de duplicação de resumos em registro no Repositório Institucional da UNICAMP	280
Figura 55. Exemplo de visualização de datas no Repositório Institucional da UNICAMP	280
Figura 56. Exemplo de visualização de inconsistência em títulos no Repositório Institucional da UNICAMP	281
Figura 57. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - A	282
Figura 58. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - B	282
Figura 59. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - C	283
Figura 60. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - D	283
Figura 61. Exemplo de visualização de inconsistência em título e tipo de recurso informacional no Repositório Institucional	284

LISTA DE ABREVIATURAS E SIGLAS

AACR2	Código de catalogação Anglo-Americano
BAE	Biblioteca da Área de Engenharia e Arquitetura
BDPD	Biblioteca Digital de Teses e Dissertações da USP
BDPI	Biblioteca Digital da Produção Intelectual da USP
BN	Biblioteca Nacional
CGB-UNESP	Coordenadoria Geral de Bibliotecas da UNESP
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CONFOA	Conferência Luso-Brasileira sobre Acesso Aberto
CRUESP	Conselho de Reitores das Universidades Estaduais de São Paulo
DCMI	<i>Dublin Core Metadata Initiative</i>
DeCS	Descritores em Ciências da Saúde
DTI	Diretoria de Tecnologia da Informação
DTRI	Diretoria de Tratamento da Informação
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FCA	Faculdade de Ciências Aplicadas da UNICAMP
FCM	Faculdade de Ciências Médicas da UNICAMP
FE	Faculdade de Educação da UNICAMP
FEA	Faculdade de Engenharia de Alimentos da UNICAMP
FEF	Faculdade de Educação Física da UNICAMP
FRAD	<i>Functional Requirements for Authority Data</i>
FRBR	<i>Functional Requirements for Bibliographic Records</i>
FRSAD	<i>Functional Requirements for Subject Authority Data</i>
FSTA	<i>Food Science and Technology Abstracts</i>
GRI-UNESP	Grupo Gestor da Política do Repositório Institucional UNESP
IFGW	Instituto de Física Gleb Wathaghin da UNICAMP
IFLA	<i>International Federation of Library Associations and Institutions</i>
IMECC	Instituto de Matemática, Estatística e Computação Científica da UNICAMP
ISO	<i>International Organization for Standardization</i>
ISKO	<i>International Society for Knowledge Organization</i>

LC	<i>Library of Congress</i>
LCARB	Lista de Cabeçalho de Assunto da Rede BIBLIODATA
LCSH	<i>Library of Congress Subject Headings</i>
MARC	<i>Machine Readable Cataloging</i>
MeSH	<i>Medical Subject Headings</i>
OAI/PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
OC	Organização do Conhecimento
ORCID	<i>Open Researcher and Contributor ID</i>
PDF	<i>Portable Document Format</i>
PIBIC	Programa Institucional de Bolsas de Iniciação Científica
PPEC	Portal de Periódicos Eletrônicos Científicos
PRG	Pró-Reitoria de Graduação
PRP	Pró-Reitoria de Pesquisa
PRPG	Pró-Reitoria de Pós-Graduação
PV	Protocolo Verbal
PVI	Protocolo Verbal Individual
RDA	<i>Resource Description and Access</i>
RDF	<i>Resource Description Framework</i>
RI	Repositório Institucional
RI	Recuperação da informação
SBU	Sistema de Bibliotecas da UNICAMP
SciELO	<i>Scientific Electronic Library Online</i>
SIBi	Sistema de Bibliotecas da USP
UNESP	Universidade Estadual Paulista
TTIC	Tratamento Temático da Informação Científica
TCLE	Termo de Consentimento Livre e Esclarecido
UFSCar	Universidade Federal de São Carlos
UNICAMP	Universidade Estadual de Campinas
USP	Universidade de São Paulo
VOCAUSP	Vocabulário Controlado da USP
WOS	<i>Web of Science</i>

SUMÁRIO

1 INTRODUÇÃO	20
1.1 Antecedentes	20
1.2 Delineamento da pesquisa.....	21
1.2.1 Problema e hipótese	29
1.2.2 Tese e proposição.....	29
1.2.3 Objetivos	30
1.3 Apresentação das seções	31
2 OS REPOSITÓRIOS INSTITUCIONAIS COMO SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO NO CONTEXTO DE BIBLIOTECAS UNIVERSITÁRIAS	34
2.1 As universidades como instituições de ensino superior.....	34
2.2 As bibliotecas universitárias e seu papel no sistema de comunicação científico- acadêmico	37
2.3 O advento dos repositórios institucionais como sistemas de recuperação da informação no âmbito acadêmico	41
2.4 Os repositórios institucionais de dados de pesquisa	50
3 A INDEXAÇÃO DE ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS: DA REPRESENTAÇÃO À RECUPERAÇÃO DA INFORMAÇÃO	55
3.1 A indexação de assuntos em repositórios institucionais pela abordagem da <i>representação da informação</i>	55
3.1.1 Metadados para a representação e recuperação da informação em repositórios institucionais	56
3.1.2 A indexação documental em repositórios institucionais: conceitos, objetivos e princípios	68
3.2 A indexação de assuntos em repositórios institucionais pela abordagem da <i>recuperação da informação</i>	84
3.2.1 A recuperação por assuntos em repositórios institucionais	89
3.2.2 Revisitando os conceitos de relevância, pertinência e saliência na recuperação da informação em repositórios institucionais	92
3.2.3 Variáveis/indicadores/métricas da recuperação por assuntos em repositórios institucionais	96
3.3 Avaliação da indexação de assuntos: em busca de indicadores para a melhoria da representação e da recuperação por assuntos em repositórios institucionais	99
4 PROCEDIMENTOS METODOLÓGICOS	107

4.1 Abordagem teórico-aplicada.....	107
4.2 Abordagem exploratório-descritiva	108
4.2.1 Caracterização do universo de pesquisa: os repositórios institucionais do CRUESP	110
4.2.2 Abordagem qualitativa: entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais	113
4.2.3 Abordagem quantitativa: avaliação da indexação de assuntos	119
5 RESULTADOS E DISCUSSÃO	134
5.1 Resultados da abordagem qualitativa: entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais	134
5.2 Resultados da abordagem quantitativa: avaliação da indexação de assuntos	172
5.2.1 Resultados da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do sistema - <i>Avaliação intrínseca quantitativa mediante a interconsistência</i>	172
5.2.2 Resultados da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do usuário - <i>Avaliação extrínseca mediante a recuperação</i>	193
6 SÍNTESE DOS RESULTADOS	303
7 RECOMENDAÇÕES PARA A INDEXAÇÃO E RECUPERAÇÃO POR ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS.....	309
8 CONSIDERAÇÕES FINAIS	312
REFERÊNCIAS	316
Apêndice A. Roteiro da Entrevista de Diagnóstico Organizacional com os gestores dos repositórios institucionais	328
Apêndice B. Roteiro da Entrevista de Diagnóstico Organizacional com os bibliotecários catalogadores-indexadores dos repositórios institucionais	330
Apêndice C. Termo de Consentimento Livre e Esclarecido (TCLE) - Entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores dos repositórios institucionais..	332
Apêndice D. Modelo de Autorização para coleta de dados nas Bibliotecas participantes ...	334
Apêndice E. Termo de Consentimento Livre e Esclarecido (TCLE) - Entrevista semiestruturada de busca por assuntos em repositório institucional pela perspectiva do usuário.....	335
Apêndice F. Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem natural	337
Apêndice G. Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem controlada.....	338

Apêndice H. Tempo de duração das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP	339
Apêndice I. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Artes e Humanidades - Biblioteca da FE	341
Apêndice J. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Artes e Humanidades - Biblioteca do IFCH	343
Apêndice K. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Biomédicas - Biblioteca da FCM	345
Apêndice L. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Biomédicas - Biblioteca da FEF	347
Apêndice M. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Ciências Aplicadas - Biblioteca da FCA.....	349
Apêndice N. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Exatas - Biblioteca do IFGW	351
Apêndice O. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Exatas - Biblioteca do IMECC.....	353
Apêndice P. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Tecnológicas - Biblioteca da BAE.....	355
Apêndice Q. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Tecnológicas - Biblioteca da FEA	357
Anexo A. Notações específicas para transcrição das entrevistas.....	359
Anexo B. Resolução Nº. 6444 (USP), de 22 DE OUTUBRO DE 2012 (D.O.E. 23.10.2012)	360
Anexo C. Regulamento Interno do Repositório Institucional (UNESP)	362
Anexo D. Resolução GR-013/2015 (UNICAMP), de 06/07/2015	365
Anexo E. Parecer Final do Comitê de Ética de Pesquisa (CEP) da Universidade Estadual Paulista (UNESP/Marília)	367

1 INTRODUÇÃO

Nesta seção introdutória, são apresentados os *antecedentes* que embasaram a elaboração desta Tese de Doutorado, o *delineamento da pesquisa* e sua estrutura na *apresentação das seções*, a fim de permitir um panorama geral da mesma.

1.1 Antecedentes

A presente pesquisa é oriunda de alguns anos de estudo e atuação profissional da pesquisadora em indexação e catalogação de assuntos. Inserida na linha de pesquisa “Produção e Organização da Informação” no âmbito do Programa de Ciência da Informação da Universidade Estadual Paulista (UNESP)/FFC, Campus de Marília, SP, tem como temário central a **indexação de assuntos em repositórios institucionais**.

Nascida mais próxima do contexto contemporâneo, garantindo assim maior “flexibilidade” e “tolerância” relativa à sua consolidação científica (SIQUEIRA, 2010, p. 64), a Ciência da Informação vem se institucionalizando enquanto campo científico, caracterizada por seus objetos, métodos, técnicas, práticas, atores e, mais especificamente, por sua produção científica e avaliação pelos pares (TARTAROTTI; FUJITA, 2016, p. 137).

A crescente produção científica aliada às novas tecnologias de comunicação e informação na contemporaneidade impulsionou o desenvolvimento dos sistemas de recuperação da informação, no intuito de disseminar a informação científica. No bojo desse novo paradigma informacional, caracterizado pelo papel cada vez mais fundamental que desempenham os processos eletrônicos de análise, armazenamento e transmissão da informação, especialmente na Internet, a Ciência da Informação busca resolver os problemas relacionados à produção, tratamento e recuperação de recursos informacionais em contextos específicos. Para Hjørland (2003, p. 88), os principais atores na Ciência da Informação são os produtores de conhecimento, os intermediários e os usuários, tendo como ponto central do campo a interação destes três elementos com os registros informacionais.

Neste contexto, a Tese de Doutorado apresenta-se em continuidade à Dissertação de Mestrado (TARTAROTTI, 2014) desenvolvida no âmbito do Programa de Pós-Graduação em

Ciência, Tecnologia e Sociedade da Universidade Federal de São Carlos (UFSCar). Partindo-se da relevância de investigações teórico-metodológicas em torno do profissional que realiza a análise de assunto, a pesquisa teve como objetivo realizar um estudo comparado da atuação bibliotecária no tratamento temático da informação de bibliotecas universitárias. Para tanto, investigou a catalogação de assuntos em catálogo coletivo online e a indexação de assuntos em base de dados por meio da coleta de dados com bibliotecários catalogadores e indexadores na área de Saúde, possibilitando uma análise detalhada da atuação profissional em ambos contextos de informação. Como procedimentos metodológicos, utilizou como metodologia qualitativa a técnica introspectiva do Protocolo Verbal (PV) na modalidade Protocolo Verbal Individual (PVI) e como metodologia quantitativa a avaliação da indexação de assuntos, na modalidade *Avaliação intrínseca da indexação ou Intrabibliotecário*. De forma geral, a pesquisa contribuiu para a temática abordada na literatura e a prática profissional da indexação e catalogação de assuntos, com o intuito de aprimorar a atuação profissional no tratamento temático da informação no contexto de bibliotecas universitárias.

Nesse sentido, com o intuito de proceder à continuidade da referida pesquisa, verificou-se a necessidade de ampliar as investigações em torno da relação entre o *processo de indexação de assuntos* realizado pelo bibliotecário catalogador-indexador e a *recuperação por assuntos* no sistema de recuperação da informação, considerando-se ser a recuperação o ponto fulcral deste processo, aspecto não abarcado na pesquisa anterior. No entanto, tendo em vista as recentes mudanças no âmbito das bibliotecas universitárias e, em um contexto mais amplo, no próprio ambiente acadêmico em que estão inseridas, o catálogo online e as bases de dados vem abrindo espaço para um novo sistema de recuperação cujo foco é a produção intelectual das universidades: os *repositórios institucionais*, o que levou ao delineamento da presente pesquisa.

1.2 Delineamento da pesquisa

No intuito de contextualizar a problemática subjacente da pesquisa, cujo temário é a *indexação de assuntos em repositórios institucionais*, julga-se importante discutir alguns dos fundamentos da Organização do Conhecimento (*Knowledge Organization*) para o processo de indexação de assuntos em duas dimensões: do tratamento/da representação e da

recuperação da informação e os repositórios institucionais como sistemas de recuperação da informação no contexto de bibliotecas universitárias.

Diversos autores têm se debruçado nas questões teórico-conceituais visando à consolidação da Organização do Conhecimento como disciplina científica e domínio de conhecimento (DAHLBERG, 1993; GARCÍA MARCO, 1995; MIRANDA, 1999; BARITÉ, 2001; THELLEFSEN, THELLEFSEN, 2004; HJØRLAND 2003, 2008; SMIRAGLIA, 2014). Por definição, a Organização do Conhecimento é “um sub-domínio-chave da Ciência da Informação, que é dedicado à ordem conceitual do conhecimento” (SMIRAGLIA, 2011, p. 1, tradução nossa); “um domínio inter e multidisciplinar indispensável para as atividades científicas, que também é praticada dentro da moderna Ciência da Informação” (BABIK, 2014, p. 328, tradução nossa) ou, em uma perspectiva mais ampla, “uma subdisciplina da ciência da ciência, com campos de aplicação não somente na Ciência da Informação, mas também para todos os campos de assunto (domínios) que necessitam de taxonomias (sistemas de classificação de objetos)” (DAHLBERG, 2014, p. 328, tradução nossa).

Ao configurar-se como um espaço investigativo inerentemente interdisciplinar, a Organização do Conhecimento vem se consolidando enquanto campo científico principalmente nas últimas duas décadas, a partir da criação da *International Society for Knowledge Organization - ISKO*¹ (GUIMARÃES, 2017, p. 84; 92). Tal instituição científica desempenha um papel central na promoção de um espaço privilegiado de construção e disseminação de conhecimento na temática, promovendo a interação científica no âmbito internacional e reflexões sobre “novas necessidades e tendências no campo, evidenciando experiências e soluções viáveis e sustentáveis, capazes de responder aos diferentes desafios e às oportunidades presentes” (ISKO Internacional, 2017).

Referente à consolidação da Organização do Conhecimento no Brasil, tomando como base os Programas de Pós-graduação e os Grupos de Pesquisa da Ciência da Informação,

¹ Fundada em 1989, a ISKO - *International Society for Knowledge Organization*, discute como a Organização do Conhecimento se configura e se desenvolve cientificamente, ao institucionalizar-se visando “avançar em termos teóricos, no que tange à dimensão conceitual e metodológica da organização do conhecimento e, em termos aplicados, no que se refere a seu impacto em ambientes diversos, tais como bases de dados, bibliotecas, dicionários e na própria Internet” (GUIMARÃES, 2017, p. 86).

Bräscher (2017) destaca a importância da criação do *Capítulo Brasileiro da ISKO*² em 2005 para o fortalecimento das relações entre os pesquisadores brasileiros e seus pares em outros países. Tal iniciativa possibilitou um espaço de discussão nos eventos promovidos no país, organizados por outros capítulos e na *Conferência Internacional da ISKO*, o que leva a autora a concluir que o campo está se internacionalizando e apresenta uma infraestrutura sólida para o desenvolvimento de pesquisas no contexto brasileiro.

Ao refletir sobre uma taxonomia do campo no que tange a seus principais aspectos investigativos, Gnoli (2014, p. 328) defende a adoção do termo *dimensões da Organização do Conhecimento*. Neste sentido, Guimarães (2017, p. 90) apresenta três abordagens de pesquisa que caracterizam a construção científica da Organização do Conhecimento, o que o autor cunhou *tridimensionalidade da pesquisa em organização do conhecimento*: dimensão epistemológica, dimensão aplicada e dimensão cultural:

Na *dimensão epistemológica* tem-se as bases conceituais, históricas e metodológicas da organização do conhecimento assim como seus diálogos interdisciplinares e sua produção científica, [...] [revelada por] uma especial preocupação com a constituição e a caracterização desse campo científico, assim como com seus pressupostos, metodologias, paradigmas, princípios e escolas de pensamento. [...] Na *dimensão aplicada* estudam-se os modelos, formatos, instrumentos, produtos e estruturas em organização do conhecimento, com especial ênfase nos denominados Sistemas de Organização do Conhecimento [...]. Já na *dimensão cultural* têm lugar as questões sociais, políticas, éticas, educativas e contextuais da organização do conhecimento, com especial ênfase ao papel mediador da organização do conhecimento entre distintos contextos culturais (GUIMARÃES, 2017, p. 90-91, grifo nosso).

Tais *dimensões* ou *subtemas de pesquisa* podem ser verificados nos *Capítulos da ISKO*, que embora variem de acordo com o tema central de cada evento, mantém uma base estrutural geral. Em busca de sua identidade científica, a ISKO tem apresentado três grandes tendências de pesquisa: 1) *Epistemológica*: o olhar recai sobre como a área se configura como um domínio, onde incide uma preocupação em torno de seus fundamentos teórico-metodológicos (quem somos); 2) *Tecnológica*: volta-se para as questões em torno das novas perspectivas e ferramentas de aplicação (metadados, Web Semântica, ontologias) (como

² De acordo com a página oficial (<http://www.isko.org>) são 14 os Capítulos da ISKO: Brasil; Canadá + Estados Unidos; República Popular da China; França; Alemanha + Áustria + Suíça; Índia; Irã; Itália; Magrebe (Tunísia, Argélia e Marrocos); Polônia; Cingapura; Espanha + Portugal; Reino Unido e África Ocidental.

podemos servir); e 3) *Cultural*: como estas questões da área se colocam mediante o contexto social, ou seja, quais desdobramentos surgem nesse processo mediador da cultura de quem produz, quem organiza, incluindo os problemas éticos decorrentes deste processo (a quem servimos).

Em um cenário de crescente atividade e produção científica no âmbito da Organização do Conhecimento, é possível identificar alguns desafios e perspectivas futuras para o campo (GUIMARÃES, 2017, p. 91). No ano de 2014, em comemoração ao 25º Aniversário da ISKO foram levantadas algumas reflexões acerca das mudanças dos processos de organização do conhecimento e quais apontamentos poderiam ser ponderados em relação ao seu estabelecimento como campo científico nos próximos anos.

Para Babik (2014, p. 330; 331), durante os últimos 25 anos foi possível observar tendências explícitas de pesquisas voltadas para a automação, a globalização e a socialização dos processos de informação e de criação de conhecimento. Entretanto, na visão do autor, a Organização do Conhecimento deve ser orientada para aspectos mais humanos, considerando ser este seu maior desafio para os próximos 25 anos. Ainda para o autor, a Organização do Conhecimento possui os requisitos necessários para auxiliar na recuperação do controle sobre um mundo caótico de informações, literal e metaforicamente, especialmente quando suportado pelo conhecimento e ecologia da informação. De forma complementar, Ohly (2014, p. 328) aponta a necessidade de mais conexões do campo com abordagens computacionais e destaca a importância em se considerar os aspectos sociais dos processos de organização do conhecimento.

As principais abordagens teóricas da dimensão aplicada na Organização do Conhecimento residem na *tríade* do Tratamento da Informação, representada pelas abordagens teóricas da *Catálogoação*, *Indexação* e *Classificação* e os denominados Sistemas de Organização do Conhecimento (Listas de cabeçalhos de assunto e terminologias, Tesouros, Taxonomias, Esquemas de classificação, Ontologias, etc.). Estes definidos como ferramentas que possibilitam categorizar recursos informacionais de acordo com um esquema de organização determinado para facilitar sua posterior recuperação (MORALES DEL CASTILLO, 2011, p. 89).

Neste cenário, os maiores desafios da dimensão aplicada da Organização do Conhecimento centram-se em “[...] evitar o lixo informacional, em grande parte decorrente

do contexto tecnológico que permite, no menor tempo possível, gerenciar e identificar grandes lotes informacionais” e na “promoção cada vez mais intensa da interoperabilidade de sistemas e à promoção de uma recuperação da informação rápida, eficaz e culturalmente significativa” (GUIMARÃES, 2017, p. 92).

Historicamente e estruturalmente foi oriundo da corrente teórica da Catalogação da Biblioteconomia que a Organização do Conhecimento tem suas origens. Tomando como base uma perspectiva interdisciplinar sobre o conceito e sua utilização nos sistemas de indexação, Ohy (2014, p. 328; 330, grifo nosso) enfatiza que a importância da Organização do Conhecimento deve ir além do domínio da Catalogação, sendo mais aberto e com potencial de aplicação em outras abordagens, tais como a geração de conhecimento virtuais, dispositivos móveis, tomada de decisão e *avaliação da indexação*. Esta percepção traz uma necessidade permanente de compreensão de novas e futuras técnicas em interlocução com outros campos científicos correlacionados com a Organização do Conhecimento.

Considerando-se ser a Organização do Conhecimento um campo investigativo nuclear na Ciência da Informação ao lidar “com uma atividade mediadora essencial, que propicia a manutenção de um contínuo movimento helicoidal” (GUIMARÃES, 2017, p. 93); diante de um cenário contemporâneo de alterações significativas na forma como o conhecimento é adquirido, representado, gerido e explorado em um mundo conectado e as novas funcionalidades associadas, em que questões em torno do conhecimento no mundo digital devem ser incluídas como objeto de pesquisa da Organização do Conhecimento (DAVID, 2014, p. 329); da necessidade em se “apontar para caminhos a serem explorados num ambiente propício à fertilização cruzada de distintas áreas do conhecimento num ambiente de cunho marcadamente digital (SIMÕES; BORGES, 2017, p. 13) e da importância do *conteúdo temático* dos recursos informacionais na Organização do Conhecimento (GNOLI, 2014, p. 329), destaca-se, em sua dimensão aplicada, a relevância das investigações em torno do tratamento temático da informação, mais especificamente, da abordagem teórica da indexação de assuntos, que possui os fundamentos teórico-metodológicos para propiciar o estabelecimento deste ciclo.

Como um componente de um campo mais amplo que é a Ciência da Informação, a Organização do Conhecimento possui os fundamentos teórico-práticos necessários no que tange aos problemas de representação e recuperação em sistemas de recuperação da

informação tais como os *repositórios institucionais* no contexto de *bibliotecas universitárias*. Em sua dimensão aplicada, a atividade de representação e recuperação da informação configuram aspectos nucleares, tanto na abordagem epistemológica (teórica) quanto pragmática (prática). Este é um desafio tanto teórico quanto prático da Organização do Conhecimento, uma tarefa bastante complexa, especialmente em sistemas que servem a um grupo de usuários grande e heterogêneo e lidam com grandes quantidades de informações, tal como as bibliotecas universitárias e no ambiente web (MAI, 2000, p. 270; MAI, 2011, p. 116). Neste cenário, questões contemporâneas como repositórios institucionais, repositórios de dados de pesquisa, agregação de metadados e conteúdo mais rico suscitam debates em torno do papel das bibliotecas universitárias (HERNON; MATTHEWS, 2013, p. 23).

Dentre as diversas atividades realizadas no âmbito das bibliotecas universitárias, destaca-se a indexação de assuntos, cujo propósito em repositórios institucionais é determinar o conteúdo temático dos recursos informacionais e expressá-lo em termos de indexação, possibilitando a recuperação por assuntos.

No processo de indexação de assuntos em repositórios institucionais, o uso adequado da linguagem documental é fundamental, pois possibilita a representação dos conteúdos dos recursos informacionais compatível com as solicitações de pesquisas dos usuários em uma determinada área científica especializada. Contudo, a inadequada representação documental em repositórios institucionais refletirá na qualidade da recuperação de seus recursos informacionais, visto que a qualidade da indexação de assuntos está relacionada à capacidade de reconstruir o assunto tratado em um determinado recurso informacional em conceitos para posterior recuperação por assuntos pelos usuários. Tal como ocorre em outros sistemas de recuperação da informação, quando não há compatibilidade entre a linguagem do repositório institucional e a de busca do usuário, “a credibilidade do sistema fica abalada, ocasionada por uma representatividade não condizente com as necessidades investigativas desses usuários” (BOCCATO, 2009, p. 71).

Nesse contexto, a abordagem teórico-metodológica da avaliação da indexação de assuntos é de grande utilidade para avaliações contínuas em uma mesma biblioteca universitária, nesse caso, em um mesmo repositório institucional (GIL LEIVA, 2008, p. 387). Considera-se que a aplicabilidade da avaliação da indexação de assuntos no contexto de bibliotecas universitárias possibilita a medição e avaliação dos aspectos intrínsecos e

extrínsecos da indexação de assuntos, obtendo-se um *feedback* da qualidade da representação da informação visando à melhoria da indexação de assuntos da instituição (REDIGOLO; DAL'EVEDOVE; FUJITA; BOCCATO, 2012) e dos índices de consistência (GIL LEIVA; RUBI; FUJITA, 2008, p. 240). Além disso, “permite inquirir tanto sobre o fazer da atividade de indexação quanto sobre os sistemas de recuperação da informação, e, ao mesmo tempo, mensurar a satisfação informacional dos usuários” (INÁCIO, 2012, p. 10).

A avaliação da indexação de assuntos consiste em uma abordagem teórico-metodológica que visa à avaliação propriamente dita da atividade da indexação de assuntos em um contexto determinado, seja este institucional ou em um sistema de recuperação da informação, podendo abarcar tanto aspectos qualitativos quanto quantitativos, dependendo da abordagem metodológica adotada.

Em busca do fomento de uma *cultura* de avaliação contínua no contexto de bibliotecas universitárias, a avaliação da indexação de assuntos possibilita a obtenção de um panorama da atuação do bibliotecário e da biblioteca universitária junto à comunidade acadêmica, permitindo a adequada reestruturação e melhoria contínua da representação e recuperação por assuntos em repositórios institucionais no âmbito acadêmico. Logo, a avaliação da indexação de assuntos é um instrumento importante para determinar a eficácia da biblioteca universitária em suprir as necessidades informacionais dos usuários e, com base em seus resultados, identificar suas limitações com o intuito de propor alternativas de melhorias que permitam superá-las. Sob a ótica de Lancaster (1993, p. 200), somente por meio de uma avaliação bem planejada é possível “determinar em que medida os usuários são bem-sucedidos, que tipos de problemas enfrentam e que tipos de erros cometem, permitindo assim que sejam adotadas medidas corretivas adequadas”.

A Organização do Conhecimento brasileiro, a maior parcela das publicações que discorrem sobre a avaliação da indexação de assuntos centra-se em outros sistemas de recuperação da informação, notadamente em catálogos online ou em aspectos bibliométricos, podendo ser citados os estudos de Pinheiro (1978); Gil Leiva, Rubi e Fujita (2008); Boccato, Fujita e Gil Leiva (2011); Redigolo, Dal'Evedove, Fujita e Boccato (2012); Inácio (2012); Fujita e Gil Leiva (2014), Tartarotti (2014); Piovezan (2015); Fujita e Piovezan, 2015; e Tartarotti, Dal'Evedove e Fujita (2017). No entanto, pesquisas que abarquem

especificamente a avaliação da indexação de assuntos em repositórios institucionais são incipientes.

Reconhecendo que o objetivo do processo de indexação é permitir a adequada recuperação por assuntos em um sistema de recuperação, tornou-se necessário verificar suas duas facetas em repositórios institucionais: o processo de indexação pela *abordagem do tratamento temático da informação* e o processo de indexação pela *abordagem da recuperação da informação*. Desse modo, em sua vertente teórica investigar seus pressupostos na Organização do Conhecimento e em sua vertente prática, verificar se as necessidades informacionais dos usuários são supridas de forma adequada, pela abordagem da recuperação. Diante deste cenário, surgem as seguintes indagações no contexto de bibliotecas universitárias, que embasaram o problema de pesquisa desta Tese:

- Como a literatura da Organização do Conhecimento retrata a relação entre a representação realizada pelo bibliotecário catalogador-indexador e a recuperação por assuntos realizada pelos usuários em repositórios institucionais?
- Em repositórios institucionais a prática da indexação de assuntos realizada pelo bibliotecário catalogador-indexador é norteada por princípios/diretrizes de indexação em consonância com a teoria?
- Em repositórios institucionais a recuperação por assuntos atende às necessidades informacionais de seus usuários tanto nas áreas científicas especializadas quanto nos indicadores científicos para a gestão universitária?
- Em repositórios institucionais os índices de recuperação estão entre 20% e 80% conforme descritos na literatura geral sobre a avaliação da indexação de assuntos em outros sistemas de recuperação da informação?

Com o advento dos repositórios institucionais em bibliotecas universitárias, identificou-se uma lacuna de pesquisas em torno da indexação e da recuperação por assuntos nestes sistemas de recuperação da informação, fatores estes que conduziram à continuação da pesquisa nesta temática. Isto posto, as principais premissas da pesquisa residem em torno da:

- Relevância da recuperação da informação condizente com as necessidades informacionais dos usuários em repositórios institucionais de bibliotecas universitárias;
- Tendência e cada vez maior visibilidade dos repositórios institucionais como sistemas de recuperação da informação não apenas no contexto de bibliotecas universitárias e, em uma perspectiva mais ampla, no contexto acadêmico no qual estão inseridas, tornando-se ainda instrumentos de gestão universitária por meio de indicadores científicos;
- Necessidade de investigações teórico-metodológicas em torno do processo de indexação e da recuperação por assuntos em repositórios institucionais, respaldados por uma política de indexação própria e única que norteie as atividades documentais das bibliotecas universitárias no âmbito do Conselho de Reitores das Universidades Estaduais Paulistas (CRUESP).

1.2.1 Problema e hipótese

Partindo-se da relevância da recuperação da informação aos usuários no âmbito acadêmico, pesquisas que visam contribuir para a resolução de problemas relacionados à indexação de assuntos na Organização do Conhecimento são fundamentais. Diante desta assertiva, emerge como **problema** de pesquisa a necessidade de verificação/avaliação da indexação de assuntos na representação e recuperação da informação científica especializada em repositórios institucionais como sistemas de recuperação da informação no contexto de bibliotecas universitárias. A **hipótese** da pesquisa é de que as investigações em torno da avaliação da recuperação por assuntos em repositórios institucionais demonstrariam a necessidade de adequação destes contemporâneos sistemas de recuperação da informação, tanto na vertente do tratamento temático da informação realizado pelo bibliotecário catalogador-indexador quanto na busca e recuperação da informação realizada pelos usuários.

1.2.2 Tese e proposição

A **tese** da pesquisa é a existência de problemas na recuperação por assuntos em repositórios institucionais de bibliotecas universitárias pela ausência de formulação e aplicabilidade de diretrizes de indexação nestes sistemas de recuperação da informação. Diante desta assertiva, a **proposição** da pesquisa foi investigar/avaliar o processo de indexação de assuntos em repositórios institucionais e sua influência na recuperação da informação pelos usuários, visando o aprimoramento do tratamento temático da informação documental e da recuperação por assuntos nestes sistemas de recuperação tanto no contexto de bibliotecas universitárias quanto no âmbito acadêmico.

1.2.3 Objetivos

O **objetivo geral** da pesquisa foi contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento. No intuito de atingir o objetivo geral, para o desenvolvimento da pesquisa delinear-se os seguintes **objetivos específicos**:

1. Realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias;
2. Investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como *metodologia qualitativa* com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores, com coleta de documentação;
3. Verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de

bibliotecas universitárias, por meio da aplicação da avaliação da indexação de assuntos como *metodologia quantitativa*; e

4. Sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais, visando elaborar um modelo pertinente de avaliação da indexação de assuntos nestes sistemas de recuperação da informação.

1.3 Apresentação das seções

A fim de proporcionar uma melhor compreensão do mapa conceitual da investigação, apresenta-se a relação existente entre o desenvolvimento e a delimitação da Tese, em suas oito seções (Quadro 1):

Quadro 1. Sistematização da pesquisa

Estrutura	Delimitação
Seção 1	<p>Problema: necessidade de verificação/avaliação da indexação de assuntos na representação e recuperação da informação científica especializada em repositórios institucionais como sistemas de recuperação da informação no contexto de bibliotecas universitárias, considerando-se a relevância da recuperação da informação aos usuários e pesquisadores no âmbito acadêmico.</p> <p>Proposição: investigar/avaliar o processo de indexação de assuntos em repositórios institucionais e sua influência na recuperação da informação pelos usuários, visando o aprimoramento do tratamento temático da informação documental e da recuperação por assuntos nestes sistemas de recuperação tanto no contexto de bibliotecas universitárias quanto no âmbito acadêmico.</p> <p>Objetivo geral: contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento.</p>
Seção 2 Seção 3	<p>Objetivo específico 1: Realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias.</p> <p>Título: OS REPOSITÓRIOS INSTITUCIONAIS COMO SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO NO CONTEXTO DE BIBLIOTECAS UNIVERSITÁRIAS.</p> <p>Título: A INDEXAÇÃO DE ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS: DA REPRESENTAÇÃO À RECUPERAÇÃO DA INFORMAÇÃO.</p>
Seção 4	<p>Objetivo específico 2: Investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como <i>metodologia qualitativa</i> com os</p>

	<p>gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores, com coleta de documentação.</p> <p>Objetivo específico 3: Verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária por meio da aplicação da avaliação da indexação de assuntos como <i>metodologia quantitativa</i>.</p> <p>Título: PROCEDIMENTOS METODOLÓGICOS.</p>
<p>Seção 5</p> <p>Seção 6</p> <p>Seção 7</p> <p>Seção 8</p>	<p>Objetivo específico 4: Sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais, visando elaborar um modelo pertinente de avaliação da indexação de assuntos nestes sistemas de recuperação da informação.</p> <p>Título: RESULTADOS E DISCUSSÃO.</p> <p>Título: SÍNTESE DOS RESULTADOS.</p> <p>Título: RECOMENDAÇÕES PARA A INDEXAÇÃO E RECUPERAÇÃO POR ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS.</p> <p>Título: CONSIDERAÇÕES FINAIS.</p>

Fonte: Elaboração própria.

Estruturalmente, a Tese apresenta-se como segue. Nesta “Introdução”, delineada na **Seção 1**, foram apresentados inicialmente os pressupostos e origem do problema, a proposição, a hipótese, a tese, o objetivo geral e os objetivos específicos da pesquisa.

Para compor os fundamentos teóricos que embasaram a pesquisa, a **Seção 2** – “Os repositórios institucionais como sistemas de recuperação da informação no contexto de bibliotecas universitárias” - discute a universidade e a importância da pesquisa e comunicação científica no âmbito acadêmico, o papel da biblioteca universitária neste cenário, o surgimento das bibliotecas digitais e, mais recentemente, dos repositórios institucionais como sistemas de recuperação da informação no âmbito das bibliotecas universitárias.

A **Seção 3** – “A indexação de assuntos em repositórios institucionais: da representação à recuperação da informação” - comporta uma revisão em torno da abordagem teórica da indexação de assuntos no âmbito da Organização do Conhecimento em uma perspectiva dicotômica no contexto de repositórios institucionais: pela representação e pela recuperação da informação. Na perspectiva da *representação*, discorre sobre os metadados no domínio bibliográfico, os principais conceitos, objetivos e princípios do processo de indexação de assuntos em repositórios institucionais e suas respectivas etapas. Na perspectiva da *recuperação*, discute, no contexto dos repositórios institucionais, alguns aspectos do papel da biblioteca universitária na recuperação da informação, a recuperação da informação e a

recuperação por assuntos e, complementarmente, os conceitos de relevância, pertinência e saliência na recuperação da informação, as principais variáveis/indicadores/métricas da recuperação por assuntos e a metodologia da avaliação da indexação de assuntos e suas respectivas abordagens.

A **Seção 4** – “Procedimentos metodológicos” - configura a trilha metodológica da pesquisa, pautados em duas abordagens: a) *teórico-aplicada*: considerações teóricas em torno dos temários abordados na pesquisa e b) *exploratório-descritiva*: entrevista de diagnóstico organizacional como *metodologia qualitativa*, com coleta de documentação; e aplicação da avaliação da indexação de assuntos como *metodologia quantitativa* em duas perspectivas: a) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema* - *Avaliação intrínseca quantitativa mediante a interconsistência* e b) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário* - *Avaliação extrínseca mediante a recuperação*.

A **Seção 5** – “Resultados e discussão” - apresenta os resultados das coletas de dados e respectivas discussões das entrevistas semiestruturadas com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais, da avaliação da indexação de assuntos pela perspectiva do *sistema* - *Avaliação intrínseca quantitativa mediante a interconsistência* e pela perspectiva do *usuário* - *Avaliação extrínseca mediante a recuperação*. Em seguida, a **Seção 6** - Síntese dos resultados - expõe, de forma sintética, os principais resultados da pesquisa.

A **Seção 7** – “Recomendações para a indexação e recuperação por assuntos em repositórios institucionais” delinea algumas recomendações para a representação e recuperação por assuntos em repositórios institucionais no contexto de bibliotecas universitárias.

Finalmente, a **Seção 8** - “Considerações finais” - sintetiza os principais pontos em torno dos temários abordados considerando-se o problema, premissas, hipótese, tese, proposição, objetivo geral e específicos e resultados obtidos na pesquisa, com algumas reflexões e contribuições teóricas, metodológicas e empíricas da mesma. Adicionalmente, apresenta-se as **Referências** utilizadas para o embasamento teórico-prático da investigação, juntamente com os **Apêndices** e **Anexos** que acompanharam o desenvolvimento da mesma.

2 OS REPOSITÓRIOS INSTITUCIONAIS COMO SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO NO CONTEXTO DE BIBLIOTECAS UNIVERSITÁRIAS

Nesta seção serão discutidos, primeiramente, alguns aspectos sobre as universidades como instituições de ensino superior e o papel das bibliotecas universitárias no sistema de comunicação científico-acadêmico. Em seguida, será abordado o advento dos repositórios institucionais como sistemas de recuperação da informação e mais especificamente os repositórios institucionais de dados de pesquisa no âmbito acadêmico.

2.1 As universidades como instituições de ensino superior

A universidade, como uma instituição, é mais jovem do que a biblioteca por pelo menos dois mil anos. No auge das sociedades gregas e romanas, a universidade não existia tal como a conhecemos atualmente. A mais antiga universidade ocidental remonta ao século XII em Paris e Bolonha e caracterizou-se por dois principais aspectos: sua organização e seu currículo. Originalmente, o próprio conceito da palavra *universidade* não se referia ao *universo* ou ao *universo do ensino*, mas à totalidade do grupo, especificamente do corpo discente. Consideradas, em alguns aspectos, *unidades de negociação coletivas*, em pouco tempo as universidades passaram a ser organizadas em faculdades, com o corpo docente à frente da responsabilidade de definição de requisitos de admissão, concessão de diplomas, gestão e conteúdo curricular. Neste contexto, a educação tinha por base as sete artes liberais: o *quadrivium* (baseado na forma) da matemática, astronomia, geometria e música; e o *trivium* (baseado na interpretação de formas) da gramática, da retórica e da lógica, estes, incorporando estudo especializado em medicina, direito ou teologia. Institucionalizadas em um período que antecede à impressão, poucos eram os livros didáticos nas universidades, visto que o ensino consistia em grande parte de palestras e recitações, por vezes copiadas, a fim de possibilitar o acesso aos discentes e a abertura de um espaço para a comercialização dos livros, cuidadosamente regulamentado pelas universidades, garantindo a disponibilidade e controle de preços com base no tamanho dos materiais. Habitualmente, o aluguel dos livros pelos discentes era uma outra possibilidade, visando reduzir a quantidade de cópias necessárias (BUDD, 1998, p. 24-25).

Neste contexto, a biblioteca universitária nasce como apoio ao universo de conhecimento acadêmico que florescia nas primeiras universidades. Tal como os livros, no início da universidade os bibliotecários eram igualmente uma raridade, considerando-se que havia pouca informação registrada disponível, com exceção de materiais de natureza religiosa. Como observa Charles Homer Haskins (1957, p. 39, tradução nossa), “no decorrer do tempo, no entanto, livros foram dados para o uso de estudantes, principalmente na forma de legados às faculdades, onde poderiam ser emprestados ou consultados no local”. Datado em 1338, o mais antigo catálogo, da *Bibliothèque de la Sorbonne*, da *Université de Paris*, listava 1722 volumes.

No século XVII, as universidades da Europa expandiam seu currículo, refletindo o aumento da amplitude do aprendizado, dos níveis de alfabetização e do interesse pela erudição, incentivados pelo negócio da impressão, empreendimento cada vez mais amplo e próspero. Tais universidades forneceriam um modelo para os primeiros passos em direção ao ensino superior na América colonial (BUDD, 1998, p. 27), e, posteriormente, para a América Latina, incluindo o Brasil. A partir deste período, as bibliotecas cresceram em propriedades e importância para as universidades, paralelamente ao negócio da impressão, que garantia a pronta disponibilização dos livros, estimulava o crescimento de publicações aos autores sobre assuntos especializados/seculares e permitia a *comunicação generalizada de ideias* que seriam de interesse tanto para docentes quanto para discentes. Ademais, mais títulos eram impressos do que poderia ser incluído no próprio currículo, fortalecendo o papel da biblioteca universitária como uma importante fonte de leitura complementar individual. Neste cenário, a proliferação de livros impressos levou as universidades a assumirem a responsabilidade de promover o crescimento das bibliotecas universitárias (BUDD, 1998, p. 26).

Nas palavras de Budd (1998, p. 80, tradução nossa), “como aconteceu com todos os tipos de comunicação, a erudição poderia, de repente, ser produzida em massa em centenas, até milhares, de cópias”. Em um movimento de contínua expansão, o advento da impressão levou a comunicação e divulgação científica a um patamar nunca antes alcançado, acelerando a pesquisa científica e estabelecendo a base para a comunicação acadêmica na contemporaneidade, que vem se caracterizando como uma *rede complexa de conexões* entre acadêmicos e pesquisadores, autores e leitores, publicações formais e bibliotecas.

O expansivo crescimento tanto da comunidade quanto das atividades científicas possibilitou, em 1665, o surgimento dos primeiros periódicos científicos: o *Journal des Sçavans* e o *The Philosophical Transactions of the Royal Society*. Compostos principalmente por breves resumos de estudos publicados, foram pioneiros no estabelecimento do tipo de material mais utilizado ainda hoje no meio científico para a comunicação e atualização das pesquisas na universidade. Na visão de Budd (1998, p. 80), tal avanço teve dois principais efeitos: o mundo foi alertado para a existência do trabalho principal através de revisões e pareceres e os pareceristas poderiam acrescentar comentários a sua própria pesquisa por meio de observações e questionamentos.

Em continuidade, o final do século XIX presenciou a expansão das universidades em todo o mundo, aumento do corpo estudantil, diversificação dos currículos, institucionalização da profissão de cientista e conseqüente surgimento das primeiras associações científicas³. De forma geral, a história do ensino superior aponta que a pesquisa se tornou um dos pilares da universidade, dado, além de questões políticas, o despertar para a riqueza das áreas potenciais de pesquisa, reforçando ainda mais a necessidade de divulgação dos resultados oriundos da atividade científica (BUDD, 1998, p. 81) e o acesso às demais pesquisas, em um movimento contínuo de retroalimentação do conhecimento científico produzido.

No viés da produção científica, a dupla missão da universidade, caracterizada pelo ensino e pesquisa, pressionava o corpo docente para a comunicação de seus resultados de seus trabalhos, com base no conceito de autoria, que imprime tanto ao autor quanto à universidade na qual está vinculado, critérios subjetivos de confiabilidade e prestígio acadêmico. No despertar do século XX, a comunicação científica se moldava com as características de um *sistema de comunicação científica*, com diversos atores componentes interdependentes, em que os objetivos de um estão relacionados aos objetivos do outro (BUDD, 1998, p. 84; 85).

Neste contexto, o papel do bibliotecário acadêmico centra-se nos elementos primários do sistema de comunicação científico-acadêmico, em dois principais aspectos: no incentivo à produção de pesquisa pelo corpo docente e os meios de comunicação disponíveis para o

³ No campo da Biblioteconomia, a *American Library Association* (ALA) institucionaliza-se em 1876 como a primeira associação científica, paralelamente à criação do primeiro periódico científico: *The American Library Journal*.

compartilhamento destes resultados; elementos claramente interdependentes. Aqui, a *comunicação acadêmica é um sistema de sistemas*, com base na existência de inúmeros subsistemas na estrutura da comunicação científica que se inter-relacionam. Dentre outros fatores que fortalecem a natureza sistêmica da comunicação acadêmica, destaca-se as diferenças inerentes às disciplinas: propósito, comportamento dos pesquisadores e assunto pesquisado, que também terão padrões de comunicação diferentes, já que “problemas ou questões diferentes requerem métodos diferentes, fundamentos teóricos diferentes e diferentes modos de investigação” (BUDD, 1998, p. 84; 85, tradução nossa).

Embora cada disciplina possua determinadas particularidades no que se refere à sua configuração no campo científico, de forma geral o processo de comunicação científica está atrelado ao registro dos resultados de pesquisas, em que cada pesquisador acrescenta novos conhecimentos, pensamentos e ideias aos até então trabalhados existentes, de forma a avançar o saber científico. Como parte do processo de comunicação científica, a divulgação e disseminação da produção das ciências alimenta o desenvolvimento do saber científico e, naturalmente, contribui para a construção de novos conhecimentos (DAL'EVEDOVE; FUJITA; TARTAROTTI, 2013). Com efeito, “cada descoberta apoia-se em estudos e pesquisas efetuados no passado e serve de base para conhecimentos futuros, constituindo-se tanto no produto das atividades científicas quanto no insumo para novas investigações, num processo espiral que nunca acaba” (MESQUITA; STUMPF, 2004, p. 262).

2.2 As bibliotecas universitárias e seu papel no sistema de comunicação científico-acadêmico

Embora algumas das primeiras bibliotecas, como as de Alexandria e Pérgamo fossem sedes de aprendizado e erudição, em seus primórdios a existência da biblioteca estava disassociada de uma missão educacional, caracterizando-se pela ausência de organização formal para discentes e docentes (BUDD, 1998, p. 24). Na contemporaneidade, as bibliotecas universitárias estão inseridas em um sistema de informação que é parte de um sistema mais amplo, denominado sistema de informação acadêmico, em que a geração de conhecimentos é o objeto da vida universitária (FUJITA, 2005, p. 98), desempenhando um papel estratégico na aquisição, organização, preservação e disseminação dos recursos informacionais à

comunidade acadêmica da universidade a qual pertence, em consonância com a missão da instituição.

De acordo com seus objetivos e metas organizacionais, é fundamental a participação das bibliotecas universitárias nas ações estratégicas da universidade onde estão inseridas, “da qual dependem e estão encaminhadas a incrementar e diversificar a formação de profissionais, docentes e pesquisadores que se vinculam com as necessidades da sociedade, criando e renovando o fazer científico que demanda uma nação”. Ademais, o objetivo principal das universidades é “formar profissionais, docentes, pesquisadores e técnicos que se vinculam com as necessidades da sociedade, e gerem e renovem os conhecimentos científicos, tecnológicos, humanísticos e artísticos que requer o país” (SÁNCHEZ-AMBRIZ, 2011, p. 100-101, tradução nossa).

Inseridas no contexto universitário, ao longo de sua história as bibliotecas universitárias se adaptaram às mudanças no contexto acadêmico onde estão inseridas: expansão para atender ao crescente número de discentes e docentes; adoção de estratégias para adquirir e posteriormente, fornecer acesso à quantidade crescente de informações produzidas; adaptação aos recursos financeiros disponíveis, dificultando a escolha de serviços e aquisições; inovações tecnológicas que afetaram a produção, o armazenamento e a recuperação de informações; incorporação da informação eletrônica nas atividades tradicionais; e busca de automação em larga escala, que somadas transformaram sua organização, serviços, equipes e usuários (BUDD, 1998, p. 45).

Como observa Thomas Friedman, no livro *“O mundo é plano: uma breve história do século XXI”*, cuja ênfase recai na questão da globalização, “nunca houve tanta gente com habilidade de achar, elas mesmas, tanta informação sobre tantas coisas e sobre outras pessoas”. Nessa ótica, a biblioteca universitária tem experimentado um processo de mudança de foco nas últimas quatro décadas, acompanhando não apenas as inevitáveis mudanças nos currículos e interesses de pesquisa das universidades como também as mudanças tecnológicas na forma de acesso às informações por seus usuários, no formato mais acessível e de menor custo disponível. Embora tais fatores sempre têm influenciado as bibliotecas universitárias ao longo de sua história, seu papel principal permanece o mesmo: o de suprir a necessidade informacional a seus usuários.

No tocante à história dos recursos informacionais tradicionais das bibliotecas universitárias, o acelerado crescimento das universidades em meados do século XX e a expansão dos currículos levaram as bibliotecas universitárias à missão de atender às necessidades de seus usuários reais e potenciais por meio do modelo de desenvolvimento de coleção *just-in-case*. Tal modelo atribuiu a especialistas a responsabilidade de identificar recursos informacionais para aquisição nos diversos domínios científicos, com o objetivo de que os usuários identificassem tais recursos informacionais por meio do catálogo da biblioteca e os utilizassem. Especialistas mantinham os docentes informados sobre as novas aquisições em seus respectivos domínios de conhecimento e bibliotecários de referência proporcionavam o suporte necessário visando o uso da coleção. Todavia, desde a década de 1990, as bibliotecas universitárias têm se concentrado cada vez mais em prover serviços e competência informacional a seus usuários, em detrimento dos orçamentos de aquisições cada vez mais reduzidos, bem como da quantidade decrescente de espaço disponível para o armazenamento de coleções (CAMINITA, 2014, p. 2).

Neste cenário, o modelo tradicional de desenvolvimento de coleções vem se tornando cada vez mais insustentável à medida em que os orçamentos destinados a esse fim e o poder de compra diminuam substancialmente ao longo do tempo e em que estudos apontam que grande parte dos recursos informacionais adquiridos neste modelo de negócio nunca circularam. Somado a isso, os usuários se tornaram mais propensos à busca imediata de informações com base no modelo de acesso fornecido pelo *Google* e por outros mecanismos de busca da Internet, tornando-se também menos pacientes ao acessar informações em formato impresso em bibliotecas universitárias (CAMINITA, 2014, p. 2).

Considerando-se o nível de incerteza que o futuro reserva às bibliotecas universitárias, conforme apontam Hernon e Matthews (2013, p. 2, tradução nossa), “nenhuma biblioteca é uma ilha, mas parte de um conjunto muito maior e complexo de relações organizacionais que existe dentro de uma comunidade [acadêmica]”. Com exceção de bibliotecas especializadas que fornecem acesso a recursos informacionais raros, caros e únicos, cuja escassez de materiais e informações permitiu-lhes desenvolver ferramentas para controlar e organizar o acesso à coleção, em outras bibliotecas universitárias os papéis e responsabilidades não são tão claros e inequívocos. Neste contexto,

[...] os papéis e responsabilidades da biblioteca acadêmica [...] de hoje diferem dos de suas antecessoras, da mesma forma, os papéis e

responsabilidades das bibliotecas de amanhã provavelmente serão diferentes. A quantidade e a velocidade da mudança significam que a fundação da biblioteca está desmoronando. As fundações e estruturais físicas das bibliotecas, construídas ao longo do tempo, estão passando por mudanças dramáticas. Para pensar realisticamente sobre o futuro das bibliotecas, devemos desafiar as suposições, crenças e até a visão de mundo que dominaram por muitos anos. A ordem e a estrutura dos recursos informacionais encontrados nas bibliotecas foram úteis para muitas pessoas por mais de cem anos, mas o ambiente da informação mudou radicalmente e provavelmente continuará a fazê-lo. A Internet trouxe o caos e uma infinidade de opções (HERNON; MATTHEWS, 2013, p. 2, tradução nossa).

Para lidar com a abundância de informações disponíveis na contemporaneidade, diversas ferramentas foram desenvolvidas: mecanismos de busca (ex. *Google*); compartilhamento de informações e outros conteúdos (sites, blogs, podcasts, imagens e vídeos); construção de comunidades (Facebook); compra de bens e serviços (Ebay e Amazon); *download* de conteúdo da Apple (iTunes), streaming e vídeos (Netflix) e aplicativos para uma infinidade de propósitos (HERNON; MATTHEWS, 2013, p. 3).

As bibliotecas universitárias não têm um propósito independente, ou seja, suas funções devem refletir a missão da instituição na qual estão inseridas. De forma geral, as bibliotecas universitárias tendem a direcionar seus produtos e serviços a programas de graduação e pós-graduação, pois, à medida em que a pesquisa científica e publicação de seu corpo docente tornan-se cada vez mais relevantes para as universidades, as bibliotecas universitárias fornecem acesso a recursos informacionais apropriados. Entretanto, a maior parte do financiamento universitário apoia o ensino (RUBIN, 2016, p. 137).

Rubin (2016, p. 139) aponta algumas questões que afetam o ensino superior e as bibliotecas universitárias: crescentes custos de publicações acadêmicas e interesse em recursos informacionais de acesso aberto e depósitos digitais locais; foco na retenção de discentes para preservar e proteger a estabilidade da instituição; ênfase no acesso a coleções de recursos informacionais digitais, incluindo e-books e acesso remoto através de uma variedade de dispositivos digitais; aumento dos custos e complicações legais decorrentes das questões de direitos autorais; enfoque decrescente nos recursos informacionais impressos tal como livros e periódicos, com preocupação concomitante de que grandes coleções impressas são subutilizadas; uso decrescente da tradicional função de *referência de mesa*; expectativa de que bibliotecas universitárias demonstrem que seu valor é maior que seu custo;

crescimento de ensino à distância, criando novas demandas de acesso online a recursos informacionais, com necessidade concomitante de monitoramento de direitos autorais e direitos de propriedade intelectual; expectativas de mudança do papel do bibliotecário e maior ênfase no fornecimento de serviços personalizados para docentes, departamentos e discentes e maiores colaborações com agências e instituições externas para melhorar a produtividade e reduzir a redundância de serviços e coleções.

De acordo com Hernon e Matthews (2013, p. 5), as bibliotecas universitárias enfrentam quatro crises ao mesmo tempo: uma *crise financeira* causada pela proliferação de recursos informacionais de diversos tipos; uma *crise espacial* desencadeada pela produção maciça de impressão; uma *crise de uso* causada pela transformação nos hábitos de trabalho dos acadêmicos; uma *crise de acessibilidade* provocada por mudanças na ecologia maior de textos e leitura. Estes e outros fatores levam os bibliotecários a repensar sua atuação profissional no âmbito acadêmico, reforçando seu papel mais amplo de transição para posições com múltiplas responsabilidades, envolvendo-se ativamente nos processos de pesquisa-ensino-aprendizagem (JOHNSON, 2014, p. 45).

A despeito destas mudanças, juntamente com o crescimento e expansão da Internet, as bibliotecas universitárias passam por uma evolução também em sua coleção, levando ao desenvolvimento de um novo tipo de biblioteca: a biblioteca digital, que, conforme Gonçalves (2011, p. 711), no contexto mais amplo de Recuperação de Informação (IR) constituem um dos tipos de sistemas mais avançados e complexos. As bibliotecas digitais se tornaram a base para o desenvolvimento de outro sistema de recuperação da informação na contemporaneidade: os repositórios institucionais.

2.3 O advento dos repositórios institucionais como sistemas de recuperação da informação no âmbito acadêmico

A gestão da organização da informação em bibliotecas universitárias enfrenta novos desafios pela dimensão contemporânea dos ambientes digitais e pela possibilidade de ampliação na divulgação do conhecimento produzido pela universidade por meio de novas tecnologias e metodologias. Neste cenário, emergem as bibliotecas digitais universitárias, propondo a formação e o desenvolvimento de coleções eletrônicas inteiramente produzidas

em ambiente acadêmico (FUJITA, 2005, p. 98). De acordo com Rubin (2016, p. 145, tradução nossa), “à medida em que a relação entre editores e bibliotecas universitárias se desgastava, as universidades e seu corpo docente começaram a considerar a agregação de conteúdo acadêmico em nível universitário ou multiuniversitário”. Tais discussões levaram à criação de repositórios digitais, “coleções digitais que capturam e preservam a produção intelectual de uma comunidade uni ou multiuniversitária” (SPARC, 2002, p. 4).

Como explica Zick (2009, p. 689) as bibliotecas digitais começaram em sua forma mais primitiva na década de 1990, com coleções de fotos postadas em páginas da web locais, caracterizadas pela ausência de padrões para criação sem possibilidade de recuperação. Dada a grande variedade de materiais digitais - fotos, imagens, documentos, livros -, poucas dessas coleções eram sustentáveis ao longo do tempo. Com o desenvolvimento de padrões de metadados, de softwares e de mecanismos de busca cada vez mais sofisticados, “esses problemas foram logo resolvidos e as bibliotecas digitais cresceram rapidamente na primeira década do século XXI, particularmente em resposta ao crescimento das oportunidades de ensino à distância nas universidades”, possibilitando o acesso pelos usuários a uma biblioteca virtual, para suprir a ausência de uma biblioteca física. Neste ambiente eletrônico, o formato da coleção incluiria textos, gráficos, vídeos, áudios, imagens, conjuntos de dados e softwares.

As bibliotecas digitais visam fornecer um serviço democrático e igualitário, geralmente gratuito ou a baixo custo, e sua organização e estrutura devem ser semelhantes aos bancos de dados tradicionais de bibliotecas. Quanto ao conteúdo, variam entre sistemas de arquivos a sistemas de armazenamento distribuído e possuem outras duas características: a indexação e os metadados suportam as funções de pesquisa, juntamente com os sistemas de localização de objetos digitais e as interfaces de usuário permitem a realização de pesquisas, navegação, visualização e entrega. As quatro características importantes das bibliotecas digitais são (nessa ordem): 1) os papéis sociais que desempenham; 2) as comunidades a quem servem; 3) as coleções que coletam para essas comunidades; 4) e as tecnologias que as suportam. Enquanto os papéis sociais e as comunidades são mais propensos a se manterem ao longo do tempo, as coleções e as tecnologias são mais propensas a mudanças (CALHOUN, 2014, p. 18).

Em decorrência do desenvolvimento das bibliotecas digitais universitárias, nos últimos anos surgem os *Repositórios Institucionais* (RIs), visando à melhoria da organização e disseminação da produção científica das universidades. Embora compartilhem de objetivos

comuns (armazenar, facilitar o acesso e disseminar informações), as bibliotecas digitais e os repositórios institucionais possuem enfoques distintos: enquanto a biblioteca digital visa à criação, à seleção e ao tratamento das informações para disponibilizá-las para o público em geral, o repositório institucional possibilita a visibilidade da propriedade intelectual de instituições e/ou comunidades (CAMARGO; VIDOTTI, 2011, p. 55).

Para Jones (2007, p. 6), tanto a biblioteca digital quanto o repositório institucional representam uma coleção de obras em que o conteúdo possui links identificáveis que agregam valor ao todo. Neste conceito comum de “coleção”, as competências bibliotecárias em gerenciar coleções impressas podem ser transferidas e aprimoradas para coleções digitais. As principais diferenças entre esses tipos de repositório podem ser identificadas como a expectativa de quem irá preencher o sistema e o escopo e propósito da coleção. Enquanto nas bibliotecas digitais em grande parte as informações são adicionadas por bibliotecários especializados com um interesse profissional em descrever com precisão e consistência os recursos informacionais visando à recuperação, nos repositórios institucionais, em geral, os metadados são inseridos pelo próprio autor (autoarquivamento/autodepósito). Tais diferenças influenciam diretamente na dinâmica de entrada dos recursos informacionais nestes sistemas de recuperação e na qualidade e validação dos dados.

De forma geral, os repositórios armazenam coleções de objetos digitais e fornecem métodos básicos de depósito e recuperação, além de, em muitos casos, fornecerem recursos adicionais como segurança e um protocolo para acesso remoto e distribuído. Neste contexto, os repositórios institucionais são responsáveis por armazenar e reunir toda a produção intelectual de uma determinada instituição ou consórcio de instituições tal como as universidades, visando à preservação, acesso e distribuição ao longo prazo (GONÇALVES, 2011, p. 719).

Conceitualmente, de acordo com Leite (2009, p. 21) um repositório institucional caracteriza-se como um serviço de informação científica em ambiente digital e interoperável, que gerencia a produção intelectual de uma universidade ou instituição. Para tanto, de acordo com Dodebei (2009, p. 91), contempla atividades de reunião, armazenamento, organização, preservação, recuperação e, sobretudo, a ampla disseminação da informação científica produzida e é definido como “uma base de dados digital e virtual [...] de caráter coletivo e cumulativo (memória da instituição), de acesso aberto e interoperável que coleta, armazena,

dissemina e preserva digitalmente a produção intelectual da instituição”. Já Lynch (2003) define repositórios institucionais como um conjunto de serviços que a universidade oferece à comunidade acadêmica para a gestão e disseminação de materiais digitais produzidos pela própria instituição e por seus membros. Ainda para o autor, o compromisso da universidade com a gestão destes materiais é talvez mais importante do que o serviço propriamente em si.

Um repositório institucional é um reconhecimento que a vida intelectual e acadêmica das universidades será cada vez mais representada, documentada e compartilhada em formato digital, e que a responsabilidade primária das universidades é exercitar a administração destas riquezas: tanto para torná-los disponíveis quanto para preservá-los. Um repositório institucional é o meio que cada uma das universidades irá afirmar esta responsabilidade tanto para os membros de sua comunidade como para o público em geral. Este é um novo canal para estruturação da contribuição da universidade para o mundo mais amplo, como tal, convida a avaliação política e cultural dessa relação (LYNCH, 2003, tradução nossa).

No contexto científico, o conceito de acesso aberto tem por base que os resultados da pesquisa financiada por fundos públicos devem estar disponíveis gratuitamente, por exemplo, por meio do depósito do texto completo do trabalho em um repositório publicamente acessível (JONES, 2007, p. 7). Como um tipo de ambiente informacional digital, o repositório institucional possibilita “a interoperabilidade de dados, o controle e o armazenamento da produção científica, a preservação da informação a longo prazo, o *autoarquivamento* do recurso informacional, o acesso livre, a recuperação e a disseminação da informação científica”, possibilitando ainda visibilidade à produção científica possibilita e diminuição dos custos de produção (DODEBEI, 2009, p. 59, alteração nossa). Ademais, “centralizam, preservam e tornam acessível o capital intelectual da instituição”, e constituem parte “de um sistema global de repositórios distribuídos e interoperáveis que fornecem a base para um novo modelo desagregado de publicação acadêmica (SPARK 2002⁴, p. 6 apud RUBIN, 2016, p. 145, tradução nossa).

No âmbito acadêmico os repositórios institucionais vem se tornando importantes sistemas de recuperação da informação, na medida em que, além de armazenar e disponibilizar a produção científica e intelectual da universidade, tornam-se instrumentos de

⁴ SPARC. *The case of institutional repositories: a SPARC position paper*. Prepared by Raym Crow. Washington, DC: SPARC, 2002.

gestão e tomada de decisão em nível universitário por meio de indicadores científicos, diferenciando-se dos catálogos online de bibliotecas universitárias, que, de forma geral, possibilitam apenas a recuperação dos recursos informacionais e indicam a localização dos mesmos no acervo físico das bibliotecas.

Segundo a ótica de Crow (2002), os repositórios institucionais: promovem um componente crítico na reforma do sistema de comunicação científica, ampliando o acesso à pesquisa; reafirmam o controle sobre o saber no ambiente universitário; aumentam a concorrência e reduzem o monopólio dos periódicos científicos; trazem maior relevância para as instituições e suas respectivas bibliotecas universitárias; e podem servir como indicadores tangíveis da qualidade de uma universidade de demonstrar a relevância científica, social e econômica de suas atividades de pesquisa, aumentando a visibilidade, o status e o valor público da instituição.

Contudo, corroboramos com a visão de Leite (2009 p. 98; 13), para quem “a construção de repositórios institucionais de acesso aberto requer esforços que precedem e vão muito além da simples instalação e configuração de um software”. Nessa perspectiva, para que os repositórios institucionais no contexto de bibliotecas universitárias desempenhem de forma satisfatória seus papéis, é imprescindível que os gestores tenham dimensão das implicações contextuais, teóricas e práticas que envolvem o seu planejamento, implementação e funcionamento no âmbito das universidades. Além disso, ressaltamos a visão de Targino, Garcia e Paiva (2014, p. 131), para quem *os repositórios institucionais devem estar vinculados às bibliotecas universitárias* “e por elas geridos, porquanto as bibliotecas constituem o segmento universitário por excelência responsável por estocar, organizar, gerir e disseminar o conhecimento produzido nas organizações”. Por meio do armazenamento e disseminação dos registros informacionais, os repositórios institucionais promovem maior acesso à informação científica e visibilidade dos trabalhos disponibilizados. Adicionalmente, a disseminação e implantação dos repositórios institucionais tem levado as instituições de pesquisa a pensar na importância do estabelecimento de políticas de informação acadêmicas (KURAMOTO, 2009, prefácio).

Quanto aos *atores* envolvidos nos repositórios institucionais como sistema de recuperação da informação, de acordo com Jones (2007, p. 14), estes podem ser classificados em quatro grandes categorias; *usuários* da informação (usuários finais); *provedores* de

informação; *mediadores* da informação; e *usuários* de metainformação (informação sobre informação). Enquanto para os três primeiros a preocupação centra-se nos diferentes aspectos do conteúdo intelectual, o último volta-se para o impacto do processo e não aos resultados. Mesmo que um ator desempenhe mais de uma responsabilidade, cada papel pode considerar diferentes perspectivas em torno dos mesmos problemas.

O primeiro grupo de atores envolvidos no repositório institucional são os *usuários finais da informação*, grupo de pessoas que precisam acessar o conteúdo do material acadêmico como apoio à pesquisa, ensino ou aprendizagem, tendo como requisito a competência informacional baseada na capacidade de identificar, localizar e ler informações relevantes e os diversos tipos de materiais existentes. Considerado um grupo de indivíduos e heterogêneo e não de organizações, possui necessidades e preferências de informações diferentes, seja pela disciplina ou estilo de aprendizagem pessoal e, mais particularmente, pela experiência de busca de informações, que afetará os métodos usados para recuperar material relevante. Cabe ressaltar que, “em um contexto de repositório institucional, a maior parte dos usuários finais do conteúdo será externa à organização, enquanto a comunidade suportada será usuário final de outros repositórios” (JONES, 2007, p. 14, tradução nossa).

Os *provedores de informação*, segundo grupo de atores envolvidos no repositório institucional, é composto por indivíduos e entidades corporativas: autores, revisores de pares, editores, bibliotecas e serviços de informação e bibliotecas nacionais, pautados em uma base comum de *disseminação da informação*, seja produzindo-a ou fornecendo os mecanismos para seu acesso. Este cenário é caracterizado por um movimento de simbiose, em que os autores dependem de um mecanismo de publicação como principal método de disseminação, fornecido pelos editores, estes dependem de pessoas e organizações para adquirirem os materiais publicados, em geral, por meio de sua biblioteca, que depende do material disponível e de suficiente qualidade para ser útil para as comunidades de usuários (JONES, 2007, p. 15).

Todo o paradigma de publicação científica é baseado nos *autores*, que produzem os resultados escritos para posterior divulgação. Em um ambiente acadêmico, a quantidade e a "qualidade" da publicação refletem a carreira dos autores, influenciados, ainda, por outros fatores como o financiamento de suas pesquisas ou os direitos autorais. De outro modo, os autores dependem de outras publicações de seus pares para coletar informações sobre um

determinado assunto (JONES, 2007, p. 15-16). Em um emaranhado complexo de reconhecimento pessoal através do trabalho intelectual, no contexto do repositório institucional os autores desempenham um papel-chave na geração do conteúdo para seu povoamento. No entanto, como alerta o autor: “uma das questões mais difíceis de superar é influenciar e persuadir esse grupo de pessoas a mudar suas atitudes e comportamentos para permitir que o repositório mantenha seu conteúdo”.

Os *revisores de pares* caracterizam-se como indivíduos especialistas reconhecidos em seu domínio que contribuem para a disseminação dos resultados produzidos pelos autores por meio da revisão por pares, visando garantir a verificação da validade científica do conteúdo intelectual do material a ser avaliado. Embora forneçam seus serviços gratuitamente, ganham reconhecimento dentro de seus domínios por sua expertise. Como a publicação é o principal processo pelo qual a experiência em um assunto é reconhecida, os revisores também podem desempenhar o papel de autores (JONES, 2007, p. 17).

Os *editores* são “as organizações que obtêm a produção intelectual escrita de um autor e a transformam em uma mercadoria de informação que é adquirida por, ou em nome de, usuários finais”. O termo *editor* abrange uma ampla gama de organizações, com uma variedade de razões para realizar o processo de publicação, comercialmente ou por acesso aberto. No contexto do repositório institucional, o compromisso de um editor com a publicação de acesso aberto ditará sua posição quanto ao tipo de conteúdo que gostariam de ver mantido pelos repositórios (JONES, 2007, p. 17, tradução nossa).

As *bibliotecas e serviços de informação* são as organizações que apoiam a pesquisa, o ensino e a aprendizagem e possibilitam o acesso aos recursos informacionais aos usuários. Considerando-se que os recursos eletrônicos se tornaram mais comuns, seu papel é garantir sua autorização e autenticação. Como as bibliotecas são os principais consumidores da publicação acadêmica, as mudanças de paradigma terão impacto no serviço prestado e nos recursos e competências de seus colaboradores. Em repositórios institucionais, é provável que a expertise da equipe da biblioteca seja reconhecida e usada para garantir que o conteúdo seja verificado e validado. Ao prover recursos informacionais no contexto da universidade, a biblioteca reitera seu papel-chave na gestão de repositórios institucionais (JONES, 2007, p. 18; 19).

O terceiro grupo de atores envolvidos no repositório institucional são os *mediadores da informação*, organizações ou indivíduos que contribuem para o fornecimento de conteúdo e de métodos alternativos de localização: agregadores de conteúdo e motores de busca. Os *agregadores de conteúdo* são organizações que se sobrepõem a editores específicos para fornecer recursos informacionais de forma agregada e mais ampla do que de um assunto específico, como o fornecido por periódicos específicos. Utilizados pelos usuários finais como uma ferramenta de localização para conteúdo que ainda não descobriram, mas acreditam estar disponível, a maioria dessas ferramentas não possui conteúdo de texto completo, mas aponta para o conteúdo em outras formas de material. Podem, ainda, agregar valor adicional por meio de indicadores científicos. Como muitos desses serviços são pagos, geralmente são mediados pela biblioteca da universidade (JONES, 2007, p. 20).

Os *motores de busca* possibilitam a localização de recursos informacionais na Internet e seguem desenvolvimentos como o *Google Acadêmico* e as mudanças na forma como os editores disponibilizam seus periódicos eletronicamente. São frequentemente o ponto de partida para muitos usuários na busca de recursos informacionais acadêmicos, visto que “o desenvolvimento de conteúdo aberto e acessível a partir de repositórios institucionais ampliou os resultados potenciais para os mecanismos de busca”.

Por fim, o quarto grupo de atores envolvidos no repositório institucional são os *usuários de metainformação* (informação sobre informação), representados pelos financiadores, instituições e organismos nacionais. Este grupo centra-se no uso de informações sobre o conteúdo e não no conteúdo em si, para processos internos de gestão. Com grande influência, tendem a ser detentores de orçamento suficiente para apoiar a pesquisa, o ensino e o aprendizado. Os *financiadores*, a exemplo das agências de fomento, fornecem apoio financeiro a instituições, projetos e programas para ampliar o conhecimento em domínios específicos, informações que, mais recentemente, têm sido incorporadas aos trabalhos publicados dos autores. Nos repositórios institucionais tal informação tem sido indexada para a recuperação (JONES, 2007, p. 21;22).

As *instituições* referem-se ao órgão que apoia financeiramente autores, leitores e serviços de biblioteca e informação. O desempenho acadêmico da universidade pode ser medido por níveis e qualidade dos resultados de pesquisa, podendo afetar o financiamento externo e a qualidade de docentes e discentes potenciais (JONES, 2007, p. 22). Nesse sentido,

o repositório institucional desempenha um importante papel na universidade visando à medição de tais índices, possibilitando o aumento da visibilidade de publicação científica produzida na instituição. Por fim, os *organismos nacionais* (BRASIL - IBICIT), visam fornecer liderança para o uso inovador de informações em um determinado país, sendo referência e influenciando o modo como as instituições respondem ao ambiente de informações em constante mudança (JONES, 2007, p. 23).

De forma geral, quanto aos *tipos de recursos informacionais* abarcados em um repositório institucional, Ware (2004) cita os *preprints*, *post-prints*, conjuntos de dados, teses e dissertações eletrônicas, objetos de ensino-aprendizagem e relatórios técnicos. A produção de diferentes tipos de resultados de pesquisa em uma universidade pode abarcar “desde dados coletados de um experimento, passando por software até publicações escritas e apresentações pontuais nas artes”. Enquanto em algumas disciplinas o acesso ao registro da pesquisa é um processo direto e bem compreendido, em outras é mais complexo (JONES, 2007, p. 91).

Apesar de seu potencial e importância, existe uma variedade de fatores que inibem a adoção mais ampla de repositórios institucionais no âmbito acadêmico pelo corpo docente, como apontadas por Rubin (2016, p. 145) com base em Quinn (2010)⁵: resistência por parte do corpo docente no preparo do recurso informacional para submissão no repositório institucional; falta de clareza sobre as vantagens de participar de um repositório; receio de que o trabalho depositado seja plagiado; falta de conhecimento técnico sobre como depositar; receio de violar restrições de direitos autorais do periódico; receio de que o depósito antecipado desqualifique o trabalho para publicação em outro lugar, dentre outras. No entanto, considerando-se os crescentes custos de acesso a periódicos eletrônicos ou impressos aliado ao interesse da comunidade acadêmica pelo acesso aberto ao conhecimento, “os repositórios institucionais podem se tornar um componente integral das bibliotecas universitárias nos próximos anos” (RUBIN, 2016, p. 145, tradução nossa). Certamente, para o sucesso do repositório institucional nas universidades, tais desafios, dentre outros, precisam ser considerados.

⁵ QUINN, B. Reducing psychological resistance to digital repositories. *Information Technology and Libraries*, jun. 2010, p. 67-75.

Na trilha do desenvolvimento dos repositórios institucionais, no âmbito acadêmico emerge um novo tipo de repositórios institucionais: os *repositórios institucionais de dados de pesquisa*.

2.4 Os repositórios institucionais de dados de pesquisa

No contexto acadêmico, os repositórios institucionais de dados de pesquisa surgem para apoiar a produção, armazenamento, uso, reuso, acesso e compartilhamento dos conjuntos de dados gerados durante as várias etapas do processo da pesquisa científica. Por parte dos editores científicos, há uma tendência “em exigirem que os dados subjacentes das publicações sejam preserváveis, detectáveis e acessíveis em um repositório, com links recíprocos entre o artigo publicado originalmente e os dados”. Os repositórios institucionais de dados de pesquisa abrangem os *conjuntos de dados de pesquisa*, “coletados como parte de um projeto de pesquisa, descritos e disponibilizados para uso por outros pesquisadores”, que não se limita à descoberta de informações e acesso, exigindo um envolvimento mais profundo com o processo de pesquisa (RICE; SOUTHALL, 2016, p. 15-16, tradução nossa). Os *dados de pesquisa* são definidos como “aquilo que é coletado, observado ou criado em formato digital, por propósitos de análise para produzir resultados originais de pesquisa⁶, tais como questionários, protocolos experimentais, cadernos de laboratório, gráficos ou informações sobre as configurações de um determinado equipamento, dentre outros.

Os repositórios institucionais de dados de pesquisa visam manter e compartilhar dados de pesquisas produzidos dentro de uma determinada universidade, diferenciando-se dos repositórios institucionais em geral, que centram-se na produção científica e, por vezes, abarcam outros tipos de produção de uma determinada instituição: artística, cultural, administrativa, intelectual, etc. Cabe ainda ressaltar que mesmo que um repositório institucional de bibliotecas universitárias incorpore também os dados de pesquisa como recursos informacionais em seu corpo de recursos informacionais, continuará sendo denominado de repositório institucional e não repositório de dados, considerando-se seu

⁶ The University of Edinburgh. *Information Services: Research Data Service*. Disponível em: <http://www.ed.ac.uk/information-services/research-support/data-library/data-repository/definitions>. Acesso em: 27 nov. 2018.

objetivo fulcral de reunir e possibilitar o acesso, uso e reuso das produções científicas, acadêmicas, artísticas e técnicas de uma determinada comunidade acadêmica que juntas conceituam o termo *produção intelectual*. Estas novas concepções atrelam-se às discussões da Organização do Conhecimento em torno dos métodos e práticas empregadas para a análise de dados de pesquisa, especialmente no ambiente web. Neste cenário, se inicialmente os serviços bibliotecários foram direcionados em meios físicos e métodos de *entrega*, com o advento das tecnologias o tema dominante tornou-se *acesso*. Concomitantemente ao surgimento dos repositórios institucionais de dados de pesquisa nas bibliotecas universitárias, emergem dois conceitos: *Biblioteconomia de dados* e *bibliotecário de dados*.

Na Biblioteconomia, tradicionalmente o conjunto das atividades de uma biblioteca universitária abrangem procedimentos que visam permitir que os recursos informacionais tanto *impressos* quanto *digitais* de uma determinada coleção possam ser organizados e disponibilizados à comunidade acadêmica.

Catálogo e organização de materiais é uma área de trabalho em andamento que forma uma fundação de grande parte da biblioteconomia. Preservação e curadoria é outra responsabilidade fundamental - especialmente quando o acesso tem que ser mantido material que é mais difícil de encontrar ou não está mais impresso. Referência e serviços ao usuário são um recurso comum na maioria das bibliotecas que se baseiam na manutenção de coleções. Workshops de consultoria e treinamento que buscam apoiar os leitores na análise de problemas, enquadrando questões de pesquisa e trabalhando com recursos informacionais de uma forma significativa também são (RICE; SOUTHAL, 2016, p. 2, tradução nossa).

Neste contexto contemporâneo de repositórios institucionais de dados de pesquisa, a perspectiva da *Biblioteconomia de dados* coloca a Organização do Conhecimento como protagonista no desenvolvimento de produtos e serviços de informação científica relacionados ao uso e consumo de dados de pesquisa no ambiente web. Nas palavras de Rice e Southall (2016, p. 2, tradução nossa), “a Biblioteconomia começa então a ser entendida não simplesmente como algo que suporta a descoberta e o acesso a publicações títulos ou recursos informacionais, mas também como algo que se envolve com a conduta de pesquisa e investigação acadêmica”.

Tomando este cenário como ponto de partida, surge, então, um novo papel para a atuação do bibliotecário - o *bibliotecário de dados*. Os bibliotecários de dados que atuam em bibliotecas universitárias tornam-se especialistas em informação científica, capazes de reunir

e assessorar os pesquisadores no desenvolvimento de coleções de conjuntos de dados de pesquisa, uso e preservação dos dados, no acesso a dados já arquivados ou na criação de novos dados, como parte de sua prática de pesquisa. Nesse sentido, “bibliotecas e seus bibliotecários de dados são idealmente colocados para atender a essas necessidades”, criando o que Rice e Southall (2016, p. 16, tradução nossa) denominam de “*um novo mapa de suporte e serviços para pesquisadores*”.

Este é apenas o ponto de partida para uma compreensão básica do que constitui biblioteconomia de dados. É uma mistura em evolução das habilidades tradicionais de biblioteconomia, como aconselhando na descoberta de recursos, por um lado, e fornecendo informações em questões específicas de formatos de dados ou obsolescência digital, por outro. Em alguns casos, pode exigir a aplicação de conhecimentos especializados detalhados sobre coleções ou análise de empreendimentos. Em outras ocasiões, pode ser necessário fornecer conselhos sobre direitos autorais, gerenciamento de dados como parte do processo de pesquisa ou as ferramentas de análise de dados mais apropriadas (RICE; SOUTHAL, 2016, p. 16, tradução nossa).

Além da atuação especializada, torna-se necessário que o bibliotecário de dados tenha familiaridade com a gama de problemas do campo da comunicação científica, considerando-se que “a demanda por suporte de pesquisa não se limita à descoberta de informações e acesso, mas exige um envolvimento mais profundo com o processo de pesquisa” (RICE; SOUTHAL, 2016, p. 16, tradução nossa).

O trabalho diário do bibliotecário de dados pode ser bastante semelhante ao dos bibliotecários universitários tradicionais. Pode envolver trabalhar dentro de sistemas de bibliotecas, adquirindo recursos e desenvolvendo relações de trabalho que permitam promover o papel de sua biblioteca. O fato de trabalharmos com dados de pesquisa ao lado de periódicos, livros e outras publicações não devem fazer muita diferença de como o nosso trabalho é visto, mas existe uma série de razões que faz. A palavra ‘dados’ em si é desanimadora para alguns bibliotecários e pesquisadores acadêmicos tradicionais e causa alguma ansiedade. Para alguns, é porque parece pertencer a outras disciplinas e ter pouca relação com seu próprio trabalho. Outros a veem como sendo uma palavra tão comum a ponto de ser quase indistinguível da ‘informação’ (RICE; SOUTHAL, 2016, p. 19, tradução nossa).

Referente a este (novo) papel do bibliotecário de dados em repositórios institucionais de dados de pesquisa, de forma geral o bibliotecário de dados apoia os pesquisadores em todos os aspectos dessa peculiar classe de informação digital - seu uso, preservação e

curadoria e na produção e utilização de outros dados de pesquisa pelos pesquisadores em volumes cada vez maiores, visando à criação de novos conhecimentos (RICE; SOUTHALL, 2016, prefácio).

Aqueles que trabalham dentro de bibliotecas fazem uma contribuição valiosa para apoiar a pesquisa e o ensino, bem como moldar o caráter e a vida intelectual de instituições individuais. Se uma universidade foca nas humanidades, nas ciências físicas, nos clássicos ou em qualquer outro número de disciplinas, o bibliotecário trabalha em última análise para apoiar a aprendizagem e a disseminação do conhecimento. Isso pode assumir muitas formas estabelecidas, mas cada vez mais existe a necessidade de suportar novos formatos de informação. Os dados digitais são um novo formulário específico. No caso de coletas de dados e criação de dados de pesquisa, isso também levou ao surgimento de um novo tipo de profissional de bibliotecas: o bibliotecário de dados. Mas até que ponto neste fato de um novo papel e de que maneiras ele difere da biblioteconomia tradicional? Seu papel é manipular, interpretar, analisar, observar ou, mais geralmente, "usar" os dados (RICE; SOUTHAL, 2016, p. 1, tradução nossa).

O bibliotecário de dados desempenha um papel importante em descobrir exatamente o que está sendo oferecido a seus pesquisadores e demonstrar o diferencial de seu repositório institucional de dados de pesquisa frente a outras opções de depósitos de dados, incentivando os pesquisadores a realizar o depósito dos dados de pesquisa no repositório de dados e criando um clima de confiança dentro da comunidade de pesquisa da instituição em relação ao repositório de dados (RICE; SOUTHALL, 2016, p. 106). Nesse sentido, o bibliotecário de dados atua "trabalhando com leitores para acessar, manipular ou compartilhar dados de pesquisa", demonstrando "capacidade de resposta às necessidades acadêmicas" (RICE; SOUTHAL, 2016, p. 2, tradução nossa).

Dentre as diversas atribuições do bibliotecário de dados no repositório institucional de dados de pesquisa, Rice e Southall (2016, p. 106) destacam sete questões-chave referentes à implantação e gestão destes contemporâneos sistemas de recuperação da informação, que também podem ser aplicados em repositórios institucionais: escopo/cobertura; a escolha do esquema de metadados; gerenciamento do acesso; a revisão da qualidade dos dados; plano de preservação digital ao longo de tempo e do espaço; promovendo repositórios digitais confiáveis e a necessidade de interoperabilidade.

Entretanto, cabe destacar os desafios deste novo profissional em torno das questões envolvidas referentes à descrição dos *metadados de dados de pesquisa* - tanto descritivos

quanto temáticos - como novos recursos informacionais e de como se relacionam com os demais recursos informacionais do repositório institucional de dados de pesquisa. De acordo com Rice e Southal (2016, p. 24) os metadados também podem desempenhar um papel importante na gestão de dados de pesquisa. No contexto dos repositórios institucionais de dados de pesquisa, os metadados configuram-se como descrições técnicas altamente estruturadas que possibilitam documentar informações sobre como os dados foram criados, originalmente usados e o relacionamento entre os arquivos (por exemplo, entre um determinado artigo de periódico e o questionário utilizado na pesquisa). Nesse caso, metadados e documentação tornam-se sinônimos.

Destaca-se os desafios da Biblioteconomia em torno das questões envolvidas referentes à descrição dos *metadados de dados de pesquisa* - tanto descritivos quanto temáticos - como novos recursos informacionais e como se relacionam com os demais recursos informacionais do repositório institucional. Nesse sentido, os repositórios institucionais e os repositórios institucionais de dados de pesquisa compartilham do mesmo desafio em torno da representação e da recuperação por assuntos nestes sistemas de recuperação da informação no âmbito acadêmico. Dentre os diversos desafios enfrentados tanto pelos repositórios institucionais de dados de pesquisa quanto os repositórios institucionais em geral, ressalta-se a questão do tratamento da informação, e, mais especificamente do tratamento temático da informação na abordagem teórica da indexação de assuntos, temário discutido a seguir.

3 A INDEXAÇÃO DE ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS: DA REPRESENTAÇÃO À RECUPERAÇÃO DA INFORMAÇÃO

No intuito de contextualizar o processo de indexação de assuntos em repositórios institucionais no âmbito de bibliotecas universitárias, a seção está estruturada em duas principais abordagens. Primeiramente, na perspectiva da indexação de assuntos pela *representação*, são apresentados os metadados no domínio bibliográfico e discutidos os principais conceitos, objetivos e princípios da indexação de assuntos, bem como as etapas subjacentes ao processo. Em um segundo momento, na perspectiva da indexação de assuntos pela *recuperação* são discutidos, inicialmente, alguns aspectos do papel da biblioteca universitária na recuperação da informação, a recuperação da informação em geral e específica de assuntos em repositórios institucionais. Na sequência, são abordados os conceitos de relevância, pertinência e saliência e as principais variáveis/indicadores/métricas da recuperação da informação e, finalmente, a metodologia da avaliação da indexação de assuntos e suas respectivas abordagens de aplicação.

3.1 A indexação de assuntos em repositórios institucionais pela abordagem da *representação da informação*

O recurso informacional, principal protocolo utilizado pelos seres humanos para comunicar informação e conhecimento, tem sido escrito no decorrer da história da humanidade em superfícies tão diversas quanto pedra, madeira, papiro, papel de arroz, em muitas formas e línguas em todos os cantos do planeta, em uma variedade de mídias, tanto impressas quanto digitais (NAVARRO; ZIVIANI, 2013, p. 187). Considerando-se o inevitável e explosivo crescimento da quantidade de informação disponível na contemporaneidade, “o desafio de organizar toda essa informação e torná-la acessível é intimidante”. Ampliado particularmente pela web, o ambiente da informação se torna mais complexo a cada ano, à medida em que a quantidade, tipos e formatos de informação proliferam, tornando cada vez maior o desafio de organizá-la e acessá-la (RUBIN, 2016, p. 339, tradução nossa).

Em uma perspectiva histórica sobre seu papel em relação aos recursos informacionais, Tarver e Phillips (2016, p. 133) afirmam que as “bibliotecas se concentraram não apenas na criação de coleções, mas na criação de um meio de descrever esses recursos informacionais

para torná-los detectáveis e úteis”. Ao longo do tempo, o processo de catalogação - ou, mais amplamente, a *criação de metadados* foi padronizado de várias maneiras estruturadas, visando à garantia da *consistência* em nível local e a capacidade de compartilhar facilmente descrições de recursos informacionais entre instituições. Na contemporaneidade,

[...] juntamente com a ascensão da Internet, as tecnologias baseadas na web permitiram a criação e publicação de informações em massa através de uma plataforma de baixa barreira de entrada - qualquer um que seja capaz de usar um editor ou capturador de texto ou imagem agora pode criar documentos ou objetos digitais e publicá-los diretamente na web. Isso democratizou enormemente a publicação e disseminação de informações e resultou em um aumento exponencial no volume e na complexidade dos recursos digitais. [...] A Internet e a web, de muitas maneiras, tornaram-se os novos catálogos de bibliotecas, indexando bancos de dados, dicionários, enciclopédias, jornais, escolas, museus, centros de entretenimento, agências de viagens, centros comerciais e muitas outras fontes e locais aos quais costumávamos fisicamente apenas acessar (ZENG; QIN, 2014⁷ apud RUBIN, 2016, p. 339, tradução nossa).

Nesta vertente, o objetivo fulcral da biblioteca universitária tem sido “adquirir, armazenar, organizar, preservar, disseminar e fornecer acesso a materiais e informações já produzidos” por sua comunidade acadêmica, por meio de sistemas de recuperação da informação tais como os catálogos online e, mais recentemente, os repositórios institucionais (RUBIN, 2016, p. 305, tradução nossa). Sob a ótica de Rubin (2016, p. 339), neste cenário de novos desafios para o controle do universo de informações em constante expansão, os *metadados* apresentam-se como a chave para este propósito, permitindo às bibliotecas universitárias desempenharem um papel fundamental na estruturação dos recursos informacionais em repositórios institucionais.

3.1.1 Metadados para a representação e recuperação da informação em repositórios institucionais

Conforme Navarro e Ziviani (2013, p. 189), “muitos documentos e coleções de textos possuem metadados associados. Os metadados contêm informações sobre a organização dos dados, seus domínios e relacionamentos, ou seja, metadados são *‘dados sobre os dados’*”, Embora este conceito seja amplamente difundido na literatura de Ciência da Informação, de

⁷ ZENG, M. L.; QIN, J. *Metadata*. 2nd ed. Chicago: ALA, 2014.

forma mais específica os metadados são definidos como “informações estruturadas que descrevem, explicam, localizam ou facilitam a recuperação, uso ou gerenciamento de um recurso informacional” (NISO 2004 p. 1, tradução nossa) ou “termos descritivos que são aplicados a recursos informacionais principalmente com o propósito de facilitar a recuperação”. Utilizados de várias formas para a descrição de tais recursos informacionais, os metadados potencialmente desempenham diferentes funções na organização do conhecimento (SMIRAGLIA, 2014, p. 65, tradução nossa).

Metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação; [...] são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação (ALVES, 2010, p. 47).

De acordo com Rubin (2016, p. 339, tradução e grifo nosso), “os metadados podem ser aplicados a um *único* recurso, a um *conjunto* de recursos ou a *partes* de um único recurso” e estão em toda parte, não necessariamente em ambientes digitais. Para Smiraglia (2014, p. 67, tradução nossa), os metadados desempenham dois papéis: “como identificadores de recursos que são artefatos do conhecimento registrado e como dispositivos de pedidos, seja para esses artefatos ou para o material conceitual que transmitem”.

Segundo Gonçalves (2011, p. 716, tradução nossa), “a propriedade fundamental que caracteriza os metadados é um relacionamento de proximidade com algum outro recurso (por exemplo, um objeto digital, coleção ou mesmo um serviço)”. Em um contexto mais amplo da recuperação da informação, a principal função dos metadados é não apenas reunir, desambiguar e apontar para os trabalhos, que são sínteses individuais de conhecimento registrado”, mas desempenharem uma função social de “reunir obras, sob suas âncoras históricas socialmente construídas [...] e disseminar a cultura” (SMIRAGLIA, 2014, p. 75, tradução nossa).

Adicionalmente, os metadados executam funções semelhantes às fornecidas por um catálogo tradicional de bibliotecas universitárias: criam pontos de acesso para os recursos informacionais e facilitam sua descoberta de pelo menos quatro maneiras: “permitindo recursos para serem encontrados por critérios relevantes; identificando recursos; distinguindo diferentes recursos e provendo informações de localização”. No geral, os

metadados permitem descrever os recursos informacionais e organizá-los de acordo com critérios definidos; agregar recursos informacionais semelhantes e fornecer caminhos para a localização da informação desejada; facilitar a troca de metadados e permitir a interoperabilidade; permitir a identificação e descrição digital para arquivamento e preservação dos recursos informacionais (NISO 2004, p. 1-2).

Sob a ótica de Smiraglia (2014, p. 75, tradução e grifo nosso), se em um primeiro momento o foco das pesquisas em torno de padrões de metadados estava no desenvolvimento, em um segundo momento voltou-se para as questões sobre como as interações compreensíveis por máquina poderiam ser compreensíveis por humanos, “evidenciando também a tão necessária contribuição da *evidência empírica* para o desenvolvimento e uso do esquema de metadados, bem como o uso de muito dos recursos informacionais”. Na evolução dos metadados e do controle bibliográfico em uma perspectiva histórica, muitos dos desenvolvimentos paralelos e entrelaçados levaram à busca atual pelo aumento da interoperabilidade (SMIRAGLIA, 2014, p. 76, grifo do autor).

Entretanto, na visão do autor, os bibliotecários continuam desempenhando um importante papel na criação de metadados, porém, como parte de um domínio muito maior de criadores de metadados, que não apenas criam mas contribuem com padrões e práticas relacionados a metadados. Atualmente no ambiente web rico e digital, outros atores - produtores de bancos de dados, criadores de sites, editores, fornecedores, organizações científicas e técnicas, desenvolvedores de software e usuários experientes da web também contribuem com a descrição de metadados.

A organização de informações sempre foi uma tarefa complexa, e uma série de ferramentas, técnicas e padrões, de códigos de catalogação, esquemas de classificação e listas de cabeçalhos de assuntos a formatos bibliográficos, tiveram que ser desenvolvidos para enfrentar esses desafios. Estes têm sido utilizados com sucesso no mundo da biblioteca para organizar, acessar e compartilhar informações (CHOWDHURY; CHOWDHURY, 2007, p. 213, tradução nossa).

Em bibliotecas universitárias, os *tradicionais* modos de organização, tais como sistemas de classificação bibliográfica, vocabulários controlados (incluindo tesouros e cabeçalhos de assuntos), catálogos online e suas regras (AACR2 e RDA) vem abrindo espaço para os *emergentes* modos de organização do conhecimento: *Requisito Funcional para Registros Bibliográficos* (FRBR), índices, resumos e bibliografias, a organização do conhecimento

utilizando-se sistemas de codificação MARC e MARC21 e a organização do conhecimento na Internet, incluindo discussões sobre catálogos da próxima geração, *Linked Open Data* e Web Semântica (RUBIN, 2016, p. 305).

Outra forma de metadados comumente utilizada na organização do conhecimento são os denominados *esquemas de metadados*, que fornecem padrões para os dados coletados, utilizados para, segundo Smiraglia (2014, p. 67), formatar outros metadados com o objetivo de povoar sistemas de recuperação. Como exemplos de esquemas de metadados utilizados no contexto de bibliotecas universitárias podem ser citados o MARC (*Machine Readable Cataloging Record*), voltado para a descrição de recursos informacionais em catálogos online e, mais recentemente o *Dublin Core*, padrão de descrição de recursos informacionais em repositórios institucionais.

No contexto das bibliotecas universitárias, o MARC é considerado o formato de metadados mais utilizado em registros bibliográficos, composto por diversos campos para os diferentes atributos de uma entrada bibliográfica, tais como título e autor (NAVARRO; ZIVIANI, 2013, p. 189). Conforme Smiraglia (2014, p. 67), um registro MARC contém pelo menos três conjuntos de dados. O *primeiro conjunto* contém dados descritivos derivados do próprio item, de acordo com as Regras de Catalogação Anglo-Americanas (campos 100 ao campos 504) e são complementados com um *segundo conjunto*, composto por descritores⁸ de assunto e números de classificação bibliográfica (campos 650 e 090), geralmente, da Classificação Decimal de Dewey. Os dados são formatados por convenções MARC que servem para facilitar a manipulação de dados por máquinas. No registro, “os campos do MARC são estruturados com *tags* de campo de três dígitos, indicadores de nível de campo de um ou dois dígitos e indicadores de subcampos para delimitar o campo”. Os demais campos, que compõem o *terceiro conjunto* de dados são chamados de campos controle, pois contêm o número do ISBN, por exemplo. Para o autor, “este é um processo muito complexo e reflete séculos de evolução na estrutura dos catálogos de bibliotecas”. Além disso, “com a tendência na descrição de recursos informacionais para usar XML como padrão para recursos baseados na web, o MARC21 foi adaptado para o padrão XML como MARCXML”.

⁸ Cunhado por Calvin Mooers, o conceito de *descriptor* é definido como o assunto de cada recurso informacional, caracterizado ou descrito por meio de um conjunto de ‘descritores’ extraído de um vocabulário formal de termos (MOOERS, 1951, p. 25).

O *Dublin Core Metadata Element Set* (DCMES) tem sido considerado, sob a concepção de Jones (2007, p. 86), a língua franca do mundo da web, em resposta às complicações e custos associados a padrões complexos como o MARC. Foi desenvolvido pelo impulso da comunidade de Ciência da Informação pela DCMI (*Dublin Core Metadata Initiative*), cujas atividades visam promover “um amplo consenso internacional aberto, neutralidade e um enfoque interdisciplinar que promove a descoberta e o gerenciamento efetivo de recursos por meio de metadados em todo o mundo” (DCMI, 2014). São diversos os indivíduos e organizações que participam da iniciativa, podendo ser citados a *Online Computer Library Center* (OCLC), “arquivos e museus, instituições educacionais, bibliotecas digitais, agências governamentais, redes, editores e gerentes de conhecimento” (RUBIN, 2016, p. 342, tradução nossa).

Basicamente, o padrão de metadados Dublin Core visa descrever recursos informacionais baseados na Internet (GONÇALVES, 2011, p. 716). Para Smiraglia (2014, p. 67, tradução nossa) o objetivo do Dublin Core é “fornecer uma melhor representação de metadados baseada na web”, sendo considerado um esquema de ‘propósito geral’, ou seja, um “sistema de metadados com elementos padronizados que possibilita a muitos indivíduos sem a experiência de um catalogador profissional descrever seus objetos digitais e atribuir termos que possam ajudar significativamente outros a acessá-los” (RUBIN, 2016, p. 341, tradução nossa). Por esse ponto de vista, no domínio bibliográfico

[...] os mesmos dados do registro MARC são representados e identificados usando convenções específicas *DC name* nos invólucros DC. As formas de organização do conhecimento são MARC ou Dublin Core (ou RDF ou outros esquemas de metadados). Quando são usados como pano de fundo para o ordenamento conceitual de registros bibliográficos, são claramente uma forma de organização do conhecimento. E por serem onipresentes na recuperação de informações, sua estrutura afeta a eficácia de outras formas de ordenação conceitual (SMIRAGLIA, 2014, p. 67, tradução nossa).

Na visão de Rice e Southall (2016, p. 15), o “padrão de metadados Dublin Core foi desenvolvido com o objetivo principal de manter os elementos de metadados simples para que os autores possam criar seus próprios metadados para os recursos que eles criam”. Na prática, como explica Gonçalves (2011, p. 716), o *Dublin Core Metadata Element Set* propõe 15 elementos que podem ser utilizados para descrever qualquer tipo de objeto digital/recurso informacional online: sete para descrever conteúdo (título, assunto, descrição, fonte, idioma, relação e cobertura), quatro para lidar com questões de propriedade intelectual (criador,

editor, contribuidor e direitos), e outros quatro para lidar com propriedades de instanciações/manifestações de objetos digitais (Quadro 2):

Quadro 2. O padrão de metadados Dublin Core

Elemento	Esquema de codificação	Definição
1 Title (Título)	<dc:title>	<i>Nome atribuído ao recurso informacional</i> (nome formalmente conhecido).
2 Creator (Criador)	<dc:creator>	<i>Entidade primariamente responsável (indivíduo ou organização) pela criação do recurso informacional.</i>
3 Subject (Assunto)	<dc:subject>	<i>Assunto do recurso informacional.</i> Geralmente, o assunto é representado por palavras-chave, frases-chave ou códigos de classificação. A melhor prática é que utilize um vocabulário controlado.
4 Description (Descrição)	<dc:description>	<i>Descrição de um recurso informacional</i> , que pode incluir, mas não está limitado a: um resumo, uma tabela de conteúdos, uma representação gráfica ou uma descrição livre do recurso informacional.
5 Publisher (Publicador)	<dc:publisher>	<i>Entidade responsável (indivíduo ou organização) pela disponibilização do recurso informacional.</i> Geralmente, o nome do publicador é utilizado para indicar uma entidade.
6 Contributor (Colaborador)	<dc:contributor>	<i>Entidade responsável (indivíduo ou organização) por contribuições em um recurso informacional.</i> Geralmente o nome de um contribuidor deve ser usado para indicar uma entidade.
7 Date (Data)	<dc:date>	<i>Momento ou período de tempo associado a um evento no ciclo de vida do recurso informacional.</i> A data pode ser utilizada para expressar informação temporal. Recomenda-se a utilização de esquemas de codificação, como o perfil <i>World Wide Web Consortium Date and Time Formats</i> (W3CDTF).
8 Type (Tipo)	<dc:type>	<i>Natureza ou o gênero do recurso informacional.</i> Recomenda-se a utilização de um vocabulário controlado como o DCMI (DCMI <i>Type Vocabulary</i>). Para descrever o formato de arquivo, meio físico, ou as dimensões do recurso informacional, utilizar o elemento 9 (Formato).
9 Format (Formato)	<dc:format>	<i>Formato de arquivo, meio físico ou as dimensões do recurso informacional.</i> Exemplos de dimensões incluem tamanho, dimensão, extensão, formato e duração. Recomenda-se o uso de vocabulário controlado pela <i>Multipurpose Internet Mail Extensions</i> (MIME).
10 Identifier (Identificador)	<dc:identifier>	<i>Referência inequívoca do recurso informacional em um determinado contexto.</i> Recomenda-se a identificação do recurso informacional relacionado em um sistema formal de identificação.
11 Source (Origem)	<dc:source>	<i>Recurso informacional relacionado a partir do recurso informacional descrito que é derivado.</i> O recurso informacional descrito pode ser derivado a partir do recurso informacional relacionado, no todo ou em parte. Recomenda-se a identificação de um recurso informacional

		relacionado por meio de uma sequência de acordo com um sistema de identificação formal.
12 Language (Idioma)	<dc:language>	<i>Idioma do recurso informacional.</i> Recomenda-se a utilização de um vocabulário controlado como o <i>Request for Comments 4646 Guidelines (RFC 4646)</i> .
13 Relation (Relacionamento)	<dc:relation>	<i>Recurso informacional relacionado.</i> Recomenda-se sua identificação por meio de uma sequência de acordo com um sistema de identificação formal.
14 Coverage (Cobertura)	<dc:coverage>	<i>Tema espacial ou temporal do recurso informacional, a aplicabilidade espacial do recurso informacional ou da jurisdição em que o recurso informacional é relevante.</i> Associado à extensão de aplicação do conteúdo do recurso informacional, ou seja, à cobertura do recurso informacional, que pode ser tanto espacial (nome de um lugar) quanto temporal (período ou data).
15 Rights (Direitos)	<dc:rights>	<i>Informação sobre os direitos detidos do recurso informacional e sobre o recurso informacional.</i> Em geral, incluem uma declaração sobre diversas propriedades de direitos relacionadas com o recurso informacional, incluindo propriedades de direitos intelectuais.

Fonte: Adaptado de DCMI (2014) e Arakaki (2015).

O uso obrigatório do Dublin Core para descrever recursos informacionais em um repositório institucional é uma das bases para a interoperabilidade implementada pelo *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)*. Embora existam muitos padrões de metadados disponíveis na atualidade, o padrão Dublin Core é a linha de base para a interoperabilidade neste campo (JONES, p. 86).

Para Rubin (2016, p. 342, tradução nossa), “os elementos do padrão Dublin Core foram formalmente endossados como um padrão NISO (Z39.85-2007) e estabelecem a base para a norma internacional (ISO 15836: 2009)”. Além disso, o Dublin Core “fornece os elementos básicos para a descrição de recursos em um ambiente eletrônico, bem como para aplicativos e padrões especializados para descrever tipos específicos de objetos”, que incluem, por exemplo: *Content Standards for Digital Geospatial Metadata* para objetos como mapas e dicionários geográficos *Categories for the Description of Works of Art (CDWA)* para descrever objetos de arte, *Visual Resources Association (VRA)*, *Core Categories for Visual Resources* e *Learning Object Metadata (LOM)* para objetos de aprendizagem, como programas, notas de aula, simulações e kits educacionais.

No que tange às principais críticas em relação ao uso do Dublin Core para descrever conteúdo no domínio bibliográfico, Jones (2007, p. 87, tradução nossa) aponta a ausência de

expressividade suficiente para descrever o conteúdo com precisão. Por exemplo, o campo “colaborador” não descreve o tipo de contribuição realizada: é um autor, um editor ou outro autor? O campo “data” também não é suficientemente preciso: à qual data se refere, à data de criação do recurso informacional ou a outra? Para o autor, a solução deste e outros problemas está na utilização de extensões que possibilitem melhor cobertura e gerenciamento de coleções, de direitos e relacionamentos com outros materiais. Em repositórios institucionais, o desenvolvimento de um perfil de aplicação para trabalhos acadêmicos definiria o uso dos campos Dublin Core com mais precisão, pautado nos conceitos do FRBR. Além disso, informações sobre o *status* de revisão por pares e financiamento poderiam agregar valor aos recursos informacionais. Porém, na visão do autor, “é improvável que o rigoroso Dublin Core ou mesmo o perfil de trabalhos acadêmicos capturem todas as informações necessárias para a funcionalidade completa de um repositório”, considerando-se que algumas das informações adicionais podem ser específicas da biblioteca universitária e que, até o momento, “não há um padrão geral detalhado em uso para todos os repositórios, como existe nos catálogos de bibliotecas com o MARC”.

Para além das questões relacionadas ao conteúdo descritivo dos recursos informacionais, como apresentado no Quadro 2, no padrão Dublin Core o campo específico para a descrição temática do recurso informacional está descrito no terceiro elemento: *Subject* (<dc:subject>). Partindo-se da premissa de que um repositório institucional no contexto de bibliotecas universitárias precisa representar o conteúdo intelectual dos recursos informacionais para prover acesso aos mesmos baseando-se no seu *assunto*, torna-se relevante tecer algumas reflexões em torno dos *metadados temáticos* em repositórios institucionais, visando à contribuição na resolução de problemas relacionados à representação e recuperação por assuntos nestes sistemas de recuperação da informação.

3.1.1.1 Metadados temáticos: possibilitando a representação e a recuperação por assuntos em repositórios institucionais

Em busca de uma categorização dos metadados temáticos no amplo universo dos metadados, há uma diversidade de tipos “que variam de acordo com o recurso informacional a ser representado, com domínio de aplicação e com as necessidades dos usuários” (ALVES, 2010, p. 48). Embora inexista um consenso na literatura em torno das tipologias dos

metadados, de forma geral em ambientes digitais os metadados podem ser classificados em três tipos principais: *metadados descritivos* - descrevem um recurso informacional para fins de descoberta e identificação e podem incluir elementos como título, resumo, autor e palavras-chave; *metadados estruturais* - indicam como os objetos compostos são reunidos, por exemplo, como as páginas são ordenadas para formar capítulos; e *metadados administrativos* - fornecem informações visando o gerenciamento de um recurso informacional e outras informações técnicas a quem pode acessá-lo, por exemplo, permissões de acesso, direitos autorais, etc. (NISO, 2004, p. 1).

De forma mais específica, Gilliland-Swetland (1998, p. 1) considera que os metadados podem ser categorizados em: *metadados administrativos* - gerenciam e administram os recursos informacionais; *metadados descritivos* - representam os recursos informacionais; *metadados de preservação* - preservam os recursos informacionais; *metadados técnicos* - proporcionam informações sobre o funcionamento dos sistemas ou o comportamentos dos metadados; e *metadados de uso* - proporcionam informações sobre o nível e tipo de utilização dos recursos informacionais.

Em ambas taxonomias os *metadados descritivos* incluem tanto os *metadados descritivos* quanto os *metadados temáticos*. Esta relação dicotômica entre a representação descritiva e a representação temática é reforçada pela *Declaração dos Princípios Internacionais de Catalogação* da *International Federation of Library Associations and Institutions* (IFLA) que classifica a catalogação em duas principais abordagens: *catalogação descritiva* - “a parte do processamento bibliográfico que fornece quer os dados descritivos quer os pontos de acesso que não são de assunto” e a *catalogação de assunto* - “parte da catalogação que fornece termos de assunto controlados e/ou números de classificação” (IFLA, 2009, p. 9). No referido glossário, a *indexação por assuntos* apresenta-se em remissiva para o termo *catalogação de assuntos*.

De acordo com Alves e Santos (2013, p. 6-7), o surgimento de novos modelos de representação oriundos da integração estratégica dos métodos tradicionais de tratamento descritivo da informação com as tecnologias computacionais configuram um novo paradigma representacional no domínio bibliográfico, dentre os quais podem ser destacados a própria revisão dos princípios de catalogação da IFLA; o desenvolvimento dos modelos conceituais FRBR - *Functional Requirements for Bibliographic Records* (Requisitos Funcionais para

Registros Bibliográficos), FRAD - *Functional Requirements for Authority Data* (Requisitos Funcionais para Dados de Autoridade), FRSAD - *Functional Requirements for Subject Authority Data* (Requisitos Funcionais para Dados de Autoridade de Assunto) e a revisão/atualização dos códigos de catalogação, juntamente com o desenvolvimento de um novo esquema de descrição de recursos informacionais baseado na estrutura dos modelos conceituais - o RDA - *Resource Description and Access* (Descrição de Recursos e Acesso). Além disso, emergem os denominados *Esquemas de codificação* em duas principais frentes: *Esquemas de codificação de conteúdo* (padrões de estrutura de dados e padrões de intercâmbio de dados); *Esquemas de codificação de valores* (padrões de conteúdo de dados e padrões de valores de dados); e os *Padrões ou Esquemas de metadados*.

Em relação à representação temática neste novo ambiente digital, destacam-se como *modelo conceitual* o FRSAD - *Functional Requirements for Subject Authority Data* (Requisitos Funcionais para Dados de Autoridade de Assunto) e como *esquema de codificação de conteúdo* os *Padrões de valores de dados*, que englobam os tradicionais instrumentos de tratamento temático da informação: tesouros, taxonomias, ontologias e esquemas de classificação.

Conforme apresenta Gonçalves (2013, p. 219, tradução e grifo nossos), historicamente, foi por volta de 300 a.C., desde os tempos do início da Biblioteca de Alexandria que bibliotecários lidavam com a questão do armazenamento de recursos informacionais para futura recuperação. Ao longo do tempo, o aumento das coleções tornou o problema de representação e recuperação cada vez maior. Nas palavras do autor: “procurar por um livro em particular dentre centenas de livros tornou-se uma tarefa tediosa, demorada e impraticável”. Para resolver a questão, *rótulos* foram adicionados aos recursos informacionais, permitindo, além do fornecimento de *metainformação* ao seu conteúdo, a organização, busca e recuperação com mais rapidez. Não obstante, a atribuição de um identificador único para cada recurso informacional “resolvia o problema sempre que o usuário soubesse dos identificadores dos livros que queriam, mas não resolvia o *problema mais genérico de encontrar documentos sobre um assunto ou tópico específico*”. Assim, a solução natural foi “agrupar os documentos por tópicos comuns e nomear cada grupo com um ou mais rótulos significativos”.

No domínio bibliográfico, “relacionado a coleções de bibliotecas, arquivos, museus e outras comunidades de informação” (IFLA, 2009, p. 13), os metadados geralmente vinculados a textos incluem dados como autor, data de publicação, fonte da publicação, tamanho (em páginas, palavras, bytes) e tipologia documental (livro, artigo, tese), informações que definem os *metadados descritivos*, ou seja, *metadados extrínsecos* ao significado do recurso informacional, relacionados à criação do registro. Já os *metadados temáticos* caracterizam o *assunto* que pode ser encontrado no *conteúdo temático* do recurso informacional, ou seja, *metadados intrínsecos* ao recurso informacional. Nesse sentido, os metadados descritivos e os metadados temáticos caracterizam uma *relação dicotômica* entre *forma* e *conteúdo*. De acordo com os objetivos e funções do catálogo, delineados pela IFLA (2009, p. 3),

[...] o catálogo deve ser um instrumento efetivo e eficiente que permita ao utilizador (usuário): *Encontrar* recursos bibliográficos numa coleção como resultado de uma pesquisa, utilizando atributos e relações entre recursos: Para *encontrar* um determinado recurso; Para *encontrar* conjuntos de recursos representando; todos os recursos que pertencem à mesma obra; todos os recursos que representam a mesma expressão; todos os recursos que exemplificam a mesma manifestação; todos os recursos associados a determinada pessoa, família ou colectividade (entidade); todos os recursos sobre um determinado assunto; todos os recursos definidos por outros critérios (língua, lugar de publicação, data de publicação, tipo de conteúdo, tipo de suporte, etc.), normalmente como uma delimitação secundária de um resultado de pesquisa (IFLA, 2009, p. 3, grifo nosso).

Para que o objetivo delineado pelos *Princípios* de “*encontrar* conjuntos de recursos representando [...] todos os recursos sobre um determinado assunto” possa ser realizado de forma efetiva não apenas em catálogos online mas em repositórios institucionais, os metadados temáticos são fundamentais para a qualidade da representação dos assuntos dos recursos informacionais, possibilitando a adequada correspondência entre a representação e a recuperação da informação por assuntos pelos usuários.

No atual contexto dos repositórios institucionais de bibliotecas universitárias, os metadados temáticos dos recursos informacionais (como dissertações, teses e artigos de periódico) referem-se aos campos específicos para a *descrição dos assuntos*, que podem ser *palavras-chave* (linguagem natural) ou *termos/descriptores* (linguagem controlada, tais como os tesouros ou taxonomias de domínios específicos), fornecidos pelo próprio *autor*, *catalogador/indexador* ou ainda oriundo de um processo de *indexação social* caracterizado pelas folksonomias.

De acordo com Navarro e Ziviani (2013, p. 189), as linguagens controladas são “estruturas hierárquicas compostas por termos que descrevem certos tópicos de conhecimento”, no intuito de padronizar os termos semânticos. De forma geral, os metadados temáticos caracterizam-se como os metadados utilizados para ordenação conceitual ou de assuntos. No caso das folksonomias, segundo Smiraglia (2014, p. 66, tradução nossa), “a atribuição de termos tem poucas ou nenhuma regra, o que significa que os termos são coletivamente menos discretos do que aqueles atribuídos usando vocabulários controlados pré ou pós-coordenados”, em uma dicotomia entre a liberdade de atribuição de conceitos associada às *tags* sociais e a falta de precisão.

Esta *atribuição do conteúdo temático* de um determinado recurso informacional em repositórios caracterizam uma forma de metadados em si, isto é, os *metadados temáticos*, usados para reunir registros bibliográficos de recursos informacionais com outros recursos informacionais semelhantes em termos de tratamento temático. Além disso, “todos os descritores de assunto são considerados como um tipo de organização do conhecimento, porque todos podem ser utilizados para identificar os *aspectos conceituais* do conhecimento registrado e depois agrupar o conhecimento de acordo com os conceitos identificados” (SMIRAGLIA, 2014, p. 67, tradução nossa).

Na representação de um determinado recurso informacional em repositórios institucionais, tem-se um conjunto de metadados que o representam, tornando-o unívoco na representação descritiva e, ao mesmo tempo, possibilitando o compartilhamento de semelhanças temáticas por meio dos metadados temáticos deste conjunto de metadados, ou seja, possibilitando a reunião de diversos recursos informacionais por assuntos. Todavia, difere-se do catálogo online, cuja reunião temática centram-se em duas principais correntes teóricas: a catalogação de assuntos/indexação de assuntos que propicia esta reunião temática, abordagem também compartilhada pelos repositórios institucionais; e a classificação, que propicia a localização física dos recursos informacionais nas estantes do acervo da biblioteca universitária. Entretanto, esta localização física em repositórios institucionais não ocorre, ao lidar exclusivamente com um conjunto de objetos digitais que podem ser acessados diretamente no próprio repositório institucional.

Embora os muitos avanços no campo da organização de recursos informacionais em ambiente digital, em repositórios institucionais de bibliotecas universitárias a questão da

descrição dos metadados temáticos continua em aberto na Organização do Conhecimento. Assim, parte-se para a abordagem teórica da indexação documental, a fim de trazeremos seus pressupostos teórico-metodológicos para o contexto dos repositórios institucionais.

3.1.2 A indexação documental em repositórios institucionais: conceitos, objetivos e princípios

Na visão de Kuramoto (2006, p. 117), “o tratamento e a disseminação eficaz da informação tem se constituído em um dos grandes desafios do homem ao longo dos anos. Equipamentos, métodos, padrões, técnicas e teorias foram desenvolvidos com esse propósito”. Naves e Kuramoto (2006, p. 3) consideram que “tanto no contexto tradicional quanto no digital, o tratamento da informação permanece como uma atividade crucial, pois cumpre a função básica no trabalho de facilitar o acesso à informação”.

O desenvolvimento científico e conseqüente crescimento da produção científica nas últimas décadas levaram à necessidade de criação de instrumentos que possibilitassem tanto a gestão quanto uma melhor recuperação dos recursos informacionais em sistemas de informação, especialmente por assunto, o que leva a duas dimensões: a *representação* e a *recuperação* temática da informação. Na opinião de Fujita (1992, p. 17), “a principal preocupação da maioria dos sistemas de informação envolvidos com o tratamento e disseminação dos documentos, pelos quais se responsabiliza, residirá na representação temática do conteúdo desses documentos”.

Esta atividade, denominada *indexação*, é percebida como um processo intelectual realizado pelo indexador, que tem por base a compreensão do texto e a representação dos recursos informacionais. Como abordagem teórica da Organização do Conhecimento, a indexação viabiliza a representação temática dos recursos informacionais, com fins de recuperação por assuntos em repositórios institucionais (DAL’EVEDOVE; TARTAROTTI; FUJITA, 2018).

A complexidade inerente tanto do fazer profissional (pragmático) quanto do construto epistemológico (teórico) reside justamente no tratamento temático da informação, que especificamente no contexto científico pode ser denominado de Tratamento Temático da Informação Científica (TTIC) caracterizado pelo ciclo da informação científica: produção científica, promoção de bem estar social, produção de conhecimento, publicação

científica, tratamento temático da informação científica (TTIC), recuperação da informação científica (TARTAROTTI, 2014, p. 135).

No patamar da tríade do *Tratamento Temático da Informação Científica* (TTIC) a indexação é definida classicamente como “a ação de descrever e identificar um documento de acordo com seu assunto” (UNISIST, 1981, p. 84) ou a “[...] técnica de caracterizar o conteúdo de um documento [...] retendo as ideias mais representativas para vinculá-las a termos de indexação adequados” (PINTO MOLINA, 1993, p. 208, tradução nossa). Nesta concepção, o processo de indexação envolve a determinação do assunto e tradução do assunto para uma linguagem de indexação (MAI, 2000, p. 277), ou seja, representação por conceitos e representação por linguagens construídas. De modo geral, o processo de indexação contempla a “a análise de assunto, quando ocorre a extração de conceitos que possam representar o conteúdo de um documento, expressos em linguagem natural, e a tradução desses termos para termos de instrumentos de indexação, que são as chamadas linguagens de indexação” (NAVES, 2000, p. 192).

Conceitualmente, a indexação de assuntos é um processo formado por subprocessos ou etapas que identifica o conteúdo de um determinado registro informacional por meio de uma metalinguagem construída - a linguagem documental -, visando a efetiva recuperação da informação (TARTAROTTI; BOCCATO, 2013, p. 40). Considerada uma atividade de natureza pragmática (CUNHA, 1989), “de raízes tecnicistas que congrega um conjunto de normas e instrumentos próprios, cujo processo operacional perpassa questões subjetivas. Tal característica imprime a esta atividade uma peculiaridade própria diante de outros processos destinados à representação temática da informação (TARTAROTTI; DAL’EVEDOVE; FUJITA, 2017, p. 111).

Em termos teóricos, qualquer recurso informacional pode ser indexado/representado por meio de conceitos em busca da satisfação de necessidades informacionais específicas (GIL LEIVA, 2008, p. 64), seja um livro, um artigo de periódico, uma partitura, um objeto tridimensional, um artefato museológico, um recurso informacional arquivístico, um recurso informacional adaptado para necessidades especiais dos usuários ou mesmo, mais recentemente, um dado oriundo de um conjunto de dados de pesquisa acadêmica. De modo geral, o propósito da indexação de assuntos em repositórios institucionais é determinar o conteúdo temático de recursos informacionais e expressá-lo em termos de indexação. Logo,

a indexação vai além da representação temática da produção científica - foco dos produtos e serviços - podendo ser aplicada, em um contexto mais amplo, a qualquer objeto passível de ser denominado recurso informacional.

Em uma concepção mais ampla por considerar seu objetivo principal, que é a recuperação da informação, Gil Leiva (2008, p. 64, tradução nossa) define a indexação como “um processo executado nos objetos suscetíveis de ser representados mediante conceitos e os pedidos dos usuários” que visam satisfazer necessidades informacionais. Tal atividade é intermediada por um sistema de recuperação da informação, definido como “um ambiente linguístico cuja eficiência depende de um controle adequado da linguagem de representação dos itens de informação e das requisições de seus usuários” e “um agente mediador na comunicação entre um estoque de informação e os seus potenciais requisitantes” (FERNEDA, 2013, p. 12). Assim, um repositório institucional como um sistema de recuperação da informação documental precisa representar o conteúdo intelectual dos recursos informacionais para prover acesso aos mesmos baseando-se no seu assunto.

No cenário dos repositórios institucionais, Hjørland (2003, p. 88) considera que o *indexador* desempenha um papel-chave na identificação das potencialidades dos recursos informacionais. Tomando como exemplo os índices de citação, afirma que

[...] pode haver várias razões para citar um documento. Para o grupo de autores que citam um mesmo documento pela mesma razão que este documento tem a mesma função e o mesmo “assunto”. A indexação é um processo no serviço de ajudar futuros autores a identificar documentos que possam ser citados. Isso é feito identificando-se exatamente os atributos que os tornam citáveis. As conexões entre artigos citados e artigos citantes são conexões de relevância, e são também (ou deveriam ser) as conexões entre os termos de indexação e os documentos indexados (HJØRLAND, 2003, p. 88, tradução nossa).

Neves, Dias e Pinheiro (2006, p. 141) esclarecem que as pesquisas sobre a indexação têm evidenciado três pontos: compreensão de texto (identificação do conteúdo de um recurso informacional); produção de texto (tradução do conteúdo identificado para a linguagem de indexação e/ou sua expressão na forma de um resumo); e a representação do conteúdo (criação de linguagens, como os tesouros).

Ao apresentarem os conceitos de dados, de conjunto de dados e de granularidade no domínio da Ciência da Informação, Santos e Sant’anna (2013, p. 206; 207) afirmam que a

granularidade de um conjunto de dados está vinculada ao *número de atributos* que compõem um determinado conjunto de dados e a diversidade de seus conteúdos, impactando diretamente nos processos de acesso e de tratamento. No conjunto de metadados, “a definição do menor subconjunto passível de ser identificado define a granularidade dos metadados utilizados”. Além disso, “conjuntos de dados com granularidade fina podem ser subutilizados se as visualizações não contiverem funcionalidades que permitam a consulta aos subconjuntos”. Porém, “conjuntos de dados com granularidade grossa diminuem as possibilidades de consulta e podem reduzir a relevância de recursos de visualização” (SANTOS; SANT’ANNA, 2014, p. 208; 207).

Nesse sentido, a *granularidade* na *representação descritiva* em repositórios institucionais está ligada ao nível de detalhamento da descrição do recurso informacional digital: quanto *maior* o nível de detalhamento, significado, descrição do atributo ou possibilidades de recuperação do recurso informacional de um repositório institucional, mais fina é a granularidade; do contrário, quanto *menor* o nível de detalhamento, do significado do atributo ou possibilidades de recuperação do recurso informacional de um repositório institucional, mais grossa é a granularidade. Em ambientes digitais, ao imprimir o significado a um conjunto de dados por meio da composição da tríade - entidade, atributo, valor $\langle e, a, v \rangle$, o conjunto de dados passa a ser passível de ser analisado como elemento que pode ser diretamente tratado por instrumentos digitais, configuração base para a recuperação da *informação* e interoperabilidade em repositórios institucionais.

O conceito de granularidade na *representação temática* em repositórios institucionais, pode ser vinculado aos conceitos de *exaustividade* e *especificidade*, elementos relacionados à qualidade da indexação de assuntos. De acordo com Gil Leiva (2008, p. 70, tradução nossa), a *exaustividade* está ligada à “quantidade de conceitos que caracterizam o conteúdo íntegro do recurso informacional e nunca com o número de descritores atribuídos a um objeto indexado”. No entanto, cabe ressaltar que em toda representação fica subentendido um nível de *redução informacional* e que em toda redução existe um nível de *perda informacional*. Assim, a representação do conteúdo íntegro é um ideal, mas não uma realidade possível.

Há vários fatores para que os sistemas de recuperação tal como os repositórios institucionais ampliam o número de descritores sem abarcar conceitos contidos no recurso informacional: generalidade ou especialização da informação tratada, exigências dos usuários

ou pela concepção de que quanto maior o número de descritores atribuídos a um recurso informacional, maior será a possibilidade de recuperação, embora evidentemente o nível de precisão diminua (GIL LEIVA, 2008, p. 71).

Já a *especificidade*, conceito cunhado por Charles Ammi Cutter nos primórdios da catalogação com a publicação da *Rules for a Printed Dictionary Catalogue* em 1876, que orientava que os recursos informacionais fossem indexados por cabeçalhos de assuntos específicos, evitando a atribuição de assuntos gerais, refere-se à “exatidão com que um termo de indexação representa fielmente um conceito particular” que aparece no recurso informacional em análise. Para que tal nível de especificidade na indexação e recuperação por assuntos em repositórios institucionais possa ser alcançado, três aspectos precisam ser considerados: a utilização de linguagens controladas que abarquem o nível de especificidade exigido na representação temática em ambientes de domínios específicos, a experiência do indexador e a política de indexação da instituição (GIL LEIVA, 2008, p. 72). Nesse sentido, o indexador deve utilizar termos de indexação que traduzam mais próximo possível o conteúdo do recurso informacional, rejeitando termos de indexação muito gerais ou muito particulares em relação às noções expressas no recurso informacional (MOREIRO, 2002, p. 58).

[...] qualquer palavra-chave constitui [um] ponto de acesso, mesmo que apenas extraída do texto dos documentos. Mas apenas um grande número de palavras-chave em linguagem natural descontrolada não é garantia de uma alta revocação por si mesma. Um documento sendo indexado com um único descritor altamente específico [...] é mais provável de ser recuperado do que se fosse indexado com diversas palavras-chave mais genéricas, embora o número de 'pontos de acesso' neste caso seja consideravelmente maior (FUGMANN, 1993, p. 202, tradução nossa).

No entanto, vale destacar que a questão da recuperação por assuntos por termos genéricos ou mais específico está diretamente ligada ao nível de conhecimento do usuário no assunto de interesse. Um usuário com pouco conhecimento no assunto pesquisado tende a usar termos genéricos em sua busca. Já um especialista tende a usar termos específicos, já que domina o vocabulário de sua área.

O nível de exaustividade ou especificidade a ser utilizado em um repositório institucional é uma decisão política, visto que a exaustividade ou especificidade durante a indexação para a representação do conteúdo de um recurso informacional são determinantes para a escolha de termos gerais ou específicos que serão, em seguida, traduzidos por descritores igualmente

gerais ou específicos para que na busca, o usuário possa representar com termos gerais ou específicos sua necessidade de informação (GIL LEIVA; FUJITA, 2012, p. 13). No entanto, tendo em vista as áreas de especialidade no âmbito acadêmico, em repositórios institucionais de bibliotecas universitárias torna-se necessário um maior nível de detalhamento na indexação dos assuntos, ou seja, um *maior nível de especificidade dos assuntos* ou de *granularidade fina*, em consonância com os objetivos da instituição de prover a informação científica em domínios específicos.

3.1.2.1 As etapas da indexação de assuntos em repositórios institucionais

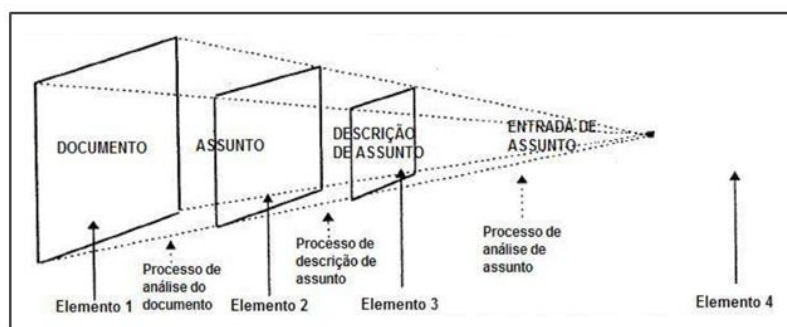
No âmbito da Organização do Conhecimento, a leitura documental é pressuposto básico das atividades desempenhadas pelos profissionais que lidam com o tratamento temático da informação. O objetivo deste tratamento temático é possibilitar a representação por assunto de recursos informacionais e sua posterior recuperação pelos usuários nos sistemas de recuperação da informação: catálogos online, bases de dados ou mais atualmente em repositórios institucionais ou repositórios de dados de pesquisa no contexto de bibliotecas acadêmicas, dentre outros. Assim, uma das características imprescindíveis em um repositório institucional de bibliotecas acadêmicas é representar o conteúdo intelectual dos recursos informacionais para prover acesso aos mesmos baseando-se no seu *assunto*.

Somada à complexidade inerente à indexação de assuntos, as diferentes concepções teóricas presentes na literatura internacional de Biblioteconomia e Ciência da Informação sobre as etapas do processo de indexação de assuntos (ISO 5963, 1985; LANGRIDGE, 1989; CLEVELAND e CLEVELAND, 1990; VAN SLYPE, 1991; NBR 12676, 1992; CHU e O'BRIEN, 1993; CHAN, 1994; MAI, 1997a; 2000; LANCASTER, 2004; TAMAYO e VALDEZ, 2008) ampliam, sobremaneira, as dificuldades da representação do conteúdo temático dos recursos informacionais, no ensino e na prática profissional. Apesar da variação no número de etapas, "é unanimidade que a análise de assunto se efetiva como o iniciar do processo, sendo sistematizada conforme abordagem teórica de cada autor" (SOUSA; FUJITA, 2014, p. 20).

Embora não consensuais entre os autores da literatura especializada no temário, o processo de indexação de assuntos é composto pelas etapas de: *leitura documental*; *análise de assunto* (também denominada de identificação de conceitos); *seleção de conceitos* e

tradução de conceitos. Todavia, a abordagem teórica de três etapas da indexação é a que melhor representa o processo, pois as etapas de análise de assunto/identificação de conceitos e a seleção de conceitos são realizadas durante a leitura documental. Neste viés, de acordo com Mai (1997b, p. 55), o processo de indexação pode ser desconstruído revelando três etapas: *processo de análise do documento*, *descrição do assunto* e *processo de análise de assunto*; e quatro elementos: *documento*, *assunto*, *descrição do assunto* e *entrada do assunto*, assim representada geometricamente (Figura 1):

Figura 1. O processo de indexação de assuntos em repositório institucional



Fonte: Mai (1997a, p. 61) adaptado de Miksa (1983)⁹.

Ao imprimirmos tais etapas e elementos no domínio dos repositórios institucionais, na primeira etapa do processo de indexação de assuntos é realizada a análise do recurso informacional digital ou, na terminologia do autor - objeto digital - visando à descrição temática, denominada de *processo de análise*. O primeiro elemento é o *recurso informacional digital* que está sendo analisado. A segunda etapa é a formulação de uma frase de indexação ou descrição de assunto, denominada de *processo de descrição*, uma formulação mental ou escrita do assunto pelo indexador, tendo como segundo elemento o *assunto* do recurso informacional digital, que pode estar presente apenas na mente do indexador. Já na terceira etapa ocorre a tradução da descrição de assunto em uma linguagem de indexação ou esquema de classificação, denominada de *processo de análise de assunto*. O terceiro elemento nesta etapa é a *descrição formal do assunto*, que pode ser escrita. Já o quarto elemento, denominado de *entrada de assunto*, é o produto da tradução da descrição formal do assunto em uma determinada linguagem de indexação do repositório institucional (MAI, 1997b, p. 55).

⁹ MIKSA, F. *The subject in the dictionary catalog from Cutter to the present*. Chicago: American Library Association, 1983.

Cabe esclarecer que o autor propõe o uso do termo “análise do documento” para a primeira etapa do processo e “análise de assunto” para a última etapa do processo. Na literatura percebe-se a existência de mais de uma denominação para o termo *análise de assunto*, podendo ser denominada ainda de *análise temática* ou *análise conceitual*. Entretanto, entendemos que o termo análise de assunto é o que melhor representa o processo de análise do conteúdo temático do recurso informacional, considerando-se seu sólido estabelecimento e aceitação da terminologia na literatura científica da Organização do Conhecimento.

Apesar das diferenças em termos teóricos e práticas adotadas, conforme Fujita, Rubi e Boccato (2009, p. 24), as etapas do processo de indexação não precisam ser necessariamente realizadas de maneira sequencial, visto que o indexador profissional já familiarizado com o processo pode realizá-las simultaneamente e até de maneira automática, situação corroborada na prática em estudo realizado com bibliotecários catalogadores e indexadores por Tartarotti (2014, p. 198), para quem quanto mais familiarizado com a área científica especializada no qual atua, mais estas etapas são superpostas durante o processo de indexação pelo profissional. Além disso, de acordo com Mai (2000, p. 280), a ação tomada por indexadores com diferentes níveis de experiência pode não ser exatamente a mesma, mas as etapas e elementos do processo de indexação de assunto são considerados fundamentais para qualquer processo de indexação.

A análise de assunto, que se caracteriza pelo conjunto dessas subetapas, é considerada na literatura e na prática profissional como a mais importante de todo o processo de indexação, pois resultará na representação do assunto principal do recurso informacional, tornando-se a “operação base para todo o procedimento de recuperação de informações” (CESARINO; PINTO, 1980, p. 32). A importância da análise de assunto é reforçada por Fujita (2003, p. 78), para quem “o assunto ou tematicidade do documento é o cerne principal e mais carente de esclarecimentos dentro dos estudos em análise documental”. Entretanto, as variáveis na indexação são muitas e pouco conhecidas, considerando-se as inúmeras possibilidades de significados e ideias contidos nos recursos informacionais em cada elemento no processo, além da presença da subjetividade, que depende particularmente do indexador (MAI, 1997a, p. 65).

Nesse sentido, no processo de indexação a análise de assunto é considerada a etapa mais complexa e mais importante da indexação, pois determina os resultados das estratégias de busca realizadas pelos usuários do sistema, no momento da recuperação da informação. Durante a análise de assunto, o indexador examina o recurso informacional, procurando compreender o texto por meio da identificação do assunto, com o objetivo posterior de selecionar os conceitos indexáveis. Chu e O'Brien (1993) consideram a análise de assunto como fase inicial do processo de indexação que decidirá sobre os principais tópicos do assunto de um recurso informacional, precedendo a fase de tradução desses tópicos de acordo com a linguagem documental adotada pelo sistema.

Segundo Taylor e Joudary (2009, p. 305), a “análise de assunto é a parte do processo de criação de metadados que identifica e articula o assunto do recurso informacional que está sendo descrito”. O processo de análise de assunto abarca três subetapas: 1) análise conceitual para determinar de que trata o item; 2) descrição da tematicidade em uma declaração por escrito; e 3) utilização dessa declaração de tematicidade para atribuir termos de vocabulário controlado e/ou notações de classificação. Já para Raju e Raju (2006, p. 14), teóricos da abordagem teórica da catalogação de assunto, a análise de assunto consiste no “processo de analisar o assunto contido em um item antes da tradução desta análise conceitual em uma linguagem”. Em síntese, a análise de assunto implica em determinar a tematicidade do recurso informacional mediante a identificação e seleção dos conceitos que compõem o assunto ou temas principal e secundários.

No contexto da análise de assunto realizada pelo indexador, a leitura documental possui um papel central, pois “para o indexador, seu cotidiano é concentrado no ato da leitura, de forma a viabilizar o acesso à informação contida nos documentos aos usuários dos sistemas de informação” (NEVES; DIAS; PINHEIRO, 2006, p. 141). Conforme Fujita (2003, p. 94), é a leitura do recurso informacional que dá início à atividade de indexação, diferindo de uma leitura normal ao exigir procedimentos diferentes, “ainda que os conhecimentos necessários para um bom entendimento de um texto sejam comuns a ambas”. Aqui, “esse leitor-indexador tem objetivo definido: identificação e seleção de conceitos que representem o conteúdo do texto e que coincidam com as necessidades informacionais da comunidade usuária do sistema de informação (LARA, 1993, p. 50).

Todavia, é notável a dificuldade apresentada pelo indexador ao ler um recurso informacional com o objetivo de identificar e selecionar os conceitos representativos do assunto do recurso informacional, considerando-se que, na maioria das vezes, não é um especialista sobre o assunto que indexa. Por outro lado, o indexador bibliotecário pode ser tornar um especialista na área em que atua devido à experiência adquirida ao longo do tempo na prática na atividade de indexação, além de cursos especializados (FUJITA, 2003, p. 84).

Considerando-se que a construção de qualquer texto de um recurso informacional é baseada em um ou mais temas, é por meio da leitura profissional na representação temática da informação que ocorre a busca pela tematicidade documental, revelada por meio da análise de assunto realizada pelo indexador - leitor profissional que interage com o texto para cumprir o objetivo da indexação. Sendo realizada durante a leitura documental, a análise de assunto é subdividida em outras três etapas: compreensão do conteúdo do recurso informacional; identificação dos conceitos que representam este conteúdo; e seleção dos conceitos válidos para recuperação. Se a análise de assunto, como primeira etapa da indexação, é realizada através da leitura - uma leitura documental -, é no momento da leitura em que os conceitos tratados em um recurso informacional são identificados e selecionados, para posterior representação em termos de indexação. Nesse sentido,

[...] a representação por conceitos na análise de assunto para identificação e seleção de conceitos é diferente da representação realizada na tradução porque, em primeiro lugar, é realizada durante a análise de conteúdo do documento, ou seja, com o documento na íntegra e em segundo lugar, porque utiliza conceitos com os quais são representados os termos e assim identificada a tematicidade intrínseca do documento. Na tradução, de outro modo, a representação se dá com os termos extraídos do documento analisado, portanto, fora do contexto *documental* em que o autor desenvolveu o conteúdo significativo e utiliza uma linguagem *documental* que representa o vocabulário de áreas de especialidade com a finalidade de compatibilizar os termos identificados com os termos buscados pelo usuário (FUJITA, 2013, p. 50, alteração nossa).

Em uma abordagem ideal da prática da indexação, as etapas de leitura documental, análise de assunto/identificação de conceitos e seleção de conceitos precisam ocorrer durante a leitura, enquanto que a etapa de tradução dos conceitos que representam os conceitos em descritores da linguagem do sistema deve ser feita após a leitura do recurso informacional. Esta visão permite que a análise seja realmente conceitual e compreensiva, pois a preservação do conteúdo do recurso informacional é uma garantia de relevância de recuperação, objetivo

da eficiente indexação de conteúdo (DAL'EVEDOVE, 2002, p. 58).

Referente às subetapas da análise de assunto, a *compreensão de leitura* é considerada uma condição necessária à leitura, ou seja, “não existe leitura sem compreensão. Então, quando falamos em leitura para indexação, podemos dizer que o indexador necessita compreender o texto para identificar e selecionar conceitos, pois somente o fará a contento se houver compreensão” (FUJITA, 2003, p. 82). Sobre isso Farrow (1991, p. 151) adverte que o indexador compreende o texto basicamente do mesmo modo que um leitor fluente, porém, sob a influência de condições de tempo, objetivo definido, modelo a ser produzido e áreas temáticas definidas com estrutura textual padronizada dos recursos informacionais, podendo induzi-lo a um processo repetitivo e automático.

Na *identificação de conceitos*, o indexador, “após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo”. Esta etapa depende “da tematicidade do texto e está atrelada à leitura do indexador e às suas concepções de análise de assunto adquiridas pela sua formação, objetivos e políticas de indexação” Deste modo, a identificação dos conceitos tratados num recurso informacional envolve como aspecto principal a compreensão do conteúdo do recurso informacional pelo leitor, de forma a representar e propiciar a seleção dos conceitos válidos para recuperação. Já a etapa de *seleção de conceitos*, que ocorre após a identificação dos conceitos, está diretamente relacionada aos “objetivos para os quais as informações são indexadas”, sendo que “nem todos os conceitos identificados serão necessariamente selecionados” (FUJITA, 2003, p. 64; 85).

Como visto, a análise de assunto é subdividida nas etapas de compreensão do conteúdo do recurso informacional; identificação dos conceitos que representam este conteúdo; e seleção dos conceitos. Todo texto possui uma *tematicidade intrínseca* - que podemos denominar de tematicidade básica, ou seja, um tema nuclear inerente ao texto e ligado à produção do mesmo, identificada na etapa de análise de assunto ou identificação de conceitos. De acordo com Todd (1992, p. 102), o nível de relação entre tematicidade e significado varia, pois depende “do uso que a pessoa pode encontrar da tematicidade do documento numa certa época, e o mesmo documento pode vir a ter diferentes significados para o mesmo leitor em diferentes épocas, entretanto o documento possui uma atinência fundamental”. Esta etapa se refere à identificação de conceitos.

Por outro lado, a *tematicidade extrínseca* - que podemos denominar de tematicidade profunda - está diretamente relacionada com o contexto de produção e de recepção do texto, com a necessidade informacional dos usuários do sistema de recuperação da informação e com a política de indexação da instituição, ligada à recepção do texto. Estes fatores, quando conhecidos pelos profissionais que executam a atividade de indexação, são decisivos para a adequação temática ao contexto e respectiva recuperação temática. Esta etapa se refere à seleção de conceitos.

Na tematicidade do recurso informacional, destaca-se o conteúdo relativamente permanente do recurso informacional (*aboutness*), o que leva à tematicidade intrínseca textual, enquanto que o significado compreendido pelo usuário (*meanings*) refere-se à tematicidade extrínseca textual. Segundo Fujita (2003, p. 80), Begthol (1986)¹⁰ fez distinção entre *aboutness* e *meanings*: “o *aboutness* é o conteúdo intrínseco do documento, que independe do uso temporal que um indivíduo possa fazer do mesmo em análise e que o faz possuir uma tematicidade relativamente permanente”; já o *meanings* (significados) pode ser medido de acordo com o uso particular do recurso informacional tendo em vista os usuários.

Ainda para Fujita (2003, p. 79), não existe um consenso entre os autores em torno do termo *aboutness* em português, cunhado em 1969 por Fairthorne (1969). Enquanto para alguns significa “do que trata um texto”, outros utilizam “atênência” ou ainda “tematicidade”. Wellish (2000, p. 5) define *aboutness* como “o conjunto de assuntos ou tópicos tratados em um documento, inclusive a intenção do autor e o possível uso pelos leitores.” Para Albrechtsen (1993, p. 220), tal conceito passou a ser pesquisado em substituição ao conceito de *subject* ou “tema”. Nesse sentido, conceitualmente podemos considerar a tematicidade como sinônimo de tema. De acordo com Fujita (2001, p. 61), o tema possui uma estrutura temática composta por conceitos ou categorias ou facetas. A identificação das categorias ou conceitos na estrutura textual do recurso informacional ocorre por meio da análise conceitual e o conjunto das categorias identificadas formulará o tema do recurso informacional. Segundo Guimarães (2009, p. 35), o tema “é uma representação abstrata da estrutura global do significado de um texto”, compreendido ainda como um

[...] núcleo informativo fundamental ou elemento em torno do qual se estrutura a mensagem, sua identificação permite ao receptor considerar

¹⁰ BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, London, v. 42, n. 2, p. 84-113, 1986.

“entender” o texto, daí poderá partir para a elaboração de resumo do texto, ou para exercício de retextualização sob diversas formas (paráfrases, comentários, resenhas, resenhas, resenhas, etc.). O processo seguido pelo receptor é, de certo modo, oposto ao utilizado pelo emissor. Se este procede desenvolvendo o tema, o receptor deve, reduzindo as informações que lhe são transmitidas, limitar-se ao fundamental, até chegar a esse núcleo informativo (GUIMARÃES, 2009, p. 35).

No geral, enquanto o *aboutness* refere-se ao “conteúdo relativamente permanente do documento”, o *meanings* é entendido como “o significado compreendido pelo usuário” do sistema de recuperação da informação (FUJITA, 2003, p. 80). A literatura científica sinalizava que o interesse do indexador deve limitar-se ao *aboutness*, ou seja, à tematicidade textual, tendo em vista que o indexador lida com os conteúdos permanentes dos recursos informacionais (com as ideias propostas pelo autor), representadas através da linguagem textual.

Entretanto, os estudos mais contemporâneos defendem o estabelecimento de um novo paradigma no processo de análise de assunto, pois o *meanings*, ou seja, o contexto, também deve ser considerado pelo indexador durante a atividade de indexação, considerando-se que embora a tematicidade seja o conteúdo relevante do recurso informacional no contexto próprio do autor, algumas variáveis irão influenciar na determinação desse conteúdo. Logo, a determinação do(s) assunto(s) do recurso informacional também está ligada a estes fatores. Segundo Cavalcanti (1989) a tematicidade intrínseca refere-se ao tema importante para o autor, enquanto que a tematicidade extrínseca refere-se ao tema importante do ponto de vista do leitor. Isso é denominado pela autora como “saliência autor - relevância leitor”. Em síntese, no processo de representação da informação existem dois tipos de tematicidade: a *tematicidade intrínseca* (resultado da fase de *identificação de conceitos* na análise de assunto) e a *tematicidade extrínseca* (resultado da fase de *seleção de conceitos* na análise de assunto) (Quadro 3):

Quadro 3. A tematicidade textual na análise de assunto do processo de indexação

Etapa da indexação	Subetapa da indexação	Tipo de tematicidade
Leitura documental	-	-
Análise de assunto	Compreensão de leitura	-
	Identificação de conceitos	Tematicidade intrínseca
	Seleção de conceitos	Tematicidade extrínseca
Tradução dos conceitos	-	-

Fonte: Elaboração própria.

Considerando-se que a tematicidade intrínseca e a tematicidade extrínseca são aspectos norteadores do processo de indexação, a prática ideal desta atividade realizada pelo indexador deve abarcar tanto o *aboutness* (contexto de produção) quanto o *meaning* (contexto de recepção) dos recursos informacionais.

Na prática, a indexação é analisada sob três diferentes concepções teóricas ou perspectivas de análise que refletem na atuação profissional do indexador: indexação orientada para o recurso informacional (ênfase no recurso informacional), indexação orientada para a demanda (ênfase nos usuários) e indexação orientada para o domínio (engloba o contexto, o recurso informacional e os usuários) (GIL LEIVA, 2008, p. 62). De acordo com Fujita (2003, p. 72), a concepção de leitura orientada para o conteúdo deve orientar a identificação de conceitos e a concepção orientada para a demanda, a seleção de conceitos, sendo que a concepção do indexador está “diretamente vinculada com sua formação educacional (concepção orientada pelo conteúdo) e com a postura do sistema de informação (concepção orientada pela demanda) e não pelo fato de ele ser um leitor menos ou mais habilitado”. Corroborando com esta perspectiva, Fidel (1994, p. 573) afirma que se a indexação de assuntos é realizada para facilitar a recuperação por assuntos, torna-se necessário considerar os requisitos do sistema de recuperação da informação de acordo com o perfil dos usuários que o utilizam. No entanto, por abarcar uma visão mais ampla, a abordagem da indexação centrada no domínio representa o ideal em termos de indexação, pois considera outros elementos além do recurso informacional ou o usuário.

Sob a ótica de Smiraglia (2009, p. 674), a interpretação de um recurso informacional envolve o contexto em que o recurso informacional é interpretado, visto que anteriormente

ao processo de indexação houve a criação do recurso informacional por uma determinada comunidade discursiva, resultado de diversos atos de interpretação por parte de seus atores sociais. Uma vez concluído, publicado e analisado por meio do processo de indexação possibilitando que o recurso informacional seja acessível, o objetivo é que este seja recuperado e utilizado pelos usuários da informação, estejam ou não inseridos dentro da própria comunidade discursiva que o elaborou (MAI, 2000, p. 272; 2001, p. 604).

No processo de indexação, a seleção de conceitos corresponde à etapa da análise de assunto que resultará na extração da tematicidade extrínseca textual do recurso informacional em análise pelo profissional indexador. Logo, a etapa de seleção de conceitos é influenciada pela variável contexto, que refere-se à análise de assunto em contextos de informação especializados e todos os elementos influentes no momento da leitura profissional, isto é, “todas as condições nas quais se encontra o leitor quando entra em contato com um texto” (GIASSON, 1993, p. 40). Além disso, “representa não apenas o conhecimento prévio profissional e os objetivos da atividade em questão, mas o ambiente em que é realizada esta atividade profissional em leitura documental” (DAL’EVEDOVE, TARTAROTTI, FUJITA, 2015, p. 610).

Nessa perspectiva, para que o significado do conteúdo informacional de um recurso informacional seja o menos ambíguo possível, o mesmo deve ser analisado de acordo com o contexto no qual está inserido, uma vez que tal variável tende a agregar valor em determinados conceitos, os quais passam a assumir novos significados apenas mudando-se o ambiente de análise (PINTO MOLINA, 1993). Desse modo, “o conceito somente assume sentido quando levado em consideração o contexto de situação, em que o bibliotecário indexador deve ajustá-lo as suas decisões conforme a necessidade exigida pelo meio” (DAL’EVEDOVE, TARTAROTTI, FUJITA, 2015, p. 610).

Todavia, como observa Olson (2007, p. 536), a maioria dos relacionamentos na abordagem temática da informação como os tesouros (produtos da concepção teórica da indexação), listas de cabeçalhos de assunto (produtos da concepção teórica da catalogação) ou esquemas de classificação (produtos da concepção teórica da classificação) está limitada às relações paradigmáticas, ou seja, são relações intrínsecas, que não dependem do contexto. Nesta abordagem interpretativa do processo de indexação, os principais problemas da representação de recursos informacionais estão relacionados ao significado e à linguagem,

consequentemente que o indexador é o fator mais interferente do processo de indexação, pois os referentes de cada elemento no processo são entendimentos subjetivos e, portanto, que os resultados do processo são bastante incertos, e dependendo em grande parte de quem realiza a análise (MAI, 2000, p. 270; MAI, 1997a, p. 61).

Na opinião de Neves (2006, p. 43), o momento mais crucial do processo de indexação é a análise de assunto, pois reveste-se de uma subjetividade inerente à política de indexação, que determina prioridades a alguns assuntos. Para minimizar esta subjetividade na indexação, a formação e capacitação profissional desempenham um papel-chave. Entretanto, conforme alertam Neves, Dias e Pinheiro (2006, p. 151), “apesar desse treinamento, em sua busca por termos para representar o assunto tratado no documento, o indexador baseia-se no discernimento próprio e na prática adquirida no exercício da profissão, o que, além de tornar o processo de indexação subjetivo, questiona a eficácia da formação profissional”.

Somada às características inerentes à etapa de identificação de conceitos da análise de assunto que trazem uma complexidade própria à atividade de indexação por envolver aspectos cognitivos, a etapa de seleção de conceitos é influenciada notadamente pelos elementos de política de indexação e pelo perfil dos usuários, e em uma abordagem mais ampla, pela política de informação presentes no contexto institucional em que o indexador atua. Embora o assunto seja a informação relevante abordada no recurso informacional, a seleção do assunto sofre a influência da política de indexação do sistema de informação à qual pertence, pois a instituição decidirá, dentre outros elementos, se o tema extraído do recurso informacional terá um nível mais específico ou mais genérico com base no perfil da comunidade usuária. Nesse cenário,

[...] a atuação profissional no tratamento temático da informação deve ser regida por uma política própria e única, que considere a instituição como um todo, seus usuários, colaboradores, estrutura física e financeira, entre outros. Como um processo que vai muito além de um fazer técnico - mas sim, uma atividade altamente cognitiva e humana, sendo, portanto, passível de influências sociais - a qualidade do tratamento temático da informação no contexto de bibliotecas universitárias depende ainda da coleção e da comunidade usuária local. [...] São diversos os fatores que afetam o processo de indexação, já que não existe uma única maneira de indexar corretamente um documento (TARTAROTTI, 2014, p. 228).

Assim, ao refletir os fundamentos teórico-conceituais da Organização do Conhecimento e, ao mesmo tempo, as necessidades profissionais, o estabelecimento de uma política de

indexação formalizada em um manual de indexação contribui com o processo de melhorias contínuas na atuação profissional do indexador, pois fornece as diretrizes de indexação ao profissional indexador em seu contexto de atuação.

A complexidade inerente ao processo de análise de assunto - realizada durante a leitura documental, ou seja, com o texto “em mãos” - é ampliada ao lançar-se o olhar para a etapa de seleção de conceitos, que ocorre após a leitura do recurso informacional, ou seja, não mais com o texto. Isto porque na seleção de conceitos são vários os fatores interferentes. Desse modo, a etapa de seleção de conceitos é tão ou ainda mais complexa que a etapa de identificação de conceitos, por conter ainda mais elementos e fatores interferentes no processo, fatores extrínsecos que independem da leitura documental e, portanto, do contexto onde o profissional atua. Reforça-se ainda a visão de Dal’Evedove, Tartarotti e Fujita (2015, p. 612) sobre a importância de se considerar tanto a abordagem sociocognitiva quanto a abordagem sociocultural nos estudos em torno da indexação, mais especificamente na etapa de seleção de conceitos que abriga a tematicidade extrínseca textual da análise de assunto, tanto na formação quanto na capacitação do indexador.

Considerando-se que no contexto de repositórios institucionais o objetivo fulcral da indexação de assuntos é possibilitar a descrição do conteúdo de um determinado recurso informacional para posterior recuperação, torna-se relevante o delineamento dos aspectos em torno da recuperação da informação.

3.2 A indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação

O termo recuperação da informação, introduzido na Ciência da Informação por Calvin Mooers é definido como o processo ou método pelo qual um potencial usuário de informação é capaz de converter sua necessidade informacional em uma relação real de citações a recursos informacionais que contém informações úteis a ele, isto é, o processo de *encontrabilidade* das informações armazenadas e “abarca os aspectos intelectuais da descrição da informação e sua especificação para a busca, e também quaisquer sistemas, técnicas ou máquinas” utilizadas para realizar a atividade (MOOERS, 1951, p. 25, tradução nossa).

Em termos de domínio, a Recuperação da Informação está inserida em um contexto mais amplo - a Ciência da Computação e se concentra principalmente em “prover aos usuários

o acesso fácil às informações de seu interesse”, trabalhando em duas principais frentes distintas e complementares: uma centrada no sistema (construção de índices, processamento de consultas e algoritmos de ranqueamento) e outra centrada no usuário (comportamento e necessidades informacionais na recuperação). Não obstante, ambas concepções teórico-metodológicas compartilham do mesmo objetivo: representar, armazenar, organizar e permitir o acesso a recursos informacionais (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 1), alcançado mediante o denominado *sistema de recuperação da informação*.

Um sistema de recuperação da informação é composto de pelo menos duas partes: um banco de dados e um sistema para recuperar as informações no banco de dados (RUBIN, 2016, p. 305). Considerando-se que seu objetivo principal é permitir que os usuários encontrem informações de seu interesse, neste cenário emergem duas medidas importantes: a *eficácia* e a *eficiência*. Enquanto a eficácia está relacionada à proporção de satisfação versus o esforço do usuário, a eficiência refere-se a quão bem um sistema provê as informações de forma rápida e precisa. Parafraseando os autores: “não importa o quão eficaz é o modelo de RI, o quão bela é a interface, os usuários ainda ficarão frustrados se tiverem de esperar dois minutos pelos resultados de uma única consulta” (GONZALO; NAVARRO, 2013, p. 409).

Em um livro inteiramente dedicado às diversas possibilidades com que comportamentos informacionais podem ser “definidos, explicados, observados, descritos e medidos em estudos do pensamento e da experiência humana”, Case e Given (2016, p. 7, tradução nossa) consideram que o termo *buscar* (como um estado ativo de busca) é considerado apenas uma parte da abordagem de um indivíduo ao lidar com informações, que *podem optar ou não* por procurar informação ou a informação pode simplesmente encontrá-lo. Este cenário caracteriza outro termo mais amplo na literatura denominado *comportamento informacional*, que alerta para o fato de que existem muitos e variados comportamentos e contextos que moldam a forma com que um indivíduo lida com informações. A complexidade inerente ao comportamento informacional está relacionada ao fato de que os comportamentos variam consideravelmente de indivíduo para indivíduo, dependendo das situações e objetos de interesse, muitas vezes de forma implícita, tornando sua observação prática um desafio a ser superado. Ao delinearem cinco cenários de busca de informação, Case e Given (2016, p. 23) salientam que na literatura existe um tipo de comportamento informacional empírico maior

do que qualquer outro: *pesquisas sobre usuários na busca de informações em bibliotecas*. Especificamente no contexto acadêmico

[...] quando cientistas ou acadêmicos pesquisam, eles coletam informações sobre seus objetos, por exemplo, átomos, planetas ou culturas antigas. [...] De forma geral, o estudo direto dos cientistas sobre seus objetos está fora do domínio da Ciência da Informação. No entanto, assim que os cientistas comunicam suas descobertas por meio de documentos, *seu comportamento de busca é parte do foco da Ciência da Informação* (HJØRLAND, 1997, p. 13, tradução e grifo nossos).

De fato, as bibliotecas estão entre as primeiras instituições a adotarem sistemas para recuperar informações (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 3). Embora de suma importância para o contexto de bibliotecas universitárias e para o domínio da Ciência da Informação, investigações que se concentram na utilização de sistemas como a web ou catálogos online, embora com extensa literatura, são considerados menos valorizadas na área de Recuperação da Informação, constituindo um “gênero separado nesta reflexão sobre a história da pesquisa com usuários da informação” (CASE; GIVEN, 2016, p. 49, tradução nossa).

Todavia, a busca e a recuperação da informação são conceitos centrais do domínio da Ciência da informação. Embora seja um termo padrão adotado no campo, a recuperação da informação centra-se principalmente na recuperação de *recursos informacionais* ou recuperação de *textos*, não abarcando em seu escopo a recuperação de *fatos*. Ainda para o autor, o prefixo “re”, integrante da palavra “recuperação”, indica que alguns itens são encontrados novamente, contudo, “é mais preciso dizer que alguns itens são *identificados* (em vez de recuperados)” na recuperação da informação (HJØRLAND, 1991, p. 11).

No cenário da recuperação de informação, o sistema tem como objetivo permitir a interação entre outros dois elementos que possuem um papel central: as *informações* e os *usuários*. Referente às informações, Hjørland (1997, p. 13, tradução nossa) advoga que “o conceito de documento pode ser decisivo na definição da diferença entre a busca de informação como objeto da Ciência da Informação e outros tipos de busca de informação”. Na Ciência da Informação o conceito de recurso informacional é usado de forma ampla, englobando tradicionais tipos de recursos informacionais em bibliotecas, tais como livros, artigos de periódico, manuscritos, fotos, filmes, vídeos, programas de computador, dentre outros (HJØRLAND, 1997, p. 14). Na visão de Case e Given (2016, p. 24), são vários os estudos

que consideram *informação como uma coisa*, tal como preconizado por Buckland (1991), isto é, livros, periódicos e outros como “pacotes” de informação.

Conforme Case e Given (2016, p. 23), nas últimas três décadas a tendência de pesquisas na Ciência da Informação está relacionada à busca, avaliação e aplicação de recursos informacionais digitais e, mais recentemente, sobre a interação do *usuário* em modernos sistemas de recuperação de bibliotecas, tal como bibliotecas digitais. No entanto, o uso de materiais físicos pelos usuários e a percepção da biblioteca universitária como um espaço físico continuam sendo importantes áreas da pesquisa sobre comportamento informacional. Dentre alguns desafios das bibliotecas universitárias expostos pelos autores, podem ser citados: reconhecimento de que todas as bibliotecas, exceto as menores, podem ser complexas e intimidadoras; meta ilusória de reunião dos materiais em assuntos semelhantes; extensa categorização de recursos informacionais - livros, periódicos, arquivos de computador, gravações de vídeo, dentre outras, ou seja, várias interações do conteúdo (intelectual) com a forma (física ou digital); localização de recursos informacionais de mesmo assunto em lugares diferentes; compartilhamento de materiais com outras bibliotecas, em detrimento da tentativa de coleta de “todo” o material relevante; dificuldade de pesquisa, mesmo para usuários experientes, dada a complexidade da atividade; e redução de utilização da biblioteca física pelos usuários, com preferência cada vez maior e exclusivamente de fontes online para atender às necessidades informacionais. Nas palavras dos autores: “muitos visitantes de uma biblioteca fazem sua busca prematuramente, quando se deparam com um grande edifício cheio de milhões de itens e a ferramenta imperfeita do catálogo online”. Lamentavelmente, um usuário *atípico* de biblioteca sabe usar um bibliotecário e um catálogo online (CASE; GIVEN, 2016, p.24-25; 27, tradução nossa).

Visando satisfazer à necessidade informacional dos usuários, o sistema de recuperação da informação, idealmente, interpreta o conteúdo dos recursos informacionais de uma coleção e classifica-os de acordo com o nível de *relevância* à consulta dos usuários, envolvendo “a extração de informações sintáticas e semânticas do texto do documento e sua utilização” para satisfazer tais necessidades. Usuários de modernos sistemas de recuperação da informação têm necessidades informacionais de diferentes níveis de complexidade (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 4; 2). No intuito de suprir tal lacuna, o usuário traduz esta necessidade informacional em uma determinada *consulta* a ser submetida ao sistema de

recuperação. Nesse caso, “*buscando* ou *consultando* informações de seu interesse”. Tal tradução gera um conjunto de palavras-chave ou termos de busca, que sumarizam a necessidade informacional do usuário. Considerando-se a consulta do usuário, o objetivo maior do sistema de recuperação da informação é recuperar informações que sejam úteis ou relevantes. Nesse caso, “a ênfase recai na recuperação de *informação*, não na recuperação de *dados*” (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 5).

Em termos práticos, o processo de recuperação da informação em sistemas é delineado por etapas, que consistem na eliminação de *stopwords*, radicalização (*stemming*) e seleção de um subconjunto de termos para serem utilizados como *termos de indexação*, utilizados para compor a representação do recurso informacional. A partir dessas representações, cria-se um *índice* do texto, geralmente um *índice invertido*, etapa conhecida no domínio da Ciência da Computação como *processo de indexação*¹¹, que precisa ser executada *off-line*, antes que o sistema esteja pronto para receber consultas. Após a indexação (do ponto de vista do sistema) do conjunto de recursos informacionais, inicia-se o *processo de recuperação*. O usuário primeiramente especifica uma *consulta* que reflete sua necessidade informacional valendo-se de termos de busca, analisados, com operações típicas como correção ortográfica e eliminação de *stopwords*. A seguir, a consulta transformada é expandida, modificada e processada a fim de obter o conjunto de *recursos informacionais recuperados*, composto por recursos informacionais que contém os termos de busca. Em seguida, ocorre o processamento rápido de buscas, viabilizado pela estrutura de índice construída anteriormente. Aqui, as etapas necessárias para a produção do conjunto de recursos informacionais recuperados constituem o *processo de recuperação*. Em seguida, os recursos informacionais recuperados são ranqueados de acordo com a probabilidade de *relevância* para o usuário, etapa mais crítica do processo, visto que a qualidade do resultado percebida pelo usuário depende do ranqueamento, em que os recursos informacionais no topo do *ranking* são formatados e apresentados para o usuário. Esta formatação consiste em recuperar o título dos recursos

¹¹ Aqui cabe a distinção entre o processo de *indexação* do ponto de vista do sistema (*indexing* no inglês e *indexación* no espanhol), concepção adotada na Ciência da Computação, e o processo de indexação que objetiva analisar o conteúdo de um recurso informacional e traduzi-lo em termos de indexação que possam representar o conteúdo temático de tal recurso informacional, concepção adotada na Organização do Conhecimento, isto é, *indexação de assuntos* (*subject indexing* no inglês e *indización* no espanhol).

informativos e gerar *snippets* do texto, isto é, trechos que contenham os termos de busca que serão mostrados ao usuário (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 8).

Em um sistema de recuperação da informação, o principal desafio reside em não apenas saber como extrair a informação dos recursos informativos, mas também saber como utilizá-la para decidir quanto à sua *relevância*. Nesse sentido, “a noção de relevância tem um papel central em recuperação da informação”. Cabe salientar que a relevância se caracteriza como um *juízo pessoal* que depende da tarefa a ser resolvida e de seu contexto, “podendo ser alterada com o tempo (por ex., à medida que novas informações tornam-se disponíveis), com o local (por ex., a resposta mais relevante é a mais próxima), ou até mesmo com o dispositivo (por ex., a melhor resposta é um documento pequeno que seja fácil de baixar e visualizar)”. Nesse ponto de vista, nenhum sistema de recuperação da informação pode fornecer respostas perfeitas a todos os usuários o tempo todo (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 4-5).

Os autores alertam que, em uma coleção com muitos itens, prever a relevância é muito mais difícil, pois “qualquer consulta recupera um grande número de documentos que contém seus termos, o que significa que existem vários documentos ruidosos no conjunto de documentos recuperados”, ou seja, “documentos que parecem estar relacionados à consulta, mas que, na verdade, não são relevantes, de acordo com o juízo de uma grande parte dos usuários”. Considerando-se que a decisão quanto à relevância de um recurso informativo é inerentemente *subjetiva*, avaliar a qualidade do conjunto-resposta é a chave para a melhoria de um sistema de recuperação da informação. Entretanto, por meio de um processo de avaliação sistemático é possível melhorar a qualidade dos resultados (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 12; 7).

3.2.1 A recuperação por assuntos em repositórios institucionais

As diferenças entre *recuperação da informação* versus *recuperação de dados* são descritas por Baeza-Yates e Ribeiro-Neto (2013, p. 5), para quem, em um contexto de sistema de recuperação da informação, a recuperação de *dados* consiste na “identificação de quais documentos da coleção contém as palavras-chave da consulta do usuário, o que, com frequência, não é suficiente para satisfazer as necessidades de informação”. Para o usuário de

uma biblioteca universitária que realiza uma busca em um sistema de recuperação, o interesse recai mais em recuperar *informações* sobre um determinado assunto do que em recuperar dados que satisfazem uma dada consulta. Em um sistema de recuperação de dados, “um único objeto incorreto em meio a milhares de objetos recuperados significa falha total”. Nessa lógica, “um sistema de recuperação de dados, como um banco de dados relacional, trata de dados que possuem estrutura e semântica bem definidas, enquanto que um sistema de recuperação da informação lida com texto em linguagem natural que não é bem estruturado”. Recuperar dados atende às necessidades específicas em um sistema de banco de dados, no entanto, “não resolve o problema de recuperar informações sobre um assunto ou tópico”.

Para Hjørland (1997, p. 20), há uma oposição entre *dados descritivos* e *dados de assuntos*, dicotomia perceptível nas práticas de gestão em bibliotecas, como apresenta, por exemplo, a subdivisão entre a catalogação descritiva e a catalogação/indexação de assuntos, ponto de vista não compartilhado pelo autor, que considera estes dois tipos de dados como categorias *sobrepostas*. Por outro lado, tal divisão é percebida na recuperação, pois, de acordo com o autor, a recuperação por assuntos é a *busca de recursos informacionais desconhecidos*, em oposição a uma *busca de recursos informacionais conhecidos*, em que o conteúdo pode contribuir para a solução de um problema concreto de satisfazer uma necessidade informacional. Assim, a busca por itens conhecidos está mais ligada à representação descritiva, enquanto que a busca por itens desconhecidos, à representação temática (HJØRLAND, 1997, p. 5).

No intuito de facilitar a busca de recursos informacionais *conhecidos*, por meio da catalogação descritiva os dados bibliográficos como autor ou título são descritos, permitindo a utilização destas informações na identificação de um recurso informacional de maneira não ambígua. Para tanto, foram desenvolvidos padrões de descrição como o AACR2, a fim de facilitar a entrada em bancos de dados adequada para essa busca de itens conhecidos. O desenvolvimento e aplicação deste conhecimento descritivo “constituem uma parte importante das habilidades profissionais de bibliotecários e especialistas em informação” (HJØRLAND, 1997, p. 20, tradução nossa). Já para os recursos informacionais *desconhecidos*, durante uma determinada *busca por assuntos*, seja por um bibliotecário ou o próprio usuário, há uma variedade de dados envolvidos no processo, que compõem diversas estratégias na mesma atividade. Diante da complexidade inerente à busca por assuntos, Hjørland (1997, p.

21) afirma que é fundamental considerar dois principais aspectos: estamos lidando com recursos informacionais desconhecidos, usando critérios relacionados ao conteúdo e buscando por recursos informacionais que contribuam para esclarecer um problema ou satisfazer uma necessidade informacional.

Especificamente sobre a *recuperação por assuntos*, para Hjørland (1997, p. 20, tradução nossa) na prática a maior parte das buscas de informação em bibliotecas e bancos de dados podem ser denominadas de *buscas por assuntos*, sendo também a maior parte da literatura sobre busca de informações e recuperação da informação. Nas palavras do autor: “se podemos dizer que o que chamamos de Biblioteconomia e Ciência da informação tem um problema de pesquisa básica, está ligado à busca de informações e às possibilidades de otimizar a recuperação em bancos de dados”, ou seja, “problemas relacionados à busca por assuntos”. Tradicionalmente, a busca por assuntos refere-se à “busca por assuntos (recuperação por assuntos); busca de informação por meio de dados de assunto” (ORDBOGEN, 1991 apud HJØRLAND, 1997, p. 20).

Em continuidade, procurar informações sobre um determinado livro, por exemplo, pelo nome do autor, representa uma atividade de *busca de informações*, não uma *busca por assuntos*. No primeiro caso, se o usuário desejar encontrar um recurso informacional no sistema de recuperação da informação e lembrar apenas certas coisas sobre ele, por exemplo, os dados de um *orientador de teses e dissertações*, sua busca será caracterizada como um tipo de verificação, não como uma busca por assuntos (HJØRLAND, 1997, p. 21, tradução e alteração nossa). Por outro lado, se o usuário desejar realizar uma busca exhaustiva sobre um determinado assunto, poderá se basear nos dados do orientador para realizar uma busca por assuntos, considerando-se que tal orientador é especialista em uma determinada área do conhecimento. Assim, ao utilizar o nome do orientador no campo de busca por autor, o sistema irá recuperar todos os itens publicados por aquele determinado orientador, ou seja, recursos informacionais que compartilham da mesma temática. Para o autor, “busca por assuntos, portanto, não se limita aos dados específicos que são produzidos exclusivamente para pesquisa de assunto. Mas a pesquisa por assunto e outros tipos de pesquisa usam categorias de dados sobrepostas”. Em outras situações, os títulos dos recursos informacionais podem ser utilizados como busca por assuntos em catálogos de palavras-título em busca de texto livre, e os mesmos dados também são frequentemente usados para buscar um recurso

informativa específica onde o título é conhecido. Assim, o autor conclui que “não existe um limite acentuado entre os pontos de acesso aos assuntos (ou dados de assuntos) e outros tipos de dados de acesso”. Os pontos de acesso mais relevantes dependem da natureza da questão. Nesse sentido,

Todos os tipos de dados que podem dar uma pista (mesmo que vaga) sobre a identificação ou avaliação de documentos potencialmente relevantes podem ser usados na recuperação do assunto, incluindo os próprios dados do documento (título, resumos, lista de referências, autor) ou dados diferentes do próprio documento (incluindo códigos de classificação, descritores, resenhas de livros, avaliações e citações em outros documentos) (HJØRLAND, 1997, p. 5, tradução nossa).

Um dos grandes desafios da busca por assuntos é a “formulação de teorias e conhecimentos mais explícitos sobre como diferentes pontos de acesso são utilizados e combinados pelos usuários” e como o bibliotecário indexador pode produzir informações de valor agregado para representações de recursos informacionais. Em um registro bibliográfico, cada campo de pontos de acesso por assunto “contém um significado diferente e possui um valor informativo diferente. Juntos, eles ampliam as possibilidades de pesquisa”. Entretanto, na semântica de dados “nenhum tipo de ponto de acesso de assunto pode ser avaliado isoladamente, mas deve ser avaliado no contexto de toda a gama de pontos de acesso existentes”. Acréscimo de descritores seriam necessários apenas se contribuíssem para aumentar a eficácia da recuperação (HJØRLAND, 1997, p. 23-24, tradução nossa). Em síntese, uma busca por assuntos em um repositório institucional de uma biblioteca universitária “não pode ser considerada totalmente bem-sucedida, a menos que o usuário consiga localizar o material que seja, em algum sentido, o “melhor”, isto é, o *mais completo, o mais atualizado ou o mais autorizado*” (LANCASTER, 1996, p. 201).

3.2.2 Revisitando os conceitos de relevância, pertinência e saliência na recuperação da informação em repositórios institucionais

Na área de Recuperação da Informação, fenômenos como *informação, necessidade informativa, navegação ou busca e uso* são considerados conceitos-chave de um termo mais amplo denominado *comportamento informativo* (CASE; GIVEN, 2016, p. 5, tradução nossa), que ainda abarca outros conceitos. Conforme os autores, o conceito de *informação*

pode ser qualquer aspecto ou padrão perceptível na realidade que possibilite uma mudança no mundo do indivíduo. *Necessidade informacional* é definida como um reconhecimento de que o conhecimento de um indivíduo é inadequado para satisfazer um determinado objetivo. A *busca de informação* caracteriza-se como um esforço consciente para obter informações em resposta a uma necessidade ou lacuna do conhecimento deste indivíduo. Já o *uso da informação* é “o que o indivíduo faz com a informação adquirida através da busca, serendipidade¹², ou outros meios, incluindo aplicar - ou ignorar - as informações para atender às metas ou contexto pessoal”.

O *comportamento informacional* ou *práticas de informação* refere-se à busca de informações que abarcam tanto comportamentos intencionais quanto não intencionais do indivíduo, incluindo, em um contexto mais amplo, “como lidam com as informações em suas vidas, de modo a explicar a situação, o tempo, o afeto, a cultura, a geografia e outros elementos contextuais na compreensão do comportamento informacional”. Já o termo *experiências informacionais* têm sido utilizado na literatura por pesquisadores preocupados com diferentes *competências informacionais* e abrange uma gama de conceitos semelhantes aos estudados em comportamento informacional. Nesse caso, a ênfase recai na aprendizagem cotidiana e “o que poderia ser feito para melhorar e apoiar as ferramentas de educação e aprendizado” (CASE; GIVEN, 2016, p. 5, tradução nossa)¹³.

Esta ampla gama de conceitos relacionados à recuperação da informação desaguam em outros dois conceitos amplamente discutidos na literatura: *relevância* e *pertinência*. Conceitualmente, a *relevância* é percebida como “um relacionamento próximo e lógico com um assunto, observação ou questão”, intrinsecamente relacionada ao significado e “descreve a relação de padrões e o que quer que os padrões indiquem para os ambientes cognitivos do criador e perceptivo de uma mensagem (RICHIE, 1991¹⁴, p. 20 apud CASE; GIVEN, 2016, p.

¹² O conceito de serendipidade está relacionado à busca de informação recuperada “por acaso”, embora relevante para o usuário. Na visão de Fugmann (1993, p. 46), em sistemas de recuperação a serendipidade manifesta a ausência de precisão na busca e é bem-vinda somente no estágio inicial do desenvolvimento do sistema, podendo ser um obstáculo com o crescimento do mesmo.

¹³ Outros conceitos intimamente relacionados à busca de informações são: aprendizagem, alfabetização, tomada de decisões, relevância, pertinência, relevância, exposição seletiva, navegação, serendipidade, lacunas de conhecimento, pobreza informacional, sobrecarga informacional, seletividade informacional, localização de informação, ansiedade de informação, entretenimento, compartilhamento e colaboração (CASE; GIVEN, 2016, p. 98).

¹⁴ RICHIE, L. D. *Information*. Newbury Park, CA: Sage, 1991. v. 2.

110). Neste viés, enquanto os dados são considerados padrões com potencial comunicativo, os *dados com relevância* são considerados *informações* (CASE; GIVEN, 2016, p. 110).

No âmbito da Ciência da Informação, a relevância é equiparada a *aboutness* e *topicalidade*. Um determinado recurso informacional é relevante para uma necessidade informacional se for considerado *no tópico* (HARTER, 1992, p. 602), ou seja, naquele contexto de um assunto determinado. A topicalidade serve como base de avaliação da eficácia de um sistema de recuperação e os julgamentos de relevância, isto é, o “ponto em que um recurso informacional começa a ser relevante e deixa de ser relevante” depende da visão de cada usuário (JAMES; MACKINNEY, 1992 apud CASE; GIVEN, 2016, p. 112, tradução nossa).

A relevância é um conceito que vai muito além da navegação e de outros conceitos relacionados. Embora tenha sido objeto de estudo da Ciência da Informação, a ênfase sobre a relevância tem sido nas medidas técnicas de recuperação de recursos informacionais, sendo a precisão e revocação as medidas mais comuns de avaliação desta eficácia (CASE; GIVEN, 2016, p. 110; 111). Todavia, os autores salientam que discussões que cerquem o real significado da relevância para um indivíduo são de suma importância.

De acordo com Case e Given (2016, p. 112), desde a década de 1970 a literatura da Ciência da Informação tem percebido a relevância baseada no *estado de conhecimento e intenções do usuário*, em detrimento a uma correspondência lógica da terminologia pelo sistema de informação (MIZZARO, 1998; INGWERSEN; JÄRVELIN, 2005). Esta *caracterização subjetiva da relevância* tem sido denominada *relevância situacional* (WILSON, 1973; BARRY, 1994; BRUCE, 1994), *relevância psicológica* (HARTER, 1992) ou *pertinência* (BELKIN; VICKERY, 1985¹⁵; HOWARD, 1994). Belkin e Vickery (1985, p. 46) distinguem claramente os conceitos de *relevância objetiva* e *pertinência* ou *relevância subjetiva*: “relevância é a propriedade que atribui uma resposta a uma pergunta e pertinência é a propriedade que atribui uma resposta à uma necessidade informacional”.

Conforme Case e Given (2016, p. 11), Froehlich (1994) aponta que o que a maioria dos usuários querem dizer quando utilizam a palavra relevância é, na verdade, pertinência: *suportar* ou *se conectar* a uma necessidade informacional, ao invés de uma questão declarada, isto é, *relevância subjetiva*. A *pertinência* é o significado central de um conjunto de conceitos

¹⁵ BELKIN, N.J.; VICKERY, A. *Interaction in information systems: a review of research from document retrieval to knowledge-based systems*. Boston Spa: British Library, 1985.

relacionados, que tem, no centro, o *sentido cotidiano de 'relevância'*, enquanto na periferia está o *sentido técnico de relevância* como respostas a uma consulta formal e declarada. Para os autores, independente do conceito, esta “visão subjetiva de relevância defende a importância do estado de conhecimento e das intenções do usuário no momento de encontrar informações”. Além disso, alguns estudos na Psicologia apontam que a pertinência está inerentemente relacionada ao contexto informacional do usuário (CASE; GIVEN, 2016, p. 113; 11, tradução nossa).

Segundo com Fugmann (1993, p. 46, tradução nossa), na avaliação de sistemas de recuperação da informação há uma confusão teórica entre os termos *relevância* e *pertinência*, causando conclusões errôneas sobre a qualidade da recuperação em um sistema de recuperação. Para o autor, a *pertinência* ocorre quando os resultados de busca em um sistema de recuperação trazem recursos informacionais que foram recuperados e são relevantes, no entanto, não são de interesse do usuário, por diversos motivos, por exemplo, recursos informacionais já conhecidos, porém, “precisamente do tipo especificado na solicitação de pesquisa”, ou seja, *relevante* e não *pertinente*. Nessa perspectiva, o termo pertinência é um critério de avaliação que se refere a “se a busca recuperar ou não um ou vários itens que o usuário considera úteis” visando satisfazer sua necessidade informacional e indica um recurso informacional que pode ser útil para o usuário porque contribui para a satisfação de sua necessidade informacional (LANCASTER, 1996, p. 188).

Outro termo relacionado, porém, menos discutido na literatura é o conceito de *saliência*, visto como “algo que é saliente, destaca-se, é vívido, inesperado, notável, conspícuo, proeminente” ou “desagradável, desviante, extremo, intenso, incomum, repentino, brilhantemente iluminado, colorido, sozinho” (KIESLER; SPROULL, 1982, p. 556), ou seja, recursos informacionais que se *destacam* mais entre os demais recursos informacionais recuperados em buscas por assuntos em um repositório institucional, ficando estes últimos em segundo plano. Nesse sentido, um estímulo saliente não é necessariamente relevante para uma determinada necessidade informacional (CASE; GIVEN, 2016, p. 114). Sob a ótica de Johnson (1997)¹⁶ apud Case e Given (2016, p. 114, tradução nossa), a saliência corresponde a “crenças como parte de *fatores de relevância pessoal* que antecedem qualquer atividade de busca de informação”. Em momento anterior à consideração das “características de uma

¹⁶ JOHNSON, J. D. *Cancer-related information seeking*. Cresskill, NJ: Hampton Press, 1997.

fonte de informação e sua utilidade, ocorre uma interação entre uma lacuna no conhecimento, crenças sobre esse tópico do conhecimento e a importância [...] do mesmo. [...]”. Assim, os indivíduos prestam atenção somente ao que é importante sob seu ponto de vista.

3.2.3 Variáveis/indicadores/métricas da recuperação por assuntos em repositórios institucionais

Em sistemas de recuperação da informação, a avaliação da recuperação é um ponto de suma importância, pois visa medir o quão bem o sistema de recuperação atende às necessidades informacionais dos usuários. Entretanto, “o mesmo conjunto resposta pode ser interpretado de maneiras diferentes por usuários distintos”. Apesar desta problemática subjacente, é possível definir métricas aproximadas, que, na média, tem uma correlação com as preferências de uma população de usuários. A avaliação da recuperação é “um processo sistemático no qual se associa uma métrica quantitativa aos resultados produzidos por um sistema de recuperação da informação em resposta a um conjunto de consultas de usuário” e que deve ser comparada à relevância dos resultados para os usuários. Cabe esclarecer que, nesta concepção, a avaliação da recuperação está ligada à verificação da *qualidade* dos resultados, não ao desempenho do sistema quanto ao tempo de processamento das consultas, nesse caso, denominada *avaliação de desempenho* (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 106).

Embora a aplicabilidade de tais métricas não considere a questão do julgamento de relevância pelos usuários, de forma geral o “processo de associar uma métrica numérica aos resultados da consulta continua sendo amplamente adotado por ser simples e por poder ser repetido muitas vezes a custos relativamente baixos”, possibilitando a verificação contínua da qualidade do sistema de recuperação. De acordo com a literatura, os primeiros estudos sobre avaliação da recuperação, iniciados na década de 1950 pelo bibliotecário Cyrill Cleverdon, culminaram no clássico *paradigma Cranfield*, que serviu de base para o desenvolvimento das métricas de avaliação em recuperação da informação (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 107).

Dentre as principais métricas utilizadas para avaliar um sistema de recuperação da informação, ou, nessa pesquisa, um repositório institucional, destacam-se a precisão e a revocação. A *precisão* refere-se à quantidade de ruído incluída nos resultados, enquanto que a *revocação* refere-se à proporção de recursos informacionais relevantes em um repositório institucional que podem ser recuperados com um determinado perfil de busca (HJØRLAND, 1997, p. 36, tradução nossa). A precisão e a revocação são consideradas as métricas mais importantes e utilizadas para avaliação da qualidade da recuperação de um sistema de recuperação da informação. Os valores de precisão e revocação são atualmente medidas de avaliação padrão na literatura para a avaliação de sistemas de recuperação da informação e permitem avaliar, de forma quantitativa, a qualidade dos resultados recuperados (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 115).

Na perspectiva de Hjørland (1997, p. 141; 142), na Ciência da Informação há uma “lei” controversa que postula um limite fundamental para a recuperação da informação: a relação inversa entre precisão e revocação na recuperação da informação. Postulada no início dos sistemas de recuperação automatizados, na opinião do autor é equivocada, pois o “propósito básico no desenvolvimento de sistemas de recuperação é fazer sistemas que melhorem tanto a revocação quanto a precisão”. Fugmann (1993, p. 203) também refuta a “relação inversa de revocação e precisão, pois em termos práticos, “essa ‘lei empírica’ tem sido frequentemente usada como uma desculpa para o baixo desempenho dos sistemas de informação com relação a evitar a perda de informações” ou para evitar o ruído de respostas irrelevantes.

Considerando-se que a revocação é a recuperação de recursos informacionais pertinentes do repositório institucional, o *coeficiente de revocação* é a medida com que os recursos informacionais pertinentes são recuperados. Tomada isoladamente, a revocação permite um panorama incompleto do nível de eficiência de uma busca, atrelando-a, então, a um *coeficiente de precisão* “para se ter uma indicação do grau de eficiência ou discriminação atingido por uma consulta”. No entanto, o estabelecimento de uma avaliação da recuperação da informação em repositório institucional para uma amostra de buscas pautando-se em métricas de recuperação como revocação e precisão, em si mesmas, precisam estar inseridas em um contexto mais amplo e político de implementação de melhorias, verificando os motivos de ocorrência das falhas nas buscas (LANCASTER, 1996, p. 188; 191).

Hjørland (1997, p. 140, tradução nossa) defende a questão da incerteza fundamental da busca de informações. Para o autor, a busca de informação é caracterizada na teoria ou na prática pela *incerteza*. Nesta visão, nem mesmo o perfil de pesquisa mais sofisticado pode garantir que todos os dados importantes em um repositório institucional sejam encontrados, sendo que o único método de busca que poderia garantir a adequada recuperação seria verificar todos os recursos informacionais, um após o outro, método irrealista para a maioria dos repositórios institucionais. Para o autor, “o entendimento de que a busca de informação é um processo prático e fundamentalmente incerto é baseado na avaliação de que questões não-problemáticas [...] são exceções, enquanto a maioria das questões mais importantes e que buscam informações pertencem à categoria problemática onde interpretação e a incerteza é a regra”. Contudo, embora a busca de informações em um repositório institucional seja incerta, “é um erro grave traçar uma conclusão agnóstica com base na incerteza fundamental, ou seja, deixar de trabalhar para melhorar as opções de busca de informações simplesmente porque não se pode alcançar um resultado 100% perfeito”, considerando-se que “mesmo em bancos de dados pobres, há muito que se pode fazer para melhorar as pesquisas, e há muitas possibilidades de melhorar as condições de pesquisa para futuros usuários no estágio de construção do banco de dados”.

Adicionalmente, outras medidas, denominadas de *medidas orientadas ao usuário*, foram propostas no intuito de tratar a questão de que usuários diferentes podem ter uma interpretação diferente sobre qual recurso informacional é relevante e qual não é, sendo complementares às clássicas medidas de revocação e precisão. Dentre as medidas orientadas ao usuário, a *taxa de cobertura* refere-se à “fração dos documentos conhecidos e relevantes que estão no conjunto resposta”, a *taxa de novidade* a “fração dos documentos relevantes no conjunto resposta que não são conhecidos pelo usuário”. Em continuidade, uma taxa de cobertura alta indica que o sistema está encontrando a maioria dos recursos informacionais relevantes para o usuário e uma taxa de novidade alta indica que o sistema está revelando ao usuário muitos novos recursos informacionais previamente desconhecidos. Já a *revocação relativa* é definida como “a razão entre o número de documentos relevantes encontrados pelo sistema e o número de documentos relevantes que o usuário esperava encontrar”, enquanto que o *esforço da revocação* “é a razão entre o número de documentos relevantes que o usuário esperava encontrar e o número de documentos examinados na tentativa de encontrar os documentos relevantes esperados” (BAEZA-YATES; RIBEIRO-NETO, 2013, p. 121; 122).

3.3 Avaliação da indexação de assuntos: em busca de indicadores para a melhoria da representação e da recuperação por assuntos em repositórios institucionais

O termo avaliação tem sido definido de diversas formas e por diferentes autores. Segundo o Dicionário Houaiss da Língua Portuguesa (2009, p. 227) a ação de avaliar ou o conceito de avaliação estão relacionados à verificação, à determinação da qualidade, extensão ou intensidade de algo. De acordo com Navarrete e Quechol (2011, p. 90, tradução nossa), “atualmente qualquer manual de gestão organizacional conta com duas palavras que se repetem com frequência e que estão intimamente relacionadas: qualidade e avaliação” e que nenhuma delas estão a parte do mundo das bibliotecas. De fato, há tempos os bibliotecários têm mostrado preocupação pela qualidade do funcionamento de suas instituições.

Os primeiros estudos sobre avaliação em bibliotecas remontam à década de 1950, voltados principalmente para elementos quantitativos, tal como número de empréstimos realizados. Contudo, estudos que considerassem a visão dos usuários eram inexistentes, o que não permitia conhecer o nível de satisfação dos mesmos em relação aos serviços oferecidos pelas bibliotecas. Especificamente no contexto de bibliotecas universitárias, a avaliação contempla duas principais frentes: enquanto a avaliação *objetiva* consiste na coleta de dados estatísticos sobre os processos e serviços, bem como as opiniões e sugestões de seus colaboradores; a avaliação *subjetiva* está pautada na opinião dos usuários para conhecer a percepção da mesma e de seus serviços (NAVARRETE; QUECHOL, 2011, p. 7). Considerando-se a relação estabelecida entre a sociedade e sua necessidade informacional, a biblioteca universitária desempenha um papel de intermediária entre os recursos informacionais e sua comunidade acadêmica, logo, “o objetivo de qualquer avaliação é determinar até que ponto a biblioteca cumpre esta função de intermediação” (NAVARRETE; QUECHOL, 2011, p. 3, tradução nossa).

A avaliação permite ao bibliotecário a percepção de uma visão sistêmica da realidade em foco, verificando problemas, permitindo tomadas de decisões ponderadas e o direcionamento correto da atividade investigada (MACIEL, 1997, p. 19). Além disso, os resultados da avaliação em bibliotecas universitárias são de interesse das próprias universidades nas quais estão inseridas, visto que a avaliação estratégica permite a melhoria contínua de indicadores de rendimento específicos, cujos comportamentos sugerem limites claros e precisos para a gestão da qualidade de seus produtos e serviços (NAVARRETE;

QUECHOL, 2011, p. 4). Para Lancaster (1996, p. 1), “uma avaliação é feita não como um exercício intelectual, mas para reunir dados úteis para atividades destinadas a solucionar problemas ou tomar decisões”. Nesse cenário, a biblioteca universitária constitui

[...] um centro de recursos para a aprendizagem, o ensino e a pesquisa que de maneira criativa participa no desenvolvimento das competências intelectuais da sociedade. Competitivamente identifica, reúne, analisa e sistematiza um universo de informação previamente selecionado, adquirido e disseminado em um portfólio de serviços e produtos de informação para a satisfação das necessidades expressas e implícitas de seus diversos grupos de usuários [...]. A qualidade dos serviços e produtos do ponto de vista da eficiência e da eficácia são aspectos que periodicamente devem ser avaliados, para assegurar seu impacto e relevância nos programas educativos e de pesquisa (SÁNCHEZ-AMBRIZ, 2011, p. 127-128, tradução nossa).

Corroborando com esta visão, Lancaster (1996, p. 2) afirma que a biblioteca universitária se caracteriza como uma interface entre os recursos informacionais disponíveis e sua comunidade de usuários. Nesse ponto de vista, “qualquer avaliação que a biblioteca seja submetida deve se preocupar em determinar em que medida ela desempenha com êxito essa função de interface”.

Em consonância com os primeiros estudos sobre avaliação de bibliotecas, as primeiras pesquisas científicas sobre a avaliação da indexação de assuntos remontam à década de 1950, principalmente no contexto norte-americano. Os resultados provenientes destes estudos permitem analisar a atividade de indexação em um sistema de recuperação da informação e são úteis tanto aos produtores da informação ao propiciar o conhecimento dos aspectos reais da realização da atividade e propor melhorias quanto aos usuários, ao conhecerem o real funcionamento do sistema de recuperação e utilizá-lo do melhor modo possível (ABAD GARCÍA; PÉREZ; BENAVENT, 1998, p. 397). Tais enfoques tem sido em *abordagem teórica*, em busca de respostas sobre as diferenças entre os resultados obtidos na indexação ou de *abordagem prática*, que busca a quantificação da consistência entre duas ou mais indexações por meio de fórmulas matemáticas (FUJITA; RUBI; GIL LEIVA, 2008, p. 235).

De acordo com Gil Leiva (2008, p. 70), os elementos que caracterizam a qualidade tanto do processo em si quanto do resultado da indexação de assuntos são a *exaustividade* (conceitos que caracterizam o conteúdo completo de um recurso informacional); a *especificidade* (relação exata entre a unidade conceitual e o termo escolhido para representá-la); a *correção* (ausência de erros de inclusão ou omissão); e a *consistência* (nível de

coincidência entre duas ou mais indexações). Assim, a qualidade da indexação de assuntos está relacionada diretamente com o equilíbrio conceitual entre estas quatro características.

Dentre estes aspectos, a consistência (uniformidade, coerência, homogeneidade ou concordância) é notadamente o elemento mais presente na literatura sobre avaliação da indexação de assuntos, sendo definida como “o nível de concordância na representação da informação essencial de um documento por meio de um conjunto de termos de indexação selecionados por cada um dos indexadores de um grupo” (ZUNDE; DEXTER, 1969, p. 259); uniformidade ou homogeneidade com que um recurso informacional é descrito da mesma maneira (MOREIRO, 2002, p. 59) ou “uma medida quantitativa do grau em que dois ou mais indexadores identificam os conceitos importantes contidos em um documento e representam estes conceitos usando códigos/termos idênticos” (LEONARD, 1977, p. 2, tradução nossa). Em síntese, no contexto dos repositórios institucionais a consistência é o nível de concordância dos assuntos atribuídos a um determinado recurso informacional entre dois repositórios institucionais, entre dois bibliotecários catalogadores/indexadores ou entre o mesmo catalogador/indexador em diferentes momentos no tempo.

Em repositórios institucionais, os termos de indexação desempenham o papel de representação deste conteúdo, tendo como resultado um *índice de consistência* (GIL LEIVA; RUBI; FUJITA, 2008, p. 234). O método para avaliar a amplitude destas variações consiste em obter a medida da consistência existente na realização da indexação documental. Quando as comparações são realizadas entre o resultado de um indexador na análise de um mesmo recurso informacional em períodos diferentes, é conhecido como *consistência intraindexador*; enquanto que a confrontação entre o resultado de vários indexadores na análise de um mesmo recurso informacional trata-se da *consistência interindexador* (ABAD GARCÍA; ABAD PÉREZ; BENAVENT, 1998; GIL LEIVA, 2008).

Na visão de Fugmann (1993, p. 94), a consistência é comumente considerada como um critério para avaliar a qualidade da indexação, prevalecendo quando indexadores diferentes produzem o mesmo resultado na análise de um recurso informacional ou quando o mesmo indexador analisa o mesmo recurso informacional repetidamente. No entanto, a indexação pode ser altamente consistente e, ao mesmo tempo, extremamente ruim” (FUGMANN, 1993, p. 94). Nesse sentido,

[...] a consistência de indexação não constitui um critério útil para indexar a qualidade, simplesmente porque mesmo o pior procedimento de indexação pode

ser um procedimento de consistência máxima. Em vez disso, é o grau de previsibilidade que é mais informativo sobre a qualidade da indexação (juntamente com a fidelidade representacional). [...] A consistência tem alguma relação, na verdade, apenas parcial, com a qualidade da indexação (FUGMANN, 1993, p. 97, tradução nossa).

No bojo científico-cultural da Ciência da Informação, o processo de indexação de assuntos permite representar *exatamente* o conteúdo de um recurso informacional. Sendo assim, idealmente, se dois indexadores analisam o mesmo recurso informacional, os descritores serão os mesmos para ambas indexações, ocorrendo um *alto nível de concordância ou consistência* (WOLFRAM; OLSON; BLOOM, 2009, p. 1995). Não obstante, esta concepção simplista do processo de indexação de assuntos vai ao encontro com os estudos mais recentes e uma mudança de paradigma de outra concepção, em que o processo de indexação é visto sob a ótica de um sistema complexo, permeado por diversos atores sociais que interagindo entre si podem alterar o resultado do mesmo.

Nesta ótica, o conjunto teórico-prático científico sobre a consistência entre indexadores, em torno do mesmo indexador em diferentes momentos do tempo ou, mais recentemente, de indexadores humanos com indexação automática permitiram o desenvolvimento de uma série de medidas para avaliar a consistência entre indexadores. Porém, de forma geral, os resultados destas pesquisas apresentaram elevados índices de inconsistência (WOLFRAM; OLSON; BLOOM, 2009, p. 1995). Enquanto alguns autores advogam a existência de uma “correlação experimental entre a consistência da indexação e a eficácia de recuperação dos documentos” (HUGHES; RAFFERTY, 2011, p. 9, tradução nossa), para outros a adequação e qualidade deste processo são de difíceis medições e avaliações, considerando-se a inexistência de uma forma *unívoca* de indexar devido à existência de diversas variáveis, como o próprio indexador, o tipo de recurso informacional, as características e particularidades do contexto e a política de indexação da instituição.

Tendo em vista que a solução de muitas das problemáticas que cercam a representação de assuntos perpassa pelo avanço de pesquisas acerca da consistência da indexação, entendendo-se “qualidade da indexação” em termos de eficácia da recuperação (COOPER, 1969; LEONARD, 1977), torna-se relevante o encaminhamento de pesquisas destinadas à temática, tendo a avaliação da consistência como sinônimo de confiabilidade do processo. Embora alguns estudos na área indiquem a dificuldade em medir a qualidade da indexação,

muito em razão da singularidade do processo (FUNK; REID; MNGOOGAN, 1983), outros sinalizam que alcançar a consistência é contribuir para uma recuperação eficiente (SAARTI, 2002). Com base em tal panorama, ressalta-se a necessidade de mais estudos sobre a metodologia da avaliação da indexação de assuntos, possibilitando um maior destaque na Organização do Conhecimento e contribuindo com melhorias na indexação de assuntos de acordo com o contexto elencado como objeto de estudo e os diferentes aspectos metodológicos adotados.

Referente aos diferentes métodos/abordagens metodológicas presentes na literatura em Organização do Conhecimento, estes comportam dois aspectos principais: intrínseco e extrínseco, em que cada um possui subcategorias de análise, sintetizadas por Gil Leiva (2008). No primeiro aspecto, a **Avaliação intrínseca** está relacionada com o resultado da indexação, com a finalidade de conhecer a sua qualidade:

[...] um conjunto de tarefas centradas no resultado da indexação (descritores, cabeçalhos, subcabeçalhos ou identificadores), com a finalidade de conhecer sua qualidade. A avaliação intrínseca de indexação pode ser qualitativa, isto é, por meio de valorização e consenso entre [profissionais] experientes, ou quantitativa, através de fórmulas (GIL LEIVA, 2008, p. 385, tradução nossa).

A **Avaliação intrínseca** em repositórios institucionais subdivide-se em dois tipos:

- ***Avaliação intrínseca qualitativa***: caracteriza-se por analisar os componentes inerentes que proporcionam a qualidade da indexação, como:

Exaustividade: quando se extrai todos os conceitos caracterizadores do conteúdo integral dos recursos informacionais do repositório institucional;

Especificidade: que exista uma relação exata entre as unidades conceituais escolhidas e o termo ou os termos elegidos para representá-la no repositório institucional;

Correção: que não ocorram erros de inclusão (um termo que não procede) e erros de omissão (a exclusão de um termo relevante) no repositório institucional;

Perspectiva do usuário: que se considere os interesses e perspectivas dos potenciais usuários, para o qual, tem que responder à pergunta se os termos ou assuntos designados, em seu conjunto ou separado, são legítimos para representar o conteúdo dos recursos informacionais no repositório institucional.

- **Avaliação intrínseca quantitativa**: refere-se a uma reindexação de um conjunto de recursos informacionais do repositório institucional repetindo-se, na medida do possível, as condições em que foi produzida a primeira indexação, para conseguir índices de consistência entre duas indexações por meio de métodos ou fórmulas matemáticas. A Avaliação intrínseca subdivide-se em dois tipos:

- **Avaliação intrínseca quantitativa mediante a interconsistência**: compara a indexação de um mesmo recurso informacional do repositório institucional realizada por dois ou mais profissionais (consistência interindexador);

- **Avaliação intrínseca quantitativa mediante a intraconsistência**: compara a indexação de um mesmo recurso informacional do repositório institucional realizada pelo mesmo profissional em diferentes momentos no tempo (consistência intraindexador).

Cabe destacar, nesta abordagem, o estudo de Tartarotti (2014) que investigou, por meio da *Avaliação intrínseca quantitativa mediante a intraconsistência/Avaliação intrabibliotecário* realizada quando o mesmo profissional indexa o mesmo recurso informacional em momentos diferentes no tempo, e, se possível, sob as mesmas condições em que foi realizada a primeira indexação, a comparação da atuação do bibliotecário em catálogo coletivo *online* e em base de dados especializada.

A consistência na indexação pode ser estudada como referência a um único indexador ou a vários. Quando um profissional indexa um mesmo documento, em diferentes momentos temporais, falamos de intraconsistência ou consistência intraindexador. Em segundo lugar, quando vários profissionais indexam um mesmo documento visando comparar seu resultado ou quando se enfrentam duas indexações de diferentes indexadores, falamos de interconsistência ou consistência interindexador (GIL LEIVA; RUBI; FUJITA, 2008, p. 234).

Portanto, “enquanto a Avaliação intrínseca qualitativa baseia-se em uma investigação da atuação profissional dos indexadores, a Avaliação intrínseca quantitativa caracteriza-se pela aplicação de fórmulas matemáticas” (TARTAROTTI; DAL’EVEDOVE, FUJITA; 2017, p. 4).

O segundo aspecto da avaliação da indexação de assuntos (Gil Leiva, 2008), a **avaliação extrínseca** mede o resultado da indexação para compará-lo com a indexação de outros repositórios institucionais que igualmente indexaram o mesmo recurso informacional

(interconsistência) ou mediante a recuperação (exaustividade e precisão). A Avaliação extrínseca subdivide-se em dois tipos:

- **Avaliação extrínseca mediante a interconsistência:** compara a indexação de um mesmo recurso informacional realizada por dois ou mais repositórios institucionais;
- **Avaliação extrínseca mediante a recuperação:** compara duas indexações procedentes de um mesmo repositório institucional ou dois repositórios institucionais diferentes.

De forma sistemática, apresenta-se as diferentes abordagens da avaliação da indexação de assuntos no contexto de repositórios institucionais, com destaque para as duas abordagens utilizadas na pesquisa (Quadro 4):

Quadro 4. Diferentes abordagens de avaliação da indexação de assuntos em repositórios institucionais

Avaliação intrínseca	Avaliação intrínseca qualitativa	
	Avaliação intrínseca quantitativa	Avaliação intrínseca quantitativa mediante a interconsistência
		Avaliação intrínseca quantitativa mediante a intraconsistência
Avaliação extrínseca	Avaliação extrínseca mediante a interconsistência	
	Avaliação extrínseca mediante a recuperação	

Fonte: Adaptado de Gil Leiva (2008, p. 385-392).

Em síntese, a avaliação da indexação de assuntos caracteriza-se como uma abordagem investigativa que visa à análise do processo de indexação de assuntos em um determinado contexto, pautando-se em procedimentos específicos, que podem ser qualitativos ou quantitativos, de acordo com o aspecto do processo de indexação que se pretenda verificar.

Embora a aplicabilidade tanto teórica quanto prática da metodologia da avaliação da indexação de assuntos seja relevante para a Organização do Conhecimento e para os profissionais em busca de elementos que norteiem esta atividade, propiciando um panorama da indexação de assuntos, ainda existe uma lacuna científica no rol de publicações sendo incipientes os estudos que se dedicam a esta temática, especialmente no contexto brasileiro. Tal situação é reforçada pela ausência de um manual de avaliação da indexação de assuntos

que forneça diretrizes para sua aplicabilidade e permita a verificação tanto na abordagem da indexação de assuntos pela representação quanto na recuperação de recursos informacionais em sistemas de recuperação.

4 PROCEDIMENTOS METODOLÓGICOS

No intuito de possibilitar a realização prática da presente Tese, nesta seção são apresentados os procedimentos metodológicos norteadores da mesma. Neste sentido, para que a pesquisa pudesse ser realizada de acordo com o objetivo geral - contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, procedeu-se à realização de cada objetivo específico proposto. Considerando-se a dimensão da pesquisa ao lidar com a representação e a recuperação por assuntos em repositórios institucionais, tornou-se necessário o delineamento de procedimentos metodológicos pautados em duas principais abordagens: **teórico-aplicada** e **exploratório-descritiva**, delineadas a seguir.

4.1 Abordagem teórico-aplicada

Para atender ao **primeiro objetivo específico** da pesquisa - realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias - foram realizadas considerações teóricas em torno dos temários abordados, isto é, em **abordagem teórico-aplicada**.

A revisão de literatura possibilitou esclarecer os próprios objetivos da pesquisa; proporcionar a profundidade e o conhecimento sobre o temário necessários e formar um quadro teórico para a parte exploratória da mesma, contribuindo para a compreensão do assunto investigado (PICKARD, 2013, p. 25-26). O levantamento foi realizado em fontes bibliográficas nacionais e internacionais, tais como: livros, periódicos, *proceedings*, teses e dissertações, bases de dados e Internet, destacando-se os acervos de outras instituições conveniadas à Universidade Estadual de Campinas (UNICAMP) que possibilitassem o acesso

aos materiais de interesse, especialmente a Universidade de São Paulo (USP) e Universidade Estadual Paulista (UNESP) (convênio CRUESP¹⁷).

4.2 Abordagem exploratório-descritiva

A abordagem **exploratório-descritiva** refere-se ao delineamento de informações e formulação de hipóteses sobre o temário e ao mesmo tempo identificação e análise das características que se relacionam com o fenômeno ou processo em questão.

A utilização de tal abordagem justifica-se para cumprirmos o **segundo e terceiro objetivos da pesquisa**, que são: investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como **metodologia qualitativa** com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores, com coleta de documentação; e verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária por meio da aplicação da avaliação da indexação de assuntos como **metodologia quantitativa**. Para tanto, a avaliação da indexação de assuntos foi realizada em duas principais abordagens: a) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e b) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*.

Destaca-se, neste momento, o aspecto inovador da presente pesquisa na aplicabilidade da metodologia da avaliação da indexação de assuntos como instrumento de coleta de dados para verificação da indexação de assuntos na recuperação da informação em repositórios institucionais tanto no Brasil quanto em âmbito internacional. Cabe esclarecer que, por se tratar de abordagem exploratória, a pesquisa limitou-se a investigar as questões relacionadas ao temário, realizando um estudo diagnóstico da situação da indexação de assuntos em

¹⁷ Instituído pela Resolução 149/99, o CRUESP/BIBLIOTECAS iniciou suas atividades em 1999 como Grupo de Estudos, tendo por objetivo a integração dos Sistemas de Bibliotecas da USP, UNESP e UNICAMP. O convênio permite o serviço denominado de *Empréstimo Entre Bibliotecas* (EEB), possibilitando o intercâmbio de recursos informacionais entre as 3 instituições.

repositórios institucionais no contexto de bibliotecas universitárias, sem pretender oferecer respostas finais e conclusivas em torno da questão, contribuindo para uma melhor compreensão do problema, que apresenta uma lacuna no campo científico, e sendo base para futuras investigações ao delinear uma metodologia de avaliação da indexação de assuntos em repositórios institucionais.

[...] nenhum método de coleta de dados é adequado a todos os problemas, pois cada instrumento de coleta de dados apresenta um tipo de limitação. Por essa razão, a combinação de diferentes formas de coleta de dados pode trazer melhores resultados, pois garantirá a credibilidade desses dados pela 'triangulação', isto é, pela possibilidade de confirmação de dados em outras fontes (ALMEIDA, 2005, p. 58).

De acordo com Rice e Southall (2016, p. 3), a rigorosa aplicação do método científico nas ciências sociais resultou no surgimento de métodos quantitativos aplicados a amostras de populações para descrever, explicar e prever comportamentos, tendo em vista que os resultados dos métodos qualitativos raramente são generalizáveis para uma população, já que o objeto de estudo das ciências sociais é o ser humano e seu comportamento em sociedade, sejam individuais ou coletivos, e, portanto, difíceis de definição e previsão. Neste cenário, a divisão entre os métodos quantitativos e qualitativos eram claramente visíveis. Na contemporaneidade, porém, percebe-se que ambas abordagens metodológicas são válidas e mesmo simbióticas, e, portanto, complementares. São diversas as possibilidades de avaliação dos processos e serviços no contexto de bibliotecas, com três principais abordagens: métodos quantitativos, métodos qualitativos e métodos integrais (ARRIOLA NAVARRETE; TECUATL QUECHOL, 2011, p. 14).

A seguir, apresenta-se a definição dos repositórios institucionais elencados como universo de estudo na pesquisa, seguida da entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais como abordagem qualitativa e da avaliação da indexação de assuntos como abordagem quantitativa, em duas principais abordagens: a) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e b) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*.

4.2.1 Caracterização do universo de pesquisa: os repositórios institucionais do CRUESP

Os repositórios institucionais do CRUESP, representados pelas três universidades: USP, UNESP e UNICAMP vem se consolidando no contexto brasileiro como referência em sistemas de recuperação da informação que realizam a gestão e disseminação dos recursos informacionais produzidos por suas respectivas comunidades acadêmicas, atendendo à demanda da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) junto ao CRUESP.

No intuito de caracterizar o universo dos repositórios institucionais delimitado na abordagem exploratória-descritiva, apresenta-se, de forma mais específica, o Repositório Institucional da USP: *Biblioteca Digital da Produção Intelectual da USP* (BDPI); o Repositório Institucional da UNESP: *Repositório Institucional UNESP*; o Repositório Institucional da UNICAMP: *Repositório da Produção Científica e Intelectual da UNICAMP* e, em um contexto mais amplo, o *Repositório da Produção Científica do CRUESP*, com base nas informações disponibilizadas nos próprios sites dos repositórios institucionais e nas coletas de documentação. Como plataforma de software, utilizam o DSpace¹⁸.

A *Biblioteca Digital da Produção Intelectual da USP* (BDPI)¹⁹ (Figura 2), repositório institucional da USP, criado em 22 de outubro de 2012, reúne a produção intelectual (científica, artística, acadêmica e técnica) da universidade, de acordo com os parâmetros da *Política de Informação da USP* definida na *Resolução nº 6.444* (ANEXO B), de outubro de 2013, que dispõe sobre diretrizes e procedimentos visando promover e assegurar a coleta, tratamento e preservação da produção intelectual gerada na universidade e pelos programas de pós-Graduação, bem como sua disseminação e acessibilidade para a comunidade acadêmica.

¹⁸ O DSpace é um software aberto desenvolvido por um esforço colaborativo entre a MIT Libraries e a Hewlett-Packard, cujo código-fonte é disponibilizado publicamente com o propósito de incentivar a formação de uma comunidade *open source* em torno do software. Tais comunidades são responsáveis por adicionar recursos e melhorar as diferentes funções do sistema, bem como por adaptá-lo para atender aos requisitos e necessidades de instituições específicas. A terminologia interna e o modelo de dados são baseados em conceitos fornecidos pelo Modelo OAIS (*Open Archival Information System*), em português, SAAI (Sistema Aberto de Arquivamento de Informação) Outros softwares para a criação de bibliotecas digitais são o *Greenstone*, *Fedora*, *Eprints*, ODL e o *5S Suite* (GONÇALVES, 2011, p. 727-728).

¹⁹ Disponível em: <http://www.producao.usp.br>.

Figura 2. Página inicial do Repositório Institucional da USP

Fonte: Dados da pesquisa.

O *Repositório Institucional UNESP*²⁰ (Figura 3) objetiva “armazenar, preservar, disseminar e possibilitar o acesso aberto, como bem público global, à produção científica, acadêmica, artística, técnica e administrativa da universidade” (UNESP, 2019). É norteado pelo *Regulamento Interno do Repositório Institucional UNESP* (ANEXO C).

Figura 3. Página inicial do Repositório Institucional da UNESP

Fonte: Dados da pesquisa.

²⁰ Disponível em: <https://repositorio.unesp.br>.

O *Repositório da Produção Científica e Intelectual da UNICAMP*²¹ (Figura 4) é o instrumento oficial da universidade para a coleta, organização, disseminação e preservação de todo o conhecimento produzido por sua comunidade acadêmica: docentes, pesquisadores, discentes de graduação, discentes de pós-graduação e servidores técnicos administrativos, tendo como objetivos “proporcionar acesso aberto e público à produção científica e intelectual da universidade, propiciando o aumento de sua visibilidade, acessibilidade e difusão; facilitar a gestão e o acesso à informação sobre a produção científica e intelectual da UNICAMP, por meio da oferta de indicadores confiáveis e validados; e integrar-se a um conjunto de iniciativas nacionais e internacionais, por meio de padrões e protocolos de integração qualificados e normalizados” (UNICAMP, 2019). Em 02 de novembro de 2015, ocorreu o lançamento oficial do Repositório da Produção Científica e Intelectual da UNICAMP, cuja formalização de criação está pautada na *Resolução GR-013/2015*, 06 de julho de 2015 (ANEXO D).

Figura 4. Página inicial do Repositório Institucional da UNICAMP



Fonte: Dados da pesquisa.

Em um contexto mais amplo, o *Repositório da Produção Científica do CRUESP*²² (Figura 5), lançado em 6 de outubro de 2013 durante a sessão de abertura da Conferência Luso-

²¹ Disponível em: <http://www.repositorio.unicamp.br>.

²² Disponível em: <http://www.repositorio.cruesp.sp.gov.br>.

Brasileira de Acesso Aberto (CONFOA), objetiva “reunir, preservar e proporcionar acesso aberto, público e integrado à produção científica de docentes, pesquisadores, alunos e servidores” do universo acadêmico da USP, UNESP e UNICAMP. Sob a coordenação das referidas universidades, foi idealizado para atuar, de forma compartilhada, cooperativa e padronizada com uma metodologia de trabalho comum, na ampliação da visibilidade e acessibilidade da produção científica, ampliando o intercâmbio com outras instituições nacionais e internacionais e a democratização e estímulo do compartilhamento do conhecimento gerado, estendendo e retornando à sociedade o investimento nelas realizado. Gerado a partir dos repositórios institucionais das três universidades estaduais paulistas, o Repositório da Produção Científica do CRUESP utiliza também o DSpace e adota padrões e normas internacionais de interoperabilidade e normalização. Além disso, a integração por meio do metabuscador *Primo*²³ possibilita a busca e a descoberta da produção do CRUESP a partir de uma interface única (Portal Repositório CRUESP, 2019).

Figura 5. Página inicial do Repositório Institucional do CRUESP

REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA DO CRUESP

O Repositório da Produção Científica do CRUESP (Conselho de Reitores das Universidades Estaduais Paulistas) tem por objetivo reunir, preservar e proporcionar acesso aberto, público e integrado à produção científica de docentes, pesquisadores, alunos e servidores da USP, Unicamp e Unesp.

A iniciativa amplia a visibilidade e acessibilidade aos resultados das pesquisas realizadas nas universidades, potencializando, desta forma, o intercâmbio com outras instituições nacionais e internacionais. Além disso, democratiza e estimula o compartilhamento do conhecimento gerado, estendendo e retornando à sociedade o investimento nelas realizado. Partindo de uma metodologia comum e atuando de forma compartilhada e cooperativa sob coordenação dos sistemas de bibliotecas das referidas universidades, conta com o apoio dos pró-reitores de pesquisa, seus conselheiros científicos, e incentivo da FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo).

O Repositório da Produção Científica do CRUESP foi gerado a partir dos repositórios institucionais das três universidades estaduais paulistas, os quais utilizam o software livre DSpace (desenvolvido pelo MIT e HP) e adotam padrões e normas internacionais de interoperabilidade e normalização. A integração por meio do metabuscador Primo (Web Scale Discover System da

REPOSITÓRIOS INSTITUCIONAIS USP, UNICAMP E UNESP

Biblioteca Digital da Produção Intelectual da USP

A Biblioteca Digital da Produção Intelectual da Universidade de São Paulo (BDPI), inaugurada em 22 de outubro de 2012, é o repositório institucional da produção intelectual (científica, artística, acadêmica e técnica) da USP.

Trata-se de um sistema de gestão e disseminação que tem como objetivos:

- aumentar a visibilidade, acessibilidade e difusão dos resultados da atividade acadêmica e de pesquisa da USP por meio da coleta, organização e preservação em longo prazo;
- facilitar a gestão e o acesso à informação sobre a produção intelectual da USP por meio da oferta de indicadores confiáveis e validados;
- integrar-se a um conjunto de iniciativas nacionais e internacionais, por meio de padrões e protocolos de integração qualificados e normalizados.

Além da integração com a Biblioteca Virtual da FAPESP, portais de revistas nacionais e internacionais de acesso aberto e demais portais e bibliotecas digitais do SIBUSP, relaciona-se de forma coordenada e cooperativa com a Biblioteca Digital da Produção Intelectual e Científica da Unicamp e com o Repositório Institucional da Unesp.

Fonte: Dados da pesquisa.

4.2.2 Abordagem qualitativa: entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais

²³ Serviço de descoberta em rede, da *ExLibris*. O princípio básico de um serviço de descoberta é fornecer um índice unificado de metadados previamente coletados, permitindo uma busca unificada em substituição à busca federada, isto é, à busca em cada uma das bases provedoras de conteúdos utilizadas (PAVÃO, 2014, p. 16).

A utilização de entrevista justifica-se para atingir-se o **segundo objetivo específico da pesquisa** - investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como **metodologia qualitativa** com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores, com coleta de documentação.

A entrevista caracteriza-se como uma técnica de coleta de dados em que um pesquisador realiza um rol de perguntas aos participantes²⁴, geralmente face a face. É utilizada para obtenção de dados qualitativos, descritivos e detalhados específicos de um determinado indivíduo e quando a natureza dos dados é complexa para ser feita e respondida com facilidade, permitindo, ao contrário de questionários, um maior nível de interação entre o pesquisador e o participante (PICKARD, 2013, p. 323; 196). Possibilita, ainda, o levantamento de dados relativos ao histórico da instituição, bem como à estrutura funcional e aos serviços (ALMEIDA, 2005, p. 55) e tem sido amplamente utilizada em pesquisas em torno da informação no contexto de bibliotecas (PICKARD, 2013, p. 195). Embora seja semelhante ao questionário, o maior benefício da entrevista são as pistas orais da observação do entrevistado e a maior possibilidade de coleta de dados (PICKARD, 2013, p. 199).

A entrevista pode ser de dois tipos: *entrevista estruturada* e *entrevista não-estruturada*. Cada modalidade de entrevista depende do objetivo de pesquisa, da experiência do pesquisador, da natureza do temário abordado, do acesso aos entrevistados, do tempo disponível para a coleta de dados e do tipo de dados a serem coletados, visando obter respostas à pergunta de pesquisa (PICKARD, 2013, p. 205; 198). A *entrevista estruturada* tem por base um cronograma de entrevista formal, geralmente um questionário administrado pelo pesquisador, que especifica a formulação e ordenação das questões a serem realizadas aos entrevistados e subdivide-se em: *entrevista aberta* - respostas de forma livre e *entrevista fechada* - respostas com base em um conjunto de respostas alternativas (PICKARD, 2013, p. 323; 199). Já a *entrevista não-estruturada* tem por base perguntas abertas, que possibilitam

²⁴ Nesta pesquisa optamos pelo termo “participante(s)”, em detrimento do termo “sujeito(s)” ou “sujeito(s) participante(s)”, tendo em vista a Resolução 466/2012 do Ministério da Saúde, atualmente documento regulamentador da pesquisa em seres humanos no Brasil. No referido documento, disponível em http://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html, não há menção ao termo “sujeito”, sendo utilizado o termo “participante”.

ao entrevistado considerável liberdade para expor livremente o assunto e influenciar a direção da entrevista, pois inexistente um plano predeterminado sobre as informações específicas a serem coletadas dos entrevistados (PICKARD, 2013, p. 323).

De acordo com Pickard (2013, p. 196), Kvale (1996) estabeleceu sete etapas de um processo de entrevista: *Tematização*: estabelecimento do temário a ser abordado na entrevista, de acordo com os objetivos da pesquisa; *Projeto*: definição de um roteiro (*script*) de temas/questões que se pretende abordar na entrevista; *Entrevista*: realização da entrevista propriamente dita; *Transcrição*: realização da transcrição e análise inicial das entrevistas realizadas; *Análise*: realização da análise dos dados coletados durante a entrevista; *Verificação*: confirmação da interpretação dos dados do pesquisador com o entrevistado; e *Relato*: apresentação e discussão dos resultados da entrevista. Tais etapas não são, necessariamente, realizadas de forma linear ou sequenciais.

Para a aplicação da entrevista para coleta de dados sobre os repositórios institucionais, os participantes da pesquisa foram definidos em duas principais categorias: *gestores e bibliotecários catalogadores-indexadores* dos repositórios institucionais. Como forma de facilitar a identificação das atividades realizadas em cada fase de aplicação da entrevista na pesquisa, foi elaborado um roteiro, baseado na metodologia do Protocolo Verbal (PV), para a aplicação de entrevista de diagnóstico organizacional, cuja aplicação possibilitou caracterizar, de forma comparada, o contexto organizacional e acadêmico do universo de pesquisa da USP, UNESP e UNICAMP: *procedimentos anteriores à aplicação da entrevista, procedimentos durante a aplicação da entrevista e procedimentos posteriores à aplicação da entrevista*:

I Procedimentos anteriores à aplicação da entrevista

- ***Definição do universo da pesquisa***: foram elencados os contextos organizacionais da USP, UNESP e UNICAMP para a coleta de dados sobre os repositórios institucionais por meio de entrevista semiestruturada;
- ***Seleção dos participantes***: a seleção dos participantes para a aplicação da entrevista resultou na escolha de um total de 9 participantes, a saber: 3 gestores dos repositórios institucionais; 3 bibliotecários catalogadores-indexadores, sendo um 1 participante de

cada categoria, de cada universidade; e, de forma complementar, 3 bibliotecários que participaram da implementação do Repositório Institucional UNESP²⁵. Foi realizada uma conversa informal com cada um dos participantes por telefone ou e-mail, resultando na aceitação e definição das datas para a realização das entrevistas;

- **Definição da entrevista:** anteriormente à aplicação formal da entrevista, foi realizado um *pré-teste* (piloto) de entrevista com um gestor e um bibliotecário catalogador-indexador do Repositório Institucional da Universidade Tecnológica Federal do Paraná (UTFPR), a fim de verificar se as questões elencadas inicialmente eram pertinentes à proposta de entrevista do presente estudo. De acordo com Almeida (2005, p. 60), “a aplicação do pré-teste ou teste-piloto é uma atividade preliminar à implementação do projeto e tem por objetivo testar o desenho e a metodologia do projeto”, garantindo “a adequação dos instrumentos ou processos a serem utilizados, bem como medir o tempo a ser gasto na avaliação”;
- **Coleta de documentação:** de acordo com Almeida (2005, p. 58), outras fontes essenciais para o diagnóstico organizacional podem ser extraídas de fontes diversas como relatórios, regimentos, regulamentos, planos de trabalho, etc. Nesse sentido, foi realizado o levantamento da documentação sobre os repositórios institucionais nos Portais dos Sistemas de Bibliotecas da USP, UNESP e UNICAMP, sendo, de forma complementar, solicitada aos gestores e bibliotecários catalogadores-indexadores dos repositórios institucionais;
- **Seleção do texto-base:** foi utilizado um rol de questões abertas sobre o temário abordado, divididas em duas partes, de acordo com cada perfil dos participantes, com os objetivos de: a) caracterizar os repositórios institucionais no contexto das universidades participantes no que tange à estrutura administrativa - aplicada aos

²⁵ Cabe esclarecer que a proposta inicial da pesquisa era de realização de 3 entrevistas com bibliotecários catalogadores-indexadores que atuam no repositório institucional. Entretanto, no decorrer da mesma verificou-se que, no caso da UNESP, nenhum catalogador/indexador que exerce atualmente a função participou da implantação do Repositório Institucional da UNESP, tornando-se necessária a complementação de mais 3 entrevistas com bibliotecários que atuaram em sua institucionalização.

gestores dos repositórios institucionais; b) verificar especificamente a indexação de assuntos nos repositórios institucionais - aplicada aos bibliotecários catalogadores-indexadores. A elaboração dos roteiros para a entrevista (APÊNDICE A - gestores e APÊNDICE B - bibliotecários catalogadores-indexadores) - pautou-se em três principais frentes: em diagnóstico organizacional proposto por Almeida (2005, p. 53-55); em questionário aplicado por Gomes (2015) que objetivou identificar e analisar repositórios institucionais de universidades federais brasileiras na padronização dos metadados na representação da informação; e em Dal'Evedove (2014), cuja aplicação objetivou retratar o contexto informacional de bibliotecários catalogadores-indexadores referente à política de indexação;

- **Conversa informal com os participantes:** antes da aplicação da entrevista semiestruturada, foram mencionados os objetivos da pesquisa e a importância da mesma para a melhoria da indexação e da recuperação por assuntos em repositórios institucionais no contexto de bibliotecas universitárias. Foi entregue previamente a cada gestor e bibliotecário catalogador-indexador participante o *Termo de Consentimento Livre e Esclarecido* (TCLE) (APÊNDICE C) assinado tanto pela pesquisadora quanto pela orientadora e solicitada a assinatura do documento e aceitação formal da participação na pesquisa, ficando uma cópia para os mesmos. De forma complementar, foi ressaltado que a identidade de cada um dos participantes permaneceria anônima, com o propósito de não comprometer os dados e deixá-los à vontade durante a realização da entrevista de modo mais natural possível.

II Procedimentos durante a aplicação da entrevista

- **Entrevista e gravação das falas dos participantes:** a fim de facilitar a realização das entrevistas, foi utilizado o software *Skype*²⁶ para as entrevistas realizadas na USP e UNESP, enquanto que na UNICAMP foram realizadas presencialmente. De forma complementar, foi realizada a gravação das entrevistas por meio de um aplicativo de

²⁶ Lançado em 2003, o software permite a comunicação pela Internet através de conexões de voz e vídeo de forma rápida e econômica.

gravador de voz de aparelho celular²⁷. Cabe esclarecer que, de acordo com Pickard (2013, p. 205), ainda inexistente um consenso na literatura em torno da utilização de entrevistas online. Se por um lado algumas pistas visuais, verbais e comportamentais dos participantes possam se perder durante a entrevista, a facilidade de comunicação e economia de recursos e deslocamento físico ainda são vantajosas. Ainda segundo o autor, neste tipo de entrevista existem duas abordagens: síncrona e assíncrona. Enquanto a entrevista síncrona é realizada *a posteriori*, seja por uso de e-mail ou outro recurso, enviada pelo entrevistado ao pesquisador em um determinado período de tempo designado, a entrevista assíncrona, tal como utilizada nesta pesquisa, é realizada em tempo real entre o pesquisador e o entrevistado (PICKARD, 2013, p. 203; 205). A intervenção da pesquisadora foi realizada para que o rol de questões pudesse ser abarcado durante a coleta de dados. Tanto a conversa informal quanto a entrevista propriamente dita foram aplicadas individualmente com cada participante, respeitando sua própria individualidade e necessidades.

III Procedimentos posteriores à aplicação da entrevista

- **Transcrição literal das gravações das falas dos participantes:** foram realizadas transcrições literais das falas dos entrevistados, preservando a identidade dos mesmos, por meio de siglas específicas de acordo com cada categoria - gestor ou bibliotecário catalogador-indexador e a instituição a qual pertencem. Para melhor visualização dos processos externalizados pelos participantes, para as transcrições das entrevistas adaptou-se um tipo específico de notação amplamente utilizado na Ciência da Informação para análise de dados oriundos da aplicabilidade da metodologia do Protocolo Verbal, com base em Cavalcanti (1989) e Tartarotti (2014) (ANEXO A);

²⁷ *Smartphone Samsung Galaxy J7 Neo TV Tela 5.5"*, Android™ 7.0 *Octacore*, Câmera 13Mp, 16Gb Dual Chip, que permite horas de gravação, dependendo apenas da capacidade de armazenamento do aparelho. Possui funções de reprodução e edição dos sons gravados, permitindo a configuração da qualidade de gravação para Baixa (64kbps, 44,1kHz); Média (128kbps, 44.1kHz) ou Alta (256kbps, 48kHz). Durante a gravação, permite o uso de marcadores. Após, mostra tempo do arquivo de som e possibilita a categorização dos arquivos gravados em pastas. A utilização deste aparelho justifica-se ainda pelos seguintes motivos: alta resolução de som que permite uma gravação de qualidade; facilidade para armazenamento e transmissão de arquivos, uma vez que o arquivo é gerado digitalmente e permite o compartilhamento dos arquivos de som por e-mail, *WhatsApp*, *Mensagens*, *Bluetooth*, *Skype*, *OneDrive* ou *Drive* do *Google*. O celular é de propriedade da pesquisadora.

- ***Leitura detalhada dos dados em busca de fenômenos significativos e recorrentes para elaboração de categorias de análise:*** foi realizada leitura detalhada das transcrições das entrevistas semiestruturadas, com o intuito de buscar fenômenos significativos para a elaboração das unidades e categorias de análise dos dados coletados;
- ***Elaboração de categorias de análise:*** após leitura detalhada dos dados coletados, foram elaboradas unidades e categorias de análise baseadas tanto na literatura (referenciais teóricos) quanto nas declarações realizadas pelos participantes, por meio das quais as informações obtidas pela aplicação das entrevistas semiestruturadas foram analisadas;
- ***Retorno aos dados para retirar trechos da discussão que exemplifiquem cada categoria de análise:*** foi realizada uma releitura das transcrições das entrevistas semiestruturadas visando à retirada de trechos da discussão que melhor exemplifiquem cada categoria de análise e síntese dos principais aspectos observados.

Em continuidade ao delineamento da abordagem exploratório-descritiva da pesquisa, apresenta-se, a seguir, a abordagem quantitativa da avaliação da indexação de assuntos.

4.2.3 Abordagem quantitativa: avaliação da indexação de assuntos

Com o intuito de cumprir-se o **terceiro objetivo específico da pesquisa** - verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária, optou-se pela avaliação da indexação de assuntos como **metodologia quantitativa**, visto sua relevância para medir e avaliar os aspectos intrínsecos e extrínsecos da indexação de assuntos. A avaliação da indexação de assuntos pautou-se em duas abordagens:

- A) Avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema* - *Avaliação intrínseca quantitativa mediante a interconsistência;*
- B) Avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário* - *Avaliação extrínseca mediante a recuperação.*

Na primeira parte foi realizada uma avaliação comparada dos assuntos atribuídos em repositório institucional pela perspectiva do *sistema*, enquanto que na segunda parte foram realizadas entrevistas semiestruturadas de busca por assuntos em repositório institucional pela perspectiva do *usuário*. Para que a pesquisa pudesse ser exequível, elegeu-se a UNICAMP como universo de pesquisa dentre as três universidades públicas paulistas do Estado de São Paulo: USP, UNESP e UNICAMP, para o estudo de caso da avaliação da indexação de assuntos no repositório institucional, tendo em vista a atuação profissional da pesquisadora neste âmbito acadêmico.

Fundada oficialmente em 5 de outubro de 1966, dia do lançamento de sua pedra fundamental, a UNICAMP está localizada no Estado de São Paulo, possui três campi — em Campinas, Piracicaba e Limeira — e compreende 24 Unidades de Ensino e Pesquisa, além de um amplo complexo de saúde, 21 núcleos e centros interdisciplinares, dois colégios técnicos, em um universo onde convivem cerca de 50 mil pessoas. A UNICAMP responde por 8% da pesquisa acadêmica no Brasil, 12% da pós-graduação nacional e mantém a liderança entre as universidades brasileiras no que diz respeito a patentes e ao número de artigos per capita publicados anualmente em revistas indexadas na base de dados ISI/WoS. Conta com aproximadamente 34 mil discentes matriculados em 66 cursos de graduação e 153 programas de pós-graduação. A média anual de teses e dissertações defendidas é de 2,1 mil e 99% de seus professores possuem título de doutor, liderando o ranking nacional per capita de publicações científicas nas revistas internacionais catalogadas. Se a produção acadêmica for calculada pelo desempenho de cada pesquisador, a UNICAMP é, atualmente, a mais produtiva universidade brasileira (UNICAMP, 2018).

Atualmente o SBU²⁸ é composto pelo Órgão Colegiado, Coordenadoria e 29 Bibliotecas, sendo uma Biblioteca Central (BCCL), uma Biblioteca da Área de Engenharia e Arquitetura (BAE) e as demais alocadas nas Unidades de Ensino e Pesquisa, Centros e Núcleos. O SBU está pautado nas seguintes características: *negócio*: informação científica, tecnológica, histórica e cultural; *missão*: promover informação, por meio de produtos e serviços de excelência, para as atividades de ensino, pesquisa e extensão, garantindo um ambiente de respeito à diversidade e à socialização; *valores*: excelência, cooperação, acessibilidade, comprometimento, inovação, satisfação do usuário, sustentabilidade, sensibilidade para com o outro e competência em informação; *princípios*: responsabilidade social, integridade, ética, pluralidade, imparcialidade e valorização do ser humano; e *visão*: ser um sistema de bibliotecas de referência nacional e internacional, reconhecido pela excelência de seus produtos e serviços (SBU, 2018).

Como um sistema de bibliotecas inserido em um contexto mais amplo, universitário, dentre a gama de produtos e serviços oferecidos pelo SBU destaca-se o Repositório da Produção Científica e Intelectual da UNICAMP como instrumento oficial para coleta, organização, disseminação e preservação de todo o conhecimento produzido na universidade. No entanto, inexistente uma linguagem padronizada que possa dar suporte à descrição temática da produção científica abarcada pelo repositório institucional. No catálogo online SophiA são utilizadas diversas linguagens em maior ou menor nível, tais como: a *Library of Congress Subject Headings* (LCSH), Lista de Cabeçalhos de Assunto da Rede BIBLIODATA (LCARB), DeCS (Descritores em Ciências da Saúde), dentre outros, de acordo com a área do conhecimento.

A seguir, apresenta-se, de forma detalhada, a primeira parte da aplicabilidade prática da avaliação da indexação de assuntos, representada pela *Avaliação intrínseca quantitativa mediante a interconsistência* e a segunda parte, representada pela *Avaliação extrínseca mediante a recuperação*, realizadas neste universo de pesquisa.

²⁸ O SBU foi criado oficialmente em 25 de novembro de 2003, através da Deliberação CONSU-A-30. Seu Regimento Interno está disposto pela Deliberação CONSU A-15, de 06 de agosto de 2013, disponível em: https://www.pg.unicamp.br/mostra_norma.php?id_norma=3425.

4.2.3.1 Avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do sistema - Avaliação intrínseca quantitativa mediante a interconsistência

A *Avaliação intrínseca quantitativa mediante a interconsistência* é aplicada quando busca-se conhecer o nível de semelhança entre duas indexações. Nesse sentido, quanto maior o grau ou o nível de consistência, mais semelhantes são os assuntos atribuídos aos recursos informacionais investigados. De acordo com Gil Leiva, Fujita e Rubi (2008, p. 236), quando comparadas as categorias variam de 0 a 1 ou de 0 a 100, se convertidas em porcentagem (%) e tais semelhanças ou diferenças podem ser quantificadas por meio de fórmulas matemáticas.

De acordo com a literatura, são diversas as fórmulas matemáticas de avaliação quantitativa de interconsistência desenvolvidas para proporcionar medidas objetivas, por meio das quais o nível de consistência entre duas indexações pode ser determinado. As mais utilizadas são as de Hooper (1965), mais recentemente adaptada por Gil Leiva (2008, p. 386) e a de Rolling (1981), que apresentam uma relação fracional simples entre os atribuídos em cada indexação e os termos comuns entre duas indexações. Para a aplicabilidade prática nesta pesquisa, foram utilizadas estes mesmos parâmetros metodológicos, permitindo uma comparação dos índices de consistência gerados no universo de pesquisa do Repositório Institucional da UNICAMP.

Ao serem aplicadas as referidas fórmulas na pesquisa, na fórmula de consistência de Rolling (1981), o índice de consistência do repositório institucional entre as duas indexações a serem comparadas é obtido multiplicando-se o número de termos comuns nas duas indexações por 2 (dois) e dividindo-se o resultado pela soma do número de termos atribuídos na *indexação A* (autores) e o número de termos usados na *indexação B* (bibliotecários catalogadores-indexadores). Já na fórmula de consistência de Hooper (1965) adaptada por Gil Leiva (2008, p. 386), o índice de consistência entre as duas indexações é obtido dividindo-se o número de termos comuns nas duas indexações pelo número de termos usados na *indexação A* (autores), somado ao número de termos usados na *indexação B* (bibliotecários catalogadores-indexadores) e subtraído novamente do número de termos comuns nas duas indexações (Quadro 5):

Quadro 5. Fórmulas matemáticas para obtenção do índice de consistência entre duas indexações no repositório institucional

Rolling (1981)	Hooper (1965) adaptada por Gil Leiva (2008)
$C_i = \frac{2C}{A + B}$ <p>Onde,</p> <p><i>C_i</i> = Índice de consistência do repositório institucional <i>C</i> = Número de termos comuns nas duas indexações <i>A</i> = Número de termos atribuídos na Indexação A (autores) <i>B</i> = Número de termos atribuídos na Indexação B (bibliotecários catalogadores-indexadores)</p>	$C_i = \frac{T_{co}}{(A + B) - T_{co}}$ <p>Onde,</p> <p><i>C_i</i> = Índice de consistência do repositório institucional <i>T_{co}</i> = Número de termos comuns nas duas indexações <i>A</i> = Número de termos atribuídos na Indexação A (autores) <i>B</i> = Número de termos atribuídos na Indexação B (bibliotecários catalogadores-indexadores)</p>

Fonte: Adaptado de Gil Leiva (2008, p. 386).

É importante salientar que, de acordo com a literatura, para a aplicabilidade prática das fórmulas de consistência há duas abordagens principais: *comparação rígida* e *comparação relaxada* (GIL LEIVA; RUBI; FUJITA, 2008, p. 238) ou *comparação flexível* (TARTAROTTI, 2014, p. 215). Na comparação rígida dos índices de consistência os assuntos atribuídos devem coincidir completamente, sendo atribuído o valor de “1” para cada assunto atribuído nas duas indexações analisadas, para cada recurso informacional. Já na comparação relaxada ou flexível, quando há a coincidência total entre os assuntos atribuídos, considera-se o valor de “1” ou apenas a metade, ou seja, “0,5”, quando ocorre coincidência nos assuntos somente em parte do assunto (por ex. apenas no cabeçalho ou no subcabeçalho). Para ambos os índices, quando não há nenhuma correspondência entre os assuntos atribuídos, considera-se o valor de “0” (zero).

A utilização da *Avaliação intrínseca quantitativa mediante a interconsistência*, que no universo desta pesquisa pode ser denominada de *Avaliação Inter-autor-bibliotecário*, possibilitou a comparação dos assuntos atribuídos pelos próprios autores (termos livres) e pelos bibliotecários catalogadores-indexadores (descritores), após validação/verificação de assuntos pelos profissionais de artigos de periódico no contexto específico do Repositório Institucional da UNICAMP (Figura 6).

Figura 6. Avaliação da indexação de assuntos comparada entre autores e bibliotecários catalogadores-indexadores



Fonte: Elaboração própria.

Como produtos desta avaliação da indexação de assuntos, foram obtidos *ensaios* ou *índices de interconsistência*, verificando *inconsistências*, *discrepâncias* ou *diferenças* entre as duas indexações no mesmo repositório institucional.

4.2.3.2 Avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do usuário - Avaliação extrínseca mediante a recuperação

A utilização da entrevista semiestruturada na recuperação por assuntos em repositórios institucionais (avaliação pela perspectiva do usuário) na presente pesquisa possibilitou a verificação da *eficácia* da indexação de assuntos dos recursos informacionais em repositório institucional pela abordagem da recuperação da informação pelos usuários.

Na literatura, destaca-se o estudo de Boccato, Fujita e Gil Leiva (2011), que avaliou, de forma comparada, o índice de precisão do uso da linguagem natural com duas linguagens de indexação especializadas em catálogo online: como linguagem pré-coordenada, a *Lista de Cabeçalhos de Assunto da Rede BIBLIODATA* (LCARB), utilizada no catálogo online Athena da UNESP, e como linguagem pós-coordenada, o *Vocabulário Controlado da USP* (VOCAUSP), utilizado no catálogo coletivo DEDALUS da USP. Por meio de estratégias de busca por assuntos com discentes de graduação e de pós-graduação em nível de mestrado e doutorado com entrevista estruturada, verificou-se a influência da disponibilidade de linguagens de indexação no desempenho do sistema de recuperação. Os principais resultados demonstraram que a

especificidade dos termos exigidos pelos usuários na recuperação, ou seja, o índice de precisão, não apresentou muita diferença entre a linguagem natural (7,2%) e as linguagens LCARB (6,5%) e VOCAUSP (7,2%), provavelmente pela realização de estratégias de busca em modo simples e não avançado no catálogo online.

Tomando como base o estudo supracitado, nesta pesquisa a realização da entrevista semiestruturada de busca por assuntos em repositório institucional pela perspectiva do usuário pautou-se, primeiramente, na definição das Bibliotecas dos usuários participantes. Para tanto, realizou-se o levantamento das mesmas no portal do SBU²⁹. As áreas do conhecimento delineadas na relação das Bibliotecas foram mantidas de acordo com a taxonomia apresentada no próprio portal. Após o levantamento das potenciais Bibliotecas participantes, procedeu-se ao critério de que a pesquisa pudesse abarcar duas Bibliotecas de cada área do conhecimento, a saber: *Área de Artes e Humanidades*, *Área de Biomédicas*, *Área de Ciências Aplicadas*, *Área de Exatas* e *Área de Tecnológicas*, exceto nos casos da *Área de Ciências Aplicadas*, que possui apenas uma biblioteca e na *Área de Conhecimentos Gerais*, não contemplada no estudo³⁰. O universo para aplicabilidade da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do usuário (*Avaliação extrínseca mediante a recuperação*) constitui-se de 9 (nove) Bibliotecas públicas universitárias da UNICAMP (Quadro 6):

Quadro 6. Seleção das Bibliotecas participantes da pesquisa para a *Avaliação extrínseca mediante a recuperação*

Área do conhecimento	Biblioteca participante	Localização
Artes e Humanidades	Faculdade de Educação (FE)	Campinas, SP
	Instituto de Filosofia e Ciências Humanas (IFCH)	Campinas, SP
Biomédicas	Faculdade de Ciências Médicas (FCM)	Campinas, SP
	Faculdade de Educação Física (FEF)	Campinas, SP

²⁹ Disponível em: <http://www.sbu.unicamp.br>.

³⁰ Cabe esclarecer que a *Área de Conhecimentos Gerais*, representada pela Biblioteca da Diretoria de Difusão da Informação (DINF) e pela Biblioteca de Obras Raras (BORA), por suas próprias características gerais, não possuem comissão de biblioteca, usuários-discentes e usuários-docentes específicos, tal como ocorre nas demais bibliotecas alocadas nas unidades de ensino e pesquisa.

Ciências Aplicadas	Faculdade de Ciências Aplicadas (FCA)	Limeira, SP
Exatas	Instituto de Física Gleb Wataghin (IFGW)	Campinas, SP
	Instituto de Matemática, Estatística e Computação Científica (IMECC)	Campinas, SP
Tecnológicas	Biblioteca da Área de Engenharia e Arquitetura (BAE)	Campinas, SP
	Faculdade de Engenharia de Alimentos (FEA)	Campinas, SP

Fonte: Dados da pesquisa.

No que tange à infraestrutura material para a realização das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP, as Bibliotecas atenderam aos quesitos necessários. Para tanto, foi utilizado o próprio ambiente das Bibliotecas dos usuários, com disponibilização do notebook da pesquisadora para acesso ao Repositório Institucional da UNICAMP. Considerando-se que os “participantes devem ser selecionados de forma criteriosa e cuidadosa, pois deles dependerá, em grande parte, o sucesso da coleta de dados” e que “a preferência deve ser dada àquelas pessoas que realmente se interessam em colaborar de maneira efetiva com a pesquisa” (FUJITA, 2009, p. 62), a definição dos participantes baseou-se na própria caracterização informacional dos repositórios institucionais, verificando as categorias acadêmicas que pudessem utilizar de forma mais significativa este contemporâneo sistema de recuperação da informação.

De acordo com Maciel (1997, p. 23), por meio de procedimentos estatísticos é possível reduzir a população global de indivíduos a uma amostra, sendo a mais significativa daquela população, “considerando-se todos os segmentos integrantes do universo global, estabelecendo-se as devidas proporções”. Desse modo, os participantes foram assim definidos em quatro categorias: *Discente de graduação*; *Discente de pós-graduação (mestrado)*; *Discente de pós-graduação (doutorado)* (usuários-discentes); e *Docente* (usuário-docente), sendo 1 (um) participante para cada biblioteca elencada no universo de pesquisa de cada categoria, totalizando 36 participantes/coletas realizadas.

Dentre os critérios utilizados para a escolha dos participantes, foram selecionados *docentes* coordenadores ou suplentes das comissões das referidas Bibliotecas, com pesquisa em andamento em determinada temática, de acordo com cada área científica especializada.

Para os *discentes de pós-graduação* (mestrado e doutorado), que possuem pesquisa em andamento, solicitou-se às Bibliotecas a indicação, entretanto, em alguns casos houve a necessidade de seleção aleatória dos mesmos com base nos critérios para os *discentes de graduação*. Para esta categoria, definiu-se o critério de bolsistas de produtividade PIBIC/CNPq, com projetos de pesquisa em andamento. Para tanto, foi utilizada a técnica da amostragem, cujos dados foram solicitados à Diretoria Acadêmica (DAC) da UNICAMP e confirmados no software gerenciador *SophiA*³¹. A opção por tal recurso deve-se ao fato da quantidade significativa de possíveis discentes no estudo, visto o amplo contexto acadêmico das Bibliotecas pesquisadas.

Em consonância com a entrevista de diagnóstico organizacional com os gestores e bibliotecários bibliotecários catalogadores-indexadores, foi elaborado um *roteiro de aplicação* da avaliação da indexação de assuntos no Repositório Institucional da UNICAMP pela perspectiva do usuário, tendo como base o estudo de Boccato, Fujita e Gil Leiva (2011, p. 59-60) de avaliação da recuperação por assuntos em catálogo online, a fim de permitir a identificação das atividades realizadas em cada fase de aplicação. Tal roteiro foi delineado como segue: *procedimentos anteriores às buscas por assuntos no repositório institucional; procedimentos durante as buscas por assuntos no repositório institucional; e procedimentos posteriores às buscas por assuntos no repositório institucional*:

I Procedimentos anteriores às buscas por assuntos no repositório institucional

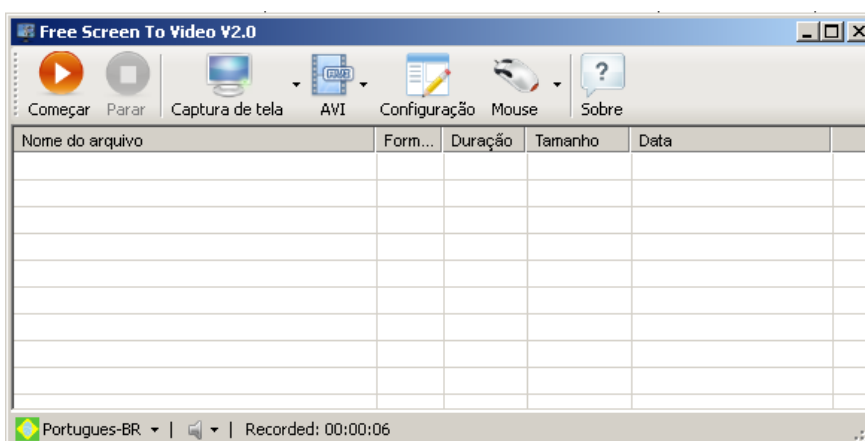
- ***Definição do universo da pesquisa***: foi elencado o contexto organizacional da UNICAMP para as buscas por assuntos no repositório institucional por meio de entrevista semiestruturada;
- ***Seleção dos participantes***: a seleção dos participantes para a aplicação da entrevista resultou na escolha de um total de 36 participantes, sendo: *Discente de graduação; Discente de pós-graduação (mestrado); Discente de pós-graduação (doutorado)*

³¹ O software, utilizado para o gerenciamento do acervo do SBU, baseia-se em padrões internacionais de catalogação e comunicação de dados (MARC 21, ISO 2709, Z39.50 cliente e servidor, XML e OAI-PMH).

(usuários-discentes); e *Docente* (usuário-docente), distribuídos em 9 Bibliotecas da UNICAMP (vide Quadro 7) e realizada uma conversa informal com cada um dos participantes por e-mail, resultando na aceitação e definição das datas para a realização das entrevistas;

- **Captura das pesquisas:** foi realizada a instalação do software *Free Screen to Video 2.0*³² no notebook da pesquisadora, para a captura das telas de busca/gravação realizadas pelos usuários (Figura 7);

Figura 7. Tela inicial do software *Free Screen to Video 2.0*



Fonte: Dados da pesquisa.

- **Conversa informal com os participantes:** antes da aplicação da entrevista semiestruturada de busca por assuntos no Repositório Institucional da UNICAMP, foram mencionados os objetivos da pesquisa e a importância da mesma para a melhoria da indexação e da recuperação por assuntos em repositórios institucionais no contexto de bibliotecas universitárias. Em seguida, foi entregue previamente a cada participante o *Termo de Consentimento Livre e Esclarecido (TCLE)* (APÊNDICE E), assinado tanto pela pesquisadora e pela orientadora e solicitada a assinatura do documento e aceitação formal da participação na pesquisa, ficando uma cópia para os

³² O software possibilita a gravação de vídeos da tela do computador, nos formatos FLV, AVI, WMV e SWF, captura de movimentos do mouse e áudio gravado com microfone, dentre outras funcionalidades. Optou-se pela gravação no formato AVI para reprodução das telas gravadas para posterior reprodução no Windows Media Player. Disponível gratuitamente na Internet (<http://www.softpedia.com/get/Multimedia/Graphic/Graphic-Capture/Free-Screen-Recorder.shtml>).

mesmos. De forma complementar, foi ressaltado que a identidade de cada um dos participantes permaneceria anônima, com o propósito de não comprometer os dados e deixá-los à vontade durante a realização da entrevista de modo mais natural possível, e informado sobre a gravação da coleta de dados;

- **Familiarização sobre o Repositório Institucional da UNICAMP e da linguagem de indexação:** foi realizada a apresentação do Repositório Institucional da UNICAMP aos participantes, com orientações básicas de como realizar as buscas, especialmente àqueles que desconheciam o sistema;
- **Familiarização sobre o roteiro da pesquisa por assunto:** para que as buscas por assuntos em linguagem natural no Repositório Institucional da UNICAMP pudessem ser realizadas, foi entregue aos participantes o *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem natural* (APÊNDICE F). Em seguida, foi solicitada a definição e anotação de 5 (cinco) palavras-chave de suas pesquisas em andamento, conforme suas respectivas áreas científicas especializadas.

II Procedimentos durante as buscas por assuntos no repositório institucional

- ***Buscas no Repositório Institucional da UNICAMP - Linguagem natural:*** foi iniciada a gravação das telas de busca pela pesquisadora e das gravações das falas dos participantes. Em seguida, foi solicitado que iniciassem as buscas no sistema utilizando-se a caixa de pesquisa na página inicial do Repositório Institucional da UNICAMP com a delimitação do campo “Assunto”, a partir de cada palavra-chave anotada no *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem natural*, limitando-se a 50 resultados por página, na seguinte ordem: a) palavra-chave em português; b) palavra-chave em português utilizando-se o recurso gráfico aspas (“), no caso de palavra-chave composta; c) palavra-chave em inglês; e d) palavra-chave em inglês utilizando-se o recurso gráfico aspas (“), no caso de palavra-chave composta;

- ***Seleção dos recursos informacionais para determinação da relevância - Linguagem natural:*** para cada uma das buscas realizadas, foi solicitado aos participantes que selecionassem os recursos informacionais relevantes recuperados, de acordo com as necessidades informacionais de suas pesquisas em andamento, anotando o número de recursos informacionais relevantes recuperados para cada palavra-chave, tanto em português quanto em inglês, utilizando-se ou não o recurso gráfico aspas (“) no *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem natural*;

- ***Consulta à linguagem de indexação:*** finalizada a primeira rodada de buscas utilizando-se as palavras-chave elencadas pelos participantes (linguagem natural), foi entregue o *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem controlada* (APÊNDICE G) e realizada a apresentação da linguagem de indexação de acordo com cada área científica especializada dos mesmos. Em seguida, foi realizada a consulta, com auxílio da pesquisadora, à linguagem de indexação definida de acordo com cada área, a partir das 5 (cinco) palavras-chave previamente escolhidas em linguagem natural e de acordo com o assunto de pesquisa em andamento de cada participante, identificando e anotando os termos correspondentes autorizados para a elaboração da estratégia de busca por assuntos a ser realizada no Repositório Institucional da UNICAMP, tanto em português quanto em inglês;

- ***Buscas no Repositório Institucional da UNICAMP - Linguagem controlada:*** foi solicitado aos participantes que iniciassem as buscas no sistema utilizando-se a caixa de pesquisa na página inicial do Repositório Institucional da UNICAMP com a delimitação do campo “Assunto”, a partir de cada palavra-chave consultada na linguagem controlada (descritores), anotada no *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem controlada*, limitando-se a 50 resultados por página, na mesma ordem da linguagem natural, ou seja: a) descritor em português; b) descritor em português utilizando-se o recurso gráfico aspas (“), no caso de descritor composto; c) descritor em inglês; e d) descritor em inglês utilizando-se o recurso gráfico aspas (“), no caso de descritor composto. Cabe esclarecer que, no caso

de termos guais tanto na linguagem natural quanto na linguagem controlada, não foram realizadas novamente as buscas, visto que os resultados seriam os mesmos;

- ***Seleção dos recursos informacionais para determinação da relevância - Linguagem controlada:*** foi solicitado aos participantes que elencassem os recursos informacionais relevantes recuperados, de acordo com suas necessidades de informação de suas pesquisas em andamento, anotando o número de recursos informacionais relevantes recuperados para cada descritor, tanto em português quanto em inglês, utilizando-se ou não o recurso gráfico aspas (“”) no *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem controlada*. Ao final de cada busca, foi finalizada a gravação das telas de busca pela pesquisadora, utilizando-se a tecla F10;

- ***Entrevista retrospectiva:*** de forma complementar, foram realizadas entrevistas retrospectivas³³ com os participantes após cada seção de coleta de dados de busca e recuperação por assuntos no Repositório Institucional da UNICAMP, a fim de verificar suas opiniões a respeito da atividade realizada, norteada por três questões:
 - Você já conhecia os recursos informacionais recuperados nas buscas?
 - Qual sua opinião sobre as buscas realizadas no Repositório Institucional da UNICAMP?
 - Qual sua opinião sobre o sistema? Alguma sugestão de melhoria?

III Procedimentos posteriores às buscas por assuntos no repositório institucional

- ***Leitura detalhada dos dados:*** foi realizada a leitura das telas gravadas, das falas gravadas, dos roteiros dos participantes e das anotações realizadas pela pesquisadora durante as buscas e levantado o número de recursos informacionais relevantes recuperados, a partir dos julgamentos realizados pelos participantes, de acordo com cada área do conhecimento no Repositório Institucional da UNICAMP, por pesquisa e

³³ As entrevistas retrospectivas referem-se a um tipo específico de entrevista realizada após a realização de uma determinada atividade de leitura, nesse caso, de busca e recuperação por assuntos em repositório institucional.

por usuário, de cada Biblioteca, elaborado em quadros de registros individuais, contendo: palavras-chave elencadas em português e em inglês; termos encontrados/correspondentes na linguagem de indexação e utilizados nas buscas; quantidade de recursos informacionais recuperados; e quantidade de recursos informacionais relevantes recuperados, em cada busca e para cada usuário. Em seguida, foi elaborado um quadro de registros geral em *Excel* de todas as buscas realizadas com o uso da linguagem natural e da linguagem controlada no Repositório Institucional da UNICAMP, contendo uma compilação dos quadros de registros individuais;

- ***Determinação do índice de precisão:*** como resultado quantitativo das buscas realizadas pelos participantes, foi realizado o cálculo do índice de precisão no Repositório Institucional da UNICAMP, tanto em linguagem natural quanto em linguagem controlada, visando conhecer com qual linguagem se conseguiram melhores índices;
- ***Transcrição literal das gravações das falas dos participantes:*** de forma complementar, foi realizada a transcrição das falas dos participantes durante a realização das buscas por assuntos no Repositório Institucional da UNICAMP. Tais transcrições foram norteadas pelos mesmos procedimentos adotados nas transcrições das entrevistas com os gestores e bibliotecários catalogadores-indexadores;
- ***Análise das gravações das falas dos participantes:*** foi realizada a leitura detalhada das transcrições das gravações, com o intuito de buscar fenômenos significativos para a elaboração de unidades e categorias de análise baseadas tanto na literatura (referenciais teóricos) quanto nas declarações realizadas pelos participantes;
- ***Retorno aos dados para retirada de trechos da discussão que exemplifiquem cada categoria de análise:*** foi realizada uma releitura das transcrições das entrevistas semiestruturadas visando à retirada de trechos da discussão que melhor exemplificassem cada categoria de análise, mediante a elaboração de quadros e

síntese dos principais aspectos observados.

Cabe esclarecer que, tal como nas entrevistas de diagnóstico organizacional com os gestores e bibliotecários catalogadores-indexadores, foi realizado um *pré-teste* (piloto) da metodologia proposta para a entrevista semiestruturada na recuperação por assuntos em repositórios institucionais pela perspectiva do usuário. Tal piloto, realizado com aluno de pós-graduação (mestrado) da *Área de Exatas*, fez-se necessário tendo em vista a aplicabilidade prática da avaliação comparada da estratégia de busca dos usuários para recuperação da informação com e sem uso de uma linguagem controlada. Assim, somente após esta aplicação e ajustes necessários foi aplicada efetivamente com os usuários-participantes da pesquisa.

As demais coletas com os usuários na recuperação por assuntos no Repositório Institucional da UNICAMP foram realizadas nos meses de julho e agosto de 2019. Tanto a coleta-piloto quanto as demais coletas foram efetuadas no próprio ambiente acadêmico/de estudo dos participantes, isto é, nas dependências das 9 (nove) Bibliotecas da UNICAMP participantes da pesquisa, previamente agendadas pela pesquisadora, de acordo com a disponibilidade de cada participante. Cabe esclarecer que oficialmente a aplicabilidade da pesquisa nas referidas Bibliotecas foram autorizadas pelos respectivos diretores (APÊNDICE D).

Tomando-se como base os procedimentos metodológicos delineados nas abordagens teórico-aplicada e exploratório-descritiva que nortearam a pesquisa, apresenta-se, a seguir, os resultados e discussões da mesma.

5 RESULTADOS E DISCUSSÃO

Nesta seção, são apresentados os resultados da abordagem qualitativa: entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais; e da abordagem quantitativa: avaliação da indexação de assuntos e suas respectivas abordagens, representadas pelos resultados da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*.

5.1 Resultados da abordagem qualitativa: entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais

A análise das coletas de dados das entrevistas realizadas com gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP pautou-se no seguinte delineamento: apresentação de cada unidade de análise e respectivas categorias; e síntese analítica com apresentação de exemplos, evidenciando trechos das declarações dos participantes, sendo cada trecho identificado por siglas, sublinhando-se os aspectos mais significativos de cada trecho referentes à indexação e recuperação por assuntos em repositórios institucionais. Com base nos referenciais teóricos, nos objetivos da pesquisa e nas entrevistas, foram elaboradas sete categorias de análise, distribuídas em três unidades de análise:

- ***Aspectos relacionados ao contexto***

Arrola sobre os aspectos decisórios do funcionamento geral do repositório institucional, com as categorias: 1) *Implementação do repositório institucional*; 2) *Política de funcionamento e documentação do repositório institucional*; 3) *O usuário do repositório institucional*; e 4) *Ações de melhoria no repositório institucional*;

- ***Aspectos relacionados ao bibliotecário catalogador-indexador***

Aborda o papel do bibliotecário catalogador-indexador no repositório institucional, com a seguinte categoria: 5) *Atuação profissional do bibliotecário catalogador-indexador no repositório institucional*;

▪ **Aspectos relacionados ao processo**

Envolve questões sobre o processo de indexação/catalogação de assuntos realizado em repositório institucional, com as categorias: 6) *Tratamento descritivo da informação no repositório institucional* e 7) *Tratamento temático da informação no repositório institucional*.

Para a análise das transcrições, foram preservadas as identidades dos participantes por meio de siglas específicas, de acordo com a instituição a qual pertencem e à categoria acadêmica - gestor do repositório institucional, bibliotecário catalogador-indexador ou bibliotecário (Quadro 7):

Quadro 7. Participantes da pesquisa e respectivas siglas para identificação e análise das entrevistas

Instituição	Categoria	Sigla
USP	Gestor	G-USP
	Bibliotecário catalogador-indexador	I-USP
UNESP	Gestor	G-UNESP
	Bibliotecário catalogador-indexador	I-UNESP
	Bibliotecário	B-UNESP-1 B-UNESP-2 B-UNESP-3
UNICAMP	Gestor	G-UNICAMP
	Bibliotecário catalogador/indexador	I-UNICAMP

Fonte: Elaboração própria.

Cabe esclarecer que a categoria de participante denominada bibliotecário refere-se a bibliotecários que atuaram no *Grupo Gestor da Política do Repositório Institucional UNESP (GRI-UNESP)* quando de sua implantação. A seguir, apresenta-se os resultados de cada unidade de análise e suas respectivas categorias.

UNIDADE DE ANÁLISE: Aspectos relacionados ao contexto de repositórios institucionais

1. Implementação do repositório institucional

A categoria “Implementação do repositório institucional” refere-se ao percurso histórico de criação e estabelecimento dos repositórios institucionais segundo os entrevistados. No caso da USP, a proposta surgiu em 2010 por ocasião de uma chamada pública sobre repositórios institucionais, com três Bibliotecas-piloto. Atendendo a uma demanda da FAPESP da obrigatoriedade de todo pesquisador que solicitasse auxílio como agência de fomento ter sua produção armazenada em um repositório, em 2012 houve a publicação da Resolução 6444 que estabelece a *Biblioteca Digital da Produção Intelectual* (BDTI), nome que será modificado em breve para *Repositório da USP*. No mesmo ano, realizou-se o CONFOA e o lançamento oficial dos três repositórios: USP, UNESP e UNICAMP, juntamente com o Repositório CRUESP.

*A primeira iniciativa de repositório aqui na universidade começou em 2010, é ... no ano que houve uma chamada pública, né, sobre repositórios institucionais e com algumas Unidades piloto, que a gente chama por sigla, né, que é a EACH, a Escola de Artes, Ciências e Humanidades, a Faculdade de Medicina Veterinária e Zootecnia (...) e a ...~~~ nossa, teve mais uma, eram três. ...~~~ Mas eram três Unidades, né, então era um projeto piloto e essas três Unidades participaram e ... é ... ah, Faculdade de Medicina. (...) Elas participaram em 2010 e aí a partir daí foram estabelecidos alguns critérios, alguns parâmetros, algumas metodologias, o que poderia vir a dar certo ou não. ... É ... feito isso, em 2012 teve a publicação da lei, é, da Resolução 6444 que é o que a gente chama de Resolução da BDTI, que é a Biblioteca Digital da Produção Intelectual que vai mudar o nome pra Repositório da USP. ... E ... em 2012 ela foi oficializada, teve um evento grande sobre acesso aberto, sobre repositório, trouxemos ícones da área (...) e lançamos, em 2012. Em 2013 teve o lançamento do Repositório CRUESP no CONFOA, que foi realizado aqui ... e ... entre 2012 e 2018 nós tínhamos um repositório mais tradicional, digamos assim, como a gente tá acostumado **(G-USP)***

*Então, surgiu em 2013, na verdade em 2013 foi a inauguração ... e um pouco antes teve um pedido da FAPESP pra que as três universidades estaduais tivessem o seu ambiente pra armazenar a produção científica das universidades, das três universidades e ... a FAPESP que ... ela ia começar a ... que todo pesquisador que solicitasse auxílio a ela como agência de fomento ... teria que ... ter sua produção armazenada em um repositório institucional. A partir disso começou a ter, a se desenvolver um trabalho conjunto com as três universidades **(G-UNESP)***

*[...] é ..., ele começou, teve início em 2013, com uma demanda que veio da ... da própria FAPESP né, essa demanda por repositório (...) e essa demanda envolveu as três universidades paulistas, que é a USP, a UNESP ... e a UNICAMP **(G-UNICAMP)***

(E em relação ao CRUESP (...) você lembra de onde surgiu a ideia de fazer o repositório do CRUESP?) É do professor Brito. (Foi do professor Brito mesmo?) É ... (...). (E quem mantém o portal do CRUESP, é a USP?) É, a USP se ofereceu no serviço de descoberta dela mas poderia ser qualquer uma das três. (...) A USP se ofereceu ... e o SIBi vinculou tudo ali. (...) ...~~~ Mas cada um ... é só uma interface de descoberta (...) cada

um é totalmente diferente. (...) Então quando um não tá funcionando, o outro funciona, entendeu? (...) Cada um é responsável pelo seu. (...) (Ele realiza a busca nos três repositórios?) ... Isso, isso, isso mesmo. (...) Mas ... a princípio foi isso, foi ... é porque na verdade como é algo das três universidades né ... então ... é ... tinha que ser algo que fizesse a busca nos três ambientes né (G-UNESP)

As principais ações realizadas para a formalização prática dos repositórios institucionais da USP, UNESP e UNICAMP centraram-se na escolha do software, no estabelecimento de metodologias e nos procedimentos de coleta e padronização de metadados das produções das respectivas universidades, a fim de cumprir o prazo estabelecido pela FAPESP. Um dos grandes problemas enfrentados nas coletas automáticas realizadas, apontado pelo gestor do repositório da UNESP, foi a ausência de padronização no nome da instituição e no nome dos autores nas bases de dados origens: *Web of Science, Scopus, SciELO, PubMed* e Currículo Lattes. Na UNESP, houve a contratação de bibliotecários que atuaram especificamente no repositório em um período de tempo determinado.

É ... uma já estava mais adiantada, as outras duas não tinham praticamente nada e ... no caso da UNESP a gente realmente, também, é ... a gente não tinha nada e a gente começou a fazer estudos, estudos de tudo, do que existia, de quais eram os softwares que funcionavam em outras instituições de fora, aqui de dentro do Brasil. ...~~ A gente começou a ver como é que a gente poderia coletar dados, porque a gente não tem um número suficiente de ... pessoas ...~~ porque a gente tinha um prazo pra inaugurar esse ambiente. (...) A gente tinha um prazo de menos de 6 meses pra colocar um ambiente desse no ar, então a gente começou, a gente sentou, centrou ... esforços ... era uma equipe multidisciplinar, então tinha gente de informática, Biblioteca, de ... do ensino à distância, tinha gente de várias áreas ... e ... começaram, a gente começou a determinar qual era o nosso ... perfil de aplicação, né, quais seriam os metadados que a gente precisava adotar ... nos registros dentro do repositório institucional. Depois a gente partiu pra ver qual disso tudo né, que a gente ia conseguir coletar né, fazer manual. ...~~ Então ... a gente viu o tipo de ... de conteúdo, de material ... tudo o que era produção da UNESP, aí a gente começou a se preparar pra vários problemas né, (...) problemas do tipo, é ... que a gente foi percebendo que as pessoas não citavam o nome da universidade de forma correta ... então a ... a forma de citação atrapalhava muito. ... A gente faz, é ... coleta automática ... então eu coloco o nome certinho então tudo que foge daquele nome (...) eu acabava perdendo de produção, né. Fora isso também tem o problema do nome dos homônimos dos autores ... (...) tem autores demais, e tal ... e é muito difícil de cruzar pra entender o que era UNESP, né. A gente coletava dessas bases que eu mencionei pra você, Scopus, Web of Science, SciELO ... a gente chegou a fazer um pouco de coleta do Lattes e ... PubMed. (...) É ... e foi assim que a gente conseguiu um número grande, não vou me lembrar agora quantos ... quantidade de registros ... mas a gente migrou pra uma quantidade boa e a gente pegou ...~~ até ... 2013, 2016, não 2013, é, pra ... pra trás. ... A gente pegou tudo o que tinha, tudo o que existia da UNESP, né. (...) Além de todo o trabalho em relação à embargo por causa das revistas, mas foi isso, aí em 2013 a gente lançou. ... (...) É ... nós tínhamos um ... um grupo de bibliotecários que trabalhavam junto ... ao repositório ... que oficializaram essa ... toda essa coleta, né, a questão do ... da folha de estilo com os metadados, né. ...~~ Por exemplo, o que tinha de tese e dissertação que estava no Aleph em formato MARC foi pra Dublin Core no repositório com o DSpace, entendeu? (...) Fez esse perfil, essa folha de estilo, né (...) e o perfil de aplicação com os metadados que a gente utilizava (...) e daí pra cá é trabalho de manutenção ... entre outras coisas (G-UNESP)

Então nesse momento, é ... como que foi feito? Foi feito uma coleta pela própria USP né, em conjunto com a UNESP, foi feito uma coleta de ... de registros, né, na época da Web of Science (...), com registros que continham dados dessas três universidades. ... Aí o que foi que feito? Então aí, é ... desses registros a USP tratou o registro dela, a UNESP o dela e nós tratamos o nosso. E aí fizemos todo o tratamento, padronização desses registros, que era 1000 e alguma coisa, ... cerca de 1500 registros. (...) E ali então

nasceu o nosso repositório com estes 1000 e tantos registros. E junto a isso nasceu o Repositório CRUESP, que aí nesse momento foi estabelecida algumas ... é ... foi estabelecida uma estrutura mínima de ... TI, (...) essa estrutura ficava lá na USP e ... inclusive na ocasião nosso repositório ele era hospedado na USP, né, mesmo sendo aqui, acesso nosso e tudo e aí, é ... esses, começou com esses 1000 e tanto registros lá na USP hospedados só que concomitantemente nasceu o Repositório CRUESP, (...) com esses 1000 e tantos da UNICAMP, e mais os outros registros, é ... aí um pouco mais que o da UNICAMP, da USP e não lembro agora a quantidade exata da UNESP viu (...), mas tinha uma quantidade lá da UNESP, aí nasceu o repositório CRUESP também e o nosso nasceu nesse momento. (...) Isso mais ou menos em 2013 inclusive nessa ocasião em 2013 teve um ... o CONFOA, que foi em São Paulo e aí nesse CONFOA que foi, é ... oficializada a criação do repositório, né, é ... foi apresentado o Repositório CRUESP e o Repositório UNICAMP, UNESP e USP. Então foi ali que deu o ... o gatilho mesmo pras questões de repositório. (...) ...~ É ... isso pegando aí esse ... esse histórico de como surgiu (G-UNICAMP)

2. Política de funcionamento e documentação do repositório institucional

A categoria “Política de funcionamento e documentação do repositório institucional” diz respeito às normas gerais estabelecidas para o funcionamento do repositório institucional e suas respectivas documentações formais e institucionais que respaldam as mesmas e a equipe atuante. No tocante à documentação, todas estão disponíveis nas páginas dos repositórios da USP, UNESP e UNICAMP (vide Anexos B, C e D).

De acordo com o gestor da USP houve a integração entre o repositório, a BDTI e o catálogo online DEDALUS e mais atualmente a documentação está sendo revisada pela equipe, para adequações de acordo com a Portaria 01/2019 da FAPESP³⁴ e as apresentações realizadas pela agência nas três universidades em 2019, especialmente referentes à disponibilização do PDF³⁵ dos recursos informacionais, tendo em vista que atualmente no repositório da USP o acesso aos mesmos é realizado por meio de links. Além disso, o *Repositório Institucional da Universidade de Dublin*, que, embora tenha sido desenvolvido em 2008 com uma interface simples do DSpace, com pouquíssimas alterações, tem sido dado

³⁴ Disponível em <http://www.fapesp.br/12632>, institui a “Política para Acesso Aberto às Publicações Resultantes de Auxílios e Bolsas FAPESP” e se aplica, conforme o Artigo 2º, à publicação de qualquer artigo ou outro tipo de comunicação científica, que contenha resultados originados de pesquisas apoiadas, parcial ou totalmente, pela FAPESP, em qualquer modalidade de apoio”, isto é, bolsas, auxílios à pesquisa e programas concedidos. No artigo 3º, inciso II, salienta-se que o acesso aberto diz respeito à “disponibilização em repositórios de acesso universal e aberto de artigos ou outros tipos de comunicação científica publicados”.

³⁵ O PDF (*Portable Document Format*), desenvolvido pela empresa Adobe, é o segundo formato mais utilizado na Internet e possui as seguintes características: mantém o aspecto original de um documento, realiza índices, hipervínculos e buscas no texto; possibilita idoneidade para a impressão de documentos e a conversão de quase todos os formatos em PDF (DIÉZ CARRERA, 2012, p. 80).

como modelo de repositório institucional pela FAPESP a ser seguido pelas três universidades paulistas.

*E a partir de 2018 houve uma integração entre o repositório mais tradicional que nós tínhamos, né, que era a BDTI, o DEDALUS, que é o nosso banco de dados bibliográficos e foi lançado uma outra versão que está disponível até hoje. (...) É... houve algumas... propostas, algumas discussões internas sobre a questão dos objetos digitais ... se eram necessários, se poderia ... mesmo sabendo que o conceito de repositório é que inclui o objeto digital ... se a gente não poderia só deixar com acesso direto pra links ... então houve essa proposta. Essa proposta foi aceita e está em curso ... só que ela foi questionada com a publicação da Portaria FAPESP. (...) E houve todo um questionamento “não, um repositório tem que ter o PDF” ... e, enfim, e agora estamos nessa fase de reestruturação para nos adaptarmos plenamente às necessidades apresentadas pela Portaria FAPESP. (...) Vai ter uma atualização agora, inclusive uma das atividades que a gente tá fazendo nessa reestruturação é essa atualização. (...) É ... olha, pra ser bem sincero já tinha que estar no ar, porque a gente já finalizou. Mas aí ... devido a essas questões de reestruturação eu diria que em breve, até o final do ano deve sair. (...) As principais mudanças são no nome, é ... em questões, é ... de abrangência ... e foco, e também a gente meio que fez uma Portaria que fosse ... é ... muito consoante com a Resolução da FAPESP. ... Acho que essas são as principais mudanças. (...) O resto, inclusive o texto é o mesmo, assim, sabe? ... Só que eu quero deixar claro que não está oficial, tá? **(G-USP)***

No âmbito da UNESP, não há previsão de atualização da documentação. Na UNICAMP, após a inauguração oficial do repositório, houve o estabelecimento de outras ações visando o aumento dos recursos informacionais inseridos nas bases de dados, incluindo o *Portal de Periódicos da UNICAMP* (PPEC). Entretanto, houve um menor nível de refinamento no tratamento dos metadados.

*Claro que depois houve aí uma série de ... de outras, é ... ações pra poder alimentar o repositório, aí passamos então a coletar todos os registros da ... da Web of Science, da Scopus, da PubMed, é ... depois agregou também os registros da ... do PPEC, aqui do nosso Portal de Periódicos e ... e SciELO. Então a gente buscou tudo o que se tinha, tudo o que a UNICAMP tinha nesses ... nesses, é ... nessas bases e inserimos no repositório, fizemos o tratamento básico, não o tratamento igual ao que foi feito com os 1000 e pouco, aquele foi um tratamento foi mais, foi um tratamento mais fino, né, foi um pouco mais, é ... apurado. É ... agora a alimentação dessas outras bases aí coletamos, aí aumentou sobremaneira os registros, porque aí coletamos cerca de ... somando tudo, acho que deu cerca, naquela ocasião, cerca de uns 40.000 registros, mais ou menos. (...) Aí fizemos um tratamento básico e inserimos no repositório ... naquela ocasião. Isso nós vamos fazendo sistematicamente essas coletas, dessas grandes bases e ... inserindo no repositório **(G-UNICAMP)***

No momento, aguarda-se a resolução de questões de direitos autorais junto à Reitoria para a aprovação da documentação. Por um lado, defende-se a concepção de que uma vez que o pesquisador tenha publicado enquanto vínculo com a universidade e não há nenhuma restrição em termos de direitos autorais, a disponibilização no repositório é obrigatória, visão defendida pela Pró-Reitoria de Pesquisa. Entretanto, para a Procuradoria Geral da universidade, embora as publicações tenham sido realizadas enquanto vínculo do

pesquisador, é necessária a assinatura de um termo de autorização para publicação no repositório, mesmo que o recurso informacional tenha sido verificado no SHERPA/RoMEO³⁶.

Então, nossa documentação ... ela existe desde 2016 ... mas é ... essa documentação ela, ela aguarda ainda um ... uma finalização da nossa Procuradoria Geral, no aspecto jurídico, porque tá enroscado numa parte que envolve ... direito autoral. (...) Então existe sim as diretrizes, né, basicamente uma política. ... Temos, isso já de maneira formal uma GR, uma Portaria GR né, e essa Portaria estabelece a função do repositório, o objetivo do repositório, é ... e os ... os atores do repositório e aquilo estabelece também ... é ... a responsabilidade dos ... dos pesquisadores vinculados à UNICAMP de se depositar, de se disponibilizar, ... de inserir as suas produções no repositório. Essa ... essa Portaria GR estabelece isso (G-UNICAMP)

Referente à equipe atuante nos repositórios institucionais, na USP a gestão é realizada de forma centralizada pelo Departamento Técnico do Sistema Integrado de Bibliotecas da USP (SIBi), enquanto que a inserção de dados, de metadados e de objetos digitais é descentralizada, isto é, realizada pelas Bibliotecas. A equipe atual é composta por três bibliotecários, sendo um gestor, que atuou na primeira fase do repositório entre 2012 a 2014 e retomou a frente do repositório em 2019 após uma reestruturação no SIBi; dois catalogadores e um bibliotecário-analista, atuando na informática. Aqui, o papel dos bibliotecários-catalogadores é atender dúvidas principalmente sobre objetos digitais a serem incorporados no repositório, direitos autorais, dentre outras e verificação da consistência e qualidade dos dados. O sistema de comunicação é centralizado no e-mail institucional para o recebimento de demandas, dúvidas, sugestões ou críticas, podendo também ser realizado o atendimento por telefone.

Na UNESP, a equipe atuante na gestão do repositório é composta por uma gestora e duas bibliotecárias, que atuam diretamente no atendimento por meio de um chat, pois, de acordo com a gestora, o volume de dúvidas é grande, especialmente após a implantação do autoarquivamento para alunos de pós-graduação. A descentralização para as Bibliotecas ocorreu há pouco tempo, sendo responsáveis por completarem os registros de teses e dissertações inseridos pelos alunos de pós-graduação no repositório institucional. Na UNICAMP, atualmente a equipe é composta pelo gestor e duas bibliotecárias, sendo uma supervisora que atua exclusivamente no repositório e uma catalogadora, que realiza ainda outras atividades, conforme a demanda, além de uma analista de sistemas, cerca de cinco

³⁶ Trata-se de um serviço online que aponta as políticas de depósito de obras com direito autoral e de acesso aberto das publicações acadêmicas. Disponível em: <http://www.sherpa.ac.uk>.

bolsistas e uma estagiária, que contribuem na inserção dos registros. Nenhum catalogador atua diretamente com a representação temática, apenas com a representação descritiva.

Quanto aos tipos de conteúdo dos repositórios, na USP são artigos de periódico, capítulos de livros, resenhas, dentre outros, inseridos na BDPI, sendo somente nas versões print ou pós-print, supondo-se a verificação da qualidade dos mesmos por meio da avaliação pelos pares, além das teses e dissertações já inseridas na BDTD. Na UNESP, os principais conteúdos são as teses e dissertações e alguns artigos de periódico inseridos no início da implantação do repositório. Relativo à inserção de outros materiais ou a produção retrospectiva, ainda não há previsão. Na UNICAMP, além de teses e dissertações, os principais conteúdos são artigos de periódico e patentes, abarcando ainda outros tipos de materiais.

É ...~~ as tipologias documentais depositadas no nosso repositório são variadas e também passarão por uma reformulação, porque, é ... tem algumas tipologias que eram muito utilizadas no passado e que nessa revisão a gente verificou que ... não faz mais sentido ter aquela diferenciação ou não faz mais sentido inclusive ter aquela tipologia. ... Então tem ... periódicos, artigos de periódico ... é ... capítulos de livros ... resenhas ... e outras materiais, sempre focando mas o principal assim, o principal, a tipologia foco mesmo são os artigos de periódico. E na versão ... é ... print ou pós-print. A gente por enquanto não está utilizando a versão pré-print porque a gente acredita que a Portaria é muito clara e expressa ao informar em ser um material foi, é ... eles escrevem assim, manuscrito aprovado. (...) Manuscrito aprovado ele passou por um crivo ... digamos, dos pares, né (...) e o pré-print não necessariamente enquanto a gente não deposita. Claro que isso pode mudar, inclusive porque a gente tem observado que ... em algumas palestras do Professor Brito tem preconizado que seja colocado. Mas até o momento a gente não deposita versão pré-print no repositório visando à qualidade e ... a revisão pelos pares (G-USP)

(Os tipos de recursos informacionais então ... é ... a priori são as teses e dissertações e alguns artigos de periódico que foram naquela época em 2013?). Não, eles estão sendo feitos até agora. (Até agora?) (...) Até 2018. (...) (Então dessas bases de dados que você citou no período de 2013 a 2018). Não não. Eu tenho desde o início da universidade (...) até 2018. ... (...) (Então vocês chegaram a trabalhar com os materiais retrospectivos dos docentes ou não?) É, porque quando eu faço a coleta automática ... eu vou coletar tudo o que tem da UNESP, pode ter coisa antiga, coisa nova, tudo misturado. (...) Eu puxei pelo nome UNESP. (...) O que veio, é ...~~ não me importa se é velho ou novo é produção da universidade. (Ah sim, mas só das bases de dados, que vocês fazem esse levantamento? Por enquanto). É, das bases que eu mencionei pra você. (Isso, porque por exemplo, se o docente publicou um livro impresso em 1980, é nesse sentido que eu pergunto do retrospectivo, ainda não chegou nessa fase, né?) Se ele foi indexado por alguma dessas bases que eu mencionei ... ele está dentro do repositório. (Sim, é ... mas se ele não foi, ainda não chegou nessa fase de digitalização desse material, é ... descrição dos metadados pra inserir no repositório, né?) Não porque eu trabalho completo automático. (...) Então a princípio nem sei se a gente vai chegar nesse ponto, a gente trabalha com bases onde já existem os materiais já feitos com ... descrições (G-UNESP)

3. O usuário do repositório institucional

A categoria “O usuário do repositório institucional” versa sobre a perspectiva dos profissionais que atuam no repositório institucional em relação à comunidade acadêmica que o utiliza, especialmente no que se refere à recuperação dos recursos informacionais. Na USP,

de acordo com o gestor houve a realização de um estudo sobre as buscas dos usuários no repositório entre 2014/2016, ocasião em que identificaram problemas na recuperação.

É ... fizemos. Isso nós fizemos, faz tempo que nós fizemos, nós fizemos em 2014 eu acho ... 2016, a gente vai ter que rever esse trabalho. Mas nós ... fizemos e ... bom, aí a gente identificou esses problemas (G-USP)

Referente à atual percepção dos usuários em relação às buscas realizadas no DSpace, há uma concordância por parte do gestor da USP e da bibliotecária catalogadora-indexadora da UNICAMP de que há problemas na recuperação pelos usuários, devido às próprias características da ferramenta na política do estabelecimento do nível de precisão e revocação do sistema. Entretanto, de acordo com o gestor, na USP existe um maior refinamento, pois possibilita buscas específicas nos campos “Título” e “Autores”, enquanto na UNICAMP, conforme verificado pela pesquisadora, não há nenhuma diferenciação entre as opções de busca: “Título”, “Autor”, “Assunto” ou “Data de publicação”.

Isso é uma reclamação geral, né. (...) É bem limitado. (...) Passamos e estamos desenvolvendo um ... um desenvolvimento do repositório mesmo pra que ele não faça ... dessa forma. (...) Então acho que a gente ainda tem ... softwares deficientes, digamos assim. (...) É a forma como ele foi estruturado, eu diria assim. Porque a gente tem que pensar que a proposta do DSpace foi fazer a pesquisa inclusive no objeto digital. (...) Então assim ... é ... pra não limitar eles acabaram abrindo muito, né, e aí os resultados acabam se tornando ... menos precisos, eu diria. (...) ... É, então ... não chega a ser um problema mas é uma ... uma limitação ((RI)), é uma limitação por ter proposto ser ... muito abrangente. (...) Ele acabou se tornando ... é uma característica ... eu colocaria melhor dessa forma. É uma característica do ... do software. (...) Que é totalmente ... customizável. E aí tem todos esses problemas da customização, né. E não necessariamente é tão simples assim. (...) O nosso ele já faz um refinamento. (...) Assunto a gente não conseguiu ainda (G-USP)

Qualquer coisa, qualquer informação que você coloca no DSpace se você colocar lá palavra solta ele busca. (...) É por isso que ele também ... ele não fala assim, eu quero tudo o que tá no resumo dessa palavra. (Não tem campo específico pra busca). Não. (Igual a uma base de dados?) Exato. (Seria o ideal). Exato. É isso que ele ... ele estoura. (...) Então vamos supor você faz uma busca de um autor (...) ele traz na primeira página algumas coisas do autor depois ele coloca 4000 negócio lá (...) páginas. (...) Então você nunca sabe o que realmente está buscando. É pior que Google, ((RI)) se for ver a busca do DSpace. Ele não é preciso em nada (I-UNICAMP)

Conforme observa a bibliotecária catalogadora-indexadora da UNICAMP sobre a ausência de precisão em geral no DSpace, pelos testes realizados pela pesquisadora verificou-se que a utilização do recurso gráfico aspas (“”) possibilita uma melhor precisão nas buscas. No entanto, assim como na UNICAMP, na USP ainda inexistia a possibilidade de buscas refinadas no campo “Assunto”, porém, com possibilidades de melhoria por meio de customização e desenvolvimento na ferramenta.

E uma das atribuições nossas no momento é essa (G-USP)

Cabe destacar que esta impossibilidade de refinamento nas buscas por assuntos é compartilhada pelo *Repositório Institucional da Universidade de Dublin*, tomado como exemplo de repositório institucional nas palestras proferidas pela FAPESP na USP, UNESP e UNICAMP em 2019.

4. Ações de melhoria no repositório institucional

A categoria “Ações de melhoria no repositório institucional” diz respeito àquelas ainda não realizadas nos repositórios que, na visão dos participantes, poderiam contribuir para a melhoria da padronização da representação descritiva e principalmente da representação temática. O gestor da USP observa a necessidade de melhoria contínua no repositório, mas de forma geral acredita que estão realizando um bom trabalho, sendo necessário alguns pequenos ajustes em relação aos assuntos. Embora a atuação do *Grupo Gestor do VOCAUSP* seja presente, considera ser trabalhoso devido à heterogeneidade própria do âmbito da universidade, em comportar várias áreas do conhecimento e os especialistas. Uma possível melhoria seria a inserção dos assuntos no VOCAUSP em tempo real.

Bom, é, no repositório de uma forma geral tem várias coisas que a gente vai ter que fazer pra melhorar. (...) Com relação à questão foco sua, né, que é a representação temática ... é ... eu acredito que a gente tem feito um bom trabalho, eu acho que a gente precisa acertar pequenas coisas, porque sempre dá pra melhorar, né. Mas ... de um modo geral ... tá funcionando bem. (...) É ... os termos novos que são sugeridos, a não ser que fosse mais rápido, né, a não ser que fosse em tempo real, mas tem todo um crivo, tem todo um grupo que precisa se reunir, analisar, e é um grupo formado por especialista, e às vezes ... você deve viver isso na UNICAMP também, é ... universidade é muito heterogênea, né. (...) Então tem preconceito, às vezes tem algumas brigas, “não pode ser assim, aqui não é assado”, é ... então isso também é um pouco difícil de gerenciar, né, no sentido ... não é que é difícil ... é ... até que é fácil, mas ... é trabalhoso. Né, muito trabalhoso e requer muitas articulações é ... (...), do tipo ... “não, nessa área é assim”, “na outra área é assado” mas precisamos entrar em um denominador comum e às vezes isso não é tão fácil. É ... agora do ponto de vista de implementação e de ... é ... sistemas eu acho que ele tem reproduzido o que é definido pelo grupo gestor, né (G-USP)

A gestora da UNESP considera que um dos grandes desafios é que os repositórios internacionais já estão trabalhando com repositórios de dados e no contexto brasileiro ainda há diversas melhorias a serem realizadas em repositórios institucionais. No tocante à representação temática, considera relevante que os assuntos utilizados nos recursos informacionais, isto é, a linguagem controlada dos especialistas, precisam estar interligados aos assuntos utilizados pelos usuários, isto é, à linguagem natural, nas buscas realizadas no

repositório institucional, proporcionando uma recuperação mais efetiva. Salienta a participação da UNESP no Comitê do Dublin Core, visando à melhoria na padronização dos metadados.

*Ah, é sem fim. (...) ((RI)) Você vê, ... infinito é pouco né, (...) é difícil porque ... é ... tem cada vez mais ... é ... mais coisas novas, né, você tem que ser mais ágil pra responder a esses ambientes né. Primeira coisa eu acho que o grande desafio é quando você vê que o mundo todo já tá trabalhando com pesquisas voltados pra dados né, (...) ... banco de dados, né, então, é ... só nesse ponto você já tem que ver como você precisa ... é ... melhorar seu ambiente em termos de tecnologia ... de recursos, né, de funções que dê conta de atender às necessidades dos usuários, né. Então, por exemplo ...~~ cada vez mais as buscas precisam ter assuntos, os assuntos precisam estar, é ... interligados com a ... como é que eu vou dizer ... interligados, tipo ao vocabulário da pessoa, né (...) mais vocabulário do ...~~ do vocabulário, é, não é do especialista, do vocabulário controlado nosso né, (...) das nossas ferramentas, com o vocabulário do especialista da área, entendeu, (...) a forma dele, a linguagem, relacionado à linguagem natural, tudo isso a gente precisa ser muito rápido pra poder trazer assunto né, ... porque isso interfere muito na questão da busca né. (...) É ... a questão dos metadados ... é ... a gente precisa melhorar cada vez mais, não deixar de lado, por isso a gente faz parte ... do Comitê do Dublin Core né, pra gente sempre estar, é ... a par das mudanças que ocorrem no ... no formato né. (...) Mais otimizado, né, então isso pra nós também é ... é bem importante que, mesmo porque tem toda aquela questão da ... de reuso, né, dos metadados (...), a gente sempre trabalha muito com isso, a gente sempre pensa muito bem em todos os campos que a gente coloca ... e que a gente trabalha. É ... em relação ... à questão ... é ... do assunto acho que é isso que eu te falei, essa ... essa ... ter cada vez mais uma maior, sei lá, integração, ligação entre a linguagem, tentar pelo menos mais se aproximar né, a linguagem natural e a ... e a do especialista, né, que é o que a gente tem aí controlada. ... (...) É ... ah, eu acho que em relação a metadados e a assunto ... eu acho que é isso. (...) Ah, os metadados até você amplia porque a gente falou de descrição, né, ... (...) mas hoje em dia tem muitas coisas quando você começa a falar desses ambientes, então você tem que falar, tem metadado de preservação, tem, tem um monte de coisa que a gente ainda não tem, (...) no meu caso da UNESP, né. Então, tem ... muita coisa que caminhar, mas isso daí a gente vai ... aos poucos né **(G-UNESP)***

Um dos bibliotecários que atuou na implantação do repositório da UNESP (B-UNESP-2) expressa a importância do papel do bibliotecário na visibilidade e conscientização por parte da instituição sobre o repositório institucional para a reunião e disponibilização da produção científica.

*Ah, eu acho que é assim, é ... as universidades elas precisam, a gente como bibliotecário a gente precisa vender um pouco melhor essa ideia porque essa questão de repositórios institucionais não é uma questão das Bibliotecas, né, é uma questão institucional. (...) ... Então eu acho que talvez a gente como bibliotecário precisa, é ... ir em reitoria, procurar as pessoas que mandam mesmo nas coisas pra vender um pouco essa ideia porque é algo pra instituição, né, não é algo assim que dá visibilidade pra Biblioteca, é o que dá visibilidade pra instituição como um todo. (...) E ... eu percebo que essa questão só tá avançando um pouco agora ... por conta das exigências das agências de fomento. ...~~ Então, assim, as agências de fomento já estão exigindo repositórios de dados científicos e a gente ainda tá engatinhando no repositório institucional. Então a gente precisa ... ir atrás mesmo de quem comanda as coisas e tentar vender essa ideia né. (...) ... E ... também vender essa ideia pros bibliotecários, porque há uma resistência muito grande na área de ... não sei se é uma resistência ou se é um medo, mas ... uma resistência mesmo de conhecer mesmo o que é, pra que serve, sabe, de ir atrás mesmo, dessas informações **(B-UNESP-2)***

Na visão do gestor da UNICAMP, independente do software utilizado, existe a necessidade de melhoria da integração com o catálogo online SophiA. Caso seja utilizado o

catálogo online para inserção, buscas e recuperação no repositório, uma das principais vantagens seria a possibilidade de utilização das tabelas já disponíveis no catálogo online, especialmente a Tabela de Assuntos, visando uma melhor padronização dos registros inseridos no Repositório Institucional da UNICAMP.

É, o que a gente precisa hoje é ... é vincular, seja se a gente fosse adotar o próprio DSpace mesmo ou seja qualquer outro software que venha ... que venha entrar no lugar do Dspace, o que nós precisamos é melhorar essa integração junto com o SophiA ou que se permaneça só o SophiA pra isso, pra que a gente possa utilizar as próprias tabelas já disponíveis pelo SophiA pra padronizar todo esse, fazer todo esse controle de ... né, da padronização dos dados em si, né, vinculados a esses metadados (G-UNICAMP)

Entretanto, o bibliotecário (B-UNESP-2) defende a utilização de softwares livres para os repositórios institucionais e expressa sua preocupação na obsolescência dos softwares privados e preservação dos dados, a exemplo do catálogo online *Pergamum*, utilizado em sua realidade profissional em outra universidade e do próprio SophiA utilizado pela UNICAMP.

eu acho que assim também ... sempre se preocupar como instituição em utilizar plataformas abertas, né, porque querendo ou não a gente tá ... disponibilizando uma produção ali que é ... institucional e que deve ser sempre aberta. (...) E ... às vezes a gente vai pra soluções privadas porque é mais fácil ... porque tem gente pra trabalhar. ... Mas eu acho que se a gente trabalhar com softwares livres e ... soluções livres, ... é ... a gente consegue ter assim, ah, uma forma de ... não ficar dependendo de empresas pra cuidar de uma produção institucional que é assim, que é pública, né. (...) Então assim, eu vejo assim, houve uma discussão no início lá na PUC de porque o Pergamum não poderia fazer isso, porque a gente não poderia comprar um software pra fazer isso. (...) É ... e aí foram levantadas questões do tipo, “ué, o Pergamum atende hoje, mas e daqui 10 anos?”... (...) É uma produção institucional que ... vai precisar de um certo cuidado maior, né. (...) Porque assim, querendo ou não os registros bibliográficos a gente pode, é ... coletar de outras fontes, há muitos materiais já catalogados, e tal, mas e a produção institucional? A gente não tem em outro lugar. Então, assim ... a gente ... foi levantada essa questão e aí foi decidido usar em forma de DSpace mesmo. Mas o grande problema do DSpace é que tem muitos poucos profissionais pra usar né, principalmente profissionais de TI (B-UNESP-2)

UNIDADE DE ANÁLISE: Aspectos relacionados ao bibliotecário catalogador-indexador em repositórios institucionais

5. Atuação profissional do bibliotecário catalogador/indexador no repositório institucional

A categoria “Atuação profissional do catalogador/indexador no repositório institucional” refere-se aos cursos e treinamentos realizados pelos bibliotecários catalogadores-indexadores das Bibliotecas para atualização profissional nos repositórios institucionais.

É, nós temos treinamentos, nós tivemos um primeiro treinamento agora em ... em junho, nós tivemos um grande ... digamos assim tira dúvidas ... tivemos alguns treinamentos no passado. Nós preparamos e nós ministramos esses treinamentos e nesse momento do treinamento também a gente conta com a ajuda [de outra colaboradora] que é responsável por uma parte de acesso aberto e aí nesses treinamentos também essa temática é abordada e é feito com o apoio dela (G-USP)

Olha só, eu peguei um grande problema, um fiquei com um gargalo, né, como eu abri pros bibliotecários eu tive que ... que passar estratégias de como manter o principal, o ambiente ... então ... considerando que mesmo eu tendo as duas bibliotecárias que me ajudam, se elas não tem o conhecimento, teriam que ter o conhecimento da sintaxe também. ... Então, por exemplo, quando eu vou fazer coleta pra aumentar o número de ... de registros no repositório ... eu faço tudo automático e é a máquina quem diz pra mim, ah... esse é, esse não é, esse foi publicado, ... isso é assim, isso é assado, entendeu? (...) E ... o que eu faço assim ... o que geralmente eu tenho mais problema, é ... eu percebo ... é o meu termômetro ... é quando ... começa a ter muitas dúvidas das pessoas, ou seja, dos alunos de pós-graduação, das secretarias de pós ou das Bibliotecas aí eu vejo que é o momento de ... ou soltar alguma informação, de chamar um vídeo e fazer alguma capacitação. (...) Caso contrário, o procedimento é sempre o mesmo, o fluxo é o mesmo, se eu não mudo o fluxo (...) eu não tenho que ficar, é ... dando capacitação toda hora. (...) Mas por exemplo, se eu tenho que mudar algum procedimento, eu vou agora inserir um procedimento pra patente ... tá, então essas duas bibliotecárias já fala com o pessoal da agência de inovação, tem as instruções pra patente, eu vou juntar a Pró-Reitoria de Pós-graduação pra também trabalhar algumas coisas relacionadas à patente aí eu vou escrever um procedimento ... vou chamar todas as Bibliotecas pra fazer uma vídeo e capacitá-las com informação e também as secretarias da pós. (...) Pra que todo mundo fale a mesma língua mas ... a não ser que eu tenha alguma coisa nova, diferente. São feitas em momentos diferentes, mas a gente chama todo mundo. O material não é ... na verdade não é da Biblioteca, né, a produção, ela é diretamente da Pós-graduação, então mesmo que a gente fizesse eles seriam chamados porque ... é ... diz respeito a eles (G-UNESP)

E para os bibliotecários nós fizemos, nós demos dois treinamentos, uma capacitação para os bibliotecários das Unidades naquele trabalho que estão fazendo ... com os assuntos, com todo o tratamento de tudo que nós temos hoje no repositório, isso vem ocorrendo de fevereiro de 2018 pra cá. (...) A ideia foi dar um outro treinamento reforçando alguns, alguns aspectos e dando treinamento pra aqueles que não, que ainda não fizeram (G-UNICAMP)

Nós fizemos dois treinamentos com ... alguns catalogadores (...) ... das faculdades, institutos e núcleos que se ofereceram a vir fazer, é, nós ... disponibilizamos pra quê? Começamos a arrumar os artigos. Qual é o grande problema? Tá tudo lá ... mas a gente não consegue fazer a pesquisa. Por quê? Não está catalogado, não está normalizado. (...) Padrão é padrão das bases de dados, não é padrão da UNICAMP, nada até hoje. Isso eu estou falando de artigos. (...) Teses e dissertações estão 100%. ... Então ... é ... nós disponibilizamos ... todos os artigos, que são 85.000 num drive ... pra ... esses bibliotecários, técnicos que fizeram o treinamento com a gente e nós estamos trabalhando de que forma? É ... os bibliotecários eles pegam por ... docentes as publicações deles ... e normalizam. (...) Normalizam todos os campos: título, autor, assunto, editor, todos os campos que tem ... nos artigos, eles normalizam, mandam para a DTRI, a gente coloca, vê se está tudo no padrão mesmo, o que não estiver a gente arruma aqui ainda e aí nós vamos subir pra ficar tudo no padrão, pra não ter mais esse problema (I-UNICAMP)

De modo geral, verifica-se que nas três universidades existe a preocupação com a formação e capacitação contínua dos profissionais que atuam em seus repositórios institucionais, ministrados pelas equipes gestoras. Na USP, as capacitações são voltadas para os bibliotecários catalogadores-indexadores que atuam nas Bibliotecas do SIBi e ministradas em parceria com a equipe responsável pelas questões de acesso aberto. No caso da UNESP, destaca-se a realização dos treinamentos não apenas para as Bibliotecas, mas também para

as secretarias de pós-graduação, tendo em vista o compartilhamento destes setores no fluxo de teses e dissertações na universidade. Os treinamentos são realizados sob demanda, de acordo com a percepção da gestão sobre as dúvidas e as dificuldades dos alunos de pós-graduação, secretarias de pós-graduação ou dos bibliotecários ou a necessidade de alteração do procedimento já bem estabelecido. Já na UNICAMP, os treinamentos são voltados para os profissionais que lidam com a padronização dos recursos informacionais, tanto bibliotecários quanto técnicos, tendo em vista a descentralização desta atividade para as Bibliotecas desde 2018.

**UNIDADE DE ANÁLISE: Aspectos relacionados ao processo documental
em repositórios institucionais**

6. Tratamento descritivo da informação no repositório institucional

A categoria “Tratamento descritivo da informação no repositório institucional” refere-se aos principais aspectos da representação descritiva no repositório: inserção de materiais no repositório, autoarquivamento, manual de representação descritiva e padronização dos nomes dos autores.

Na USP, são utilizadas duas bases de dados para o cadastro dos recursos informacionais: a *Biblioteca Digital de Teses e Dissertações* (BDTD), que contém apenas as teses e dissertações produzidas no âmbito da universidade e a *Biblioteca Digital da Produção Intelectual* (BDPD), com os demais registros de produção. Os dois bancos não estão interligados, sendo apresentados na interface de Busca Integrada, que possibilita a visualização dos metadados. Na UNESP, as Bibliotecas realizavam a catalogação dos registros de teses e dissertações no catálogo online Athena e os registros eram importados no repositório. Atualmente, o processo é inverso, isto é, primeiramente os metadados são inseridos no repositório institucional pelos alunos de pós-graduação no autoarquivamento por ocasião de sua defesa e em seguida são importados para o catálogo online Athena, sendo então validados pelo bibliotecário. Já na UNICAMP é recente a preocupação com a padronização dos metadados descritivos e temáticos, embora também utilize, assim como a USP e a UNESP, o esquema de metadados Dublin Core.

[...] são bases diferentes. Pra fazer essa junção a gente utiliza o Portal de Busca Integrada. (...) E aí ele pesquisa todos os nossos recursos informacionais. (Então não há nenhuma integração das duas bases em relação às teses e as dissertações?) ... Por enquanto não. O que talvez possa encontrar, que a gente andou fazendo alguns testes é que, por exemplo ... é ... quando o arquivo é fruto de uma dissertação, né, de uma tese, tem um link dentro dos metadados remetendo pra BDTD. (...) Mas fazer a pesquisa por teses e dissertações na BDPI você não vai encontrar. (Então, porque aqui no repositório ele tem assim, selecione a base, produção científica e teses e dissertações). Isso, aí vai pra outra base. (Aí ele varre nessa outra base, mas ele traz os resultados dentro aqui da ...). Isso, isso, só que ele traz os resultados como interface, só, a busca é na BDTD. Isso é só pra facilitar. (...) Pra integrar mesmo **(G-USP)**

(Sempre do DEDALUS pro repositório, nunca o contrário, certo?) Isso, nunca o contrário **(I-USP)**

Então ... é ... antes a gente tinha uma biblioteca digital de teses e dissertações que na verdade era o nosso próprio catálogo, né. (...) E a gente trabalhava com ... procurando o PDF e adicionava na biblioteca digital. Então na verdade as dissertações estão no repositório, mas antes era assim, as Bibliotecas catalogavam no catálogo ... e a gente puxava esses registros pro repositório. (...) Hoje é o contrário, como eu trabalho com autoarquivamento o procedimento começa no repositório, os alunos fazem o autoarquivo ... e de lá a gente puxa o registro já validado e completo para o Aleph **(G-UNESP)**

Os metadados são feitos ... existe um controle, existe um ... eles foram construídos, o padrão de metadados que a gente adota é o Dublin Core e o perfil de aplicação, construímos um perfil de aplicação ... pra atuar no ... pra disponibilizar pro repositório. (...) Então existe aí um perfil de aplicação que a gente segue em relação aos metadados, agora a padronização em si dos dados em si ... ela deve ser feita né. ... A gente tem iniciado esse trabalho de padronização, tanto é ... tanto assim em nível do ... de autoria como ... é ... como de assunto **(G-UNICAMP)**

Como mencionado, na UNICAMP existe a possibilidade de não utilização do DSpace para a descrição dos metadados no repositório institucional, utilizando-se, além do catálogo online SophiA, uma aplicação de repositório desenvolvido pela mesma empresa, Prima. Com isto, será possível o gerenciamento integral dos metadados, padronização e utilização de todos os controles de autoridade já utilizados pelos bibliotecários por ocasião da catalogação dos registros no SophiA, incluindo a Tabela de Assuntos. Além disso, possibilitará buscas mais precisas em campos específicos como “Título”, “Autor” e “Assunto”, o que não ocorre com o DSpace. Outra proposta em estudo é a substituição do DSpace pelo *Digital Commons*, plataforma de repositório institucional privada.

E ... agora ... pra 2019 o projeto é: ... os ... periódicos, os artigos de periódico, passar por ... também pro SophiA como analítica. Estamos estudando isso, essa possibilidade, se isso acontecer mesmo provavelmente nós vamos acabar com o ... o DSpace, nós vamos ficar com o SophiA. Se der certo, isso é uma proposta, então ainda não é 100% nem nada, só estamos em estudo. Mas ... de qualquer forma mesmo que fique com o DSpace a gente quer ver se o SophiA faz analítica pra gente, pra não precisar ficar mexendo no DSpace e pra gente ter tabelas de assunto, tabelas de autoridade, que é o que ... para os artigos também. E pra gente acabar com esse negócio de duplicidade, né, porque o ... SophiA nos fala ... se tem duplicidade ou não, coisa que o DSpace não nos traz. E o agravante: que o DSpace, ele, por mais que ele esteja tudo arrumado, (...) ... ele não traz a pesquisa ideal pra gente, mas isso é uma deficiência do DSpace. Todos reclamam disso. É ... eu fui ... no CONFOA (...) e foi a reclamação de todo mundo. É um programa livre, tá todo mundo usando (...) só que ele não nos retorna a pesquisa que a gente pede. E isso é o que os usuários mais reclamam. Então esse é um agravante pra todos, 100%. Então pra gente acabar com esse impacto a gente quer passar pro SophiA. Pelo menos todo mundo buscando no SophiA e vai ter

o retorno correto. Se formos continuar no DSpace ou não. Estamos analisando a proposta deles. É ... ou a gente vai fazer igual teses e dissertações (...) ou a gente vai acabar com tudo e eles ficam responsáveis, a gente faz um repositório pelo SophiA. Então a gente tá vendo qual ... fica melhor. E a gente já estaria usando o que é nosso, né. (...) Pelo menos pra ficar ... como estão as teses e dissertações, pelo menos essa parte. Ele virar analítica pra ficar igual as teses e dissertações. E continuar no DSpace ou ... eliminar o DSpace e ficar só no SophiA, a proposta é uma das duas (I-UNICAMP)

Referente às teses e dissertações, na USP são inseridas diretamente na BDTD pelos alunos de pós-graduação e posteriormente os registros são importados para o catálogo online DEDALUS e validados pelos bibliotecários das Unidades. Entretanto, de acordo com a experiência da bibliotecária catalogadora-indexadora, na prática as teses e dissertações estão sendo inseridas na BDTD pela própria Biblioteca.

*Olha, na verdade o que a gente faz aqui, eu acredito que a USP toda funcione do mesmo jeito, mas lá na Biblioteca a gente faz o cadastramento do material no banco, no DEDALUS, né. (...) Da produção, todo o material a gente faz no DEDALUS, né e do DEDALUS é que ele vai para o repositório, que ele é relativamente recente, né. ... Teve um ... há alguns anos atrás teve uma iniciação nisso, agora foi retomado, né, então ele vai do DEDALUS para o repositório, então quando a gente pensa no tratamento, a gente não pensa no repositório, a gente pensa de fazer isso pro DEDALUS que é a nossa base de dados, né. (...) Então ele vai ... dessa forma, como ele é feito no DEDALUS ele é ... ele é migrado pro repositório. (E isso pra todos os materiais?) *Pra todos os materiais.* (E os registros que são catalogados no DEDALUS eles sobem exatamente do jeito que está, do jeito que foi catalogado pro repositório ou tem alguma adaptação?) ... Não, que eu saiba né, o processo ele sobe exatamente como ele é feito, e toda modificação, a gente percebe às vezes, eu tenho trabalhado muito com isso porque agora ... o que eu tenho feito com relação ao repositório agora é fazer os uploads dos artigos de acesso aberto pro repositório e ... (...) e aí toda a inconsistência que eu percebo no DEDALUS eu ... conserto e isso automaticamente, acho que depois de uns minutos ele já é atualizado lá no repositório. (Então mesmo pros artigos vocês fazem a catalogação no DEDALUS primeiro depois). ... Depois que ele é migrado pro repositório (I-USP)*

[...] como a gente tem a Biblioteca Digital de Teses e Dissertações ... a gente ... lá na minha Biblioteca em particular a gente fez um trabalho retrospectivo. A gente ... foi atrás de alunos, de professores, tudo pra conseguir autorização das antigas, né, pra poder colocar e a gente fez um trabalho e colocamos todas as nossas teses, são muito poucas as que não estão na BDTD, né, que é a Biblioteca Digital de Teses e Dissertações. ... Hoje em dia a gente sempre faz o trabalho de ir lá, de cadastramento lá da BDTD para o DEDALUS ... porque ... é uma norma da USP, né, que tudo tem que ir pro DEDALUS (...) a gente exporta pro DEDALUS. Na verdade quem teria que fazer isso era o próprio aluno, mas lá na minha Unidade isso não acontece, então a gente tomou pra gente essa responsabilidade, que na verdade ou é do aluno ou é a pós-graduação, mas aqui, lá na minha Biblioteca quem faz somos nós, a gente faz todo o processo de ... de submeter a dissertação e de ... depois fazer a catalogação no DEDALUS. (...) Porque o processo de submissão na verdade teria que ser o aluno e aí a ... a pós-graduação faria a revisão e a gente faria a catalogação, mas isso não ocorre. Então como a gente ... depois que a gente fez aquele trabalho de colocar todo, fazer todo um trabalho retrospectivo a gente faz tudo (I-USP)

Na UNESP a inserção dos metadados das teses e dissertações é realizada pelos próprios alunos de pós-graduação diretamente no repositório institucional e depois exportados ao catálogo online Aleph, onde ocorre a verificação/correção pelos bibliotecários catalogadores-indexadores.

Então a pessoa faz o autodepósito ... as teses chegam aqui ... pra gente ... eu ... abro o registro, né, localizo aquele registro daquela tese ou dissertação específica (...) ... é .. abro também no repositório, então eu to trabalhando no Aleph né nesse primeiro momento, abro também uma cópia dele no ... no repositório então trabalho ... com a versão em papel, ... com a planilha, com o bibliográfico no Aleph e ... o documento eletrônico que foi feito o depósito no repositório. (...) (Quando o aluno faz o autodepósito automaticamente esses metadados que estão no repositório eles já caem no Aleph também?) Sim, é ... se não me engano foi feito algum aplicativo que ele, ele ... importa ... é ... a partir do momento que ele faz o autodepósito ele já importa pro Aleph. (...). Então, assim, muita coisa ali já tá de acordo, ... o nome, algumas coisas assim. Mas outras a gente tem que preencher, que é um trabalho braçal mesmo. (Quando o aluno faz o autodepósito gera esses dois registros, um no repositório e um no Aleph?) Sim. (E você trabalha com esses dois registros ao mesmo tempo, ou primeiro você trabalha com o registro no Aleph?) Sim, é ... eu abro, eu trabalho no MARC, o que eu faço? Eu vejo o ... to com a tese, o documento em papel, às vezes acontece de alguma diferença do título por algum motivo então a gente tem que ter esses dois abertos pra definir, é ... o válido ... é o papel. Esse é o documento ... matriz, vamos dizer. Então se precisar mudar alguma coisa vai ser mudado, com as devidas ... permissões no documento eletrônico. (...) Então o papel está comigo, o documento eletrônico está aberto ... e o trabalho é feito ... no MARC, no bibliográfico do Aleph. (Você lembra porque que teve essa opção, por exemplo, por que não foi optado). Sim. (É ... por que não utilizar somente um registro igual tanto no repositório quanto no Aleph? Porque acaba ficando dois registros diferentes, né). É. ...~~. É, não, assim, quando ele ... quando chega ... quando o indivíduo faz o ... o registro dele no repositório e gera pra gente eu trabalho encima daquele, ele já tem um número ... um 'UM' que a gente fala né, que é o CPF daquele registro (...), de verificação única dele ... então ele vai ficar somente um registro. (...). Não sei se é isso, não sei se eu entendi o que você perguntou, então eu trabalho encima do que ele fez, ele vai gerar, é ... ele vai ser modificado mas vai ser único aquele. Não vão ser dois registros (I-UNESP)

Na UNICAMP, as teses e dissertações são catalogadas primeiramente no catálogo online SophiA pelos bibliotecários catalogadores-indexadores e posteriormente importadas para o repositório institucional, exatamente com os mesmos metadados, tanto descritivos quanto temáticos.

Tese e dissertação, o que que aconteceu? O ano passado nós pedimos pro suporte, nosso suporte do SophiA, que é o que nós usamos ... é ... a gente fazer uma parceria. ... Nós tiramos tudo que estava no Nou-Rou que nós usávamos de tese e dissertação e também estava catalogado no nosso acervo no SophiA e ... hoje eles conversam, SophiA conversa com DSpace. Então nada, nada nada que trata de tese e dissertação a gente pode fazer no DSpace, nada. Tudo nós fazemos no SophiA. Qualquer ponto, uma vírgula que nós mudamos no SophiA ele muda ... no DSpace. (...) É automático (I-UNICAMP)

A inserção de registros de outros tipos de materiais no repositório é realizada por meio de coletas automáticas nas três universidades, importados das principais bases de dados *Web of Science*, *Scopus*, *SciELO* e *PubMed* e validados pelos bibliotecários antes da inserção nos respectivos repositórios. Na UNICAMP está sendo estudada a coleta no Currículo Lattes, principalmente para contemplar a *Área de Artes e Humanidades*. Enquanto na USP estes registros são distribuídos às Bibliotecas, de acordo com a produção de seus usuários, na UNESP e na UNICAMP esta validação das coletas automáticas é realizada pela equipe de gestão do repositório.

*A gente faz uma coleta automática e principalmente dos metadados em diferentes fontes só que isso não fica disponível até que tenha passado pelo crivo de um bibliotecário. (Aí no caso do bibliotecário é aí do SIBi ou das Bibliotecas?) Das Bibliotecas a maioria. Do SIBi normalmente a gente só faz quando tem alguma dúvida, algum problema e a gente tem que verificar ... se o número do ORCID³⁷ não está correspondente ... seria mais na parte de consistência mesmo. (Mas quando vocês fazem as coletas automáticas vocês distribuem pras Bibliotecas de acordo com ... com qual critério?) Docente e pesquisador, ou aluno de pós-graduação ... ou pós-doutorado, é com esse critério que a gente utiliza porque a gente já tem mapeado, na grande maioria, tá, existem alguns que não ainda, mas na maioria a gente tem uma integração do nosso sistema com a base institucional. (...) Então a pessoa entra, é apresentado uns menus, ela entra pra essa base institucional ... e nós já fazemos esse mapeamento automático. ... Por exemplo, então um aluno de graduação publicou um paper. ... Esse paper será ... esse paper saiu assim na Web of Science. Então os metadados estarão, vai ter uma informação que “olha, tem um aluno que publicou que é da sua Unidade”, né, e as Bibliotecas tem isso mapeado (...) “e ele publicou esse artigo aqui, você não quer colocar ele na BDPI, você pode colocar ele na BDPI?”, aí vem o bibliotecário. (...) (E aí eles fazem essa validação pra depois poder esse material ser definitivamente inserido ali no repositório?) Exato **(G-USP)***

*Então ... é ...~~ na verdade, a gente trabalha com coleta automática, né. Então ... é ...~~ os únicos que preenchem metadados mais são os alunos de pós-graduação. Então a gente trabalha com o autoarquivamento, né. ...~~ Os outros tipos de materiais a gente ainda não liberou. Ou seja, outros tipos de materiais, ou outros materiais ... que não estejam indexados na Web of Science, Scopus ...~~ SciELO ... PubMed. ...~~ Acho que a gente tem umas quatro fontes de coleta, a gente coleta de forma automática, né **(G-UNESP)***

*O Web of Science, Scopus e PubMed eu faço a coleta ... tá ... nas bases. A SciELO quem faz a coleta é a TI (...) porque existe um script que a SciELO disponibiliza pra UNICAMP e é coletado tanto é ... a bibliografia, os metadados, como os PDFs, só que mesmo assim essa documentação sobe pra DTRI, a gente arruma, normaliza, organiza tudo e volta pra subir pro repositório. Então ... tudo é feito por coleta, pegamos todos os PDFs tirando a SciELO ... e ... arrumamos pra DTI subir via script, tudo é via script. (...) Duplicidade entre essas bases e entre elas mesmas é feita pelo gestor ... através de ... do Excel, não existe um programa. (Em lote?) Em lote. Então é comparado ele compara o que já tem e o que ... nós estamos subindo porque entre elas há duplicidade, então a gente tira também. Tira 100%? Não. Tem muita coisa duplicada. Então esse foi o processo até hoje. (...) ... E nós também temos as patentes. As patentes, é ... quem nos fornece é a Inova. ... Então ela nos fornece tanto os metadados como os PDFs também, então o que já pode ser publicado no repositório é eles que nos dão, a gente nem precisa ir atrás. Tá, então a Inova ela pode publicar tá aqui com a gente, a gente coloca lá. (E os que não podem ainda ser publicados, essas patentes, vocês colocam só os metadados ou não colocam?) A gente não coloca nada porque a gente não tem nem acesso, eles deixam fechado mesmo. (...) Então essa política não tem, é só quando eles vão liberando eles vão nos entregando. Aí a gente já sobre pra lá. E quem sobe tudo, a gente não mexe no DSpace, quem sobe tudo é a DTI, é só o TI que mexe, então nós fazemos essa política. Nós estamos estudando também o Lattes, com o agravante de que ... é ... a gente só vai colocar enquanto ele está aqui, então nós vamos ter que fazer essa limpeza né. E aí a hora que a gente, é ... pegar o período que o professor tá aqui, o aluno está aqui na UNICAMP, e funcionários, é ... aí a gente corre atrás da publicação. Então a gente não vai alimentar tudo de uma vez, a gente vai fazer por blocos tá, esse trabalho vai ser feito dessa maneira **(I-UNICAMP)***

³⁷ O ORCID (*Open Researcher and Contributor ID*) é um identificador digital persistente e gratuito que identifica e padroniza um autor, permitindo armazenar e gerir informações relativas à formação e à produção acadêmica. O ORCID vincula o pesquisador às publicações de sua autoria e/ou coautoria e torna a busca por autores mais precisa e menos sujeita a erros, ao passo que elimina o problema de homônimos e de variações nos nomes dos pesquisadores. É composto por um código numérico de 16 dígitos (ORCID iD), por exemplo, <https://orcid.org/0000-0001-2345-6789>, compatível com a norma ISO 27729/2012 - *Information and documentation – International Standard Name Identifier* (ISNI). Até o momento, existem 7.526.446 de registros ORCID no mundo (dados extraídos em 19/11/2019) (OPEN RESEARCHER AND CONTRIBUTOR ID, 2019).

Na UNICAMP há uma proposta de inserção de todos os recursos informacionais retrospectivos dos últimos cinco anos por meio de criação de analíticas no catálogo online SophiA, com atualização automática no repositório institucional. Entretanto, ainda são aguardadas decisões políticas, principalmente referente aos direitos autorais.

*[...] nós vamos colocar ... tudo. 100% de tudo o que acontece na UNICAMP em questão de publicações. Mas nós vamos, É ... nos concentrar nos cinco últimos anos sempre, por que? Por causa das métricas dos professores, eles precisam disso pra ... pra apresentar pra FAPESP. (Para os indicadores mesmo?) Isso, pra apresentar pras agências de fomento. Então ... como eles precisam de 5 anos nós vamos concentrar nos últimos 5 anos, de 2019 pra trás, depende de quando a política realmente estiver 100%. Por que isso? Quando ela estiver 100% vai ser lá na reitoria, vai ser ... são eles falando ... é assim. Por enquanto a gente não pode colocar por causa de restrições, né. Mesmo que a gente fale está fechado, mesmo assim a gente não pode colocar. ...~~ (Mas pros capítulos de livros seriam como ... os artigos de periódico, seria a avaliação individual de cada capítulo ... com os assuntos). ... Sim, seria uma analítica do ... do livro né. (Todos os recursos informacionais seriam feitas analíticas). Isso. Exato **(I-UNICAMP)***

Um ponto a ser destacado é a ausência de padronização nos registros dos recursos informacionais oriundos das bases de dados, tanto no que se refere aos aspectos descritivos como filiação e autores quanto aos aspectos temáticos. Em relação ao autoarquivamento para teses e dissertações, na USP o processo já é bem estabelecido entre os alunos de pós-graduação, que realizam o autoarquivamento na Biblioteca Digital de Teses e Dissertações (BDTD). No entanto, para os demais materiais houve uma tentativa de implementação desta funcionalidade, tanto realizado pelo próprio autor, sem validação/verificação pelo bibliotecário quanto com validação/verificação.

Pela experiência observada pelo gestor, o autoarquivamento sem validação/verificação abre espaço para a ausência de padronização dos metadados, comprometendo a qualidade dos mesmos, além de problemas sobre as questões legais envolvidas. Atualmente, com a orientação da FAPESP de que a responsabilidade pelo upload da produção seja das Bibliotecas, acredita que a possibilidade de autoarquivamento pelos autores é mais remota. Para a bibliotecária catalogadora-indexadora da USP, o procedimento não funciona na prática, pois os usuários não costumam assumir o papel de inserção dos recursos informacionais por meio do autoarquivamento dos demais materiais, no contexto de sua Biblioteca.

*(Essa questão do autodepósito vocês pretendem implementar?) Olha ... eu não sei mas pela minha experiência, pelo menos na minha Unidade é uma coisa muito difícil de ser feita, viu? Professor eles não têm essa ... ((RI)), essa disposição de ... de colocar. Ou é a Biblioteca que faz ou os professores não vão fazer não **(I-USP)***

Então, isso é um problema, tá ((RI)). Inclusive quando a gente lançou, o meu mestrado foi uma proposta de submissão ... de promoção de autoarquivamento. ...~~ Só que nunca deu certo de forma ... holística, né. ... Tem dados que deu muito certo, funcionam muito bem mas não foi a maioria delas e ... aí também

... principalmente agora com a portaria da FAPESP nós ... vimos que o Professor Brito deixou bem claro lá, né, ele falou "olha, quem vai ter que fazer esse trabalho de inserção dos objetos digitais, da produção são as Bibliotecas ... e ... porque eu não quero dar nenhum trabalho extra pro pesquisador". Então se já era difícil antes agora se tornou assim mais complicado. (...) Então a gente gostaria de implementar o autoarquivamento, a gente tem algumas Unidades que demandam isso de nós ... mas ... de forma ... holística, assim ... acho que vai demorar bastante. Infelizmente. (Mas aí a proposta seria que esse autodepósito seja feito pelas Bibliotecas ou pelo próprio autor?) ... Seria pelo próprio autor a proposta. Mas as Bibliotecas tem perfil pra ... depositar pelos autores, né, das suas respectivas Unidades. (E no caso se fosse o autodepósito pelos próprios autores, ... ele já iria cair diretamente no repositório ou passaria por algum tipo de validação pelo bibliotecário?) ... Olha, lá no passado a gente fez os dois caminhos. ... É ... nós fizemos o autodepósito com o ... é ... declaratório, né, do próprio autor, sem nenhum tipo de validação por parte dos bibliotecários e nós fizemos com validação. É ... eu acho que ... se um dia nós viermos ... isso é muito instável, muito mutável eu diria ... mas se um dia nós viermos a ... é ... implementar totalmente o autoarquivamento ... provavelmente vai ter um trabalho aí de validação pelo bibliotecário ... porque das experiências anteriores a gente viu que ... por não ser da área, né, e também tem muito da importância do cuidado dos metadados. ... (...) É ... tivemos alguns problemas que no final a gente depois. ... E um dos maiores problemas é com relação a ... aos direitos autorais, né (G-USP)

Na UNESP, não há um profissional específico que realize a catalogação de materiais no repositório. Os metadados de teses e dissertações são inseridos por alunos de pós-graduação diretamente no Repositório da UNESP por autoarquivamento, por ocasião de defesa de dissertação e tese. Em seguida, os metadados dos registros das teses e dissertações são exportados para o catálogo online Athena, onde ocorre o processo de validação/verificação/correção/ampliação dos metadados de pelos bibliotecários catalogadores-indexadores das Unidades, após o recebimento dos registros importados do repositório. A quantidade de materiais processada está ligada diretamente ao número de Programas de Pós-graduação e ao número de alunos de pós-graduação formandos. Como observado pelo bibliotecário catalogador-indexador da UNESP, o processo, implementado há aproximadamente 5 anos por uma necessidade da comunidade acadêmica, tem facilitado a atividade de catalogação no que se refere à digitação dos metadados, pois os alunos de pós-graduação já realizam uma pré-catalogação das teses e dissertações.

(É ... você lembra quando que foi implementada a questão do autodepósito?) O ano exato eu não me lembro, mas ... mas faz mais ou menos uns ... quatro ou cinco anos atrás, talvez seis. (...) Era uma necessidade né, uma necessidade de ... necessidade em geral, da comunidade acadêmica em geral mas a UNESP largou um pouco atrás nessa questão de repositórios né ... e ... conseguiu um esforço grande do pessoal aqui da universidade e ... fazer um harvesting em geral de arquivos e tudo o mais e ... estabelecer essa política ... das teses e dissertações pra se tornar mais visível e pra povoar mais também o ... o repositório, né. Mais ou menos isso, 5 ou 6 anos. (E na época teve algum treinamento pros bibliotecários do sistema?) Eu não me recordo exatamente, eu acredito que não, porque ... aí eu dou uma opinião particular né, que é uma coisa relativamente natural, não foi algo assim, tão ... diferente do que já se fazia. (...) É ... até a gente ganha um pouco de tempo porque antes a gente descrevia do zero né e agora o autor já faz uma ... já descreve tudo embora nem tudo tenha a qualidade do padrão mas ... simplificar um pouquinho. (E você lembra se teve boa aceitação por parte ... das Bibliotecas da Rede?) É, com relação à questão da ... eu não posso falar, eu não conversei tanto com as pessoas sobre isso, mas eu acredito que sim, não é algo que interferiu negativamente no trabalho (I-UNESP)

Mas hoje ...~~ nós não temos um ... um bibliotecário ... é ... que faça catalogação no repositório ... a gente trabalha no repositório com coleta automática. ... (...) E quando se trata de tese e dissertação, é ... a gente pede ... o aluno de pós ... ele preenche um conjunto de metadados livres e esses metadados eles vão parar na Biblioteca e um bibliotecário valida nesses catálogos, tá. Então existe a validação por qualquer, é ... bibliotecário da Rede. É ... todo mundo, todas as Bibliotecas ... na verdade é ... onde tem programa de pós-graduação (...) ...~~ tem uma Biblioteca que alguém recebe pra validar os registros de tese e dissertação, porque, com ... um número mínimo de registro de metadados ... e aí a Biblioteca fica por ... é ... ampliar o número de ... de campos e verificar se o aluno preencheu corretamente. É ... testamos bastante antes de colocarmos ele em ... em produção, né? E ... assim, a quantidade depende de quem tem muito programa de pós tem muito aluno de pós, logo terá muitos ... muitos ... muitas teses e dissertações. Então se você pega programa né, uma faculdade que tem bastante programa de pós-graduação você vai ter ... por exemplo, Sorocaba acho que tem muito programa. ...~~ Marília deve ter uns 5 ou 6 ... é ... o IBIOS ... Rio Preto deve ter mais, lá tem as três áreas do conhecimento mais um monte de coisas, entende, então assim, isso vai ... depender da diversidade dos programas, assim, a quantidade, né. (...) ...~~ (E ... esse bibliotecário que faz a validação dos dados, é um bibliotecário de referência ou é um catalogador?). Ele é catalogador. Ele é catalogador. ... Na verdade ele valida, é ... registro bibliográfico. (...). (É um em cada Biblioteca?) ... Ahh não, eu sempre falo pra elas, que se uma pessoa morrer (...) a loja tem que continuar aberta ((RI)). Eu ...~~ sempre sugiro pra por duas pessoas, mas eu não sei como é que tá, elas reclamam muito que ... é mais serviço, mais trabalho, então, mas eu já avisei pra elas e elas sabem disso, que o aluno depende dessa validação ... pra ...~~ conseguir enviar um documento que eles tem lá pra pós, pra secretaria da pós pra pegar o título, entendeu? Então elas sempre correm pra fazer esse tipo de coisa, né, elas tem que agilizar. (...) Duas pessoas se está de férias, né (G-UNESP)

Outro ponto a ser destacado é que a tipologia documental compartilhada tanto no repositório quanto no catálogo online Athena em maior quantidade são as teses e dissertações, isto é, é possível encontrar o registro de uma determinada tese ou dissertação tanto no repositório quanto no catálogo. Além das teses e dissertações, livros editados pela Editora da UNESP e e-books também podem ser encontrados em ambos bancos de dados. Quanto aos assuntos, verifica-se que os livros e e-books são catalogados no catálogo online Athena e posteriormente exportados para o repositório, permanecendo com os mesmos assuntos já tratados pelos bibliotecários catalogadores-indexadores. Entretanto, no caso das teses e dissertações verifica-se atribuição de assuntos diferentes no catálogo online Athena e no repositório, já que no catálogo ocorre a validação/verificação dos assuntos pelos bibliotecários de acordo com o vocabulário da UNESP e no repositório os termos são livres, elencados pelos próprios alunos de pós-graduação no momento do autoarquivamento.

(O único tipo de recurso informacional que tem tanto no repositório quanto no Aleph são as teses e dissertações?) *Eu acredito que sim, muita coisa que tem às vezes é até de eventos, que eles vão fazendo ... pegando aí pela web né, fazendo o harvesting da coisa né, eu acredito que sim. Eu acredito que seja isso. Bom, pode ser que tenham livros. ... (...) Livros às vezes da Editora da UNESP, publicação às vezes de e-books então eu acho que não são só as teses e dissertações não. (Aí no caso dos livros ...) É, livros é ... produzidos por ... pessoal da UNESP né. Professores ... pesquisadores. (Isso, aí os livros eles são catalogados no Aleph e importados pro repositório?) Isso eu não sei te falar. Isso eu não sei dizer. A gente não faz esse trabalho na Biblioteca. (...) Às vezes chega um livro em papel aqui por exemplo e ... a gente já cataloga, às vezes já existe uma versão eletrônica dele. (...) É ... se é pela Editora da UNESP eu acredito que a própria Editora da UNESP faça esse tipo de ... (...) Eu não, aí eu to supondo, que eles façam esse ...*

esse trabalho. (Tudo bem. A Biblioteca fica responsável só pelas teses e dissertações no repositório, né?)
Sim. (Os demais materiais provavelmente é a CGB que faça). ... *Exato (I-UNESP)*

Já na UNICAMP a proposta é a implementação do autoarquivamento assim que as políticas e diretrizes do repositório forem aprovadas. Será iniciado um trabalho junto às Bibliotecas de coleta da produção retrospectiva, desde a fundação da universidade e os próprios autores terão autonomia para depositar os recursos informacionais diretamente no repositório, em um sistema baseado no atual sistema de elaboração de fichas catalográficas. Tal procedimento poderá ser realizado pelo próprio usuário, secretarias de pós-graduação ou Bibliotecas. Entretanto, no caso de autores e secretarias de pós-graduação, os registros serão validados pelos bibliotecários catalogadores-indexadores, visando à qualidade dos metadados.

*Então, passada essa ... essa fase do ... é ... né, de ... da política, das diretrizes, aí vamos iniciar um trabalho com, junto às Bibliotecas ... né, inicialmente de coletar o passivo, tudo o que nós temos, é ... tudo o que nós fizemos ao longo do ... dos anos, aí, tudo o que nós, a UNICAMP publicou ao longo dos anos, desde sua fundação. Isso é um trabalho que vai ser feito em conjunto com as Bibliotecas. Todavia, por exemplo, quando realmente ... é ... é ... superarmos essa fase do ... dos antigos, daqui pra frente, os documentos que vão sendo publicados daqui pra frente o ... cada autor, cada, é ... pesquisador vai fazer sim uma espécie de um autodepósito, né. (...) Pra nós, principalmente de seus metadados, e esses metadados vão ser validados, ele não vai entrar diretamente no repositório, ele vai entrar num, num banco paralelo, intermediário, pra depois ele ser inserido no ... no repositório em definitivo, após o nosso tratamento, tratamento feito pelos bibliotecários aqui. (E seria utilizado um sistema como o de ficha catalográfica hoje?) Perfeito, um sistema como o da ficha catalográfica. (...). A ideia é essa mesmo, desenvolver um sistema como o da ficha catalográfica pra ... pra coletar esse ... entre aspas esse autodepósito, os metadados dele e o seu documento e aí a gente faz a validação e insere no repositório. (E a validação seria feita pelos bibliotecários das Unidades?) A ideia sim. (Ou seria centralizado?) Não, a ideia é ser feito pelo ... é descentralizar nesse momento. Só vai centralizar se a gente conseguir uma equipe, que é aquilo que a gente requer. Do contrário as próprias Bibliotecas farão essa ... essa validação. Mas assim, ... é importante registrar. O nosso sonho mesmo é que ... tenha uma equipe centralizada, única trabalhando com isso. ...~~ (No caso se tivesse equipe ... os bibliotecários não iriam mais ...). Não iriam mais atuar nesse sentido. O repositório, toda a gestão toda a inserção, toda a operacionalização do repositório seria feito por ... pela DTRI, né. Um órgão centralizado **(G-UNICAMP)***

Na visão da bibliotecária catalogadora-indexadora da UNICAMP, a grande vantagem da implementação do autoarquivamento na universidade seria o acesso às publicações da comunidade acadêmica, em detrimento da descrição dos metadados pelos usuários, atividade que seria realizada pelos próprios bibliotecários catalogadores-indexadores.

E aí, o que eu acho mais vantagem é o seguinte: os bibliotecários eles vão ter acesso, coisa que a gente não tem hoje. Então vai fazer um autoarquivamento no sistema que a gente tá pensando em fazer igual o da ficha catalográfica ... mas não preencher tudo aquilo, se o professor, se a secretária for colocar o PDF pra gente, (...) não precisa fazer mais nada, só pôr o PDF aí o bibliotecário vai lá e faz a catalogação ... é o suficiente pra nós, não precisa de mais nada, pelo menos a gente tem a documentação. (...) Então a gente tá brigando por isso, então ele faz o autoarquivamento ali e já vem tudo, pro bibliotecário catalogar no SophiA, que ele já tem familiaridade com isso. (O usuário não precisaria preencher os metadados?) Nada.

(Só colocaria o PDF). ... *E a gente faz o restante. A gente não perde as publicações e assim, o que a gente quer pegar é o seguinte, toda vez que um professor, um aluno ... sai da UNICAMP pra fazer alguma coisa ele tem que prestar contas pra sair ... então esse aprovar e sair ele já vai entregar o trabalho, porque ele tem que fazer isso mesmo, ele tem que entregar pelo menos o resumo. Então esse entregar o resumo já tem que ir pra Biblioteca ele está indo para um congresso, apresentar ... sei lá, ... fazer uma palestra, então já tem que vir pra cá, ele vai pra algum evento, então ... ele publicou o artigo ... na Web of Science, então ele já coloca pra gente, entendeu? Ele já deposita ali aberto ou fechado ele vai ter essa opção mas já vai ficando no repositório. (Mas teria que ser uma questão política?) Não, essa questão está sendo tratada com a ... PRDU, vai sair de lá pra eles fazerem, já ficar lá no começo, entendeu? (...) A gente não vai exigir, vai ser uma exigência deles pra gente já colocar aqui. (...) É o mesmo que tá acontecendo com ... que vai acontecer, que está acontecendo com o ORCID. A PRDU tá pedindo, tem que fazer e já mandar pra Biblioteca, pra gente fazer essa varredura né, o robô faz, a gente não precisa ficar correndo atrás, então se todos colocam fica mais fácil. Nós vamos deixar livre ... porque ... geralmente quem faz é a secretária, né, mas nós vamos deixar livre. (...) Ele tá fazendo, ele tá pedindo, ele quer sair, ele quer colocar tudo bem, ele vai pedir pra secretária colocar? Não tem problema. A gente vai deixar livre pelo seguinte, é ... a gente precisa da documentação. Tem alguns que vão entregar na Biblioteca, sem problema, pode entregar. Desde que eles entreguem, a política vai ser essa. (A validação será feita pelos bibliotecários das Unidades?) Pelos bibliotecários. (Não será concentrado na DTRI?) Não, aí vai ser pros bibliotecários. O que a gente vai ver é o andamento do trabalho. (...) Como a gente dá o apoio ao andamento do trabalho. (A qualidade). ... **Exato (I-UNICAMP)***

Quanto ao manual que norteie a atividade dos bibliotecários catalogadores-indexadores que atuam no repositório institucional, na USP verifica-se a existência de um manual de catalogação no DEDALUS, composto por um tutorial para inserção dos materiais e procedimentos de solicitação de abertura de novos termos ao *Grupo Gestor do VOCAUSP*. Na UNESP também verifica-se a existência de um manual, assim como no contexto da UNICAMP.

Como o primeiro cadastramento é feito no DEDALUS, que utiliza formato MARC ... tem um manual desse cadastramento inicial em formato MARC e o tutorial para os uploads e os acertos quando necessários na BDPI, no repositório (G-USP)

Eu tenho uma ... uma apostila né pra ... pra livros, teses, dissertações, TCCs então a gente usa isso como uma ferramenta de auxílio pra ... ajudar na memória de preenchimento de alguns campos aqui (I-UNESP)

No que se refere à padronização dos nomes dos autores, na UNESP a autoridade dos autores é padronizada somente no catálogo online Athena, quando o bibliotecário catalogador-indexador utiliza a Tabela de Autores da própria UNESP.

(Vocês tem, vocês usam ... alguma base de dados pra controlar a autoridade dos autores?) É ... tem a ... da própria UNESP. Na própria UNESP a gente tem, né, no campo ... sim, tem. ...~~ A gente faz autoridade, descreve o autor, se ele já tá descrito a gente aproveita, se não há descrição a gente cria um registro de autoridade pra ele. (Tem uma tabela dentro do Aleph só de autoridade de autores?) Tem. (...) (É ... essa mesma autoridade é utilizada no repositório?) É ... no repositório, você diz, pelo ... pelo autor da tese ou dissertação? (Isso). Eu ... é ... ele tem um ... ai, me foge agora o ... o termo, aqui, mas ele tem ali ... um ... um tesouro. Mas o tesouro não sei, não creio que a base de autoridade esteja interna ao tesouro, deve ser somente pros termos, pros campos 600, o que seria o equivalente aos campos 600 né, o autor não sabe que ele tá fazendo isso, mas eu acredito que autoridade não (I-UNESP)

A bibliotecária catalogadora-indexadora da UNICAMP expõe que é realizado um trabalho de padronização dos nomes dos autores no catálogo online de acordo com a preferência do autor, utilizando-se também remissivas para outras possíveis entradas. Verifica-se ainda o andamento da implementação do ORCID na USP, UNESP e UNICAMP.

*O que a gente usa aqui pra autor é o seguinte ... é ... ele vai chamar o que ele quiser (...), tá... é o que vai tá ... o que a gente chama no 100, metadados do autor. E a gente faz as remissivas pra todos os que estão no Lattes. E a gente faz uma remissiva hoje é o nome dele por extenso. E ... e usamos data, tá. Ele não quer data a gente coloca num campo que é o 041 (...) a data dele pra gente saber que é ele e usamos qualificador se ele não quer a data pra não confundir com outros ... outros autores. Mas é isso que a gente usa. É o que ele quer. Se ele mudou o nome, é ... casado, depois separou, depois ... é o que ele quer usar como publicação dentro da UNICAMP a gente deixa e o restante é remissiva. (E aí lá no repositório se eu buscar pela remissiva eu vou encontrar?) **Encontra (I-UNICAMP)***

[ORCID] *Já tá. Ele vai subir já com ele completo. E ele vai passar né no repositório pegando. (Ele vai puxar informações, o repositório vai puxar informações do ORCID pro professor atualizar?) Os dois, um vai alimentar o outro. Nós vamos fazer as duas coisas pra ... pra não ter mais esse negócio, ah, ali tem, aqui não tem (I-UNICAMP)*

*Estamos implementando ... o ORCID é uma demanda da Reitoria, ele vai ser implementado em todos os nossos sistemas ... Na verdade ele já foi ... é ... a gente precisa só de uma ... de um acerto com a parte da informática que isso tá demorando um pouquinho mas ... já tá bem adiantado assim. (...) Mas precisa de uma parte burocrática, de iniciativa, na verdade mais legal que está demorando um pouco porque houve um questionamento por parte do pessoal do Direito ... que não queria ... deixar tão exposto algumas das suas produções, acharam que o ORCID poderia ... Mas eu diria que ... até o meio do ano que vem, mais tardar estará totalmente implementado e automatizado **(G-USP)***

7. Tratamento temático da informação no repositório institucional

A categoria “Tratamento temático da informação no repositório institucional” refere-se aos aspectos que envolvem a representação de assuntos no repositório: catalogação de assunto/indexação de teses e dissertações, catalogação de assunto/indexação de outros materiais, política de indexação, manual de indexação, linguagem documental utilizada, atualização da linguagem e número de termos atribuídos.

Referente à catalogação de assunto/indexação de teses e dissertações, considerando-se a inserção diretamente por autoarquivamento na *Biblioteca Digital de Teses e Dissertações* (BDTD), no tocante aos assuntos verifica-se que são utilizadas as palavras-chave ou termos livres elencados pelos próprios alunos de pós-graduação. Em seguida, na Biblioteca ao qual estão vinculados, há a verificação dos registros e dos assuntos do registro no catálogo online DEDALUS, ou seja, a compatibilização/tradução destes termos livres de acordo com o VOCAUSP.

A questão dos assuntos é assim, a gente tem um vocabulário controlado, né, institucional. (...) ... Então nessa validação feita pelos bibliotecários eles já verificam os assuntos que são colocados e já fazem se for necessário a ... tradução dos assuntos do vocabulário controlado (G-USP)

Na USP, a indexação de teses e dissertações é realizada no catálogo online DEDALUS utilizando-se o VOCAUSP, após importação dos metadados inseridos pelos alunos de pós-graduação na BDTD. A ficha catalográfica é realizada em um sistema de solicitação online, onde é disponibilizada a consulta ao VOCAUSP. Entretanto, pela percepção da bibliotecária catalogadora-indexadora da USP, de forma geral os termos não são consultados pelos alunos, preferindo-se a linguagem natural. No catálogo online DEDALUS, os termos livres permanecem no campo 952 do MARC, enquanto que os termos já verificados e padronizados pelos bibliotecários catalogadores-indexadores de acordo com o VOCAUSP permanecem no campo 650. Tanto a linguagem natural quanto a linguagem controlada são recuperáveis.

[...] a indexação na verdade é feita no DEDALUS mas na BDTD, como tem as palavras-chave, essas palavras-chave elas vem para o DEDALUS e ficam num campo de palavras que a gente depois no vocabulário a gente faz uma análise até pra ... incluir essas palavras-chave no vocabulário. (...) Ou seja, palavras-chave que a gente não conseguiu ... ou que não estão no vocabulário USP ou que não tem o sinônimo lá no vocabulário a gente no campo, num outro campo lá que é pra ... se alguém pesquisar conseguir ... recuperar. Mas na parte de assunto do DEDALUS são palavras que a gente tira do vocabulário (I-USP)

(E no caso em que momento é elaborada a ficha catalográfica?) Olha, lá na minha Biblioteca a ficha catalográfica ... os alunos vêm no último instante pedir, a hora que a tese tá toda pronta eles vem. ... Agora a gente tem um recurso que eles mesmos podem fazer, então já perdeu um pouco o controle de que em que momento que eles fazem ((RI)), mas pela experiência é quando eles vão entregar o volume final eles resolvem fazer a ficha catalográfica. ... E a nossa ficha ela tá pra eles fazerem online e nesse formulário que tem online tem um link do lado pra eles consultarem o vocabulário USP pros termos. (...) Mas o que eu percebo é que não é consultado. (...) Geralmente não é consultado, geralmente eles põem os termos das palavras-chave deles mesmos. (E acaba entrando esses termos das palavras-chave na biblioteca digital de teses). ... Sim. (E vocês fazem essa verificação). ... Isso. (Junto ao vocabulário da USP, é isso)? Isso, é isso aí. (Então na Biblioteca Digital de Teses e Dissertações ficam os termos livres, vamos dizer assim, que eles colocaram, mas no DEDALUS tem essa validação do bibliotecário, são os termos). ... Fica tanto os termos que eles colocaram como os termos do vocabulário, que a gente faz é ... não deixa de ter o campo das palavras-chave, tira esses termos desse campo de palavras-chave, deixa no campo de assunto nosso, né, e ... como a gente usa o MARC, nosso campo de assuntos é o 650 e o da palavra-chave é o 952, que é palavras-chave sugeridas por eles. Então as palavras que constam no vocabulário ficam no campo 650 e as palavras que não constam ficam no ... 952 (I-USP)

Por serem bancos de dados diferentes, pode ocorrer diferenças entre os assuntos atribuídos na BDTD e a BDTI. Em alguns testes realizados pela pesquisadora, verificou-se estas inconsistências, principalmente em relação aos subcabçalhos/qualificadores dos assuntos. Em buscas realizadas na BDTI e no DEDALUS para a mesma tese, verifica-se que no DEDALUS os assuntos estão um pouco mais completos do que na BDTI.

[...] porque são bases diferentes e esse metabuscador que a gente colocou ele ... a parte do vocabulário ... que tá na BDPI respeita os critérios do vocabulário controlado e o da BDPD não necessariamente. O repositório utiliza. (...) A BDPD que não. (...) Isso eu diria que ... 20% dos registros pode acontecer. (...) Não, ele comporta, é que ele é mais ... digamos assim, é mais controlado ((RI)) (G-USP)

Olha, no caso, o que eu faço lá, né. Eles colocam essas palavras-chave tudo separadas. Por exemplo, lá na minha Biblioteca tem muito local, né? Eles estudam muito ... a geologia de um determinado lugar, então eu faço a junção do local com o assunto que eles estão estudando. Então o meu termo ele fica ... fica um termo mais ... composto, vamos dizer assim, né. (...) E como também às vezes eles juntam, jogam uma palavra que a gente coloca ... agora me fugiu, que a gente coloca como um complemento lá e eles usam como termo separado. O que acontece é que a gente junta tudo isso. É, o qualificador. (...) Tem o qualificador e o local, aí fica um ... um assunto ... que engloba mais, uma coisa só quando às vezes ele tá separado nas palavras-chave. Mas quando isso acontece ... eu tiro lá debaixo das palavras-chave dele, porque de uma certa forma eu usei isso no campo 650. (Então é por isso que está aparecendo diferente aqui no repositório, porque esse ... porque esses metadados que estão aparecendo aqui no repositório ele puxou a Biblioteca Digital de Teses e Dissertações, que foram os termos que os alunos colocaram, é isso?) ... Não, se ele não for do DEDALUS. Está tanto os termos que eu coloquei quanto os termos que o aluno colocou. ... Vai os dois (I-USP)

Na UNESP, tendo em vista que os registros informacionais das teses e dissertações são inseridos no repositório por autoarquivamento, os assuntos são elencados pelos próprios alunos de pós-graduação e inseridos diretamente no repositório institucional. Entretanto, existe validação/verificação por um bibliotecário catalogador-indexador de cada uma das Bibliotecas da UNESP quando os registros são importados para o catálogo online Athena utilizando-se o *Vocabulário Controlado da UNESP*. No entanto, os assuntos contidos nos registros das teses e dissertações e sugeridos pelos alunos de pós-graduação permanecem os mesmos no repositório, isto é, termos livres/linguagem natural, enquanto que no catálogo online Athena existe a verificação pelo bibliotecário catalogador-indexador e padronização dos assuntos de acordo com uma linguagem controlada. Na visão da gestora da UNESP, tal verificação no vocabulário controlado e ampliação dos termos sugeridos pelos alunos de pós-graduação são importantes.

Olha, eu acho que elas devem pegar mais ou menos o que vem, né? Os alunos devem inserir lá como ... recomendação ... agora ... eu não sei te dizer como é que as Bibliotecas, se elas adotam, se elas incluem mais, se elas olham o vocabulário, isso aí acho que você vai descobrir. Eu não vou saber te dizer. (...) É, aí eu já ... então eu não sei te dizer, eu acho que elas deveriam, usar né, ou incluir mais, né, não sei. Eu incluiria mais, eu pegaria e aproveitaria o deles ... e iria pro vocabulário também (G-UNESP)

Referente ao processo de catalogação de teses e dissertações na UNESP, verifica-se que as solicitações de fichas catalográficas são realizadas em um sistema online. O bibliotecário catalogador-indexador da UNESP acredita que em alguns casos a definição das palavras-chave é um desafio aos alunos de pós-graduação, pela ausência de formação específica, elencando palavras-chave mais livres. Além disso, especificamente na Biblioteca da UNESP de Marília, as

teses e dissertações são classificadas, por vezes tomando-se como base aquelas já classificadas anteriormente.

Um ponto a destacar é que todos os assuntos sugeridos pelos alunos no momento do autoarquivamento são carregados no catálogo online Athena no campo 650, independentemente do tipo de autoridade: termo tópico, pessoa, local geográfico, dentre outros, havendo a necessidade de correção pelo bibliotecário catalogado-indexador e compatibilização dos termos de acordo com a linguagem do *Vocabulário Controlado da UNESP*. Logo, são duas verificações: primeiro, se o termo existe/aprovado na linguagem controlada. Se pertinente, é mantido. Caso não, verifica-se a possibilidade de termos correspondentes ou mesmo sugeridos para abertura ao *Grupo de Trabalho em Política de Indexação da Rede de Bibliotecas UNESP*. Pela percepção do profissional, é necessária a alteração na grande maioria dos assuntos sugeridos no momento do autoarquivamento.

Bom, antes do repositório a gente recebia ... a tese e a dissertação em papel né, muitas vezes já havia feito a ... a ficha catalográfica, então o processo de indexação já começava ali, né, com a escolha das palavras-chave ali, pelo ... pelo vocabulário UNESP, né, fazendo a ... tendo o cuidado de tratar o assunto ali fechado, usar essas palavras pra fechar o assunto de forma que a recuperação fosse, ... tivesse a melhor precisão possível. (...) Com o repositório, o aluno faz o autodepósito, né, ... é ... recentemente é ... a UNESP disponibilizou na Rede, CGB disponibilizou na Rede um aplicativo que permite que a pessoa ... é ... elabore a sua ficha. Então muitas vezes o ... o candidato a mestre ou doutor ele não tem esse cuidado, na verdade não ele tem a formação, nem sempre é fácil você definir as palavras por conta dos assuntos, né, nosso conhecimento prévio. (...) Então pra ele é muito mais simples simplesmente colocar aquelas palavras que ele ... entende que são as melhores pra fechar o assunto dele e ... preenche, elabora a ficha ... copia e cola no verso da página de rosto e ela chega aqui pra gente, então ... ~ é ... tem que ter esse trabalho a mais, não a mais, né, porque a gente já fazia, mas um trabalho em uma outra etapa, às vezes a gente fazia antes, (...) que era a ficha catalográfica destacada da ... do processo de catalogação ali no ... no final, vamos dizer assim, né. ... E agora ela é feita mais ou menos em ... junto. (...) E entra aquela coisa um pouco mais intelectual que é a questão da definição do campo 082, que é o assunto (...) e os campos ... 600 né, que são ... entidade, o nome pessoal ou o próprio assunto tópico né. (...) Então eu trabalho desde o campo 001, preenchendo o campo de UN³⁸, enfim, todas essas coisas que são ... entre aspas braços né (...) um copia e cola às vezes ou uma designação do código que vai pra um campo ou outro e vou seguindo o bibliográfico até chegar nessas ... nessas questões aí do campo 082 e dos campos 600. O 082 a gente tem, trabalha com a CDD ... é ... tem aqui, né, já uma planilha, assim, do que é mais usado aqui no ... no campus, né, daqueles assuntos relacionados aos ... às pesquisas elaboradas aqui, que são retirados obviamente da ... da CDD. (...) Então a gente para, pensa um ali um pouco, né ... às vezes o assunto é muito direto, muito fácil definir por conta do conteúdo, do conhecimento prévio igual eu mencionei ... e de como a pessoa descreveu isso daí, tanto o título ... o resumo ... e as palavra-chave juntos né. (Então vocês usam classificação pras teses?) ... Sim, aqui nós usamos, sim, aqui em Marília sim elas são classificadas. (É uma coisa que é da UNESP de Marília ou é padronizada pra todas as UNESP?) Não, não. Até noto que às vezes a gente tem uma dificuldade, então a gente busca alguma tese com um assunto, com alguma palavra similar pra ver se a gente acha um ... um ... um fio solto pra gente aproveitar naquele trabalho de catalogação de definição do assunto (...). Por exemplo, Araraquara ... não trabalha, Letras eu lembro da época que eu trabalhava, às vezes a gente procura, acha, de lá, olha já sei que eles ... tem uma numeração serial, um, dois, três e outras também, não posso afirmar que todas sejam assim, (...) mas ... Marília é um critério de Marília, já, antes de mim aqui pra ... da gente trabalhar pra ... determinar o assunto da tese ou

³⁸ Número de controle de dissertações e teses no Catálogo Athena da UNESP (campo 001 do MARC).

dissertação. (E o 090 vocês não usam?) O 090 ... sim, sim. O 090 tem que é o número de chamada local mas eu to vendo aqui a ... tomei o cuidado de abrir, ele ... é o código padrão. (...) Né, tem ... não é colocado nada ali ... é ... isso é padronizado pra UNESP (I-UNESP)

(Você chega no campo 600 e começa a fazer a verificação das palavras-chave?) Sim, sim, é, exatamente, ... vou preenchendo, chega no 600, 610, 650, é ... normalmente quando chega pra gente tudo é 650, né, o ... o registro bruto que o ... que candidato preencheu, ele vem tudo 650, então às vezes tem lá ... é ... Mercosul né... é um organismo, então ele deveria ser 610 mas tudo aparece como 650 então eu eu verifico se aquele termo ... ele existe, né, nos ... nas bases de ... nas autoridades de termos lá na UNESP e se ele é ... conveniente né à ... descrição do material. Às vezes ele pode até ser da Área mas de repente o autor ... julga erroneamente, então a gente faz essa verificação, então são duas verificações, a verificação da existência deles, se ele existe, ... e é pertinente ele é mantido. (...) Se ele ... não existe eu preciso encontrar ... algum sinônimo ou alguma outra, ou outras palavras, às vezes mais de uma que ajuda a ... a compor o assunto né, que vão fechando o ... o termo ... vamos dizer assim (...) e ... vou preenchendo dessa forma. Quer dizer, a verificação da consistência do ... do termo e se ele é pertinente ao ... ao trabalho. (Você nota assim na sua experiência, por exemplo, quando você vai fazer, você tem que alterar, tem alterado bastante essas palavras-chave ou geralmente as palavras-chave que os alunos colocam atendem mais ou menos ali). ... É difícil dizer um número mas, assim, a maioria, a grande maioria, a gente tem que ... vamos dizer, refazer, porque é como você disse, é a linguagem natural, não só a linguagem natural mas ... é... dependendo da área, assim, ou ... e da especificidade do trabalho tem umas palavras que são muito ... particulares e aquelas não são atendidas, né, às vezes quase um neologismo. (...) Então a gente acaba tendo que trabalhar sempre, é raríssimo ter um ... um trabalho que a gente ... não mexa (I-UNESP)

Na UNICAMP, o gestor observa que para os demais recursos informacionais inseridos no repositório institucional, será adotado o mesmo procedimento atual para as teses e dissertações, isto é, a indexação dos assuntos realizada pelos bibliotecários catalogadores-indexadores, pautando-se nas linguagens controladas adotadas de acordo com a respectiva área do conhecimento.

[...] a gente faria nos moldes das teses e dissertações, ou seja, todo assunto, cada assunto vinculado a ... a um documento ele seria feita a padronização, todo o tratamento antes de entrar no repositório, que é aquela validação que eu comentei agora há pouco (G-UNICAMP)

Em relação à linguagem natural, a proposta é de que a representação temática dos demais recursos informacionais inseridos no repositório contemple não apenas a linguagem controlada mas também a linguagem natural, ou seja, os termos livres ou palavras-chave elencadas pelos autores dos recursos informacionais, embora sem a utilização de remissivas do tipo “ver” e “ver também”, utilizadas na linguagem controlada. Somente após a validação/verificação é que os termos livres passam a ser descritores, como os demais da linguagem controlada. Para o gestor, tal medida vem para ampliar as possibilidades de recuperação dos recursos informacionais no repositório.

Existe um metadado que ... que permite a gente inserir a linguagem natural, seria os termos livres né, linguagem natural mesmo, mesmo aquelas palavras-chave soltas que são vinculadas aos documentos, ficaria só eles mesmo. Eles ficariam ali apenas pra ... é ... um plus de recuperação, mas o que iria validar mesmo com remissiva e tudo o mais seriam os descritores formados ali no assunto, agora o ... os termos

livres que a gente vai adotar é aquilo que vem no documento muitas vezes. ... (As palavras-chaves do próprio autor?) Do próprio autor, seja ... estando ou não em uma linguagem, é ... natural ou não. Pode ser que aquilo que foi inserido ali não esteja necessariamente em uma linguagem natural, já numa linguagem documentária, porém ... a gente chamou aquilo de termo livre porque é aquilo que já estava no documento, enfim. (Mas ele é recuperável?) É recuperável. Mas aquilo que está no ... é ... aquilo que está na base também, então não necessariamente é só o que tá no documento. Então digamos que eu peguei um registro lá de uma base 'x'. (...) Aqueles assuntos que a base trouxe eu estou chamando aqueles assuntos não tratados de termos livres. (...) E aí a partir do momento que ele é tratado e tudo o mais ele passa a ser um descritor, né, com os assuntos controlados mesmo (G-UNICAMP)

A partir dos dois treinamentos e do trabalho iniciado pelos bibliotecários catalogadores-indexadores das Bibliotecas da UNICAMP foi que se iniciou a padronização dos assuntos contidos nos demais recursos informacionais, por meio da compatibilização da linguagem natural (termos livres) com a linguagem controlada, permanecendo dois campos de assuntos nos registros. Entretanto, o número ainda é muito incipiente pelo volume de recursos informacionais a serem verificados e pelo número de profissionais que realizam tal atividade, paralelamente a outras atividades profissionais.

Inclusive, é ... é ... pouco nós temos trabalhado com assunto no repositório, assim. A gente ... a gente é ... costuma ... trabalhar ... é ... costuma manter os assuntos que já vem com nos próprios documentos. (...) Não é feito um trabalho mesmo com, com tradução na linguagem documentária, não, essa é uma etapa posterior ainda, a gente não faz esse tratamento ... fino dos assuntos. O ... os assuntos por exemplo, a partir de fevereiro do ano passado, aí sim com a ajuda de ... das Bibliotecas, dos demais profissionais das outras Bibliotecas nós temos feito esse ... essa padronização de assuntos, essa entrada mesmo, essa busca e ... e ... e tratamento dos assuntos no repositório, mas hoje não, ele não reflete mesmo, tá ... o tratamento afinado dos assuntos (G-UNICAMP)

Referente à questão da validação/verificação dos assuntos na proposta de autoarquivamento pelos autores, no âmbito da UNICAMP seria realizada pelos bibliotecários catalogadores-indexadores em procedimentos semelhantes ao realizado atualmente para teses e dissertações quanto à representação descritiva e temática.

[...] vamos fazer uma validação como é feita hoje pra teses e dissertações, que diga-se de passagem também é importante voltar, na parte quando a gente alimentou, depois de um tempo também nós, nós inserimos as teses e dissertações, cerca de 45.000 teses e dissertações na ocasião, teses e dissertações foram inseridas junto com os artigos que foram coletados das bases e aí, é ... então nos moldes como é hoje na tese a gente vai, é ... vai fazer pros ... pros documentos do repositório. Então vai haver uma validação, como é feito hoje com as teses e dissertações. (E aí fariam analíticas, pros capítulos de livros). ... Analíticas. Isso, capítulo de livro, artigos de evento, faria tudo no ... no nível de analíticas. (E o vocabulário controlado seria do mesmo modo como está ...). Que está no SophiA (G-UNICAMP)

No tocante à política de indexação já estabelecida para o catálogo online Athena, a gestora do repositório da UNESP acredita que pode ser utilizada futuramente também para o repositório institucional.

Então a política até onde eu sei é pra ... é pro catálogo mas eu acredito que ... ela ... eu acho que isso pode ajudar no ... no repositório. É porque ... é produção da UNESP né (G-UNESP)

Na USP, verifica-se a existência de uma política de indexação formalizada em um manual de indexação, disponível em uma área técnica de acesso pelos profissionais. Embora o manual esteja sendo reformulado, verifica-se que é voltado para o catálogo online DEDALUS e que não existe um manual específico para o repositório institucional, já que toda a descrição dos metadados pelos bibliotecários catalogadores-indexadores é realizada no DEDALUS e pelo fato do repositório ser um sistema de recuperação recente na universidade.

Então, a gente tem isso na área mesmo do vocabulário (...) ... ele tá numa área técnica aqui da USP, né, onde tem o manual de indexação, ... a gente tem acho que dois manuais, que é o ... como usar nossa base de sugestão, como fazer a solicitação do termo e tem a ... o de indexação que inclusive a gente estava fazendo uma revisão ... e acho que ainda não ... não colocamos. A gente tem um manual de indexação sim. Como solicitar o termo, né, e tal. ... (E no caso essa política de indexação é geral, não tem uma específica pro repositório, né?) Não, a gente não pensou isso como repositório, o repositório aqui é agora que ele tá começando ... a criar corpo mesmo né ((RI)). A gente não tinha ... o repositório, agora que ... que tá tendo, agora que a gente ... tá começando a colocar, a fazer os uploads, tudo e ... é uma coisa mais recente. (...) Então ... nunca foi pensado como repositório, sempre foi pensado para o banco DEDALUS (I-USP)

Para a bibliotecária catalogadora-indexadora da USP, a política de indexação e o vocabulário controlado vigentes atendem às necessidades da instituição. Entretanto, reforça a necessidade de maior participação das áreas na solicitação de novos termos.

Olha, eu acho que atende, mas ... tudo depende muito de cada ... cada Unidade fazer a solicitação dos termos que não tem, né. O que a gente percebe é que tem determinadas áreas que as pessoas se contentam com aquilo que tá lá, ou pessoas que estão acostumadas a fazer no geral e aí tá bom e se contentam. Mas tem outras ... que não e que pedem bastante porque ... não, isso não tem, a gente precisa e a gente vai colocando, então eu acho que ... atende. Se a pessoa diz que não atende é porque ela não tá pedindo. (...) Depende da solicitação das áreas, porque a gente não conhece, né, a necessidade de todas as áreas, então eles pedem ... pra gente colocar isso no vocabulário, né (I-USP)

Na visão da bibliotecária catalogadora-indexadora da UNICAMP, o vocabulário controlado utilizado para a padronização dos assuntos atende satisfatoriamente às necessidades do repositório.

(E ... você acha que esse vocabulário controlado, essa linguagem utilizada nos assuntos atende às necessidades do repositório?) Acredito que sim, porque a gente trabalha assim pra livros, pra teses e dissertações e (...) temos sucesso, até onde eu sei ((RI)) (I-UNICAMP)

Na UNESP, verifica-se a existência de uma política de indexação direcionada para utilização no catálogo online Athena e é formalizada em um manual de indexação, que apresenta diversos exemplos para nortear a prática da indexação pelos profissionais. O

bibliotecário catalogador-indexador considera que a política atende satisfatoriamente no que se refere à padronização dos assuntos dos recursos informacionais.

(Vocês têm a política de indexação, né?). *Sim.* (Que vocês utilizam no Aleph). *Sim.* (É ... mas assim, essa política por enquanto ela só tá ... formalizada pra utilização no Aleph, né, no repositório não tem nada definido ainda). *Não, eu desconheço, eu acredito que não.* (E essa política de indexação que vocês têm está regulamentada em um documento? Tem um manual?) ... *Tem. Nós temos assim uma ... uma apostila, que ela é baseada nos critérios definidos por uma comissão já há algum tempo né, acho que a última versão é de 2016 ou 14 se eu não me engano. Então essa apostila ela vem com, com muitos exemplos né, pra fechar, tentar fechar ao máximo as possibilidades ali ... e ... mais ou menos isso.* (Essa política de indexação no Aleph atende às suas necessidades profissionais? Você sente que ela é ... é adequada, precisaria de alguma ... é ... reformulação). ... *Não, eu não tenho nenhuma crítica maior, assim eu ... as dificuldades que às vezes surgem são muito mais, assim de ... de pessoal, minha, no caso, assim, às vezes de entender melhor o assunto, porque às vezes um trabalho de Filosofia, que é um pouco distante da realidade da gente, por exemplo, então a gente tem dificuldade às vezes em ... em entender o que a pessoa está querendo ... o que ela trabalhou na dissertação, na tese dela e pra gente identificar o assunto, né, o 082 e o 600. Mas com relação à ... à documentação que existe não, a documentação é muito boa ... nada é perfeito né, mas ela é muito boa e atende bastante sim às ... às necessidades (I-UNESP)*

No caso da UNICAMP, verifica-se que a inexistência de uma política de indexação formalizada em um manual de indexação, entretanto, há diretrizes internas que são seguidas pelos profissionais apreendidas em treinamentos ou na prática profissional. Ainda não há uma parceria formalizada com docentes especialistas, nem a criação formalizada de um grupo de trabalho que possa compor e formalizar a política de indexação no contexto da universidade.

(A instituição dispõe de uma política de indexação regulamentada?) ... *Não ((RI)).* (Tem uma política mas ela é ... do dia a dia né?) *É interna ... é o que a gente faz aqui ... é ... vamos supor, tem a área igual, Medicina e Química. (...) Batem. ... Algumas palavras, é ... assuntos batem. Então uma conversa com a outra, entra num acordo, vamos usar tal. (...) Mas ... não é uma política de indexação. ... (Formalizada, num manual né?) Não, isso ... infelizmente ainda não ((RI)). ...~ (Então o nível de especificidade vai ser norteado de acordo com cada área, né?) Com cada área. Por isso que todos, quem tá fazendo são os bibliotecários das áreas, (...) não somos nós aqui. (Os especialistas) ... *Exato. Eles que estão decidindo o que eles vão colocar. A DTRI não está trabalhando com isso. ... (E ... ainda não tem ... nenhuma parceria com os docentes especialistas, pra trabalhar os assuntos?) Então, tem algumas Unidades que ... que procuram os professores ... quando são assuntos muito diferentes eles procuram por conta de definição, ... aí ver com o professor se é viável ou não. (...) Acontece muito, muito na Engenharia e ... na Medicina isso. Bastante. Eles procuram bastante os professores. Mas é uma política interna deles, como eles fazem pra teses e dissertações também (I-UNICAMP)**

Quanto ao manual de indexação, verifica-se a existência de diretrizes para padronização dos demais recursos informacionais inseridos no repositório institucional, utilizados pelos bibliotecários catalogadores-indexadores das Unidades. Entretanto, em *Excel*, juntamente com as planilhas com os registros a serem padronizados, sem uma formalização oficial. Com a possível integração entre o catálogo online SophiA e o repositório, será necessária a

alteração dos procedimentos, tendo em vista que serão utilizadas as Tabelas de Autoridade do próprio SophiA.

(Tem alguma documentação, tem algum manual ... interno de trabalhar mesmo no repositório?) *Pra trabalhar no repositório? Nós temos esse depois do ... treinamento, que é do treinamento. Porque agora todo mundo que começar a trabalhar no repositório vai ter que seguir. Só que ... é assim. Esse é ... manual é dentro do Excel, (...) se ele realmente passar pro SophiA ele vai ... mudar porque aí vai ser em MARC (...), que vai nos facilitar muito a vida. (...) ... Então aí ele vai ... não é que ele morre, vai ser, são as mesmas informações mas vai facilitar porque a gente tem as tabelas, (...) então seria só puxar, vou pôr cidade, eu não preciso pensar como eu vou colocar, é só colocar, vou colocar o editor, que seria nossa editora como livro ... é só puxar. ... (Seria até bem mais rápido né, padronizado e tudo). Muito mais. O assunto ... eu já vou olhar lá se ele tá revisado, bonitinho. Tá? Então é ele mesmo. É o que hoje a gente faz ... quando elas vão fazer né, vão arrumar os artigos, tem que abrir o SophiA e tem que ficar no Excel, então vai acabar isso. ... Pelo menos é o que nós esperamos, que não tenha mais isso, esteja tudo dentro do SophiA (I-UNICAMP)*

Referente à linguagem documental utilizada, o VOCAUSP é a linguagem padronizada construída pela USP para padronização dos termos. Na UNESP, utiliza-se o *Vocabulário Controlado da UNESP* ou, na área do bibliotecário catalogador-indexador entrevistado, recorre-se à BN ou LC quando trata-se de um termo novo ou é solicitado ao *Grupo de Trabalho em Política de Indexação da Rede de Bibliotecas UNESP*. Já na UNICAMP inexistente um vocabulário próprio e único da instituição, recorrendo-se a diversas fontes, como BN, LC ou tesouros mais específicos como DeCS, *Ei Compendex* e *FSTA Thesaurus*.

A gente tem o nosso vocabulário USP, né. (...) É ele que a gente usa, o vocabulário USP ele desde 2000 que ele tá ... que a gente usa e ... e ele é atualizado, é dinâmico, né, a gente tem um grupo de gestão do qual inclusive eu faço parte também ... e ele é atualizado à medida que as pessoas vão percebendo assuntos novos ou assuntos que não tem no vocabulário eles são inseridos ... dentro de uma hierarquia e é encima desse vocabulário que a gente faz a indexação, é com ele, a gente trabalha com ele. (E aí no caso dos artigos, por exemplo, quando vem os assuntos em inglês vocês fazem a tradução pro VOCAUSP?). ... *Não, então, ... os assuntos quando vem em inglês a gente faz, procura os semelhantes ... Lá na GEO a gente chegou a fazer uma época a gente fez essa tradução dos nossos assuntos pro inglês, inclusive a gente ia mudar a nossa plataforma do vocabulário pra outra que iria contemplar isso, só que não foi levado a diante ainda. Então a gente usa dentro da área da gente o conhecimento da tradução ... na verdade nem é a tradução, é a compatibilização desses assuntos em inglês com os assuntos em português. (...) Quando a gente percebe que o assunto está sendo bastante estudado, bastante citado, assim, nos arquivos a gente já pega e coloca no vocabulário (I-USP)*

(E vocês têm uma tabela de assuntos dentro do Aleph né?) *Sim. (Que você faz essa verificação). Exato. (E ... quando você não encontra esse assunto ... novo, às vezes é muito específico né). Sim (Ali na base e vocês recorrem a, em qual fonte pra estar ...). É, é possível ... é possível o seguinte, eu encontro, se eu não, existe um termo, aquele termo não é válido, ou eu procuro um outro similar, ou outros né, que fazem isso. (...) Ou então é possível ... é ... a gente procurar às vezes na ... na Biblioteca Nacional, se esse termo existe, né (...) esse que na nossa base ainda não ... não existe na nossa base, pode ser que ele tenha na Biblioteca Nacional, na LC, (...), né, na Library of Congress ou então a gente solicitar ao Comitê ... não me lembro exatamente o nome mas é um Comitê que acho que a CGB gerencia que ele ... ele ... vamos dizer, mediante critérios ele ... ele aceita né, ele determina que aquele termo passe a ser válido pro ... pra nossa base de dados. (Além da BN e da LC quais as fontes que ... você lembra de alguma outra fonte que você teve que recorrer?) Pra tese e dissertação (...) eu não me recordo agora, só essas duas que a gente ... pra outras atividades, a gente ... não pra teses e dissertações a gente usa o BIBLIODATA por exemplo pra verificar se*

aquele registro já existe, não o registro do termo, mas a descrição. (...) Mas pra tese e dissertação, os campos 600 assim ... que eu me lembro a BN e a LC. (E vocês colocam no Aleph o mesmo, o termo correspondente em inglês daqueles que vocês colocaram em português?) É, nós escolhemos um, às vezes não necessariamente esse ter, precisa ter ... ser, é ... descrito na LC também. (...) Não é um termo qualquer. Às vezes tem uma tradução pro inglês, e aquela tradução ... vamos dizer assim não é autorizada pela LC, (...) não foi importada de lá, então a gente ... precisa ser um que seja de acordo com esse critério também (I-UNESP)

Ah, as teses o que a gente faz é pegar do Pesq, da BN, OCLC e tem uma base da ... FEA ... de alimentos que eles pegam. E alguns pegam de tesouros ou a gente ... é... faz um assunto, colocando definições e ... referências positivas, né. (...) A mesma coisa está acontecendo pro artigo. É por isso que foi passado pras Unidades, porque o que eles estão fazendo é a mesma política (I-UNICAMP)

Quanto à atualização da linguagem, na USP assuntos não contemplados no vocabulário controlado são inseridos e ao mesmo tempo encaminhados ao *Grupo Gestor do VOCAUSP* para análise e aprovação. O grupo, composto por representantes das Bibliotecas, será reestruturado e embora não tenha realizado reuniões frequentes, continua desempenhando o papel de inserção e de análise de termos. Além disso, está sendo analisada a possibilidade de aumento da automatização do processo.

Nós temos esse vocabulário controlado que ... é ... ele já vem ... salvo engano pelo menos duas décadas ... ~ ou algo próximo disso, é ... ele já é consolidado. Tem um grupo de trabalho e o grupo gestor é bem atuante, inclusive é uma das minhas responsabilidades também como chefe da divisão ... A gente tá nesse momento de reestruturação, então provavelmente a gente ... vai mudar alguns integrantes, talvez mudamos algumas políticas ... e ... é ... estamos automatizando, a gente já automatizou muita coisa mas a gente quer automatizar ainda mais pra verificar a ... uma maior agilidade na inserção de termos novos ... e ... ~ eu diria que temos tido bons resultados, a gente tem problemas, né, isso é um trabalho contínuo de melhora, de aperfeiçoamento (...), mas a gente tem tido bons resultados nesse tempo aí e inclusive nosso o vocabulário ele é muito utilizado por outras instituições, tanto as parceiras, como a UNICAMP, a UNESP quanto outras que ... até pediram acesso a nossas inserções. (...) E a gente falou que num primeiro momento isso não seria possível, que isso seria possível ter acesso ao que já é disponibilizado, mas a gente tá revisando isso também, se não é melhor deixar uma coisa ... até pra fazer um trabalho mais colaborativo com outras instituições também (G-USP)

Segundo a bibliotecária catalogadora-indexadora da USP, que atuou na coordenação do *Grupo Gestor do VOCAUSP* por 6 anos, o grupo é composto por bibliotecários representantes das três áreas do conhecimento e docentes especialistas das áreas. Dentre as atividades realizadas, ocorre uma verificação de cada assunto solicitado para ser incorporado no vocabulário e uma divulgação frequente às Bibliotecas dos novos termos incorporados.

Então, a gente tem no grupo, nessa gestão do vocabulário a gente tem representantes de três áreas, que é Humanas, Exatas e Biológicas. (...) Os assuntos são pedidos através da ... né, de uma plataforma, e são solicitados pelas Bibliotecas. Cada área tem dois, de dois a três representantes que analisam esse pedido ... vão consultar as fontes, porque eles tem que vir com a solicitação e com uma nota de escopo a gente vai verificar se ele, de repente ele é um assunto pra ser inserido mesmo ou se ele não é uma remissiva de um outro já existente, alguma coisa assim, a gente faz essa verificação e depois tem o coordenador que ... que faz a validação, né, do assunto e aqui no SIBi tem a equipe de apoio que é que vai introduzir o assunto.

Aí a gente mês a mês divulga todos os assuntos que ... novos que entraram no vocabulário para as equipes das Bibliotecas ficarem cientes desses assuntos. (...) Eu fui coordenadora ... é ... até o fim do ano passado ... agora mudou a gestão e eu to como representante da área de Exatas, não mais como coordenadora, só que como ainda não foi validada essa nova gestão, eu ainda to atuando como coordenadora, na verdade não como coordenadora, só pra validar os registros, né, porque a pessoa ainda não ... não assumiu isso (I-USP)

A criação do VOCAUSP, em 2000, foi baseada em linguagens já utilizadas pelas Bibliotecas e com os docentes especialistas, sendo que o *Grupo Gestor do VOCAUSP* foi implementado em seguida, tendo em vista a necessidade de atualização e verificação constante dos termos, de acordo com a dinâmica da produção científica das áreas.

Olha, quando eu entrei na universidade, há 22 anos atrás isso já estava praticamente ... pronto. Eu na verdade eu entrei no lugar da ... bibliotecária anterior ... que foi a que montou nossa ... porque cada Biblioteca, cada área montou a sua área junto com os especialistas das áreas, né e ... junto com os seus vocabulários que eram usados antigamente ... foi a base daquilo e aí junto com os docentes especialistas foi ... foi validado, cada área foi validado assim. E como ... a ideia era ser um vocabulário que fosse válido pra USP inteira ... porque o outro usado era bem ... geral né, era bem geral, eram termos ... eu não sei direito da onde vinha, mas só sei que a indexação ficava a desejar porque eram assuntos muito gerais. (...) E com a ampliação do ... eram metas que tinha antigamente na USP, né. E essa criação do vocabulário era uma meta que foi desenvolvida e em 2000 foi quando ... acho que em 2000 ele foi ... que ele foi já implementado pra uso pra todo mundo. (...) Então desde então que a gente ... usa. ... O grupo de gestão mesmo acho que foi um pouquinho mais pra frente, porque aí a gente foi vendo a necessidade de ter as pessoas pra ... pra estar acompanhando e pra estar alimentando os novos termos que fossem ... solicitados, porque ... não poderia parar naquilo, que a gente começou acho que ... 20 e poucos mil termos, hoje a gente tem, sei lá, mais de 40.000, não sei direito. (...) Não vi isso aí, entendeu? (...). E aí foi criado e eu sempre participei como representante de área ... e aí há oito anos que eu fiquei coordenadora, 8 anos não, acho que uns 6 anos, fiquei 6 anos como coordenadora do ... do grupo e agora a gente trocou, passamos pra ... pra outra. (E na época você sabe se as Bibliotecas montaram seus próprios vocabulários, né, junto com os especialistas mas você sabe se foi com base em alguma outra linguagem, se foi a LC, BIBLIODATA ... ou não teve um ...). Olha, eu acho que ... ((RI)) aí cada área tem as suas, por exemplo, o que eu posso te dizer da área médica ... eles usam muito o DeCS. (...) Então acho que na área médica foi muito baseado ... hoje em dia eu tenho certeza que é né, porque a gente recebe os termos, é muito DeCS. ... E assim, a princípio eu não sei se era só o DeCS ou tinha outras coisas, mas sim, tesouros da área ... é ... bibliografia da área, não é só o especialista, né. E eu sei que a área de Engenharia, por exemplo, foi muito acima da LC. Então cada área tem, tinha os seus ... o seu embasamento, né (I-USP)

O *Grupo de Trabalho em Política de Indexação da Rede de Bibliotecas UNESP* é composto por integrantes da Coordenadoria Geral de Bibliotecas (CGB) e por bibliotecários representantes das Unidades. As solicitações de abertura de novos termos são avaliadas pelo grupo e, caso aprovadas, passam a incorporar o *Vocabulário Controlado da UNESP*.

(Então se esse termo não nem na LC nem na BN ... vocês têm que abrir uma solicitação ... de aprovação de termo novo?). *Sim, é uma possibilidade, sim, sim. (E ... esse Comitê que você fala é só a CGB que gerencia ou tem participação de bibliotecários de área?) Não, não, é ... eu não tenho certeza como está atualmente. (...) Mas são pessoas da Rede também, não é só pessoas da CGB, são bibliotecários da Rede. (E ... demora? Como que é na prática?) Ah, eu não me lembro, acho que até participava, mas eu não ... a professora Mariângela com certeza vai ter esse nome, nem sei se é um Comitê. (Será que não é o grupo de política?) É ... exatamente, exatamente. (Então é o grupo que faz a aprovação ou não dos assuntos). Sim, exato. (E esse assunto fica assim meio que em off-line ali, à espera dessa aprovação). É, se eu fizer*

essa solicitação eu acredito que ... seja avaliado e aceito ou não em dois ou três dias, vamos dizer assim. (E pela sua experiência geralmente é aceito ou não?) Normalmente sim. A gente quando faz o pedido a gente já indica uma fonte (...) pra embasamento, vamos dizer, da qualidade do ... do termo, né. (...) Então eles costumam aceitar sim. (E vocês têm que fazer a ... tem que descrever certinho o termo no Aleph, fazer as remissivas, indicar a fonte). Sim. (A árvore de autoridade do assunto?) Sim. Seria nesses casos, seria isso daí (I-UNESP)

Na UNICAMP, verifica-se a existência de abertura para novos termos, desde que sejam corretamente referenciados na Tabela de Autoridades de Assuntos, com as respectivas fontes e remissivas.

O que já tem no SophiA eles estão usando o que tem no SophiA que já tá padronizado, o que não tem no SophiA se preciso eles estão criando ou indo atrás de referências pra colocar, com as definições. ... Então nós estamos usando a mesma política (I-UNICAMP)

Referente ao número de termos atribuídos, na USP a orientação é de que não se ultrapasse 6 termos, ou seja, possui uma variação de 1 a 6 assuntos atribuídos, enquanto que na UNESP a política estabelecida é de no mínimo 3, sem limites de termos atribuídos, por considerarem que isto depende do conteúdo do trabalho. O bibliotecário catalogador-indexador da UNESP afirma que, obrigatoriamente, são utilizados tanto os termos em português quanto a versão dos termos em inglês, no registro MARC do catálogo online Athena. Na UNICAMP a mesma política aplicada para teses e dissertações está sendo aplicada na padronização dos demais recursos informacionais no repositório institucional, realizada pelos bibliotecários catalogadores-indexadores das Unidades, isto é, mínimo de 1 e máximo de 6 assuntos atribuídos aos recursos informacionais, mesmo número da USP.

Olha, na verdade a gente orienta que não ... que não ultrapasse 6 ... 5, 6 termos ... mas ... não tem limite, a pessoa pode pôr quantos assuntos ela achar necessário (I-USP)

(E no autodepósito eles têm uma quantidade limite de termos pra estar colocando?) Não, não, isso ... eu desconheço esse critério, mas eu acredito que não, às vezes tem ... às vezes as pessoas colocam em ... em português, em espanhol, em inglês, por exemplo, então isso não, a gente trabalha com esses três termos em língua portuguesa e ... pelo menos, pelo menos um não, um em língua inglesa, que é o 650 zero, né, então a gente ... acaba tendo que colocar esse também. (Então o limite de termos no Aleph são três, é isso?) É, é o mínimo. (...) (E o máximo, vocês têm um máximo?) Não ... vai depender muito do ... do trabalho. Às vezes tem mais, um trabalho com mais termos, às vezes 4, 5 ou 6, né. (...) Mas não há um máximo (I-UNESP)

Aí a gente tem ... é ... português ... de 1 a 6. Mínimo 1, máximo 6. (...) As mesmas, os mesmos descritores pra inglês, os mesmos (...) ... não importa se é uma publicação em inglês, se é português, se é espanhol, se é francês. O primeiro é sempre português, o segundo é ... inglês, tá? E nós temos os termos livres, nós deixamos esses três campos. Os termos livres ... é ... os bibliotecários eles estão vendo o que tá no artigo, o que ... os ... os docentes, discentes colocaram no artigo, ... ver se dá pra ser colocado, se cabe e aí eles colocam os termos livres, pra não perder essa informação. (...) Então nós vamos ter as três. (Então os termos, as palavras-chave dos artigos que os autores indicaram ... entram nesses termos livres ou o bibliotecário pode excluir alguma delas?) ... Pode. É ... filtro do bibliotecário. (...) Se ele achar não, essa

palavra aqui é bem usada, eu vejo bastante, é ... acho relevante pra busca dele ele deixa. ... Senão ele tira. E ele não vai estar em tabela. Termos livres que vai estar ... no SophiA só que eles vão estar soltos. Campo livre, como se fosse nota. Busca mas ele ... não tá em tabela, não tá tabelado (I-UNICAMP)

No tocante à avaliação da indexação de assuntos, embora não seja realizada especificamente com usuários na Biblioteca em que a bibliotecária catalogadora-indexadora entrevistada atua, por sua percepção como membro do *Grupo Gestor do VOCAUSP* verifica-se que determinadas áreas não costumam solicitar abertura de novos termos, mantendo uma indexação de forma mais geral e abrangente, sem muita especificidade nos termos. Nesta situação, são realizados treinamentos e revisões de área buscando a melhoria contínua da indexação. Na UNESP, o bibliotecário catalogador-indexador afirma não ter conhecimento de testes ou ensaios de avaliação da indexação de assuntos com usuários no catálogo online Athena ou no repositório institucional. Na UNICAMP inexistem tais estudos, entretanto, a bibliotecária catalogadora-indexadora observa que todas as dúvidas ou sugestões tanto dos bibliotecários quanto dos usuários são respondidas prontamente.

(Vocês fazem algum tipo de avaliação dessa indexação ... tem algum estudo com usuário ou ... alguma coisa nesse sentido?) Olha ... a gente não faz, o que a gente faz, a gente como grupo percebe muita coisa. Então o que a gente já fez aqui por perceber que a indexação, como eu te falei, a gente percebe que tem áreas que não solicitam muito e isso reflete no registro que a gente vê que tá muito geral ... e a gente percebe alguns erros ... Toda vez que a gente nota esse tipo de coisa a gente procura o que fazer, primeiro fazer treinamentos, porque no grupo do vocabulário a gente tem uma professora que ... nos orienta também, né, que é da ECA que faz orientação da gente. E a gente já fez vários treinamentos. (...) E também a gente faz revisões de área. Já foram feitas várias revisões de áreas, porque ... teve áreas que a gente percebeu que estavam muito ... a última que foi feita agora foi a parte de engenharia naval, mas nós já fizemos revisão da área de ... de ... da parte de informática lá, de ... ciência da computação ... já fizemos revisão da área de matemática ... então ... inclusive [criamos] uma metodologia pra fazer revisão de áreas. Então toda vez que a gente percebe essas deficiências, não por parte do usuário, na verdade, a gente mesmo que faz revisões de área pra ... e treinamentos pra indexação, pra ver se procura melhorar a indexação (I-USP)

Não ... não me lembro também de ter acontecido algo do tipo assim, de ter individual, do tipo, acredito que não (I-UNESP)

Não. (...) Então ... o que a gente faz aqui é o seguinte ... é ... a gente não faz avaliação, mas vamos supor, o usuário não conseguiu chegar aonde ele queria, ele manda as reclamações, as sugestões, o que ele tá precisando por e-mail que tem no repositório. Aí eu respondo pra ele. Então a gente tem ... não é chat porque não fica online 100%, não fica 24h mas ... a resposta é rápida, é isso que a gente faz, a gente não faz avaliação (I-UNICAMP)

Em relação às possíveis melhorias apontadas pelos profissionais em relação à representação temática, a bibliotecária catalogadora-indexadora da USP acredita não haver melhorias significativas a serem realizadas em termos de vocabulário, apontando a

necessidade de uma maior contribuição dos profissionais que realizam a indexação, na atualização do vocabulário controlado ou padronização em remissivas.

Olha, eu acredito que não, em termos de melhoria é mesmo contar com a contribuição dos bibliotecários que trabalham com isso e que estão indexando e vem os assuntos novos que estão entrando ou mesmo ... um assunto que está de uma forma e hoje em dia é de outro, pra gente fazer as remissivas, fazer as modificações, é mais nesse sentido, né, de melhoria que eu acho mas ... de resto é ... é a contribuição mesmo que ... do colega, né, que eu acho que é importante pra mim (I-USP)

O bibliotecário catalogador-indexador da UNESP sugere que o *Vocabulário Controlado da UNESP* possa ser disponibilizado aos alunos de pós-graduação no momento em que realizam o autoarquivamento de teses e dissertações no repositório.

É ... é ... talvez ... né ... é uma coisa recente né ... quando do preenchimento dos metadados eles tenham à disposição do tesouro, talvez algum tipo de campanha, campanha é uma coisa muito ... genérica de dizer né, mas isso aí seria até das seções de pós pra conscientização deles de tentarem se esforçar pra ... o uso dos metadados, né ... de uso do tesouro pra ... pra uma melhoria nesse sentido. (...) Mas eu não sei, única coisa que me vem à mente agora, mas ... (...) Seria alguma coisa nesse sentido (I-UNESP)

Além disso, o bibliotecário catalogador-indexador da UNESP observa que, após a implementação do autoarquivamento no repositório institucional, há um maior número de termos elencados pelos alunos de pós-graduação, possibilitando ao bibliotecário que valida os respectivos metadados no catálogo online Athena uma visão mais ampla das temáticas abordadas, melhorando a compatibilização da linguagem natural para a linguagem controlada. Em sua visão, isto não era possível no processo anterior, em que os alunos de pós-graduação solicitavam a elaboração de fichas catalográficas, limitando-se à quantidade estipulada de termos.

Não, só que é interessante essa parte porque você ... apesar de você estar mais distante do ... do vamos dizer assim, não mais distante né, mas ... não estar tão próximo do ... do usuário, do ... do pesquisador, mas, é que você parece que tem uma visão a mais, por ele preencher os ... os metadados ali, os campos 600 principalmente às vezes você tem uma visão mais ... porque quando a gente pedia pra ... dados pra elaboração da ficha a gente pedia página de rosto. (...) É ... embora ele, né, no resumo já houvesse essas palavras ali parece que ele se solta um pouco mais, e às vezes a gente, apesar de elas não serem precisas muitas vezes a gente consegue ... ter uma ideia um pouco melhor às vezes, né, pela quantidade que ele pode fazer nesse preenchimento né. (...) É uma coisa que ... nem sempre né, nem todos os casos são assim mas muitas vezes pode acontecer (I-UNESP)

O profissional também observa a insatisfação por parte dos alunos de pós-graduação quando no momento de elaboração das fichas catalográficas, pois em diversas ocasiões os termos elencados/sugeridos não eram compatíveis com a linguagem controlada. Já no repositório institucional, há uma maior liberdade na atribuição das teses e dissertações, visto

que os termos sugeridos permanecem nos registros das mesmas no repositório institucional, enquanto que no catálogo online há validação/verificação e compatibilização dos termos de acordo com uma linguagem controlada.

(E ... você percebe assim ... que existe essa ... eles percebem quando trocam os assuntos, existe esse questionamento?) *É, isso só acontecia muito quando havia a ficha catalográfica. (...) Então a gente chegava, elaborava, eu devolvia ... é ... pra eles, né, às vezes em PDF, às vezes no Word, dependendo da época tinha uma política pra isso ... e eles falavam, mas essas não foram as palavras-chave que eu coloquei, as palavras-chave que eu coloquei, aí eu explicava olha, nem sempre é possível utilizar essa ou aquela palavra por uma questão de ... de política da instituição pra ... enfim, fazia, é ... tinha um discurso padrão pra isso daí. Às vezes ... não é que batiam o pé, às vezes insistiam que aquela palavra era muito importante, né ... então a gente tentava dar um jeito ali na ... na situação. (...) Às vezes até colocava mesmo ela não sendo válida e depois na catalogação ... a gente usava ... o que era melhor pra eles (I-UNESP)*

Por fim, pela sua percepção as sugestões de termos em linguagem natural são tanto abrangentes quanto muito específicas, isto é, termos tão específicos que ainda não são consolidados na literatura.

(E ... você percebe se eles colocam os termos mais gerais ou mais específicos em relação à linguagem?) *É, é uma mistura, viu (...) às vezes eles são tão específicos que eles são praticamente um neologismo ali no ... talvez eles nem existam ... nem no dicionário, por exemplo, né. (...) É uma coisa muito ali ... em Ciências Sociais às vezes pesquisas de ... das cidades, como seguimentos, assim, então tem um vocabulário, às vezes eles, né ... usam essas palavras desse vocabulário ali que é segmentado, de grupos, às vezes, por exemplo. (...) É uma coisa muito nova e não tem, né. (...) Então às vezes tem termo que ... que ...~~ abarca isso daí, que abrange, mas aquele termo específico não, às vezes nem gostam muito que troca, né, mas enfim ((RI)) (I-UNESP)*

A bibliotecária catalogadora-indexadora da UNICAMP acredita ser relevante o estabelecimento de uma política de indexação na UNICAMP e explica que embora tal política não esteja formalizada em um manual de indexação, está sendo adotada para os demais recursos informacionais a mesma política para teses e dissertações no que se refere à padronização dos assuntos. Enquanto a produção retrospectiva está sendo realizada pelos bibliotecários das Unidades, os assuntos dos novos recursos informacionais oriundos de coletas automáticas são padronizados por uma estagiária de Biblioteconomia.

Então, o que ... o ideal seria a gente ter uma política de indexação na UNICAMP, o que não tem ainda. Mas o que nós estamos fazendo é usar a mesma política que nós usamos pra teses e dissertações em todo o repositório (...) é o que nós estamos fazendo agora. Só que, qual é a realidade? É algo que vai demorar pra gente chegar 100%. Não é algo rápido. E uma coisa que ... eu ... nós estamos fazendo por conta do repositório: nós não estamos deixando subir nada que não esteja realmente padronizado no SophiA. (...) Então chegou um assunto aqui pra gente, olha, esse artigo está todo padronizado, eu vou lá olhar. Primeira coisa que eu olho são os assuntos. Realmente foi feita toda a indexação? Ah, foi, então realmente esse pode entrar. Não foi feita toda a indexação? Não entra. (...) Isso ... é, ah, mas isso aqui já tá no SophiA faz tanto tempo, tem sei lá, 500 pendurados, ok, tem 500 pendurados mas você vai fazer a indexação completa, você vai fazer toda a normalização, porque senão ele não vai entrar no repositório. Isso a gente tá fazendo, esse cuidado nós estamos tendo, com todos os assuntos. Ele está no SophiA, ele está ...

padronizado? Não está. Não entra. (E os outros artigos que já entraram no repositório?) Não, então, a gente tá tendo que olhar um por um. (...) E é assim, vamos supor, 2018 nós fizemos a coleta automática ... só que ele não subiu pro repositório. Por que? Ele já está sendo padronizado. Ele já vai subir padronizado, tá sendo feito todos os assuntos, esses assuntos estão sendo feitos na DTRI, porque quem está fazendo é uma estagiária de Biblioteconomia. Então nós demos o treinamento pra ela e ela está fazendo todos os assuntos. E não são poucos, porque tem muito assunto novo. Então ... ela que está fazendo, um a um, não estamos subindo mais (I-UNICAMP)

Em continuidade, apresenta-se os resultados da avaliação da indexação de assuntos enquanto abordagem quantitativa da pesquisa.

5.2 Resultados da abordagem quantitativa: avaliação da indexação de assuntos

Visando realizar a análise e discussão da recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária, nesta subseção são apresentados os resultados da aplicabilidade da avaliação da indexação de assuntos enquanto abordagem quantitativa no Repositório Institucional da UNICAMP, em duas principais abordagens: a) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e b) avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*.

5.2.1 Resultados da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do sistema - Avaliação intrínseca quantitativa mediante a interconsistência

A aplicabilidade da *Avaliação intrínseca quantitativa mediante a interconsistência*, denominada nesta pesquisa de *Avaliação Inter-autor-bibliotecário* possibilitou a comparação dos assuntos atribuídos pelos próprios autores (termos livres) e pelos bibliotecários catalogadores-indexadores (descritores), após validação/verificação de assuntos pelos profissionais de artigos de periódico no contexto específico do Repositório Institucional da UNICAMP.

Para o levantamento dos registros, foi gerado um relatório em *Excel* pela *Seção de Produtos/Conteúdos Informacionais Digitais da Diretoria de Tratamento da Informação* (DTRI) do SBU. Tais registros foram anteriormente validados pelos bibliotecários das Bibliotecas do SBU. Após a verificação inicial pela pesquisadora de que tais registros apresentavam tanto os *metadados descritivos* quanto os *metadados temáticos*, elegeu-se uma amostra de um total de 1000 registros³⁹ passíveis de análise, distribuídos em cinco áreas do conhecimento. Para tanto, foram utilizados os seguintes campos do relatório: “Termos livres” (atribuídos pelos autores) e “Assunto” (atribuídos ou validados pelos bibliotecários catalogadores-indexadores).

Para a análise dos dados, inicialmente, foram construídos dois principais arquivos, a saber: a) arquivo contendo o rol da totalidade de assuntos atribuídos por autores e bibliotecários catalogadores-indexadores e b) arquivo contendo o rol da totalidade de assuntos atribuídos por autores e bibliotecários catalogadores-indexadores de acordo com cada área do conhecimento analisada. Em um segundo momento procedeu-se à aplicação das fórmulas de consistência de Rolling (1981) e de Hooper (1965) adaptada por Gil Leiva (2008, p. 386), permitindo uma comparação dos índices de consistência gerados no universo de pesquisa do Repositório Institucional da UNICAMP.

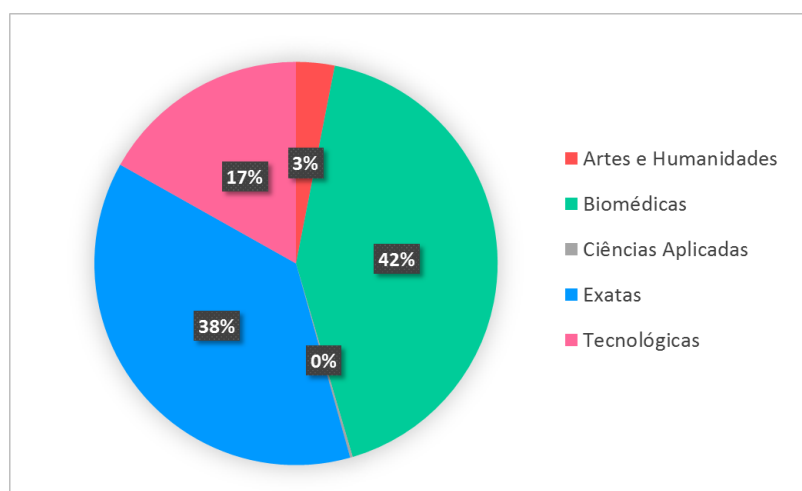
As análises possibilitaram o levantamento de três principais variáveis/indicadores, com suas respectivas subvariáveis/subindicadores: a) *Dados gerais*: total de recursos informacionais analisados no Repositório Institucional da UNICAMP por área do conhecimento; b) *Número de assuntos atribuídos*: número de assuntos atribuídos na indexação por autores e bibliotecários catalogadores-indexadores por recurso informacional; número de assuntos atribuídos em *linguagem natural* por área do conhecimento; número de assuntos atribuídos em *linguagem controlada* por área do conhecimento; e c) *Índices de interconsistência*: média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por tipo de índice; índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores por área do conhecimento; média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento.

³⁹ Cabe esclarecer que tal quantidade refere-se à produzida até junho/2019 pelas referidas Bibliotecas após capacitações profissionais, e não à quantidade total de registros do Repositório Institucional da UNICAMP.

A) Dados gerais

No primeiro indicador analisado, realizou-se o levantamento geral dos dados dos 1000 registros por meio da verificação do número dos registros referentes a cada área do conhecimento do SBU, a saber: *Área de Artes e Humanidades*, *Área de Biomédicas*, *Área de Exatas*, *Área de Ciências Aplicadas* e *Área de Tecnológicas*. Para a *Área de Conhecimentos Gerais*, não houve registros. A seguir, apresenta-se os recursos informacionais analisados no Repositório Institucional da UNICAMP por área do conhecimento (Figura 8).

Figura 8. Total de recursos informacionais analisados no Repositório Institucional da UNICAMP por área do conhecimento



Fonte: Elaboração própria.

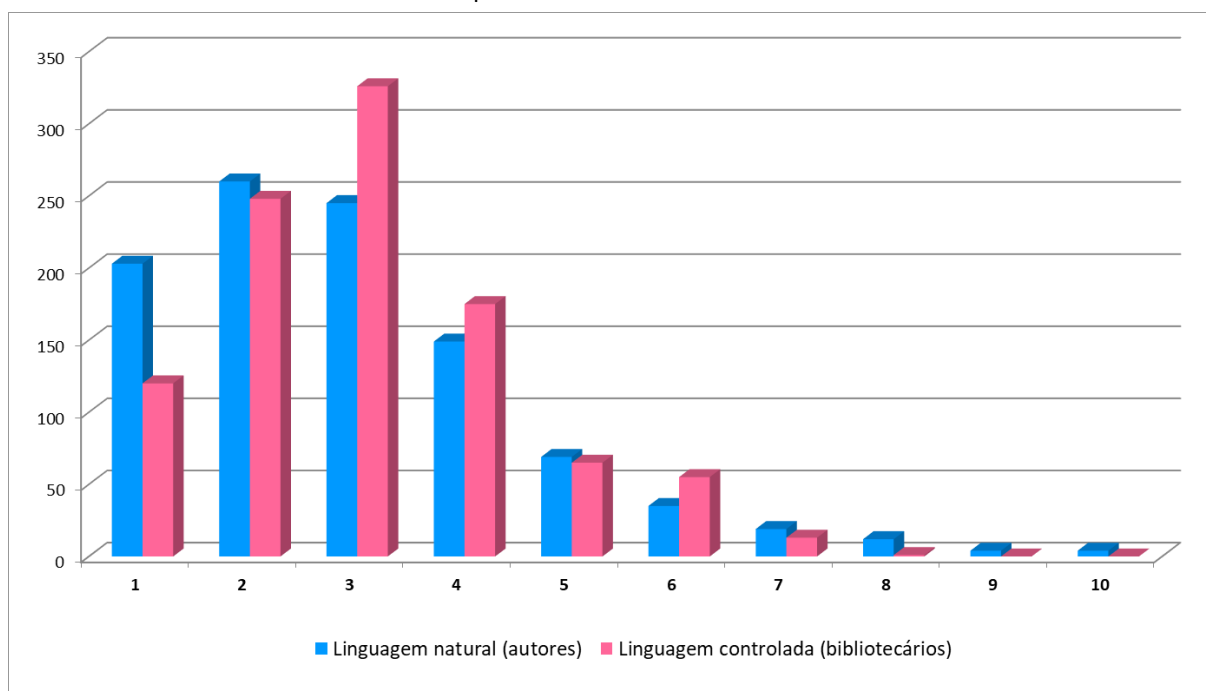
Do total do corpus de 1000 registros analisado, na *Área de Artes e Humanidades* foram analisados 31 artigos (3,1%), na *Área de Biomédicas* 425 artigos (42,5%), na *Área de Ciências Aplicadas* 2 artigos (0,2%), na *Área de Exatas* 376 artigos (37,6%) e na *Área de Tecnológicas* 169 artigos (16,9%).

B) Número de assuntos atribuídos

A segunda variável analisada - número de assuntos atribuídos – possibilitou o levantamento de duas subvariáveis: a) *número de assuntos* na indexação por autores e

bibliotecários catalogadores-indexadores por recurso informacional e b) *número de assuntos* na indexação por autores e bibliotecários catalogadores-indexadores por área do conhecimento.

Figura 9. Número de assuntos atribuídos na indexação por autores e bibliotecários catalogadores-indexadores por recurso informacional



Fonte: Elaboração própria.

Referente ao número de assuntos atribuídos por recurso informacional (Figura 9), em relação à linguagem natural (autores) verifica-se que o maior número de recursos informacionais foi atribuído pelos autores com 2 assuntos (26%), seguido de 3 assuntos (24,5%) e 1 assunto (20,3%), demonstrando certo equilíbrio entre estas variáveis. Com menores percentuais aparecem gradativamente 4 assuntos (14,9%), 5 assuntos (6,9%), 6 assuntos (3,5%), 7 assuntos (1,9%), 8 assuntos (1,2%) e 9 e 10 assuntos, com apenas 0,4% cada. Já em relação à linguagem controlada (bibliotecários catalogadores-indexadores), verifica-se que o número mais expressivo de recursos informacionais foi atribuído com 3 assuntos (32,6%), seguido de 2 assuntos (24,8%), 4 assuntos (17,5%) e 1 assunto (12%). Com menores percentuais aparecem 5 assuntos (6,5%), 6 assuntos (5,5%), 7 assuntos (1,3%) e 8 assuntos, com apenas 0,1% dos recursos informacionais. Para 9 e 10 assuntos não houve ocorrências.

Ao realizar-se uma análise comparada do número de assuntos atribuídos pelos autores e bibliotecários catalogadores-indexadores, verifica-se que enquanto para a linguagem natural (autores) houve uma tendência de atribuição dos recursos informacionais entre 1 e 3 assuntos, na linguagem controlada (bibliotecários catalogadores-indexadores) houve uma maior expressividade na atribuição dos recursos informacionais com 3 assuntos. Por outro lado, tanto em relação à linguagem natural (autores) quanto à linguagem controlada (bibliotecários catalogadores-indexadores), houve uma menor incidência na atribuição de 5 assuntos em diante. Na sequência, de forma mais específica foi levantado o *número de assuntos* na indexação por autores e bibliotecários catalogadores-indexadores por área do conhecimento (Tabela 1).

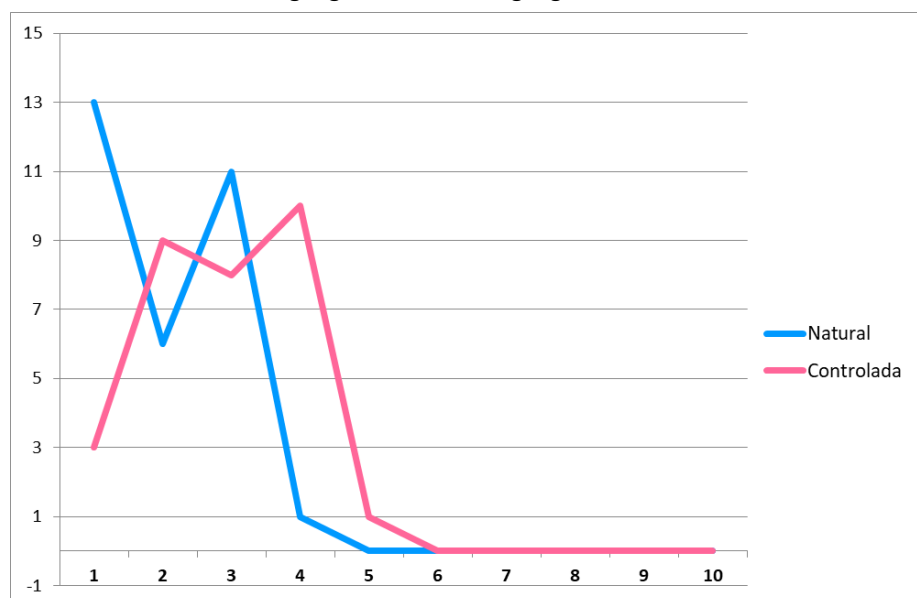
Tabela 1. Distribuição do número de assuntos atribuídos por autores e bibliotecários catalogadores-indexadores no Repositório Institucional da UNICAMP por área do conhecimento

N. assuntos	Artes e Humanidades				Biomédicas				Ciências Aplicadas				Exatas				Tecnológicas			
	Autores		Bibliotecários		Autores		Bibliotecários		Autores		Bibliotecários		Autores		Bibliotecários		Autores		Bibliotecários	
	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi	N	Fi
1 assunto	13	41,9%	3	9,7%	115	27,0%	48	11,3%	0	0%	2	100%	45	12,0%	24	6,4%	30	17,7%	43	25,4%
2 assuntos	6	19,3%	9	29%	134	31,5%	114	26,8%	0	0%	0	0%	63	16,7%	61	16,2%	57	33,7%	64	37,9%
3 assuntos	11	35,5%	8	25,8%	98	23,0%	111	26,1%	1	50%	0	0%	97	25,8%	166	44,1%	38	22,5%	41	24,2%
4 assuntos	1	3,2%	10	32,2%	43	10,1%	63	14,8%	0	0%	0	0%	79	21,0%	85	22,6%	26	15,4%	17	10,0%
5 assuntos	0	0%	1	3,2%	15	3,5%	35	8,2%	1	50%	0	0%	46	12,2%	26	6,9%	7	4,1%	3	1,7%
6 assuntos	0	0%	0	0%	9	2,1%	48	11,3%	0	0%	0	0%	20	5,3%	6	1,6%	6	3,5%	1	0,6%
7 assuntos	0	0%	0	0%	5	1,1%	5	1,1%	0	0%	0	0%	11	2,9%	8	2,1%	3	1,8%	0	0%
8 assuntos	0	0%	0	0%	5	1,1%	1	0,2%	0	0%	0	0%	7	1,9%	0	0%	0	0%	0	0%
9 assuntos	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	3	0,8%	0	0%	1	0,6%	0	0%
10 assuntos	0	0%	0	0%	1	0,2%	0	0%	0	0%	0	0%	2	0,5%	0	0%	1	0,6%	0	0%
Total	31	100%	31	100%	425	100%	425	100%	2	100%	2	100%	376	100%	376	100%	169	100%	169	100%

Fonte: Elaboração própria.

De forma mais específica realizou-se a análise comparada do número de assuntos atribuídos em linguagem natural e em linguagem controlada para cada uma das cinco áreas do conhecimento.

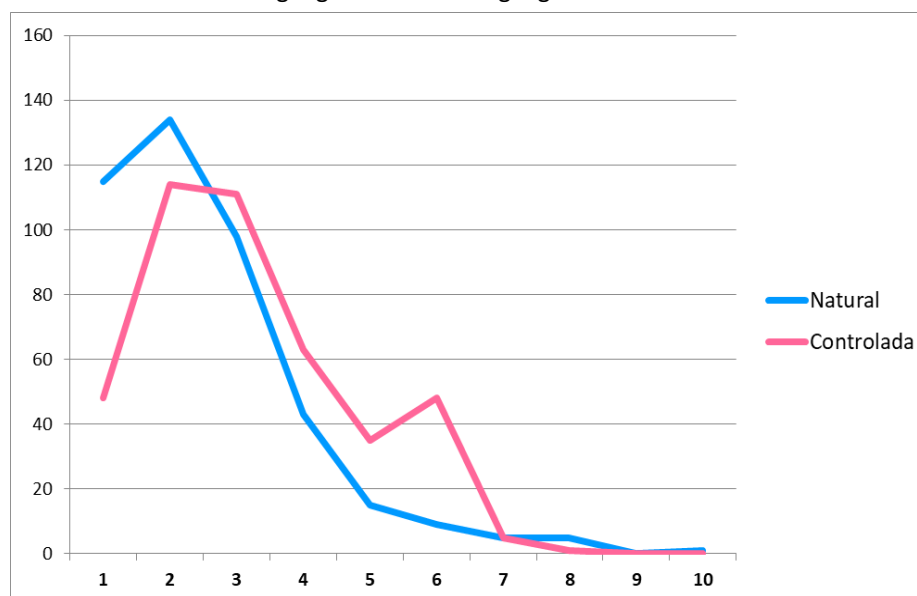
Figura 10. Número de assuntos em linguagem natural e linguagem controlada - Área de Artes e Humanidades



Fonte: Elaboração própria.

Na *Área de Artes e Humanidades* (Figura 10), com 31 registros analisados, a maior parte do número de assuntos atribuídos pelos *autores* foi de 1 assunto (41,9%) e 3 assuntos (35,5%), representando 77,4% do total dos recursos informacionais analisado. Em seguida aparecem 2 assuntos (19,3%) e 4 assuntos, com apenas 3,2%. De 5 a 10 assuntos atribuídos não houve ocorrências. Já em relação aos assuntos atribuídos pelos *bibliotecários catalogadores-indexadores*, houve um maior equilíbrio entre três variáveis: 4 assuntos (32,2%), 2 assuntos (29%) e 3 assuntos (25,8%). Com um menor percentual aparece 1 assunto (9,7%) e 5 assuntos (3,2%). De 6 a 10 assuntos não houve ocorrência.

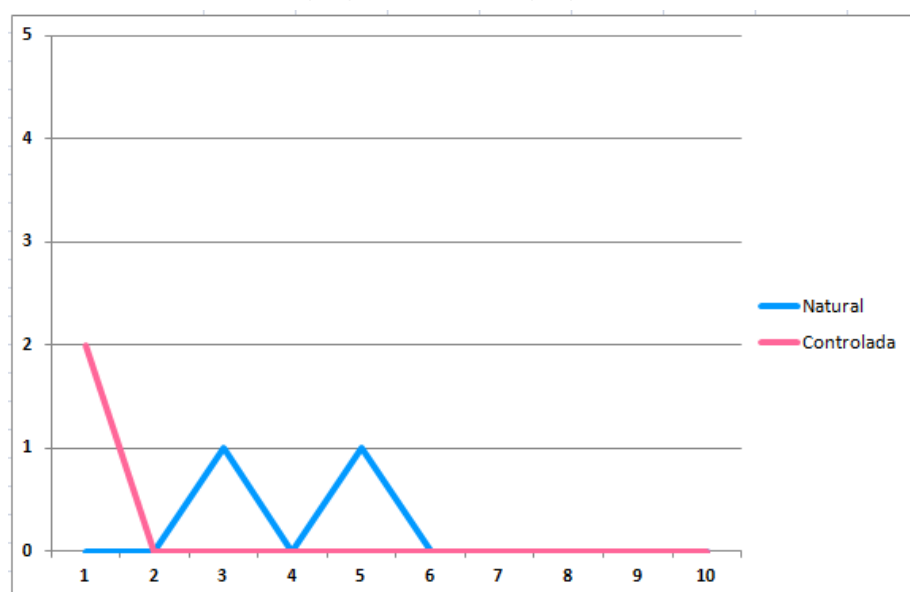
Figura 11. Número de assuntos em linguagem natural e linguagem controlada - Área de Biomédicas



Fonte: Elaboração própria.

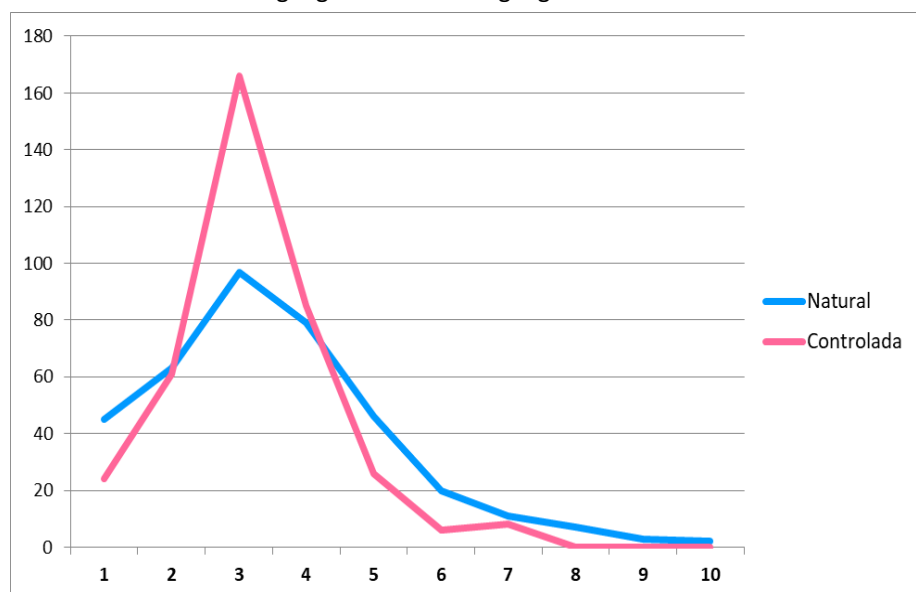
Na *Área de Biomédicas* (Figura 11), representada com a maioria dos recursos informacionais analisados, ou seja, 425 artigos, em relação aos assuntos atribuídos pelos *autores*, a maior parte foi de 2 assuntos (31,5%) e 1 assunto (27%), seguidos de 3 assuntos (23%). Com menores percentuais aparecem 4 assuntos (10,1%), 5 assuntos (3,5%), 6 assuntos (2,1%) e apenas 1,1% para 7 e 8 assuntos. Com um percentual ainda menor, aparece 10 assuntos (0,23%). Para 9 assuntos não houve ocorrência. No que se refere aos assuntos atribuídos pelos *bibliotecários catalogadores-indexadores*, houve um equilíbrio entre 2 assuntos (26,8%) e 3 assuntos (26,1%), seguidos de 4 assuntos (14,8%). Com um percentual de 11,3% aparecem 1 e 6 assuntos atribuídos, seguidos de 5 assuntos (8,2%), 7 assuntos (1,1%) e 8 assuntos, com apenas 0,23% do total. De 9 a 10 assuntos não houve ocorrência.

Figura 12. Número de assuntos em linguagem natural e linguagem controlada - Área de Ciências Aplicadas



Fonte: Elaboração própria.

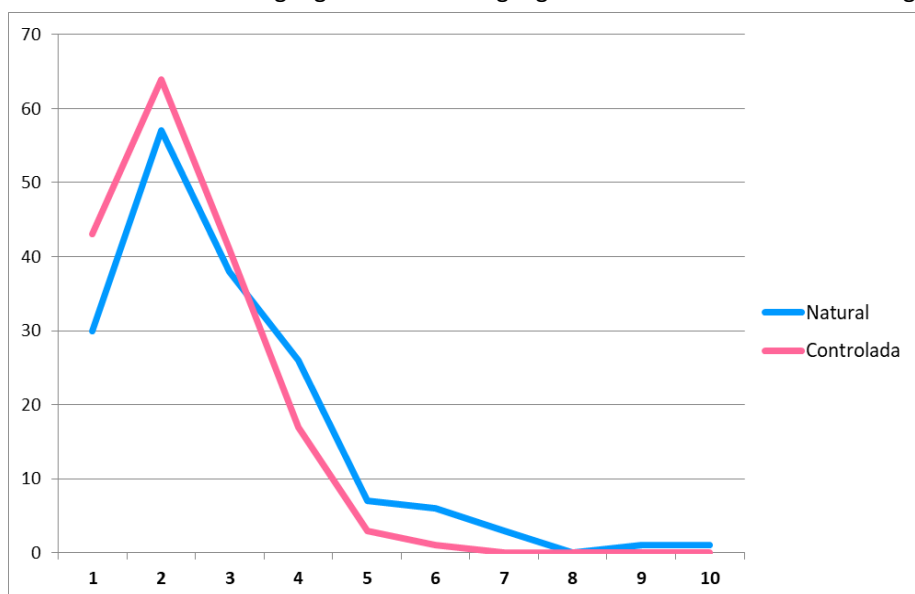
Na *Área de Ciências Aplicadas* (Figura 12), em relação aos assuntos atribuídos pelos *autores*, 50% foram 3 assuntos e 5 assuntos representam os demais 50%. Já em relação aos assuntos atribuídos pelos *bibliotecários catalogadores-indexadores*, todos (100%) foram da ordem de 1 assunto. Cabe ressaltar que apenas 2 assuntos da respectiva área foram verificados - o menor número do corpus analisado.

Figura 13. Número de assuntos em linguagem natural e linguagem controlada - Área de Exatas

Fonte: Elaboração própria.

A *Área de Exatas* (Figura 13), com 376 recursos informacionais analisados, destaca-se pela única ocorrência dentre as 5 áreas analisadas em que houve variação entre 1 a 10 nos assuntos atribuídos pelos *autores*, sendo: 3 assuntos (25,8%), 4 assuntos (21%), 2 assuntos (16,7%), 1 assunto (12%), 5 assuntos (12,2%), 6 assuntos (5,3%), 7 assuntos (2,9%), 8 assuntos (1,9%), 9 assuntos (0,8%) e 10 assuntos (0,5%). Já em relação aos assuntos atribuídos pelos *bibliotecários catalogadores-indexadores*, a maior parte foi 3 assuntos (44,1%) e 4 assuntos (22,6%), seguidos de 2 assuntos (16,2%), 5 assuntos (6,9%) e 1 assunto (6,4%). Com um menor percentual aparecem 7 assuntos (2,1%) e 6 assuntos (1,6%). De 8 a 10 assuntos não houve ocorrência.

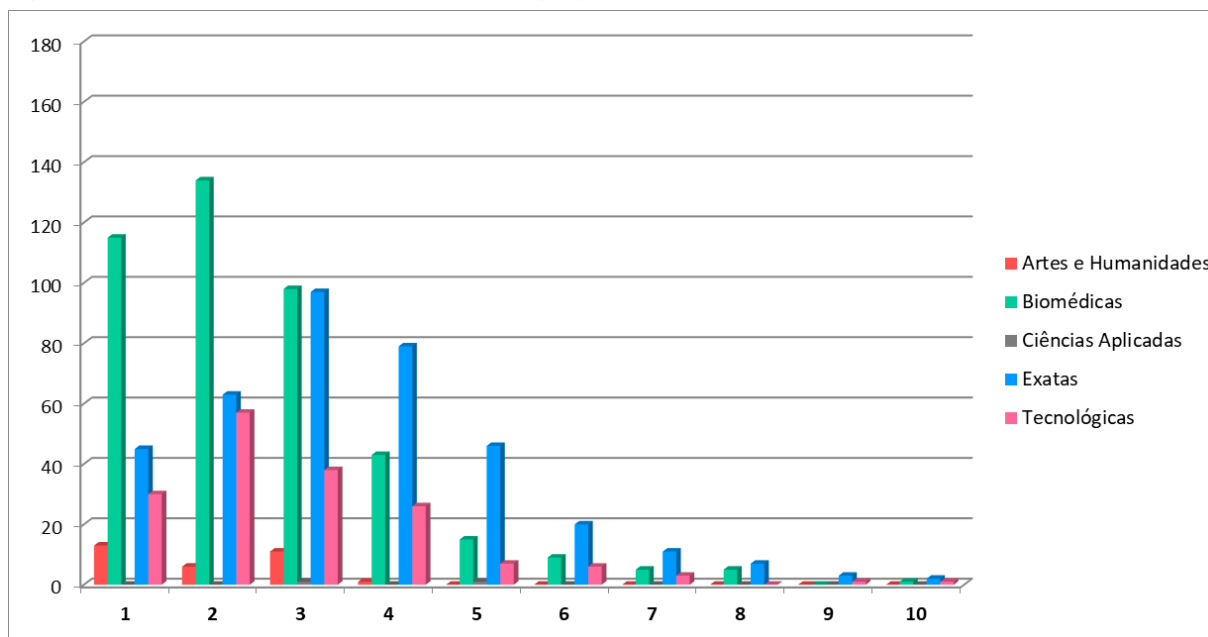
Figura 14. Número de assuntos em linguagem natural e linguagem controlada - Área de Tecnológicas



Fonte: Elaboração própria.

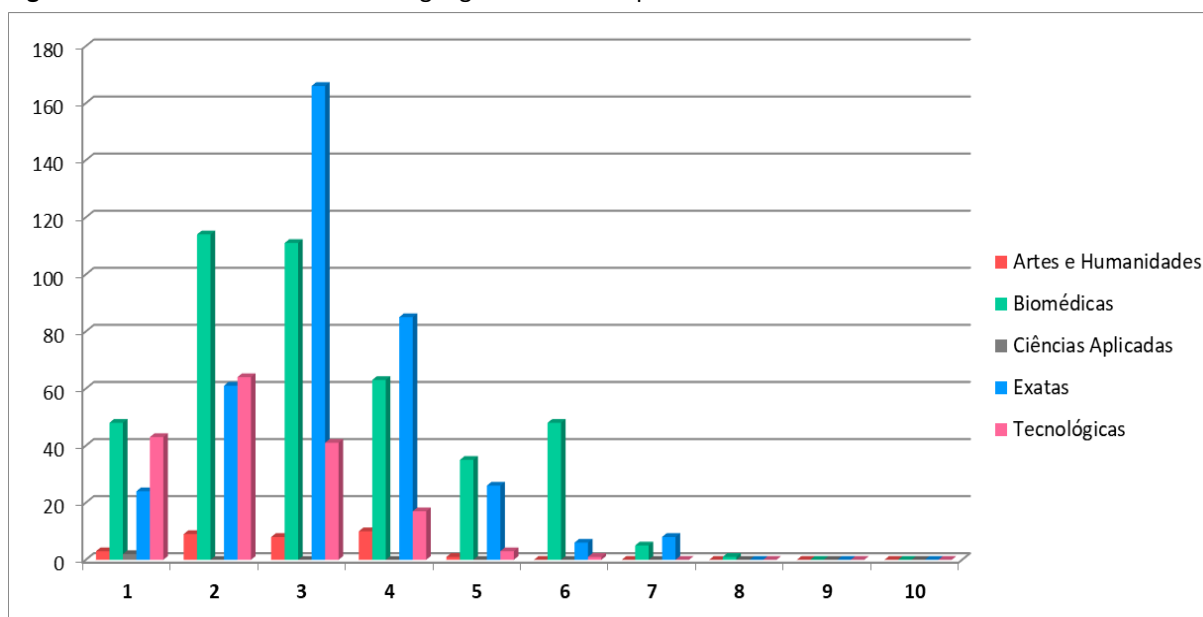
Na *Área de Tecnológicas* (Figura 14), com 169 artigos, em relação aos assuntos atribuídos pelos *autores*, a maior parte foi 2 assuntos (33,7%) e 3 assuntos (22,5%), seguidos de 1 assunto (17,7%), 4 assuntos (15,4%) e 2 assuntos (15,3%). Com um menor percentual aparecem 5 assuntos (4,15%), 6 assuntos (3,5%) e 7 assuntos (1,8%). Com apenas 0,59% aparecem 9 e 10 assuntos, sendo que para 8 assuntos não houve ocorrência. Já em relação aos assuntos atribuídos pelos *bibliotecários catalogadores-indexadores*, a maior parte foi de 2 assuntos (37,9%), seguida de 1 assunto (25,4%) e 3 assuntos (24,2%). Com menores percentuais aparecem 4 assuntos (10%), 5 assuntos (1,7%) e 6 assuntos (0,59%) do total dos recursos informacionais analisados. De 7 a 10 assuntos não houve ocorrências.

Em continuidade, procedeu-se à elaboração da representação gráfica da análise comparada do número de assuntos atribuídos em *linguagem natural* e do número de assuntos atribuídos em *linguagem controlada* por área do conhecimento (Figura 15):

Figura 15. Número de assuntos atribuídos em linguagem natural por área do conhecimento

Fonte: Elaboração própria.

No geral, ao realizar-se a análise comparada do número de assuntos atribuídos em *linguagem natural* por área do conhecimento, verifica-se que houve uma maior predominância na atribuição de recursos informacionais com 2 assuntos na *Área de Biomédicas* e na *Área de Tecnológicas*, enquanto que para a *Área de Exatas* houve maior incidência na atribuição de recursos informacionais com 3 assuntos. Já para a *Área de Artes e Humanidades*, destaca-se a maior incidência com apenas 1 assunto. A *Área de Ciências Aplicadas* aparece tanto com 3 quanto 5 assuntos. Tais resultados indicam uma variação na atribuição de assuntos pelos autores por área do conhecimento de 1 a 5 assuntos.

Figura 16. Número de assuntos em linguagem controlada por área do conhecimento

Fonte: Elaboração própria.

Por outro lado, no que tange ao número de assuntos atribuídos em *linguagem controlada* por área do conhecimento (Figura 16), nota-se também uma predominância na atribuição de recursos informacionais com 3 assuntos, fator corroborado pela *Área de Biomédicas* e *Área de Exatas*. Notadamente para a *Área de Biomédicas*, houve também uma predominância de atribuição de 2 assuntos, praticamente empatada com a de 3 assuntos. A *Área de Tecnológicas* destaca-se com uma maior incidência de 2 assuntos, enquanto que na *Área de Artes e Humanidades*, houve uma predominância de atribuição de recursos informacionais com 4 assuntos. Apenas para a *Área de Ciências Aplicadas* houve uma predominância de atribuição de apenas 1 assunto para os recursos informacionais. Tais resultados indicam uma variação na atribuição de assuntos pelos bibliotecários catalogadores-indexadores de 1 a 4 assuntos.

C) Índices de interconsistência

Já na terceira variável analisada - índices de interconsistência, foi possível verificar três subvariáveis: a) média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por tipo de índice; b) índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores por área do

conhecimento; e c) média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento.

Conforme delineado de forma detalhada nos procedimentos metodológicos, para a comparação dos índices de consistência gerados no universo de pesquisa do Repositório Institucional da UNICAMP, como parâmetro metodológico foram utilizadas a fórmula de consistência de Rolling (1981) e a fórmula de consistência de Hooper (1965) adaptada por Gil Leiva (2008, p. 386), tanto no índice flexível quanto no rígido (Quadro 8):

Quadro 8. Exemplos de aplicação das fórmulas de consistência no Repositório Institucional da UNICAMP

Exemplos		Índice flexível fórmula de Rolling	Índice flexível fórmula de Hooper	Índice rígido fórmula de Rolling	Índice rígido fórmula de Hooper
Indexação A (Autores)	Indexação B (Bibliotecários)				
<u>Pobreza*</u> Desigualdade <u>Mobilidade*</u> de renda	<u>Pobreza*</u> Igualdade Regiões metropolitanas <u>Mobilidade*</u> social	$C_i = \frac{2 * 1,5}{3 + 4}$ $C_i = 0,42 \times 100 = \mathbf{42\%}$	$C_i = \frac{1,5}{(3 + 4) - 1,5}$ $C_i = 0,27 \times 100 = \mathbf{27\%}$	$C_i = \frac{2 * 1}{3 + 4}$ $C_i = 0,28 \times 100 = \mathbf{28\%}$	$C_i = \frac{1}{(3 + 4) - 1}$ $C_i = 0,16 \times 100 = \mathbf{16\%}$
<u>Fluoride*</u> <u>Polysaccharide*</u>	Caries Dental plaque Fluorides* <u>Polysaccharides*</u> Sucrose	$C_i = \frac{2 * 2}{2 + 5}$ $C_i = 0,57 \times 100 = \mathbf{57\%}$	$C_i = \frac{2}{(2 + 5) - 2}$ $C_i = 0,40 \times 100 = \mathbf{40\%}$	$C_i = \frac{2 * 2}{2 + 5}$ $C_i = 0,57 \times 100 = \mathbf{57\%}$	$C_i = \frac{2}{(2 + 5) - 2}$ $C_i = 0,40 \times 100 = \mathbf{40\%}$
Posição <u>internacional*</u> de investimentos Vulnerabilidade externa Balanço de pagamentos Transações correntes Economia brasileira	Relações econômicas <u>internacionais*</u>	$C_i = \frac{2 * 0,5}{5 + 1}$ $C_i = 0,16 \times 100 = \mathbf{16\%}$	$C_i = \frac{0,5}{(5 + 1) - 0,5}$ $C_i = 0,09 \times 100 = \mathbf{9\%}$	$C_i = \frac{2 * 0}{5 + 1}$ $C_i = 0 \times 100 = \mathbf{0\%}$	$C_i = \frac{0}{(5 + 1) - 0}$ $C_i = 0 \times 100 = \mathbf{0\%}$
Ensino de <u>Geociências*</u> Tics no <u>Ensino*</u> <u>Formação de professores*</u>	<u>Geociências*</u> - Estudo e ensino* <u>Formação de professores*</u>	$C_i = \frac{2 * 2}{3 + 2}$ $C_i = 0,8 \times 100 = \mathbf{80\%}$	$C_i = \frac{2}{(3 + 2) - 2}$ $C_i = 0,66 \times 100 = \mathbf{66\%}$	$C_i = \frac{2 * 1}{3 + 2}$ $C_i = 0,40 \times 100 = \mathbf{40\%}$	$C_i = \frac{1}{(3 + 2) - 1}$ $C_i = 0,25 \times 100 = \mathbf{25\%}$
<u>Pão*</u> integral Análise <u>sensorial*</u> ADQ Aceitação	<u>Pão*</u> Avaliação <u>sensorial*</u>	$C_i = \frac{2 * 1}{4 + 2}$ $C_i = 0,33 \times 100 = \mathbf{33\%}$	$C_i = \frac{1}{(4 + 2) - 1}$ $C_i = 0,2 \times 100 = \mathbf{20\%}$	$C_i = \frac{2 * 0}{4 + 2}$ $C_i = 0 \times 100 = \mathbf{0\%}$	$C_i = \frac{0}{(4 + 2) - 0}$ $C_i = 0 \times 100 = \mathbf{0\%}$

Fonte: Elaboração própria.

*Termos considerados correspondentes entre a *Indexação A* (autores) e a *Indexação B* (bibliotecários).

Os resultados dos índices de consistência variam entre os valores de 0 a 1, e posteriormente são multiplicados por 100 para a transformação em porcentagem (GIL LEIVA, 2008, p. 387). Aplicando-se as referidas fórmulas entre o processo de indexação de autores e bibliotecários catalogadores-indexadores, foram obtidos os seguintes *índices* ou *ensaios de interconsistência* (Tabela 2):

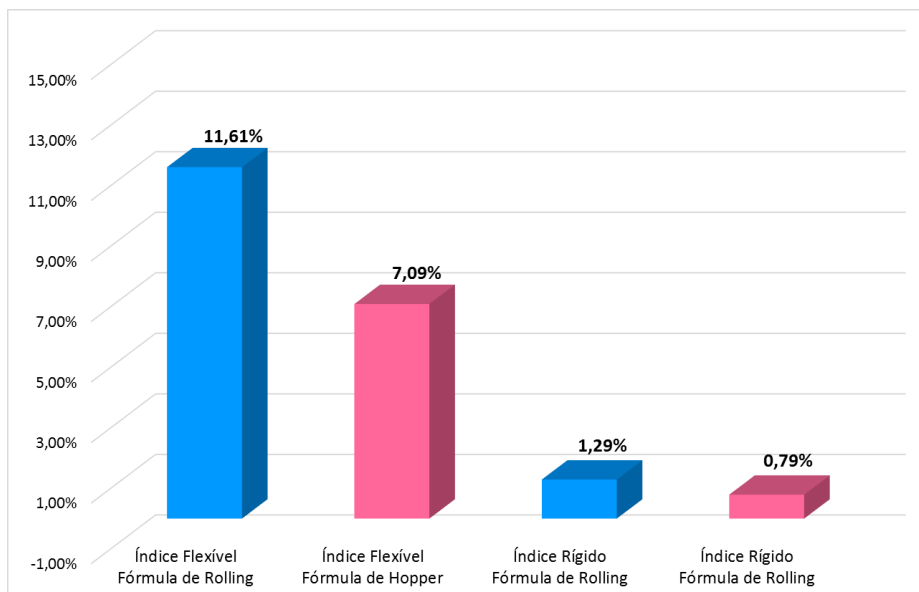
Tabela 2. Índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores no Repositório Institucional da UNICAMP

Área do conhecimento	Índice flexível fórmula de Rolling	Índice flexível fórmula de Hooper	Índice rígido fórmula de Rolling	Índice rígido fórmula de Hooper	Média por área do conhecimento
Artes e Humanidades	12,74%	7,71%	2,78%	1,62%	6,21%
Biomédicas	5,95%	3,42%	0,32%	0,20%	2,47%
Ciências Aplicadas	12,50%	7,14%	0%	0%	4,91%
Exatas	21,05%	13,89%	3,13%	2,01%	10,02%
Tecnológicas	5,83%	3,32%	0,24%	0,15%	2,38%
Média por índice	11,61%	7,09%	1,29%	0,79%	-

Fonte: Elaboração própria.

Calculando-se a média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por tipo de índice, verifica-se que o maior nível foi obtido com o índice flexível da Fórmula de Rolling (11,61%), seguido do índice flexível da Fórmula de Hooper (7,09%), do índice rígido da Fórmula de Rolling (1,29%) e do índice rígido da Fórmula de Hooper (0,79%) (Figura 17):

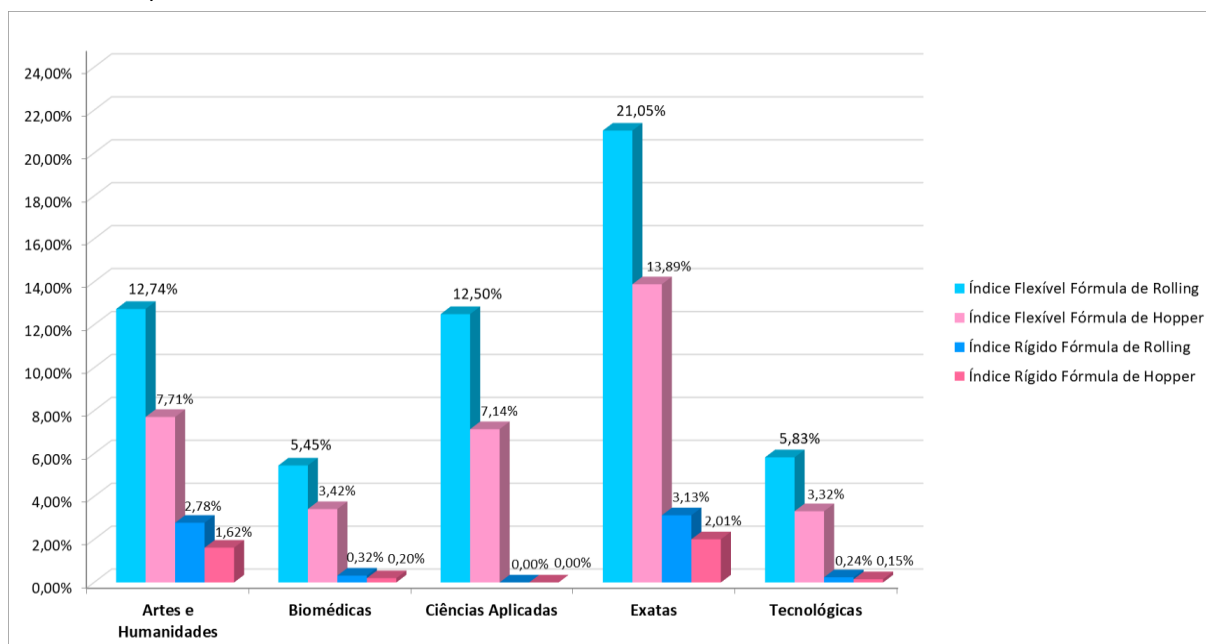
Figura 17. Média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por tipo de índice



Fonte: Elaboração própria.

De forma mais específica, ao realizar-se a comparação dos índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores por área do conhecimento, foram obtidos os seguintes resultados (Figura 18):

Figura 18. Índices de interconsistência na indexação de assuntos entre autores e bibliotecários catalogadores-indexadores por área do conhecimento



Fonte: Elaboração própria.

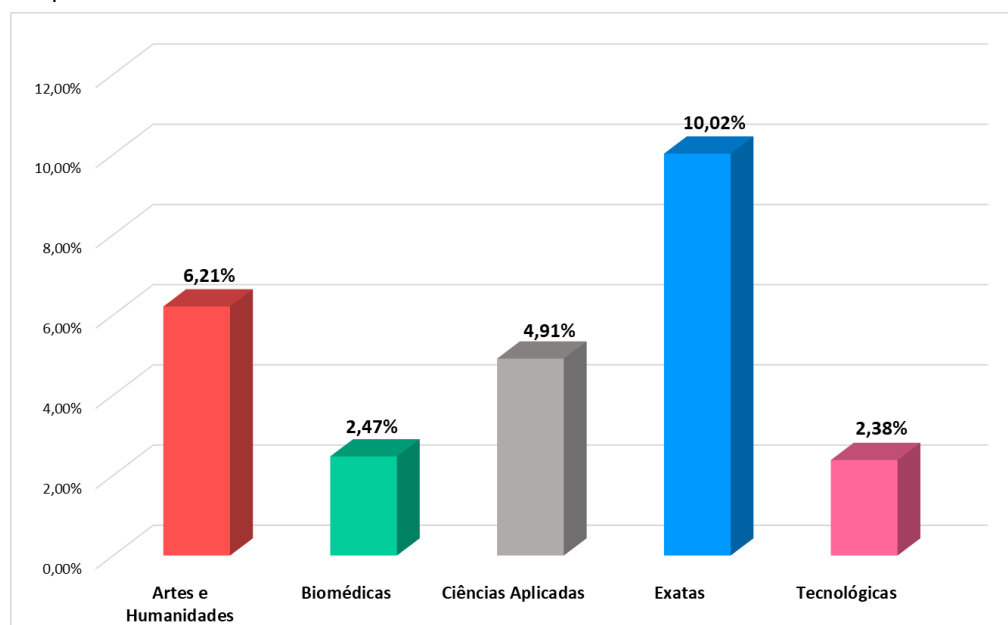
Conforme observado, na *Área de Artes e Humanidades* os índices com a Fórmula de Rolling alcançaram 2,78% para o índice rígido e 12,74% para o índice flexível. Já com a Fórmula de Hooper, os índices ficaram mais abaixo, com apenas 1,62% para o índice rígido e 7,71% para o índice flexível. Tal área equivale a 31 recursos informacionais, ou seja, 3,01% do total analisado, ficando em segundo lugar na média por área, com 6,21%. Com a maior quantidade de registros analisada, ou seja, 425 recursos informacionais (42,5%), na *Área de Biomédicas* foram obtidos apenas 0,32% para o índice rígido e 5,95% para o índice flexível da Fórmula de Rolling. Com a aplicabilidade da Fórmula de Hooper, os índices são de 0,20% para o índice rígido e 3,42% para o índice flexível.

Na *Área de Ciências Aplicadas* houve 0% de correspondência nos índices rígidos das Fórmulas de Rolling e de Hooper. Entretanto, considerando-se os índices flexíveis, os índices sobem para 12,50% e 7,14% respectivamente. Embora esteja em terceiro lugar na média dos melhores índices obtidos por área do conhecimento (4,91%), tal área corresponde a apenas 2 (0,2%) dos recursos informacionais analisados. A *Área de Exatas*, com 376 recursos informacionais analisados (37,6%) destaca-se dentre as cinco áreas do conhecimento pesquisadas com a maior média dos índices de interconsistência na aplicabilidade de ambas as fórmulas por área, com 10,02%, tendo sido obtidos 2,01% no índice rígido e 13,89% no índice flexível para a Fórmula de Hooper. Já com a Fórmula de Rolling, houve 3,13% para o índice rígido e 21,05%, para o índice flexível, o mais alto índice de interconsistência entre todos os índices levantados.

Na *Área de Tecnológicas*, que representa 169 recursos informacionais analisados (16,9%), com a aplicabilidade da Fórmula de Rolling houve 5,83% para o índice flexível e 0,24% para o índice rígido. Já com a Fórmula de Hooper, os índices ficaram entre 3,32% para o índice flexível e apenas 0,15% para o índice rígido, corroborando para que a área ficasse em último lugar na média por área do conhecimento, com apenas 2,38%.

Calculando-se a média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento, ou seja, por meio da somatória dos resultados dos quatro índices (Índice Flexível - Fórmula de Rolling; Índice Flexível - Fórmula de Hooper; Índice Rígido - Fórmula de Rolling; e Índice Rígido - Fórmula de Hooper) dividindo-se pelo total de quatro índices analisados, obtivemos os seguintes resultados (Figura 19):

Figura 19. Média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento



Fonte: Elaboração própria.

Em relação à média dos índices de interconsistência na indexação entre autores e bibliotecários catalogadores-indexadores por área do conhecimento, verifica-se que o maior índice de interconsistência foi obtido na *Área de Exatas*, com 10,02% de concordância entre os termos, seguida da *Área de Artes e Humanidades*, com 6,21% e da *Área de Ciências Aplicadas*, um pouco mais abaixo, com 4,91%. Por outro lado, os menores índices de interconsistência ocorreram na *Área de Biomédicas*, com 2,47% e, por fim, na *Área de Tecnológicas*, com apenas 2,38% de coincidência entre os assuntos atribuídos pelos autores e os assuntos atribuídos ou validados pelos bibliotecários catalogadores-indexadores.

Devido à ausência de outros estudos que abarquem a avaliação da indexação de assuntos em repositórios institucionais, não foi possível realizar uma análise comparada com outros índices neste específico sistema de recuperação da informação. Entretanto, no geral a literatura sobre a temática avaliação da indexação de assuntos na Organização do Conhecimento apresenta que as pesquisas em torno da avaliação da consistência da indexação na abordagem quantitativa revelam que os valores de consistência variam muito entre os indexadores profissionais, sendo que a média dos índices oscila entre os 25% e os 60% de coincidências obtidas na atribuição de assuntos de um mesmo recurso informacional

(GIL LEIVA, 2008, p. 76) e, em uma variação mais ampla, entre 10% e 80% (PINHEIRO, 1978, p. 109).

Em investigação comparada entre o processo de catalogação de assunto em catálogo online e o processo de indexação de assuntos em base de dados realizados pelo mesmo bibliotecário ao longo do tempo, ou seja, avaliação pela intraconsistência, Tartarotti (2014) verificou que os índices da referida investigação ficaram dentro de ambas as margens, sendo que a média dos índices de intraconsistência na catalogação de assunto das Bibliotecas universitárias pesquisadas oscilou entre 40% e 45%, enquanto que na indexação ficou mais baixo, entre 27% e 41%. Cabe salientar que estes estudos se centraram na indexação de assuntos realizada, em sua grande parte, por *indexadores profissionais*. Entretanto, ao ser comparada a indexação realizada entre autores e bibliotecários catalogadores-indexadores nesta pesquisa, os índices variaram entre 0% e 21,05%, considerando-se tanto a Fórmula de Rolling quanto a Fórmula de Hooper, para os dois tipos de índices, relaxado e rígido.

De acordo com Hughes e Rafferty (2011, p. 11), a literatura aponta que a consistência na indexação varia consideravelmente, sendo altamente improvável de ser alcançado 100% de consistência, ou seja, 100% de concordância entre os termos atribuídos. Pouco mais de duas décadas antes, Blair (1986, p. 230) salientou que nos estudos realizados até tal publicação verificou-se um maior ou menor nível de consistência, mas em nenhum deles a consistência esteve totalmente ausente ou insignificante. Entretanto, os achados obtidos nesta pesquisa indicam que, na comparação de assuntos atribuídos entre autores e bibliotecários catalogadores-indexadores profissionais, os índices ficam bem abaixo, alcançando um índice de 0% de coincidência na atribuição de termos, ou seja, índice de consistência nulo (totalmente inconsistente), exemplificado na *Área de Ciências Aplicadas* com os cálculos do índice rígido da Fórmula de Rolling e de Hooper, por certo pela inexpressiva quantidade de recursos informacionais analisada, ou seja, apenas 2 recursos informacionais (0,2%) do total.

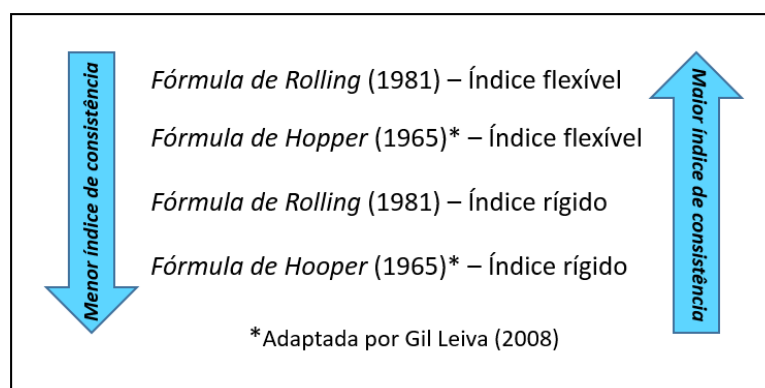
Aqui, retoma-se a perspectiva de Tartarotti (2014, p. 218), para quem, em “qualquer documento indexado, existe um conjunto mais ou menos definível de termos de indexação potenciais que serão realmente selecionados para atribuição de um documento” (TARTAROTTI, 2014, p. 218). Tal premissa, verificada na aplicabilidade de outra abordagem da avaliação da indexação de assuntos - *Avaliação intrínseca quantitativa mediante a*

intraconsistência ou *Avaliação intrabibliotecário* -, é aplicável, até o momento, somente na comparação da indexação entre profissionais utilizando-se a referida abordagem, não entre autores e profissionais, pelo menos nesta pesquisa.

Neste cenário, é pertinente a visão de Tonta (1991), para quem os índices de consistência de diversos estudos devem ser considerados separadamente, considerando-se que os valores de consistência dependem de alguns fatores em que a indexação foi realizada, como, por exemplo, as diferentes fórmulas adotadas. De forma complementar, acredita-se que além da questão das diferentes fórmulas adotadas, como observa a autora, a quantidade de recursos informacionais analisada também interfere diretamente nos resultados dos índices. Logo, a análise de mais recursos informacionais na *Área de Ciências Aplicadas* propiciaria maiores índices de consistência.

Por fim, em relação às fórmulas adotadas, os achados indicam que existe um certo nível de resultados de índices de consistência, de acordo com dois principais aspectos: a fórmula adotada e o tipo de índice, se rígido ou flexível, delineando-se a seguinte escala (Figura 20).

Figura 20. Escala dos índices de consistência



Fonte: Elaboração própria.

Nesse sentido, verifica-se que os maiores índices de consistência são obtidos aplicando-se a Fórmula de Rolling (1981) com o índice flexível, seguida da Fórmula de Hooper (1965) adaptada por Gil Leiva (2008) com o índice flexível, da Fórmula de Rolling (1981) com o índice rígido e, por fim, da Fórmula de Hooper (1965) adaptada por Gil Leiva (2008) com o índice rígido, em um *crescendum* de índices obtidos, dependendo da fórmula e do tipo de índice adotado. Conforme Tartarotti (2014, p. 172), a padronização dos métodos de avaliação da indexação de assuntos no contexto de bibliotecas universitárias possibilita “maior

fidedignidade e comparação dos índices entre os diversos estudos, detectando lacunas e proporcionando uma maior qualidade no processo de indexação”.

Os fatores que levam a tais resultados podem estar relacionados à comparação dos assuntos utilizados pelos autores na atribuição dos recursos informacionais importados para o Repositório Institucional da UNICAMP. Tais assuntos correspondem a palavras-chave livres, ou seja, à linguagem natural atribuída pelos próprios autores nos recursos informacionais. Por outro lado, os mesmos recursos informacionais foram reindexados ou validados pelos bibliotecários catalogadores-indexadores utilizando-se a linguagem padronizada de acordo com as especificidades de cada área do conhecimento.

5.2.2 Resultados da avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do usuário - *Avaliação extrínseca mediante a recuperação*

Os resultados da aplicabilidade da avaliação da indexação de assuntos pela abordagem da *Avaliação extrínseca mediante a recuperação* (avaliação pela perspectiva do usuário) foram norteados por duas principais variáveis: a) dados gerais sobre o Repositório Institucional da UNICAMP e b) entrevistas de busca por assuntos no Repositório Institucional da UNICAMP.

A) *Dados gerais sobre o Repositório Institucional da UNICAMP*

A primeira variável analisada sobre a aplicabilidade da avaliação da indexação de assuntos pela abordagem da *Avaliação extrínseca mediante a recuperação* teve como objetivo realizar um levantamento dos seguintes indicadores⁴⁰, apresentados na opção *busca facetada* disponível na página do Repositório Institucional da UNICAMP: a) total de *assuntos*; b) total de recursos informacionais por *agência de fomento*; c) total de recursos informacionais por *tipo de recurso informacional*; e d) total de recursos informacionais por *data de publicação*; e) total de recursos informacionais por *tipo de acesso*.

Referente ao total de assuntos constantes na *busca facetada*, os quinze primeiros assuntos são: Humans (3930), Male (2910), Female (2741), Animals (2278), Brazil (1834), Adult

⁴⁰ Dados extraídos da página do Repositório Institucional da UNICAMP em Agosto/2019.

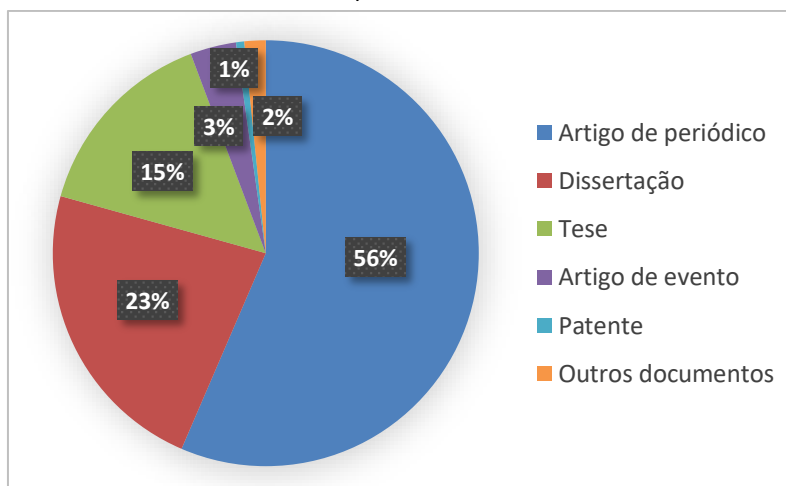
(1699), Middle Aged (1337), Rats (950), Adolescent (890), Aged (817), Educação (737), Child (654), Rats, Wistar (625), Mice (532) e Time Factors (499). No total, há 156.146 entradas diferentes de assuntos no Repositório Institucional da UNICAMP, em todos os tipos de recursos informacionais. Quanto ao total de recursos informacionais por *agência de fomento*, a maior parte dos recursos informacionais financiados foi da FAPESP (9739), CNPq (8766) e CAPES (4093). Todavia, verifica-se diversas entradas tanto para as respectivas agências de fomento quanto para outras instituições. No caso da FAPESP, por exemplo, verifica-se a denominação “Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)”; “FAPESP”, “FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo”, “FAPESP, São Paulo Research Foundation”, dentre outras. O total de entradas de diferentes de agências de fomento, incluindo as três primeiras, é de 14.911.

Referente ao total de recursos informacionais por *tipo de recurso informacional* na *busca facetada*, o Repositório Institucional da UNICAMP possui 143.048⁴¹ registros, sendo: Artigo de periódico (73.024); DISSERTAÇÃO (28.007); TESE (17.585); Artigo (6.853); DISSERTAÇÃO DIGITAL (4.638); Artigo de evento (4.593); TESE DIGITAL (3.768); Patente (891); Artigo de Periódico (745); Editorial (489); Carta (352); research-article (298); Resenha (193); Resumo (186); Capítulo de livro (178); Nota (131); Apresentação (122); Errata (113); Book (25); Notícias (22); Estudos (15); Dossiê (12); Livro (12); Resenhas (12); review article (12); Comentário (10); Relatório (9); Resumo de tese ou dissertação (9); Capítulo de livro (8); Outros recursos informacionais (8); Outros recursos informacionais (seções) (6); rapid-communication (6); Seção livre (6); COMENTARIO (5); Artigo de Segurança Alimentar e N ... (4); Artigos e Ensaios (4); Entrevista (4); Notas (4); Abstracts 93); Artes (3); Author index (3); Artes (3) Author Index CMU (3); arquivo e suas fontes (3); Estudo de caso (3); Notas de Pesquisa (3); Varia (3); carta cometario (2); Dossiê: X Jornada do HISTEDBR- "H... (2); Pareceristas (2); press-release (2); Relato de experiência (2); Atigo (91); Capítulo de e-book (1); Carta ao leitor (1); Case Report (1); Clinical Cases (1); Conceito, corpus e crític revisi ... (1); Conferencia (1); Debate (1); E-book (1); editorial (1); Ensaio (1); formação filosófica (1); leituras (1); Nota dos Editores da Zetetiké (1); Notícia (1); Prefácio (1) e Short Communication (1). Aqui, verifica-se a necessidade de adequação/reunião na descrição dos tipos de recursos

⁴¹ Para se chegar a este valor, basta realizar qualquer pesquisa na página inicial do Repositório Institucional da UNICAMP sem realizar a digitação de nenhuma informação na caixa de busca principal.

informativos nos metadados dos registros no Dublin Core, reunindo recursos informativos semelhantes e padronizando as entradas. O total de recursos informativos é apresentado a seguir (Figura 21):

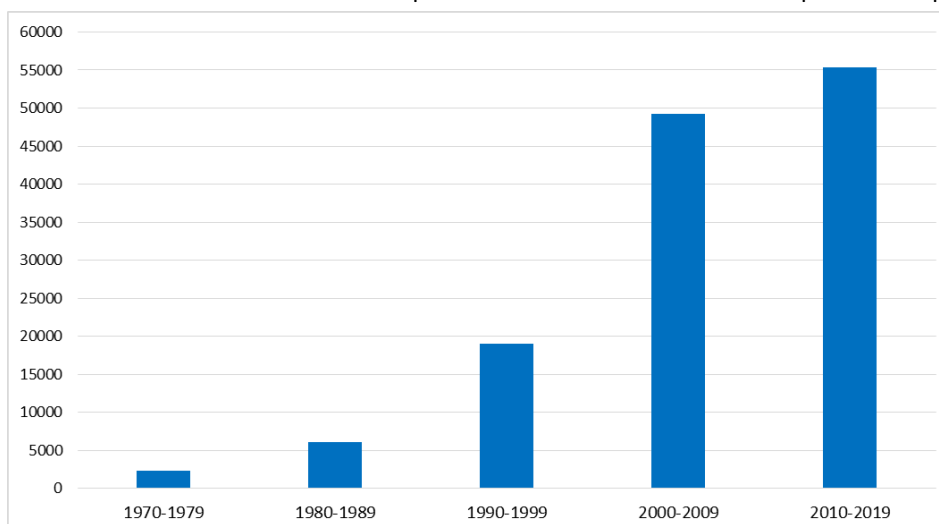
Figura 21. Total de recursos informativos no Repositório Institucional da UNICAMP



Fonte: Elaboração própria.

Quanto ao total de recursos informativos por *data de publicação* na *busca facetada*, há uma maior concentração de recursos informativos produzidos entre 2010 e 2019, com 55.344 registros; 49.203 registros entre 2010 e 2019; 19.023 registros entre 1990-1999; 6051 registros entre 1980 e 1989; 2.324 registros entre 1970 e 1979; e apenas 72 registros entre 1964 e 1969 (Figura 22):

Figura 22. Total de recursos informativos no Repositório Institucional da UNICAMP por data de publicação



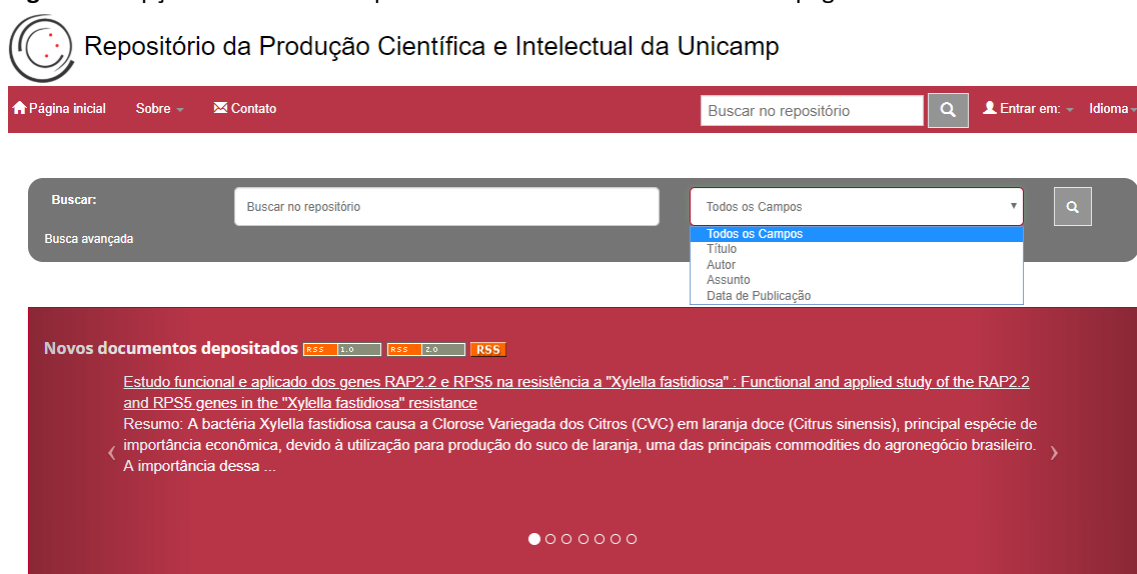
Fonte: Elaboração própria.

Por fim, do total de recursos informacionais por *tipo de acesso* na *busca facetada*, 51487 recursos informacionais possuem acesso fechado, 35.418 de acesso aberto, 2.144 de acesso em período de embargo e 1 recurso informacional possui acesso “Aberto”, totalizando 89.050 recursos informacionais, número diferente do total encontrado anteriormente, indicando que os demais recursos informacionais não possuem esta informação de tipo de acesso na descrição de seus metadados no Dublin Core.

B) Opções de busca no Repositório Institucional da UNICAMP

A variável *Opções de busca* refere-se às possibilidades de busca no Repositório Institucional da UNICAMP disponíveis aos usuários na página inicial (Figura 23).

Figura 23. Opções de busca no Repositório Institucional da UNICAMP na página inicial



Fonte: Dados da pesquisa.

As opções de busca aos usuários são: *Todos os campos*: realiza a busca em todos os campos possíveis nos recursos informacionais; *Título*: realiza a busca somente no campo *Título* dos recursos informacionais; *Autor*: realiza a busca somente no campo *Autor* dos recursos informacionais; *Assunto*: realiza a busca somente no campo *Assunto* dos recursos informacionais; e *Data de publicação*: realiza a busca somente no campo *Data de publicação*. No Repositório da USP, as principais opções são: *Termos de busca* (caixa de pesquisa com a

seguinte informação: *pesquisa por termo ou autor*), *Seleção da base* (Produção científica e teses e dissertações) e *Seleção de Unidade USP* para filtrar a busca, enquanto que no Repositório da UNESP são: *Tipo de Produção, Data do Documento, Autor, Título e Palavras-chave*. Já no Repositório do CRUESP uma única caixa de busca é apresentada, podendo ser realizada nos três repositórios ou individualmente, isto é, em cada um dos três repositórios (USP, UNESP e UNICAMP).

Ao efetuar-se uma determinada *busca por assuntos* no Repositório Institucional da UNICAMP, o usuário é direcionado à próxima página, onde é apresentada a visualização dos recursos informacionais recuperados na seguinte ordem: *Pré-visualização; Data do recurso informacional; Título; Autor(es); Orientador; e Tipo*, não sendo passíveis de ordenação (Figura 24).

Figura 24. Visualização geral dos registros recuperados no Repositório Institucional da UNICAMP

Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

Conjunto de itens:

Pré-visualização	Data do documento	Título	Autor(es)	Orientador	Tipo
	2014	The Concept Of Compensation In The Dialogue Of Vygotsky With Adler: Human Development, Education And Disability [o Conceito De Compensação No Diálogo De Vigotski Com Adler: Desenvolvimento Humano, Educação E Deficiência]	Dainez D.; Smolka A.L.B.	-	Artigo de periódico
	2014	Dysphagia After Anterior Cervical Spine Surgery: A Systematic Review Of Potential Preventative Measures	Joaquim A.F.; Murar J.; Savage J.W.; Patel A.A.	-	Artigo de periódico

Desenvolvimento
 Coordenação de...
 CNPq
 CAPES
 FAPESP - FUN AMPARO À PE
 Fundação de F...
 Tipo de Doc
 Congresso

Fonte: Dados da pesquisa.

Adicionalmente aos registros recuperados com base em um termo de busca por assuntos, na página dos resultados de busca são apresentadas outras opções de filtro, detalhadas a seguir.

Figura 26. Opções de busca no Repositório Institucional da UNICAMP por tipo de recurso informacional

Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em: por Ir Retornar valores

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Tipo de Documento Adicionar

Título
Autor
Orientador
Resumo ou Abstract
Assunto
Data de publicação
Acesso
Departamento
Agência de Fomento
Todos os campos

Ordenar registros por Ordenar

(s) Atualizar

Busca facetada

Assunto	
Políticas públicas	42
Inclusão social	37
Inclusão	34
Educação	33
Qualidade de vida	32
Inclusão digital	27
Brazil	25
próximo >	

Agência de Fomento

Fonte: Dados da pesquisa.

Outro filtro permitido são as opções: *Iguais*, *Contém*, *Identificado*, *Diferentes*, *Não contém* e *Não identificado*, permitindo uma maior especificidade nas buscas, por exemplo, por assuntos (Figura 27).

Figura 27. Opções de busca no Repositório Institucional da UNICAMP por assuntos

Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em: por Ir Retornar valores

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Assunto Adicionar

Resultados/Página Relevância Ordenar

Descendente Atualizar

Busca facetada


Assunto	
Políticas públicas	42
Inclusão social	37
Inclusão	34
Educação	33
Qualidade de vida	32
Inclusão digital	27
Brazil	25
próximo >	

Fonte: Dados da pesquisa.

No exemplo acima, ao selecionar-se os filtros *Assunto* e *Iguais*, verifica-se a apresentação de uma lista de assuntos que compõem a *busca facetada*, apresentada à direita da tela após a realização das buscas. Aqui, o termo *Políticas públicas* aparece com 42 registros

vinculados, tanto na lista quanto na *busca facetada*, sendo o primeiro termo apresentado. Ao selecionar-se a opção *Políticas públicas* no filtro e selecionar-se o botão *Adicionar*, ocorre um refinamento da busca inicial do termo *Inclusão* para 42 resultados. Sem a opção de filtro, havia 3267 registros recuperados (Figura 28).

Figura 28. Opções de busca no Repositório Institucional da UNICAMP por assunto com o recurso autocompletar



Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em:

por

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Assunto

Resultados/Página | Ordenar registros por

Descendente

Políticas públicas (42)
Política educacional (6)
Política de saúde (5)
Polimerização (4)
Política social (4)

Busca facetada

Assunto	Contagem
Políticas públicas	42
Inclusão social	37
Inclusão	34
Educação	33
Qualidade de vida	32
Inclusão digital	27
Brazil	25

[próximo >](#)

Fonte: Dados da pesquisa.

Tal número de registros recuperados corresponde aos mesmos registros apresentados ao selecionar-se o assunto *Políticas públicas* (42) na *busca facetada*. Nesse caso, o sistema adiciona automaticamente o “Assunto” ao filtro (Figura 29).

Figura 29. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Assunto



Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em:

por

Filtros correntes: Assunto

X

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Tipo de Documento

Busca facetada

Assunto	Contagem
Public policies	8
Educação	6
Educação especial	5
Inclusão digital	5
Inclusão	4
Inclusão social	4
Acesso aos serviços de saúde	2

[próximo >](#)

Fonte: Dados da pesquisa.

Em continuidade, verifica-se a existência da opção *Resultados/Página*, que filtra a quantidade de resultados a serem visualizados por página, com uma variação de 10 a 100, e a opção *Ordenar registros por*, com as subopções: *Relevância*⁴², *Data de Publicação*, *Título*, *Autor* e *Orientador* (Figura 30).

Figura 30. Opções de busca no Repositório Institucional da UNICAMP por ordenação de registros - A

Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em: **Todo o repositório**

por **Inclusão** Ir Retornar valores

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Tipo de Documento | Iguais | Adicionar

Resultados/Página: 10 | Ordenar registros por: Relevância | Ordenar

Descendente | Registro(s) | Todos | Atualizar

Relevância
Data de Publicação
Título
Autor
Orientador
???search.sort-by.dc.contributor.department_sort???

Resultado 1-10 de 3267.

Busca facetada

Assunto	
Políticas públicas	42
Inclusão social	37
Inclusão	34
Educação	33
Qualidade de vida	32
Inclusão digital	27
Brazil	25
próximo >	

Agência de Fomento

Fonte: Dados da pesquisa.

De forma complementar, há ainda a opção de ordenação *Descendente* e *Crescente*, podendo ser aplicada em *Todos* os registros ou em quantidade específica, que varia de 1 a 50 (Figura 31).

⁴² Cabe esclarecer que, embora exista tal opção de ordenação no Repositório Institucional da UNICAMP, que tem por base o software DSPACE, na prática verifica-se a inexistência de tal critério.

Figura 31. Opções de busca no Repositório Institucional da UNICAMP por ordenação de registros - B



Repositório da Produção Científica e Intelectual da Unicamp

Página de Busca

Buscar em: **Todo o repositório**

por **Inclusão** Ir Retornar valores

Adicionar filtros:
Utilizar filtros para refinar o resultado de busca.

Tipo de Documento **Iguais** Adicionar

Resultados/Página **10** Ordenar registros por **Relevância** Ordenar

Descendente Registro(s) Todos Atualizar

Busca facetada

Assunto

Políticas públicas	42
Inclusão social	37
Inclusão	34
Educação	33
Qualidade de vida	32
Inclusão digital	27
Brazil	25

próximo >

Fonte: Dados da pesquisa.

Outros filtros na *busca facetada*, localizada à direita da tela, incluem: *Agência de Fomento*, *Tipo de recurso informacional*, *Data de Publicação* e *Tipo de Acesso* (Figuras 32, 33 e 34):

Figura 32. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Agência de fomento



Repositório da Produção Científica e Intelectual da Unicamp

Resultado 1-10 de 3267.

Anterior **1** 2 3 4 ... 327 Próximo

Conjunto de itens:

Pré-visualização	Data do documento	Título	Autor(es)	Orientador	Tipo
	2003	Inclusão escolar e a identidade do professor : a escola como palco de invenção	Lima, Norma Silvia Trindade de, 1960-	Mantoan, Maria Teresa Eglér, 1943-	TESE
	2003	Inclusão social e digital : o uso da Internet como complemento da aula presencial de língua estrangeira	Pudo, Paula Barbosa	Schmitz, John Robert, 1935-	DISSERTAÇÃO

Agência de Fomento

Conselho Nacional de Desenvolvi...	113
Fundação de Amparo à Pesquisa do ...	106
Coordenação de Aperfeiçoamento de...	49
III	4
FAPESP - FUNDAÇÃO DE AMPARO À PES...	3
Fundação para o Desenvolvimento d...	3
CAPES, Conselho Nacional de Desenv...	1

próximo >

Fonte: Dados da pesquisa.

Ainda na *busca facetada*, na opção de filtro *Tipo de recurso informacional* são apresentados os tipos de recursos informacionais que compõem o Repositório Institucional da UNICAMP.

Figura 33. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Tipo de recurso informacional



Repositório da Produção Científica e Intelectual da Unicamp

Buscar no repositório

Year	Title	Author	Year	Document Type
2003	Bolsões Fechamentos E Cia	Eni Puccinelli Oriandi	-	Estudos
2003	A educação pelo outro : Lorelai, uma experiencia de inclusão	Laurindo, Tania Regina	Ayoub, Eliana, 1966-	DISSERTAÇÃO
2015	Inclusão e segmentação social no Ensino Superior público no Estado de São Paulo (1990-2012)	Ana Maria F.; Almeida; Mauricio; Ernica	-	Artigo de periódico
1998	Compostos de inclusão Beta-ciclodextrina com convidadados organicos e organometalicos	Manhães, Giacomia Frasson	Alves, Oswaldo Luiz, 1947-	DISSERTAÇÃO
1994	Estudo da inclusão molecular de produtos de interesse farmacologico utilizando ciclodextrinas.	Micaroni, Antonio	Jockes, Ines	DISSERTAÇÃO
2003	Mais falares sobre a inclusão - diferenças ou	Xavier, Evelise Cristina Couto	Souza, Regina Maria de, 1977.	DISSERTAÇÃO


Tipo de Documento	
Artigo de periódico	2117
DISSERTAÇÃO	488
TESE	324
TESE DIGITAL	125
DISSERTAÇÃO DIGITAL	102
Artigo	47
Artigo de evento	28
Patente	13
Editorial	7
research-article	7

próximo >

Fonte: Dados da pesquisa.

A busca facetada por *Data de Publicação* possibilita ao usuário o filtro nos seguintes períodos: 1964-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019. Já o *Tipo de Acesso* permite o filtro por 3 principais tipos: *aberto*: recursos informacionais com acesso na íntegra; *fechado*: acesso não permitido; e *embargo*: acesso em período transitório entre o fechado e o aberto.

Figura 34. Opções de busca no Repositório Institucional da UNICAMP por busca facetada: Data de publicação e Tipo de acesso



Repositório da Produção Científica e Intelectual da Unicamp

Buscar no repositório

Year	Title	Author	Year	Document Type
2016	O esporte nas políticas públicas de inclusão social para pessoas com deficiência, no Brasil = Sport in public policies of social inclusion for people with disabilities in Brazil	Pancotto, Heloisa Pereira, 1991-	Fuentes-Rojas, Marta, 1957-	DISSERTAÇÃO DIGITAL

Data de Publicação	
2010 - 2019	1530
2000 - 2009	1502
1990 - 1999	196
1980 - 1989	25
1970 - 1979	13

Tipo de Acesso	
aberto	2095
fechado	133

Fonte: Dados da pesquisa.

Tanto as opções de busca quanto os filtros apresentados no Repositório Institucional da UNICAMP possibilitam um maior refinamento nos resultados aos usuários, permitindo a encontrabilidade dos recursos informacionais mais relevantes para suas pesquisas. Entretanto, verifica-se notadamente a necessidade de adequações e melhorias em duas principais vertentes: tanto na ferramenta utilizada, ou seja, no DSpace, quanto na *qualidade* da descrição dos metadados tanto descritivos quanto temáticos.

C) Entrevistas de busca por assuntos no Repositório Institucional da UNICAMP

Tal como realizado nas entrevistas de diagnóstico organizacional, para a análise das coletas foram preservadas as identidades dos participantes por meio de siglas específicas, de acordo com a Biblioteca a qual pertencem e à categoria acadêmica, conforme delineado (Quadro 9):

Quadro 9. Participantes da pesquisa e respectivas siglas para identificação e análise da Avaliação extrínseca mediante a recuperação

Biblioteca participante	Categoria	Sigla
Faculdade de Educação (FE)	Discente de graduação	G-FE
	Discente de pós-graduação (mestrado)	PGM-FE
	Discente de pós-graduação (doutorado)	PGD-FE
	Docente	D-FE
Instituto de Filosofia e Ciências Humanas (IFCH)	Discente de graduação	G-IFCH
	Discente de pós-graduação (mestrado)	PGM-IFCH
	Discente de pós-graduação (doutorado)	PGD-IFCH
	Docente	D-IFCH
Faculdade de Ciências Médicas (FCM)	Discente de graduação	G-FCM
	Discente de pós-graduação (mestrado)	PGM-FCM
	Discente de pós-graduação (doutorado)	PGD-FCM
	Docente	D-FCM
Faculdade de Educação Física (FEF)	Discente de graduação	G-FEF

	Discente de pós-graduação (mestrado)	PGM-FEF
	Discente de pós-graduação (doutorado)	PGD-FEF
	Docente	D-FEF
Faculdade de Ciências Aplicadas (FCA)	Discente de graduação	G-FCA
	Discente de pós-graduação (mestrado)	PGM-FCA
	Discente de pós-graduação (doutorado)	PGD-FCA
	Docente	D-FCA
Instituto de Física Gleb Wataghin (IFGW)	Discente de graduação	G-IFGW
	Discente de pós-graduação (mestrado)	PGM-IFGW
	Discente de pós-graduação (doutorado)	PGD-IFGW
	Docente	D-IFGW
Instituto de Matemática, Estatística e Computação Científica (IMECC)	Discente de graduação	G-IMECC
	Discente de pós-graduação (mestrado)	PGM-IMECC
	Discente de pós-graduação (doutorado)	PGD-IMECC
	Docente	D-IMECC
Biblioteca da Área de Engenharia e Arquitetura (BAE)	Discente de graduação	G-BAE
	Discente de pós-graduação (mestrado)	PGM-BAE
	Discente de pós-graduação (doutorado)	PGD-BAE
	Docente	D-BAE
Faculdade de Engenharia de Alimentos (FEA)	Discente de graduação	G-FEA
	Discente de pós-graduação (mestrado)	PGM-FEA
	Discente de pós-graduação (doutorado)	PGD-FEA
	Docente	D-FEA

Fonte: Elaboração própria.

A partir dos referenciais teóricos, dos objetivos da pesquisa e das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP, a seguir são apresentadas as categorias de análise elaboradas como subsídios para a análise quantitativa. Cabe esclarecer que, embora o ponto principal das coletas de dados tivesse sido a determinação do índice de

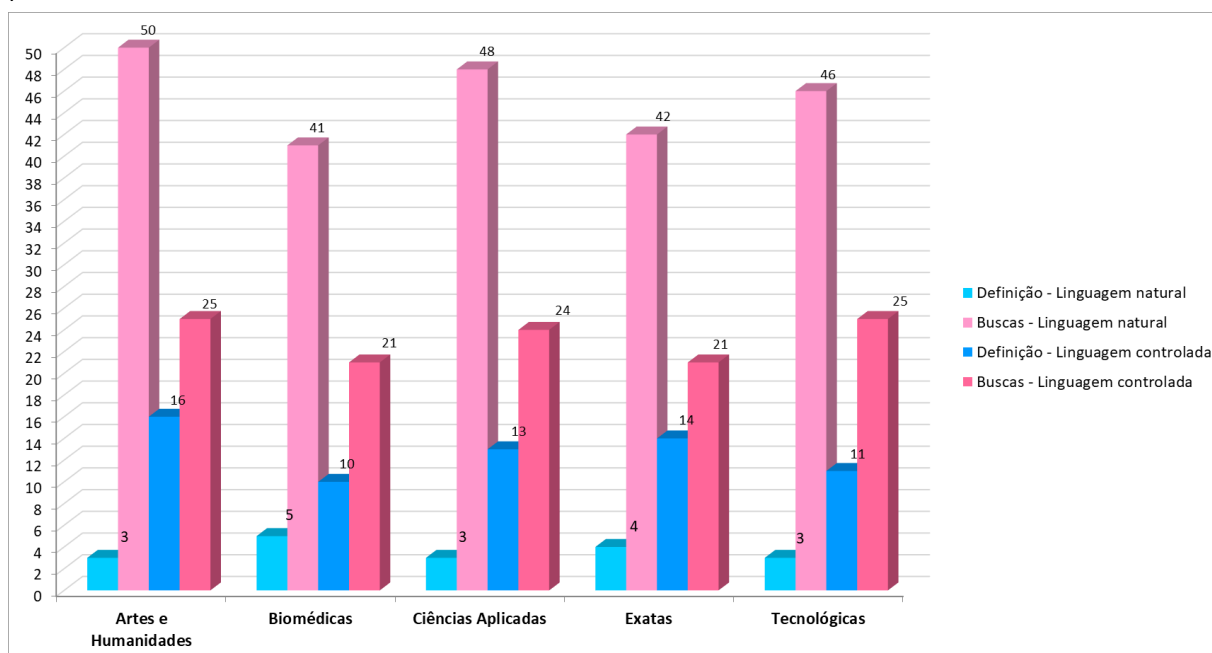
precisão dos recursos informacionais recuperados e posterior comparação dos índices entre as buscas utilizando-se linguagem natural e linguagem controlada no Repositório Institucional da UNICAMP, de forma complementar percebeu-se a necessidade de inclusão de trechos significativos das entrevistas de busca por assuntos com os usuários, levantados por meio da transcrição das falas dos participantes.

Posto isto, com base nos dados levantados foram elaboradas 10 categorias de análise para as entrevistas de busca e recuperação por assuntos no Repositório Institucional da UNICAMP: 1) *Tempo de realização da entrevista de busca por assuntos*; 2) *Conhecimento prévio sobre o repositório institucional*; 3) *Definição da linguagem natural*; 4) *Estratégias de busca por assuntos*; 5) *Avaliação da relevância - Linguagem natural*; 6) *Definição da linguagem controlada*; 7) *Avaliação da relevância - Linguagem controlada*; 8) *Determinação do índice de precisão*; 9) *Representação descritiva*; e 10) *Comentários adicionais sobre o repositório institucional*, detalhadas a seguir.

1. Tempo de realização da entrevista de busca por assuntos

A categoria “Tempo de realização da entrevista de busca por assuntos” refere-se ao tempo utilizado em cada uma das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP (APÊNDICE G), sendo divididos em quatro períodos: a) definição das palavras-chave (linguagem natural); b) busca por assuntos utilizando-se as palavras-chave (linguagem natural); c) definição dos novos termos (linguagem controlada); e d) busca por assuntos utilizando-se os novos termos (linguagem controlada) (Figura 35):

Figura 35. Tempo de duração das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP por área do conhecimento



Fonte: Elaboração própria.

Na definição das palavras-chave (linguagem natural) pelos usuários, o maior tempo demandado foi da *Área de Biomédicas*, com uma média de 5 minutos, seguida da *Área de Exatas*, com 4 minutos e da *Área de Artes e Humanidades* e *Exatas* e *Tecnológicas*, com 3 minutos cada. Em relação às buscas por assuntos utilizando-se as palavras-chave (linguagem natural), ou seja, à primeira rodada de buscas, a *Área de Artes e Humanidades* aparece com uma média de 50 minutos, seguida da *Área de Ciências Aplicadas*, com 48 minutos, da *Área de Tecnológicas*, com 46 minutos, da *Área de Exatas*, com 42 minutos e da *Área de Biomédicas*, com 41 minutos.

Já na definição dos novos termos (linguagem controlada) com base na busca das palavras-chave em linguagens padronizadas/tesauros, a *Área de Artes e Humanidades* aparece com uma média de 16 minutos, seguida da *Área de Exatas*, com 14 minutos, da *Área de Ciências Aplicadas*, com 13 minutos, da *Área de Tecnológicas* com 11 minutos e da *Área de Biomédicas* com 10 minutos. Referente às buscas por assuntos utilizando-se os novos termos (linguagem controlada), ou seja, à segunda rodada de buscas, a *Área de Artes e Humanidades* e a *Área de Tecnológicas* aparecem com 25 minutos cada, seguidas da *Área de Ciências Aplicadas*, com 24 minutos. Em seguida, estão as *Áreas de Biomédicas* e *Exatas*, com 21 minutos cada.

2. Conhecimento prévio sobre o repositório institucional

A categoria “Conhecimento prévio sobre o repositório institucional” refere-se à familiarização do usuário com o Repositório Institucional da UNICAMP antes da realização das entrevistas de busca por assuntos. Ao serem indagados sobre esta questão, verifica-se a categorização em três grupos: a) usuários que conhecem e utilizam; b) usuários que conhecem mas não utilizam; e c) usuários que não conhecem e por consequência não utilizam o Repositório Institucional da UNICAMP.

Dos 36 participantes, 14 (39%) conhecem e utilizam o Repositório Institucional da UNICAMP, alguns de forma mais frequente que outros, sendo 6 da *Área de Tecnológicas*, 5 da *Área de Biomédicas*, 3 da *Área de Artes e Humanidades* e 1 da *Área de Ciências Aplicadas*, sendo a maior parte alunos de doutorado, seguida de alunos de mestrado, docentes e alunos de graduação. Na *Área de Exatas*, nenhum usuário manifestou o conhecimento e utilização.

Conheço. Por causa de uma palestra que deram na Faculdade de Educação. (...) Mas não sei mexer muito também. (Quem deu a palestra? Foi algum docente?) *Não. Foi um ... trabalhador da Biblioteca. (...) Ele deu uma palestra na aula. (Aí ele mostrou os recursos da Biblioteca, de forma geral?) Mostrou. Todos os sites que a Biblioteca tinha, o SBU, o SophiA ... é ... mas tudo muito rápido, assim, então a gente acaba não ... sabendo mexer muito.* (Isso foi no primeiro ano?) *Foi esse ano. Mas no primeiro ano a Biblioteca também dá uma palestra na Faculdade de Educação, pra todos os ingressantes, né (G-FE)*

(Você conhece o Repositório?) *Da UNICAMP? Sim. (Já fez alguma pesquisa?) Já. (...) Eu não uso muito mas já fiz (PGD-FE)*

Sim, conheço. (Já fez alguma pesquisa nele?) Já. (Do seu interesse mesmo?) Eu às vezes pesquiso pra mostrar pra aluno buscar referências, principalmente pra buscar tese, dissertação ... eu acho um jeito ... mais agregado assim pros alunos ... os papers né (D-IFCH)

Conheço, usei algumas vezes no começo pra levantar, enfim, referencial, mas agora eu confesso que faz um tempo que eu não uso mais ((R1)) (PGD-FCM)

Sim. (Conhece? Já fez alguma pesquisa nele?) *Já, sempre eu faço pesquisas nele (PGM-FF)*

Sim (PGD-FF)

Conheço. (Você chegou a fazer alguma busca por assuntos?) *Sim, esses assuntos, a gente sempre faz, meus alunos também fazem. O último que eu fiz foi no final do ano passado ... mas pra mim foi normal (D-FF)*

Conheci enquanto ele não era tão estranho. (...) Acho ele péssimo ((R1)). Com mais 80% do pessoal da graduação da FCA ((R1)) (PGM-FCA)

É, eu já usei. (...) *Na metodologia científica, uma disciplina da Engenharia Civil. (Alguém apresentou o repositório pra você?) Não, a gente só ... achou ele ((R1)). (Ah, tá, não foi o catálogo não, foi o repositório mesmo.) ... Foi, eu acho que sim, faz tempo já que eu fiz, eu fiz no 2º ano (G-BAE)*

Conheço, conheço. (Já procurou alguma vez aqui?) Já, já. É que assim, né, eu fiz alguns levantamentos de bases de dados em geral só que assim, na hora você acaba indo no mais rápido, né (PGM-BAE)

Já, já conheço, mas eu uso pouco. (Então você já usou alguma vez?) Já, já usei (PGD-BAE)

Usei algumas vezes. Eu uso mais pra busca de trabalhos de conclusão, teses. (E artigos?) ... Não, no repositório não (D-BAE)

Usei, mas já faz um tempinho que eu não to usando. Não sei se mudou (PGD-FEA)

Eu to pensando ... porque mudou o site da Biblioteca, né ... Não sei se é o mesmo que antigamente dava pra fazer busca das teses digitalizadas ... ou era outro? (...) ...~ É ... no dia da palestra do Brito eu entrei nessa página ... e fiz uma busca, joguei qualquer coisa aí apareceu alguns artigos. (...) Foi só isso, depois eu tentei logar ... porque até então eu desconhecia como ... pra eu inserir algum artigo meu se precisava logar ou era via Biblioteca, na época eu não entendia direito (D-FEA)

Em relação aos usuários que expressam conhecer mas não utilizam o Repositório Institucional da UNICAMP, verifica-se um total de 8 (22,22%) participantes, sendo 4 da Área de Exatas, 2 da Área de Biomédicas e 1 da Área de Artes e Humanidades, sendo a maior parte docentes, seguida de alunos de doutorado, alunos de graduação e nenhum aluno de mestrado. Quanto às áreas, não houve resultados na Área de Ciências Aplicadas e na Área de Exatas.

Conheço. Você já realizou alguma busca nele? Ah, eu não ((RI)) (D-FE)

Conheço. (Você já fez pesquisa no repositório antes?) Não, normalmente eu faço nas bases de dados. No repositório eu nunca fiz pesquisa não, eu sempre faço nas base de dados (PGD-IFCH)

Que é aquele site de pesquisa ou não? (...) Ah, eu acho que eu conheço. É um que a gente teve uma aula com uma bibliotecária que foi lá na minha turma, faz um tempinho mas ... ela mostrou todas as ferramentas do site [do SBU] (G-FCM)

Sim. Ah, o repositório da UNICAMP eu não utilizo ... (Não? Mas conhece? Já ouviu falar ...). Conheço. Se tiver uma parte qualitativa aí eu diria que eu não gosto muito dele. Porque eu acho que ele duplica os trabalhos. É, porque você, em tese, encontraria nos indexadores ... é ... consagrados, né, a produção da UNICAMP. (Na Web of Science, na Scopus ...?) Isso. (...) É que eu acho um esforço institucional muito grande, né. (...) A gente já tem o repositório da FAPESP ... é, grupos de estudo do CNPq ... Currículo Lattes ... PubMed, Web of Science, Scopus ... LILACS ... EMBASE ... várias bases. No caso eu tenho um olhar porque eu participo, é, de algumas discussões do Órgão Colegiado, enquanto coordenadora docente aqui, né. (...) Então eu sei que tem uma discussão sobre ... o quanto esse repositório exige ... do esforço dos profissionais, principalmente da área de Biblioteca, né. (Agora está sendo prioridade ...). Isso, e o quanto ele ... termina por ... duplicar o esforço, né. Porque já está tão ... é ... disseminado nessas diferentes bases ... (...) (É, porque assim, o objetivo do repositório é reunir em um único lugar toda a produção da UNICAMP). ... Hoje como a gente tem acesso ... aos recursos da web, né. (Isso). É ... você acaba achando as coisas, né ((RI)). ((RI)). (É, é que em algumas áreas, por exemplo, na Humanas ... é ... a produção geralmente não é realizada nessas bases de dados). Isso, exatamente (D-FCM)

É o Acervus? (Não, é outro, outro sistema). Talvez eu nunca tenha usado. Não? Ou talvez tenha usado ((RI)). Na verdade eu liio muitas coisas que tem nesse repositório ... mas é que eu nunca entrei de fato no repositório. (Você sabe que existe então?) Sim (G-IFGW)

Conheço. (Já usa?) Não costumo usar, mas eu conheço (PGD-IFGW)

Então, eu conheço mas eu nunca usei, assim ((RI)), eu conheço por causa das atividades aqui da BIF, mas, assim, eu realmente usar o repositório pra busca alguma coisa que eu preciso eu nunca usei, tá (D-IFGW)

Às vezes eu uso mas sem saber (...) ~... Não fiz buscas (D-IMECC)

Todavia, o número de participantes que expressa desconhecer este sistema de recuperação no âmbito da universidade à qual estão vinculados foi de 14 (38,9%), sendo 4 da *Área de Exatas*, 3 da *Área de Artes e Humanidades*, *Área de Ciências Aplicadas* e *Área de Tecnológicas* e 2 da *Área de Biomédicas*, sendo a maior parte alunos de mestrado, seguida de alunos de graduação, alunos de doutorado e docente.

Eu acho que sim, o Acervus, né. (O Acervus é o catálogo). Tá. (O Repositório só tem documentos que foram produzidos pela universidade). Ah, tá, então esse eu não conheço (PGM-FE)

Não. Nunca usei na vida ((RI)) (G-IFCH)

O Acervus? (Não. O Acervus é o catálogo. O repositório é um outro sistema que só tem reunida a produção da UNICAMP) (PGM-IFCH)

Esse eu não uso, só o SophiA (PGM-FCM)

*Não. (Você conhece a página do SBU?) *Sim. Da Biblioteca?* (Isso) (G-FEF)*

*Não. (Nunca pesquisou nele ... já ouviu falar?) *Não* (G-FCA)*

*Não, costumo utilizar o da CAPES. (O Portal de Periódicos da CAPES?) *Isso, exatamente. Porque já pega a Web of Science e tal e eu não olhei esse da Biblioteca na verdade* (PGD-FCA)*

O repositório que tua falas é aquele que a gente procura livros ... ou ele é algum específico? (Ele é específico). Não (D-FCA)

É o Acervus? (Não, tem o Acervus e tem o repositório). Então eu não conheço (PGM-IFGW)

(Você conhece o repositório da UNICAMP?) Não, acho que não (G-IMECC)

Não, vamos ver se me interessa (PGM-IMECC)

Não conheço (PGD-IMECC)

*Eu já acessei vários artigos que são do repositório mas assim, entrar direto no repositório e ... procurar por ele não. (Não?) *Tipo, eu geralmente vou pelo ... Google Scholar, o ... Google Acadêmico* (G-FEA)*

Na verdade não (PGM-FEA)

Em síntese, os alunos de doutorado representam a categoria de usuários que mais conhece e utiliza o Repositório Institucional da UNICAMP, em sua maioria na *Área de Tecnológicas*. Por outro lado, os docentes representam a categoria de usuários que mais conhece mas não utiliza e os alunos de mestrado a que não conhece e consequentemente não utiliza.

3. Definição da linguagem natural

A categoria “Definição da linguagem natural” refere-se à definição das palavras-chave pelos usuários, de acordo com suas respectivas pesquisas em andamento. Cabe esclarecer que não foi solicitado anteriormente aos mesmos que elencassem as palavras-chave de suas pesquisas ou enviassem resumos das mesmas à pesquisadora, sendo solicitadas apenas no momento da entrevista de busca por assuntos no Repositório Institucional da UNICAMP.

Os resultados desta categoria possibilitaram a categorização de três grupos de usuários: a) usuários que elencaram todas as palavras-chave sem dificuldade ou consulta a outras fontes; b) usuários que manifestaram dificuldade e/ou consultaram outras fontes; e c) usuários que solicitaram a alteração das palavras-chave durante as buscas por assuntos. Verifica-se que 12 usuários (33%) elencaram todas as palavras-chave sem dificuldade aparente ou consulta a outras fontes (G-IFCH, D-IFCH, PGD-FCM, D-FCM, PGD-FEF, PGM-FCA, D-FCA, PGD-IFGW, D-IFGW, PGD-BAE, D-BAE e PGM-FEA).

(Você está pesquisando sobre o que?) *É ... VIOLÊNCIA POLICIAL CONTRA MENORES INFRATORES. (Você já entregou um relatório?) Já entreguei o relatório parcial e falta entregar o relatório final agora. (No relatório você fez um resumo, não fez?) Sim. (E no relatório você não colocou palavras-chave?) Sim. (Então ... são essas palavras-chave que você vai usar pra fazer as buscas no repositório. (...) ...~ Aqui você vai elencar as 5 palavras-chave que mais representam a sua pesquisa). Tá. (E a versão dessas palavras-chave em inglês). Tá bom. Duas delas eram expressões maiores, né, tipo SEGURANÇA PÚBLICA aí nas palavras-chave a gente sempre coloca as duas ... (..) VIOLÊNCIA, RACISMO, SEGURANÇA PÚBLICA, POLÍCIA e JUSTIÇA CRIMINAL (G-IFCH)*

(Eu vou pedir pra você pensar em 5 palavras-chave que você usaria pra fazer uma busca sobre o seu tema de pesquisa e a mesma versão das palavras em inglês). ...~ (...) (Quais seriam?) PENSAMENTO SOCIAL, GÊNERO, RURAL, CLASSE e TEORIA SOCIAL (D-IFCH)

Tá ótimo então. (...) SAÚDE DO TRABALHADOR, SILICOSE, PNEUMOCONIOSE, MINERAÇÃO, FADIGA (PGD-FCM)

Tá, HIV, CRIANÇA, ADOLESCENTE, TERAPIA ANTIRRETROVIRAL e VACINAS, RESPOSTA (D-FCM)

É ...~ ENSINO e APRENDIZAGEM eu posso colocar como uma só, ENSINO E APRENDIZAGEM? ENSINO/APRENDIZAGEM? Por que na verdade a minha pesquisa é ... (Os dois juntos?) A relação ENSINO/APRENDIZAGEM. (...) (Nas buscas você costuma usar juntas?) Uso juntas. Se eu puser por exemplo ENSINO, é muito amplo. Eu vou sair com uma pancada de teoria sobre ENSINO que não me interessa. Se eu coloco APRENDIZAGEM outro ... então nada disso me interessa, só me interessa a relação entre os dois. (Então a gente faz essa busca, como se você fosse fazer uma busca mesmo do jeito que você costuma buscar). ...~ É a mesma coisa por exemplo ESCOLAS DE CIRCO. ... (..) Se eu puser ESCOLA não me interessa ESCOLA. ...~ Não sei se ... faria diferença na busca a expressão, eu tenho esses dois eixos principais ... CIRCO e RUA que desemboca nesse e CIRCO DE RUA. ... Aí eu não sei se eu puser a 5ª como CIRCO DE RUA se vai fazer diferença, por exemplo, pro Google Scholar não faz. ... Se eu puser CIRCO e RUA ele vai buscar separado. (A gente pode tentar). Pra ver se vai dar diferença, né? ... Porque isso é uma coisa importante,

né, o Google Scholar é ineficiente nesse ponto ((RI)). (...) *É, o Google Scholar tem lá as aspas e tem o exceto, por exemplo eu quero buscar CIRCO mas não quero RUA, pelo menos o ... funciona* ((RI)). (É, no repositório nem tem essa opção ainda) ((RI)). ...~~ *Muitos trabalhos usam ... junto, eu tenho encontrado bastante ...* (...) (Mas geralmente coloca ENSINO E APRENDIZAGEM?) ENSINO/APRENDIZAGEM (...) CIRCO, RUA, ENSINO/APRENDIZAGEM, ESCOLA DE CIRCO e CIRCO DE RUA **(PGD-FEF)**

(Você já começou a escrever sua pesquisa?) ... *Já, já tenho artigo publicado, já to bem encaminhado no meu mestrado.* (Ah, tá ótimo. Então aqui a primeira parte você vai elencar pelo menos 5 palavras-chave da sua pesquisa que você usaria pra fazer uma busca qualquer, pra fazer um levantamento bibliográfico do tema da sua pesquisa). *Certo.* (E aí você elenca aqui as 5 palavras e a mesma versão em inglês dessas palavras, tá). *Ok.* ...~~ (Quais palavras foram?) *É ...* LIPASE, IMOBILIZAÇÃO, BURKHOLDERIA, ADSORÇÃO e BIOLUBRIFICANTE **(PGM-FCA)**

SECAGEM, BAMBU, BAGAÇO, CANA-DE-AÇÚCAR e ETANOL **(D-FCA)**

FÉRMIONS PESADOS, FASE NEUMÁTICA, SUPERCONDUTIVIDADE, PARES DE COOPER, TEORIA DE LANDAU **(PGD-IFGW)**

Mas é pra eu procurar o que, os meus próprios artigos, artigos de colega, ou qualquer coisa aleatória? (...) (Então aqui você vai elencar 5 palavras-chave em português e em inglês). *Tá.* ...~~ *Eu escrevo as palavras primeiro antes de fazer a busca, é isso?* (É). *Tá, eu vou pela minha área que fica talvez mais fácil.* *É.* ...~~ *Siq̃la pode?* (Pode). *Porque tem uma definição de uma teoria que chama QCD, CROMODINÂMICA QUÂNTICA.* (Coloque do jeito que ... é normal pra você). *Tá.* (Porque de qualquer forma depois vamos consultar em uma linguagem controlada e aí se indicar que é a sigla, não vamos refazer a busca). ... *Então eu vou colocar aqui a sigla, aí a tradução dela e depois em inglês, pode ser?* (Pode). ...~~ *Aí eu faço as 5 de uma vez, é isso né.* (É, já elenca as 5). ...~~ *Isso daqui, supostamente tem que ter isso por causa que tem pesquisador que faz ... coisas nessa área.* (...) QCD, FATOR DE FORMA, SUPERSIMETRIA, NEUTRINO, ESTRANHEZA **(D-IFGW)**

Tá. (Na qualificação você fez o resumo?). *Já.* (Você utilizou algumas palavras-chave? ...) Então pode utilizar essas mesmas palavras, a versão delas em português e em inglês das mesmas palavras. É com base nelas que a gente vai fazer as buscas no repositório). ...~~ (Tá, quais foram?) GASEIFICAÇÃO, SYNGAS, FISHER-TROPICH, TERMOQUÍMICA e CATALISADOR. (Esse FISHER não tem tradução pro português, é assim mesmo?) *Não, é o nome de um processo* **(PGD-BAE)**

Eu tenho praticamente três linhas, né, de trabalho. (...) (Escolhe uma delas). *É, né. Seria a parte de madeira laminada colada. Depois dela derivam as partes que estão mais voltadas ao trabalho, né, poderia ser até mais amplo, derivados de madeira onde madeira laminada colada está dentro, compósitos de madeira.* (...) *Porque ... dentro de madeira laminada colada já é um campo, né, depois ele está inserido em COMPÓSITOS DE MADEIRA.* (A gente pode tentar COMPÓSITOS DE MADEIRA também. COMPÓSITOS ... fica bem mais amplo.) *É ... se você quiser colocar um mais geral e um mais específico ...* na segunda parte a gente vai consultar essas palavras-chave em um vocabulário controlado, em um tesouro e se tiver alguma diferença com o mesmo sentido vamos refazer a busca no repositório). (...) ...~~ COMPÓSITOS DE MADEIRA, MADEIRA LAMINADA COLADA, MADEIRA LAMINADA ... REFORÇADA, PROPRIEDADES MECÂNICAS ... BARRAS DE AÇO **(D-BAE)**

SOBERANIA ALIMENTAR, SEGURANÇA ALIMENTAR, EMERGIA, MUDANÇAS CLIMÁTICAS e AGROECOLOGIA. (Então com base nessas palavras você vai fazer a busca no repositório). *Tá.* (No caso de alguns termos compostos você vai fazer a busca tanto sem aspas como utilizando aspas). *Tá.* (Porque costuma dar diferença) **(PGM-FEA)**

Em outros casos, a questão apresentada por um usuário é sobre a mudança da grafia na palavra, podendo ser tanto buscada/recuperada de mais de uma maneira (PGD-FE).

(E aí eu vou pedir pra você definir 5 palavras-chave em português e em inglês que representam o seu trabalho). *Tá.* (Que você usaria pra fazer uma busca, ou ... do seu resumo, tá. Você já deve ter ...). *Ah, sim. É o que a gente mais faz, né.* (...) ...~~~ *Aqui eu coloquei tudo no plural.* (Não tem problema). *Eu posso colocar ... porque às vezes eu procuro ALUNOS, ALUNO, PROFESSORES, PROFESSOR, que às vezes a gente encontra, né, coisas diferentes, ESTRATÉGIAS DE APRENDIZAGEM, ESTRATÉGIA DE APRENDIZAGEM.* (Não tem problema, porque depois na segunda parte o que vamos fazer? Pesquisar essas palavras-chave em um vocabulário controlado, um tesouro). *Tá.* (E se for igual não vamos refazer a pesquisa. Se for diferente, pro mesmo assunto). *Tá.* (Aí vamos refazer a pesquisa no repositório). *Ah, tá bom, então mantém assim, né.* (Isso, mantém do jeito que você ... achar melhor). ...~~~ *E aqui também, olha, às vezes eu coloco ... porque com a mudança do ... do acordo ortográfico, às vezes antes era AUTO traço, né, hífen, REGULAÇÃO. Então dependendo eu procuro também. Porque às vezes tem trabalhos anteriores, mas você acha que não precisa pôr aqui?* (Aí seria interessante buscarmos dos dois jeitos). *Pode colocar então?* (...) ... (Porque os trabalhos mais antigos né). ... *É, vão estar desse jeito.* ... (E em inglês fica a mesma coisa?) *Sim, é porque daí não muda, né.* ... *Pronto.* (Quais foram?) *ESTRATÉGIAS DE APRENDIZAGEM, AUTOREGULAÇÃO DA APRENDIZAGEM, INTERVENÇÃO, ALUNOS e PROFESSORES (PGD-FE)*

Observa-se que cinco usuários manifestaram dúvidas ao elencarem as palavras-chave em português de suas respectivas pesquisas em andamento (G-FE, PGM-FCM, G-IFGW, D-IMECC e G-BAE), em alguns casos recorrendo à outras fontes de consulta, tais como artigo publicado (PGM-FCM), pôster apresentado em evento (G-BAE) e e-mail do orientador (G-IMECC).

Nossa, eu sempre ponho só três ... tenho que pensar em mais duas ((RI)) (G-FE)

(Então ... quando você qualificou você fez um resumo da dissertação?) *Fiz.* (E tem algumas palavras-chave que você colocou?) *Sim.* (...) ...~~~ *Tem uma artigo que eu fiz, talvez eu possa pegar dele ... (É nacional?) É. ...~~~ Tinha um resumo. ...~~~ Não tinha palavra-chave lá também, na revista. (Não? Estranho, normalmente tem, né?) É ... geralmente tem no artigo. ...~~~ ENSINO ...~~~ PESQUISA ...~~~ e EXTENSÃO ...~~~ Cinco, só? (Só ((RI))).* (...) *GESTÃO, ADMINISTRAÇÃO, ENSINO MÉDICO, FCM-UNICAMP e ARQUIVO (PGM-FCM)*

Eu posso consultar ... pra confirmar os termos? (Pode). ...~~~ *Palavras-chave pode ser duas palavras?* (Pode, é um termo, né ((RI))). ((RI)). ...~~~ *(Você está consultando as palavras no seu trabalho?) É, do e-mail do meu orientador.* (Aí eu vou pedir pra você anotar a mesma versão das palavras-chave em inglês). *Tá.* (...) *ISOLANTES TOPOLÓGICOS, TÉCNICA ARDES, CLIVADOR DE AMOSTRAS, PROPRIEDADES ELETRÔNICAS, TÉCNICAS DE CARACTERIZAÇÃO (G-IFGW)*

Tá (...) 5 palavras ...~~~ *pode ser composta, né?* (Sim). ...~~~ *Na verdade em português tem dois termos, tem PROGRAMAÇÃO ou OTIMIZAÇÃO, eu não sei o que faz, se eu coloco os dois. ...* (Define um só, porque vamos fazer as buscas usando essas palavras). *Tá.* (E aí a segunda parte vamos fazer a pesquisa dessas palavras-chave em uma linguagem controlada). *Ah, tá.* (Em um tesouro). *Porque assim ... até o nome da ... disciplina, nome de livro eu uso um barra o outro porque ... pode ser os dois?* (Tudo bem, depois damos uma olhadinha como vai estar na linguagem controlada, se tiver diferente vamos refazer a busca. Se for o mesmo termo não vamos precisar refazer). (...) (Então ficou?) ... *PROBLEMA DE CORTE DE ESTOQUE, PROBLEMA DE DIMENSIONAMENTO DE LOTES, GERAÇÃO DE COLUNAS, PROGRAMAÇÃO LINEAR e PROGRAMAÇÃO INTEIRA (D-IMECC)*

(Você já entregou relatório?) *Só falta entregar o relatório final.* (E quando você fez o relatório parcial você colocou resumo e palavra-chave?) *Não. No relatório parcial não mas tem no pôster.* (...) *Já entreguei o pôster.* (Porque assim, vamos fazer buscas por assuntos no repositório da UNICAMP, então eu vou pedir pra você anotar aqui 5 palavras-chave que representam a sua pesquisa). *Tá bom.* ... (E a mesma versão das palavras em inglês). ... *Tem SMARTPHONE, pode escrever SMARTPHONE nas duas?* (Pode, é assim que

fala em português). É, então. ...~ Ai tem AVALIAÇÃO DINÂMICA, ANÁLISE MODAL, é uma palavra só isso ou põe separadas? (...) Separadas? É porque no pôster eu coloquei como se fosse uma palavra só. ... Fica ANÁLISE DINÂMICA ... e ... AVALIAÇÃO DINÂMICA E ANÁLISE MODAL. (Se você costuma usar assim pode pôr.) ... É, porque é difícil pesquisar separado. ... (Porque depois vamos buscar em uma linguagem controlada, no tesouro, se tiver alguma diferença vamos refazer a busca com essas novas palavras). ... Tá bom. (Se for igual ... fica igual). Tá bom. ...~ Eu vou pesquisar essas palavras. (...) ...~ Deve ser a sigla mesmo, eu acho que não é. ...~ (...) Ficou ... ACELERÔMETROS, SMARTPHONES, AVALIAÇÃO DINÂMICA, ANÁLISE MODAL e MODOS DE VIBRAÇÃO (G-BAE)

Em continuidade, onze usuários tiveram dúvidas na tradução das palavras-chave em português para o inglês, recorrendo a fontes externas ou mesmo à pesquisadora, sendo, a maioria alunos de mestrado e nas *Áreas de Artes e Humanidades e Biomédicas* (D-FE, PGM-FE, PGM-IFCH, G-FEF, PGM-FEF, D-FEF, G-FCA, PGD-FCA, PGM-IFGW, G-FEA e PGD-FEA).

(Você vai elencar as 5 palavras-chave ... de uma pesquisa que você tem em andamento). Não não, pra mim isso é simples, tranquilo toda essa parte. (...) Faz diferença, por exemplo, como eu ponho direto no Lattes, é ... palavras-chave que são muito parecidas mas eu insisto, por ex. FORMAÇÃO INICIAL DE PROFESSORES, o grande campo, FORMAÇÃO DE PROFESSORES ... FORMAÇÃO INICIAL e FORMAÇÃO CONTINUADA. Você tá repetindo mas são coisas distintas. (...) Então tem que ter repetição mesmo? Tem, não tem? (Sim, são diferentes). Então eu vou colocar. O sentido é outro, né. ...~ Daí eu ponho em inglês, aqui, é isso? É, a mesma versão em inglês. ...~ (...) Agora eu só não lembro em inglês, em inglês eu acho que é. ...~ Eu esqueci como é AUTORIA em inglês, daí você ... (ah, AUTHORSHIP). AUTHORSHIP? (AUTHORSHIP) (D-FE)

Agora um dos termos em inglês que eu não ... (Você pode tentar dar uma olhadinha) ... Deixa eu ver aqui. ...~ Na verdade eu ainda não fiz essa busca de ... eu ia fazer essa busca de ... de termos, minha orientadora, como ela é da Antropologia ela fala “não, não, já existe um monte de trabalho sobre isso” ela quer mesmo o campo na pesquisa, assim. (...) Pra mim tá sendo bem ... bem interessante. (...) MERCANTILIZAÇÃO e PRIVATIZAÇÃO. ...~ É COMMODIFICATION, termo de COMMODITY. (COMMODIFICATION ((RI)). É COMMODIFICATION, mesmo?) Acho que sim. ...~ (Google tradutor ((RI))). É, apareceu no Google depois eu coloquei o HIGH EDUCATION, MERCANTILIZAÇÃO, apareceu ... (MERCANTILIZAÇÃO mesmo em termo de COMMODITY. (...) Então ENSINO SUPERIOR, COOPERAÇÃO, MERCANTILIZAÇÃO, PRIVATIZAÇÃO, ENSINO SUPERIOR PRIVADO) (PGM-FE)

Não lembro como é SACRA. ...~ (SACRA de sagrado?) (...) ...~ (SACRED). ... Esse é o nome de uma pessoa. (A sua pesquisa é sobre o que?) Sobre um artista de Itu, ele chama MIGUEL DUTRA e na verdade é uma proposta de pesquisa pra montar um catálogo da ornamentação dos templos, dos projetos dele. Porque ele foi trabalhado muito como uma pessoa ... ingênua, que fazia paisagens, que ele era um caipira e aí a minha pesquisa mostra justamente o contrário, que ele era muito mais arquiteto e tinha noções básicas de desenho, de proporção, sabia dos padrões simétricos de como compor um edifício sacro. ... (...) Então a ideia é a partir do levantamento das fontes, montar um catálogo novo sobre ele. ... Ai essas palavras é justamente ARQUITETURA BRASILEIRA, ARTE SACRA, ARTE PAULISTA, ICONOGRAFIA PAULISTA e MIGUEL DUTRA, que é o nome dele. ICONOGRAFIA é justamente o ramo em que ele foi enquadrado pela crítica. (...) ARTE PAULISTA é um grande assunto dentro do qual ele é um principal proeminente representante. ARTE SACRA pra tentar entender, é ... o que que existe neste campo de ARTE SACRA sobre SÃO PAULO. E ARQUITETURA BRASILEIRA porque é a partir de alguns arquitetos que ele começa a aparecer como um arquiteto também, e aí justamente trata especificamente da ARQUITETURA NO BRASIL. Então é essa ponte entre a arte e arquitetura, né. Isso (PGM-IFCH)

Sigla também? (Pode ser. Se você souber o nome completo da sigla é melhor ainda, mas não tem problema, coloca do jeito que você faria). ...~ To com uma dúvida de como escreve COMORBIDADES, eu acho que é COMORBITY. (COMORBIDADES?) ... Isso porque eu escrevi esses dias, hein, que eu passei o meu resumo pro inglês ((RI)). ... (COMORBI ... DADES? Assim?) (...) ...~ Compensa olhar no meu resumo. (É, se

você tiver o resumo é melhor ...~~ porque às vezes o *Google* vai trazer uma outra ... o *Google* não é muito científico ((RI)), acho que é COMORBITY, mas vamos confirmar). ... *Ah, é esse mesmo* ...~~ (Conseguiu?) (...) ...~~ EPILEPSIA, COMORBIDADES, ATIVIDADE FÍSICA, SÍNDROME METABÓLICA e ELT (G-FEF)

Cinco? (Isso). ...~~ É ... então, mas as minhas palavras-chave são em francês. ...~~ Como é que a gente faz? ... (Você não consegue traduzir pro inglês?) É, então ... aqui é HANDSPORT, mas acho que em inglês também é, porque é um termo usado em francês mas acho que em inglês também é, porque é um termo tanto usado em francês quanto no inglês. (...) Acho que é a mesma coisa. (Porque no repositório não sei se tem documentos em francês). ...~~ Eu acredito que não tenha. ...~~ (É). ...~~ PARALIMPÍADAS em inglês ... (PARALIMPICS? Vamos pro nosso amigo *Google Tradutor*). ...~~ (PARALIMPICS). ... PARALIMPICS? (Isso). ...~~ E EDUCAÇÃO FÍSICA é ... em francês ÉDUCTION PHYSIQUE. ...~~ (PHYSICAL EDUCATION). PHYSICAL? (Com Y). ...~~ (REABILITAÇÃO é REHABILI ... TATION. Eu acredito que é. É ...~~ É. REHABILITATION). ...~~ ATIVIDADE MOTORA é ... (MOTOR ACTIVITY). MOTOR ... MOTOR ACTIVITY. ...~~ (Você estuda o que?) *Eu estudo o esporte adaptado na França. (E ... em qual viés?) ... As PARALIMPÍADAS. O foco é quem cuida do esporte adaptado na França, que é a Federação Francesa de Handport. (...) HANDSPORT significa ESPORTE ADAPTADO. Será que é melhor especificar FRANÇA? Porque aqui ele vai dar uma gama de coisas que não ...* ~~~ ESPORTE ADAPTADO NA FRANÇA? ... *Ou não, tá muito* ... (É, talvez restrinja bastante, mas podemos tentar, porque depois vamos checar no tesouro como é que tá e se for o caso vamos refazer). (...) ESPORTE ADAPTADO NA FRANÇA, PARALIMPÍADAS, EDUCAÇÃO FÍSICA, REABILITAÇÃO, ATIVIDADE MOTORA (PGM-FEF)

Meu assunto é único, eu trabalho com história das práticas corporais e dos esportes na natureza ... então minha palavra-chave é sempre OUTDOOR LIFE, NATUREZA, é ... ESPORTES NA NATUREZA, HISTÓRIA DO ESPORTE, HISTÓRIA DA EDUCAÇÃO FÍSICA, HISTÓRIA DA EDUCAÇÃO NA NATUREZA. É isso. Minha grande palavra-chave é EDUCAÇÃO E NATUREZA. É isso. (...) ...~~ Põe a palavra-chave em ... português. ...~~ Essa aqui é uma palavra composta (...) que é a mais importante ... EDUCAÇÃO. ...~~ Essa daqui também é outra palavra-chave. ...~~ Eu trabalho mais com a língua francesa do que com a inglesa. (...) Quantas palavras-chave? (Cinco). ...~~ Acho que está bom assim, essas daqui são as que dão. (...) ...~~ (E o inglês?) É, então, isso que eu disse ((RI)), essa primeira eu acho que é NATURE AND EDUCATION, esse foi o tema de um Congresso em Berlin no ano passado, da história da educação. OUTDOORLIFE é OUT né ... OUTDOORLIFE ...~~ HISTORY OF SPORT ...~~ SPORT AND NATURE ...~~ BODY AND NATURE. ...~~ Ok. (...) ...~~ É muito interessante como no Brasil tem pouco trabalho sobre isso, é muito interessante (...) Então é isso que eu estudo. (...) NATUREZA E EDUCAÇÃO, VIDA AO AR LIVRE, HISTÓRIA DO ESPORTE, ESPORTE E NATUREZA, CORPO E NATUREZA (D-FEF)

TECNOLOGIA SOCIAL, REDES, REAPLICABILIDADE, GERAÇÃO DE RENDA e RECICLAGEM (G-FCA)

É, na minha primeira ... eu tentei fazer uma revisão esse ano e eu tive muita dificuldade em achar os artigos porque é um tema super incipiente. ... (Tem pouca coisa?) *Tem pouca coisa, não só por ter pouca coisa mas por ser um tema pulverizado entre disciplinas, entre áreas, né, as nomenclaturas são diversas e além disso, é ... os termos em inglês e em português são diferentes, então tem mais dificuldade, então nesse sentido eu não sei até que ponto pode contribuir pra sua pesquisa. ...* ~~~ (Você está estudando o que?) *Eu estudo INOVAÇÃO NO PODER JUDICIÁRIO, nos serviços jurídicos, então ... (É interdisciplinar?) Isso, interdisciplinar. ...~~ Então primeiro eu vou começar com os termos em inglês. ...* ~~~ (...) Então é ... LAW INNOVATION ... INOVAÇÃO NA JUSTIÇA ...~~ COURT INNOVATION ... INOVAÇÃO NAS CORTES (...) JUSTICE INNOVATION ... DESEMPENHO JUDICIAL. ...~~ (Essa é a estratégia que você utiliza nas bases de dados?) *É, porque eu peço todas as ... pesquisa booleana, não sei como chama isso daqui. ... (Truncada?) Isso, aqui eu ponho um asterisco, né, pra pegar tudo o que tem INOVAÇÃO daí eu coloco dentro o que se refere ao MERCADO JURÍDICO, né, isso eu tenho pegado bastante coisa, agora eu tenho feito isso pra pesquisa nacional e internacional, só que agora eu não to pegando tanto nacional, aí quando eu vou tentar o nacional eu procuro mais por essas palavras, assim, DESEMPENHO JUDICIAL, INOVAÇÃO NAS CORTES e INOVAÇÃO NA JUSTIÇA, mas não tenho tido muito sucesso não. (...) Então eu acho que é isso, não sei se ... (E a versão de LAWTECH e LEGALTECH ficaria como eu português?) Não tem. (Não tem?). Não, é uma coisa nova né (PGD-FCA)*

(É ... pra gente começar a fazer as buscas, vai ser de acordo com o seu tema de pesquisa. Você já tem um tema definido?) ... *Tenho, eu só acho que não vai ter ... tá bem no começo, né, e assim, se eu pegar os termos que eu mais conheço, assim, né, que tá aparecendo mais não sei se vai ter.* (Então você tá

pesquisando o que?) *É ... relacionado à supercondutividade pra circuitos quânticos e supercondutividade. (...) Cinco você falou? (...) ...~ Vou pôr duas aqui, não tem como separar. (É composta, não tem problema). Não? Então vou deixar. ...~ (Se você não lembrar o inglês podemos procurar). É mais fácil eu não lembrar em português ((RI)). ((RI)). Sabe, porque fica tudo em inglês. ... (É, sua área é mais o inglês ((RI))). (...)* **BITS QUÂNTICOS, NANOESSONADORES, CIRCUITOS SUPERCONDUTORES, COMPUTAÇÃO QUÂNTICA, METAMATERIAIS (PGM-IFGW)**

Tá, então ... o KEFIR, acho que é ((RI)). ...~ É ... CASCA DE MANGA. ...~ (...) Tem um que eu não sei se é com y ou com i ... (Quer confirmar?) Eu acho que não é com y, mas ... (Vou dar uma olhada). ...~ Eu acho que é com i mesmo ... (Isso, com i). Com i? (Com i mesmo). Ah, tá. ...~ Eu vou colocar outras coisas aqui ou ... (Não, só 5). Só 5? Tá, então eu passei. (Então pode riscar ((RI)), é as principais, tá). Tá (...) **KEFIR, CASCA DE MANGA, COMPOSTOS BIOATIVOS, ANTIOXIDANTES, FERMENTAÇÃO (G-FEA)**

(O que você tá fazendo no doutorado?) *Eu to fazendo BIOFILMES na INDÚSTRIA LÁCTEA. (...) Com MICROBIOLOGIA eu mexo. Eu tenho as palavras-chave já. (Ah, então tá ótimo. Em português e em inglês?) ... Isso, em português e em inglês. (...) ...~ A quinta palavra eu vou colocar o que eu to usando agora que é MICROSCÓPIO. Então eu to fazendo a busca sempre por MICROSCÓPIO pra ver o que que tem de novo. ...~ Na verdade eu faço busca com várias palavras ... eu acho que tá certo o inglês, MICROSCOPY, acho que é (...)* **BIOFILME, BACILLUS SPOROTHERMODURANS, FORMAÇÃO DE BIOFILME, DISPERSÃO DE BIOFILME e MICROSCOPIA ELETRÔNICA (PGD-FEA)**

Outros quatro usuários apresentaram dúvidas tanto na definição das palavras-chave em português quanto na tradução das mesmas para o inglês (G-FCM, G-IMECC, PGM-IMECC e PGM-BAE).

(A sua pesquisa é sobre o que?) *SÍNDROME DE TURNER. (...) Não tenho a versão em inglês mas eu posso procurar o inglês. (É que com essas palavras vamos fazer a busca no repositório). Ah, tá ... tá, beleza. Deixa eu achar aqui. ...~ (...) Ah, precisa de mais duas. Mas ... ah, na verdade eu tenho, é que pro PIBIC o banner pediu três mas são cinco, né? (É, você pode pensar em outras que você acha que encaixaria na sua pesquisa). Tá, não, mas eu acho que tem, de qualquer forma. (...) ...~ Tá em algum lugar aqui. ...~ Vou ter que pensar em mais uma ((RI)), só tinha quatro. (Dá uma olhadinha no resumo). É, vou dar uma olhada aqui. ...~ Olha ... eu vou pôr uma ...~ é, eu acho que pode ser. (Imagina que você vai fazer uma pesquisa, um levantamento bibliográfico pra compor o seu relatório). ... Tá. Tem um aqui que é muito geral, mas eu acho que tem muito a ver, né. (Depois vamos consultar em uma linguagem controlada, pode ser que refine um pouquinho mais. Não tem problema deixar assim). ...~ (Então ficou SÍNDROME DE TURNER, né) (...), **CARDIOPATIAS CONGÊNITAS, VÁLVULA AÓRTICA BICÚSPIDE ... COARCTAÇÃO DE AORTA ... e DISMORFISMOS (G-FCM)***

Pode ser mais de uma palavra, né, no caso? (Pode. Composta você fala? Pode ser). Tipo, PROBLEMA DE CORTE DE ESTOQUE. (Pode. ...~ É sobre o que a sua pesquisa?) É PROBLEMA DE CORTE DE ESTOQUE. Que é ... (...) eu vejo a melhor maneira de cortar uma peça maior em itens menores pra que perca menos material e ganhe mais tempo da máquina. Bem resumidamente, assim. ...~ Não lembro se tem um 't' a mais ou não. Que palavra é? É CUTTING STOCK PROBLEM. ...~ Se você puder ver também pra mim, é PROGRAMMING eu não lembro se tem um ou dois 'ms'. (Dois 'ms'). (...) PROBLEMA DE CORTE DE ESTOQUE, OTIMIZAÇÃO MULTIOBJETIVO, PROGRAMAÇÃO LINEAR INTEIRA, MÉTODO DAS SOMAS PONDERADAS, MÉTODO DO ϵ -RESTRITO (G-IMECC)

Palavras-chave do meu tema de pesquisa, né, porque eu ainda to definindo, né, mas tudo bem. (...) Eu até fiz um negócio desse, mas sei lá, como as coisas ainda não estão bem definidas ... porque assim, eu tenho que escrever ... as palavras-chave que me vem à cabeça (...), então seria algo direcionado à resposta que eu quero? (Por exemplo, você quer saber sobre ... computação alguma coisa. Aí você vai ... colocar essa palavra-chave como se você fosse mesmo fazer um levantamento bibliográfico). Tipo num Google qualquer? (Isso, ou numa base de dados) ... Perfeito. (Independente do sistema isso é ... o que você faria em qualquer outro sistema). Joia. (Pra fazer o levantamento bibliográfico da sua pesquisa). (...) ...~ São

5, né. (É). ...~~ É, eu acho que com 3 funciona ... já é suficiente. (Mas ... eu preciso de 5). ...~~ Eu não sei o que fazer agora, eu preciso da sua recomendação. (Qual o tema de pesquisa que você tá pesquisando?) ... Tema seria algo de grande área ou seria algo ... (A sua linha de pesquisa ou ... qual o objetivo da sua dissertação, você já pensou em alguma coisa? Quando você entrou no processo seletivo você apresentou um projeto?) Então, aqui não. Aqui eu primeiro fiz, é ... passei pelo processo seletivo e agora eu defini. Então assim ... (É um assunto que você quer pesquisar. Na dissertação). Eu acho que com três, assim, essas palavras são indissociáveis, então eu não sei se isso configura uma palavra-chave, porque pra mim isso é indissociável no meu tema. (...) (E a versão delas em português qual seria?) Ah, vou dar uma olhada aqui. Acho que não vai caber aqui. (...) É que eu não sei a tradução mais adequada pra isso ... eu tenho algumas mas acho que nenhuma delas fica boa. (...) Então vamos lá. ...~~ Ah, tá ... acho que ... fica mais completo. (...) Então todo mundo entende isso até em português, então eu realmente traduzo? Porque ... (...) (Você tem a versão completa dela em inglês?) Essa é em inglês, então assim, se eu colocar isso em português fica só MCMC, mas não sei se alguém usa isso, eu não conheço ninguém que use isso então eu não sei o que eu faço. (...) Eu deixo assim ou eu ... (É, se é um termo bem usado até em português deixa assim). Pode deixar assim então? (É). Joia. ...~~ Olha, eu acho que assim tá legal. (...) Tá bacana (...) **PREDIÇÃO ESPACIAL, ABORDAGEM BAYESIANA, PROCESSO DE DIRICHLET, MCMC e CLASSIFICAÇÃO ESPACIAL (PGM-IMECC)**

Tá, tá bom então. As cinco por exemplo, eu tenho um termo que eu uso, ESTRATÉGIA PROJETUAL, ele conta como uma palavra-chave ... (Isso, porque é o termo, né). Por exemplo, ARQUITETURA PAULISTA aí também é outro termo, né. (...) (A sua pesquisa é sobre o que?) É sobre a ARQUITETURA RESIDENCIAL do DECIO TOZZI, que é um arquiteto paulista, então envolve toda essa questão e tem também um viés teórico, porque eu to definindo um pouco melhor ... que é associar essa arquitetura dele à corrente fenomenológica. (...) Então. ...~~ Por exemplo é uma coisa assim ... eu sei que do DECIO TOZZI não existem já trabalhos. Então é uma coisa assim, que eu sei que se eu ... pesquiso no ... eu geralmente não coloco como palavra porque se eu pesquiso no repositório eu sei que eu não vou encontrar, entendeu? Que é um arquiteto que ainda não foi estudado ... então eu uso outros termos, não tem problema, né? (Não, mas se você quiser colocar como palavra adicional só pra testar se realmente não tem nada). ... Ah, tá, perfeito. (Porque as buscas vão ser feitas no repositório da UNICAMP. *Sim ...* Então de repente se alguma outra pesquisa abordou em algum momento). ...~~ (...) ARQUITETURA RESIDENCIAL, ARQUITETURA PAULISTA, DECIO TOZZI, ESTRATÉGIAS PROJETAIS e ANÁLISE GRÁFICA. (...) (E aí você anota a versão das palavras-chave em inglês). Posso roubar se eu não lembrar de todas? ((RI)). ((RI)). ...~~ (Conseguiu anotar?) **Consequi (PGM-BAE)**

Embora não tivessem manifestado dificuldade ao elencarem as palavras-chave em português, quatro usuários solicitaram a alteração das palavras-chave durante as buscas por assuntos (PGD-IFCH, PGD-BAE, D-BAE e D-FEA).

É pra escrever uma em cada linha? (Isso). ...~~ A gente trabalha na verdade em português com um termo em inglês, que é GLOBAL SOUTH mas ... é que nas publicações brasileiras eles usam o termo em inglês. (...) CONSUMO, JUVENTUDE, MEIO AMBIENTE, CHINA, GLOBAL SOUTH. [busca por JUVENTUDE] Na verdade tem dois outros termos que seria mais interessante, tipo, pensando que a produção é da UNICAMP, seria mais interessante ... de procurar seria BRICS e CHINA. Eu acho que ... filtrar por maior onde tem uma produção menor nesses dois temas ia trazer uma coisa um pouco mais ((RI)) na área, porque ... é porque como eu procuro na base de dados como BRICS aparece coisas demais, sabe? Mas se a gente tá procurando dentro do universo da UNICAMP que ... os estudos sobre CHINA não são tão antigos ... tipo BRICS porque a nomenclatura foi criada em 2001 e isso traria coisas muito mais contemporâneas e ... porque assim, o volume de coisas acho que um pouco menor, sabe. (Você quer trocar as palavras-chave?) Posso? (Pode). ... Eu vou trocar VALORES por CHINA. GLOBAL SOUTH tem uma coisa que engloba ... teria que ser mais específico. [alteração para BRICS] Mas tem problema que os dois são iguais? [inglês e português] (Não, não). ...~~ MEIO AMBIENTE eu acho que vai ser um pouco mais tranquilo (PGD-IFCH)

[busca por GASEIFICATION] E no caso de palavras que sejam iguais em português e em inglês? (Aí não vamos repetir a busca). (...) Eu posso colocar ao invés de SYNGASS, GÁS DE SÍNTESE, que é a tradução? [alteração para GÁS DE SÍNTESE] (...) Porque aí o resultado seria o mesmo, né. (É, no caso de FISHER-TROPCH sim) **(PGD-BAE)**

[busca por PROPRIEDADES MECÂNICAS] Esse daqui eu acho que ... (muito amplo?) Muito amplo, acho que não vou conseguir chegar. ...~~ (É que ... são as PROPRIEDADES MECÂNICAS DA MADEIRA?) É ... (...) [alteração para PROPRIEDADES MECÂNICAS DA MADEIRA] (...) [busca por BARRAS DE AÇO] Esse também ficou bem amplo aqui ... BARRAS DE AÇO. Cai muito pra parte de estruturas de concreto. ... (Seria BARRAS DE AÇO em que sentido?) É ... pra reforço da madeira. ... É, mas ele fica bem restrito que basicamente aqui no Brasil sou só eu que trabalho. Acaba caindo só ... e pelo visto nós não temos informações externas. ... (Então ...~~ como a gente poderia inserir MADEIRA?) ...~~ PARA REFORÇOS ... PEÇAS DE MADEIRA, fica meio estranho também ... vai pegar diversas palavras ...~~ talvez ... (E MADEIRA?) Também ... acho que ficaria, né. ...~~ (...) [alteração para BARRAS DE AÇO E MADEIRA] [busca por BARRAS DE AÇO E MADEIRA] ...~~ (É que o problema é que não daria pra colocar duas palavras separadas ... porque vai dar influência na pesquisa. (...) Teria que deixar ou BARRA DE AÇO mesmo ou BARRA DE AÇO EM MADEIRA). EM MADEIRA. Tá. ...~~ (Talvez se fosse MADEIRA COM BARRAS DE AÇO). Tudo bem, pode ser. (...) MADEIRA COM BARRAS DE AÇO **(D-BAE)**

Entendi. Tá. (...) ...~~ MICROENCAPSULAÇÃO, SECAGEM POR PULVERIZAÇÃO, HIDRÓLISE PROTÉICA, EMULSÃO e PROTEÍNAS VEGETAIS. (Agora eu vou pedir pra você fazer uma busca dessas palavras-chave aqui no repositório (...) usando o campo assunto). (...) [busca por MICROENCAPSULAÇÃO] (...) É, agora eu to pensando ... colocado mais palavras pra dar uma ... tinha que ser mais específica, em vez de usar um termo geral, podia ter feito uma busca mais específica, né. (Você pode refazer a busca também). Ah, eu posso? (Pode. ... Porque na prática é isso que acontece, né, ((RI)), é ... tem vários registros com esse termo, que você reparou?) Tem. [alteração para MICROENCAPSULAÇÃO DE ÓLEO] [busca por MICROENCAPSULAÇÃO DE ÓLEO] Nossa, o que aconteceu? Parece que piorou ((RI)). ...~~ Agora foi pra ... 140.000. 14.000 ... não, 140.000. Ah, quando eu deixei mais específico ... espera aí. ...~~ Ah, será que é porque ele busca a palavra MICROENCAPSULAÇÃO + ÓLEO? **(D-FEA)**

Apenas um usuário alterou as palavras-chave tanto antes quanto durante as buscas por assuntos (PGM-FCM).

Pode riscar aqui, nas palavras-chave? (Pode). Depois o inglês, to pensando. (Você iria mudar qual?) PESQUISA. E EXTENSÃO. (...) (Ah, tá, você ia colocar qual?) FCM UNICAMP, acho que tem mais a ver. (Então pode colocar). Pode colocar? (Pode) (...) A quinta palavra eu coloquei mas não ... talvez não vai achar nada. Vou ter que pensar na quinta, na verdade. [busca por FC-UNICAMP] ...~~ FCM-UNICAMP pode considerar como uma palavra-chave ou não? (Sim, se faz parte da sua pesquisa). ...~~ Posso mudar essa daqui? Ainda não pesquisei. (Qual você vai por?) ARQUIVO **(PGM-FCM)**

4. Estratégias de busca por assuntos

A categoria “Estratégias de busca por assuntos” refere-se às estratégias utilizadas pelos usuários para determinação da relevância dos recursos informacionais durante as buscas por assuntos no Repositório Institucional da UNICAMP. Cabe esclarecer que, por orientação da pesquisadora, todos os usuários utilizaram como primeira estratégia de busca a caixa de pesquisa disponível na página inicial do repositório, a fim de padronizar a busca inicial. Ao

realizarem a busca pelo primeiro assunto, o sistema possibilitou a visualização de outra página com mais opções de busca e filtros, não sendo orientada pela pesquisadora a utilização de limitadores na *busca facetada* por *Assunto*, *Agência de Fomento*, *Tipo de recurso informacional*, *Data de Publicação* ou *Tipo de Acesso*.

Para a ordenação dos recursos informacionais, utilizou-se a própria opção padrão do Repositório Institucional da UNICAMP, com ordenação por relevância. Entretanto, foi solicitado aos usuários a seleção do filtro quantidade de resultados/página para 50 registros, tendo em vista a possibilidade de uma quantidade considerável de registros recuperados em cada uma das buscas, inviabilizando-se a avaliação da relevância pelos usuários. Após, os mesmos permaneceram livres para realizar a estratégia de busca desejada. Com base nos resultados das buscas por assuntos pelos usuários, as principais estratégias utilizadas foram categorizadas em seis principais aspectos: caixa de pesquisa, digitação dos termos, nível de leitura, posicionamento do cursor, nível de detalhamento e outras estratégias.

Em relação à caixa de pesquisa utilizada, a maioria dos usuários permaneceu utilizando a principal na página dos resultados de busca. Um usuário, ao finalizar a avaliação da relevância na primeira busca, realizou a próxima busca na caixa superior à direita na página de visualização dos resultados, todavia, no restante realizou a busca retornando à página inicial do repositório (PGD-IMECC). Percebeu-se também outras estratégias: opção de continuar as buscas na caixa superior à direita da tela (G-BAE); clique por engano no botão ao lado da caixa de pesquisa denominado “Retornar valores” (D-IMECC e D-BAE); e utilização da caixa de pesquisa principal apenas em parte das buscas (G-FE e PGM-FE).

Quanto à digitação dos termos, os usuários realizaram a busca utilizando-se todos os termos em letras minúsculas (PGM-IFCH, G-FCM, G-FCA, D-FCA, PGM-BAE, PGD-BAE, D-BAE, G-FEA e D-FEA); todos os termos em letra maiúscula (PGD-IFCH, D-IFCH, PGM-FCM, PGM-FEF, PGD-FEF, G-IMECC, D-IMECC, G-BAE, PGM-BAE, PGM-FEA e PGD-FEA); e alternando entre os dois tipos nas buscas (G-IFCH, PGD-FCM, PGM-FCA, PGD-FCA, PGD-IFGW, D-IFGW e G-BAE).

Quanto ao nível de leitura, os usuários a realizaram de forma atenta na visualização dos resultados de busca, passando pelos registros calmamente e pausando a tela em alguns momentos para que o nível de atenção fosse ainda maior (D-FE, G-IFCH, PGD-IFCH, G-FCM, PGM-FCM, PGD-FCM, PGM-FEF, PGD-FEF, G-FCA, PGM-FCA, PGD-FCA, G-IFGW, D-IFGW, G-IMECC, PGD-IMECC, G-BAE, D-BAE, G-FEA e D-FEA). Outros usuários expressaram por meio da

vocalização da fala, a leitura de títulos e autores (PGD-IFCH, D-FE, G-FCM, PGD-FCM, PGD-FEF e PGD-FCA). Já outros realizaram uma leitura mais rápida de forma geral, atentando-se poucas vezes em trechos específicos (PGM-IFCH, D-IFCH, PGM-IFGW, PGD-IFGW, D-IMECC, G-BAE e PGM-FEA). Um usuário primeiramente arrolou rapidamente todos os resultados de busca até o final da página, para, em seguida, realizar novamente uma leitura mais atenta desde o início dos resultados de busca (G-BAE).

Quanto ao posicionamento do cursor do mouse, a maioria o manteve sobre os títulos dos registros (D-FE, PGM-IFCH, PGM-FCM, G-FCM, G-FEF, PGM-FEF, G-FCA, PGD-FCA, D-FCA, G-IFGW, PGM-IFGW, PGD-IFGW, D-IFGW, G-IMECC, PGM-IMECC, D-IMECC, G-BAE, PGM-BAE, D-BAE, G-FEA, PGM-FEA, PGD-FEA e D-FEA); seguida de autor(es) (D-FE, PGM-IFCH, PGM-FCM, PGD-FCM, G-IFGW, D-IFGW, PGD-IMECC, D-IMECC, G-BAE, G-FEA, PGM-FEA e PGD-FEA); orientadores, no caso de teses e dissertações (D-IFCH, PGM-FCM, G-FCA, D-FCA, G-IMECC, D-IMECC, PGM-FEA e D-FEA); tipo de recurso informacional (PGD-IFCH, D-IFCH, D-IMECC, PGD-FEA e D-FEA) e data do recurso informacional (D-FE, G-IMECC e PGD-IMECC), sendo que alguns usuários posicionaram o cursor em mais de um tipo de informação. Já outros não mantiveram o cursor em informações específicas: utilização do *scroom* do mouse (G-IFCH, G-FCA e PGD-FCA); utilização da barra de rolagem da página (PGM-FCM e PGD-FEF); utilização de ambos os recursos (G-BAE); e não utilização de mouse (PGM-FCA).

Quanto ao nível de detalhamento do recurso informacional, categoriza-se em três níveis: a) nível de visualização dos resultados; b) nível de visualização dos registros; e c) nível de visualização do texto completo (PDF). A maioria dos usuários permaneceu no nível de avaliação básico, isto é, apenas nas informações de pré-visualização: data do recurso informacional, título, autor(es), orientador e tipo. Já outros avançaram no nível de detalhamento, em busca principalmente dos resumos dos recursos informacionais, conforme detalhado a seguir.

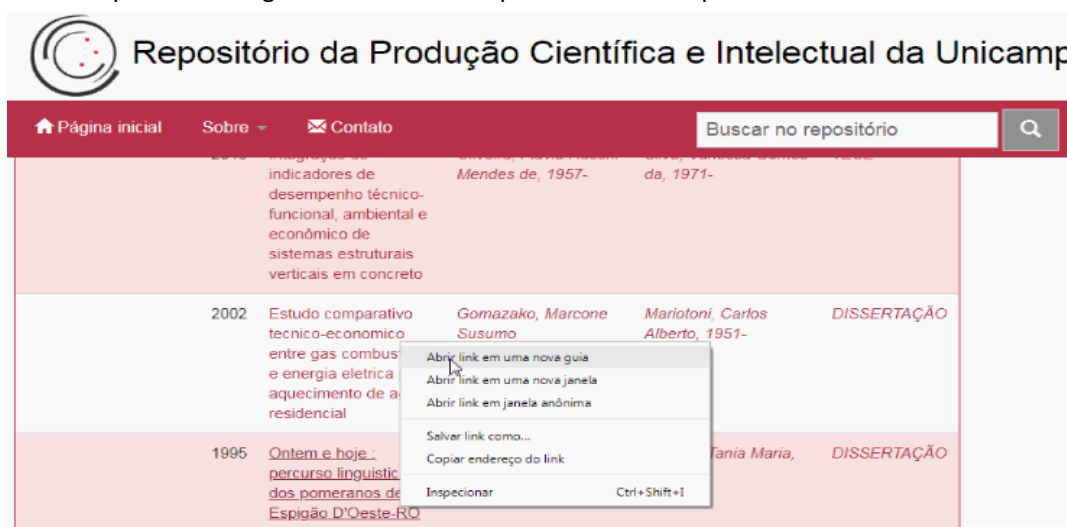
Alguns usuários, ao acessarem o registro do recurso informacional, verificaram rapidamente apenas o resumo/abstract, retornando em seguida à página anterior com a visualização dos resultados, nos seguintes recursos informacionais: artigo/artigo de periódico (G-FCA, G-IFGW e PGD-IMECC); e tese/dissertação (PGD-FCM, PGD-FCA, G-IFGW, PGM-IFGW, PGM-BAE e G-FEA). Em alguns casos houve tentativa de visualização, porém, o artigo (G-FCM e G-FCA) ou dissertação (PGM-BAE) não tinham resumos. Já outros usuários, ao entrarem no

registro do recurso informacional, realizaram a leitura atenta do resumo/abstract, retornando em seguida à página anterior com a visualização dos resultados, nos seguintes recursos informacionais: artigo/artigo de periódico (PGD-FCM, PGM-FCA, G-IFGW e G-FEA); tese/dissertação (PGD-FCM e G-IFGW); e artigo de evento (G-IFGW). Aqui, somente um usuário passou rapidamente pelas palavras-chave do registro (G-IFGW).

No terceiro nível, visualização do texto completo, os usuários tiveram acesso ao texto na íntegra para um maior detalhamento das informações, possibilitando maior segurança na avaliação da relevância (PGM-FCA, G-IFGW e G-FEA). Um usuário acessou uma dissertação na íntegra pois o registro não continha o resumo. No texto, passou pelo sumário, atentando-se ao resumo. Tratando-se de uma dissertação e considerada antiga pelo mesmo, retornou à página anterior. Em outro caso semelhante, passou rapidamente pelo resumo e retornou à página anterior (PGM-FCA). Outro usuário também verificou que uma tese era datilografada. Entrou em um artigo, passando rapidamente pelas informações e retornou à página anterior (G-FE).

Um usuário acessou tanto dissertação quanto artigo. No artigo, percorreu a introdução, desenvolvimento e retornou. Em outro, atentou-se à introdução, percorreu o artigo até o final, atentou-se às figuras, rapidamente ao resumo e retornou. Em uma tese, percorreu o sumário, retornou ao registro e atentou-se rapidamente ao resumo (G-IFGW). Já outro utilizou tal recurso apenas para confirmação se o texto na íntegra continha a palavra-chave digitada, pois o título do recurso informacional lhe parecia estranho (PGD-FCA). Além disso, houve tentativa de acesso ao texto na íntegra (PDF) do recurso informacional (D-FCA e G-FEA). Entretanto, devido às questões de direito autoral, o acesso aos mesmos estava fechado. Em todos os níveis, somente um usuário entrou em um registro de dissertação utilizando-se os termos de busca em linguagem controlada (G-FEA), sendo que os demais usuários apenas em linguagem natural. Verifica-se ainda a utilização de outras estratégias de busca: um usuário, conforme avançava na visualização das informações, utilizou o recurso de duplicar as páginas com os registros que interessavam (PGM-IFGW). Um usuário utilizou a opção “abrir link em uma nova janela” para visualização posterior dos registros (PGM-BAE), enquanto outros abriram novas guias com o botão direito do mouse (G-FCA, PGD-FCA e G-IFGW) (Figura 36).

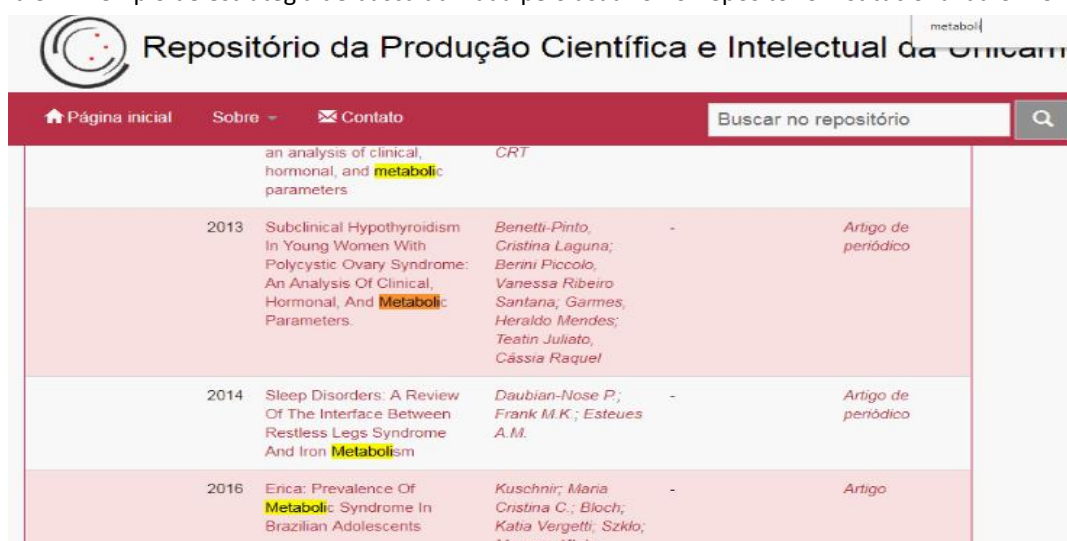
Figura 36. Exemplo de estratégia de busca utilizada pelo usuário no Repositório Institucional da UNICAMP - A



Fonte: Dados da pesquisa.

Já na página com as informações detalhadas do registro (resumo e palavras-chave), um usuário utilizou o recurso CTR+F para localizar a palavra-chave “Antioxidante”, realizando a busca por “anti” no resumo de um artigo e “man”, para localizar a palavra-chave “Manga” em uma tese, porém não sendo possível pelo fato de a mesma estar datilografada (G-FEA). Outro usuário utilizou o mesmo recurso para a palavra-chave “Arpes”, realizando a busca por “arp” em uma tese (G-IFGW). Tal estratégia foi compartilhada ainda por outro usuário para localizar a palavra METABOLIC, realizando a busca por “metaboli”, porém, no nível básico de visualização dos resultados (G-FEF) (Figura 37):

Figura 37. Exemplo de estratégia de busca utilizada pelo usuário no Repositório Institucional da UNICAMP - B



Fonte: Dados da pesquisa.

As estratégias de busca utilizadas pelos usuários no Repositório Institucional da UNICAMP demonstram certa familiaridade com outros sistemas de recuperação da informação. Salienta-se a importância tanto de títulos quanto de resumos que representem adequadamente os recursos informacionais, pois em sua maioria, a avaliação da relevância pelos usuários e de suas necessidades informacionais perpassam principalmente por estes dois substitutos de informação. Nesse sentido, torna-se relevante que todos os recursos informacionais contidos neste sistema de recuperação possuam resumos, tendo em vista que alguns não possuíam tal informação. Apenas um usuário passou rapidamente pelas palavras-chave do recurso informacional (G-IFW), situação preocupante que indica que os assuntos dos recursos informacionais não são relevantes para os usuários para a avaliação de suas necessidades informacionais, pelo menos não nas buscas realizadas neste universo de usuários e no contexto específico do Repositório Institucional da UNICAMP.

5. Avaliação da relevância - Linguagem natural

A categoria “Avaliação da relevância - Linguagem natural” refere-se aos comentários dos usuários sobre os recursos informacionais recuperados com seus temas de pesquisa utilizando-se a linguagem natural, isto é, as palavras-chave elencadas para a primeira rodada de busca por assuntos no Repositório Institucional da UNICAMP.

Os usuários manifestam dúvidas gerais sobre as buscas: indicação de quais recursos informacionais são relevantes ou quantos (D-FE, G-FCA, D-FCA, PGD-IMECC, G-BAE, PGD-BAE e D-FEA) ou relevantes por tipo de recurso informacional (PGM-FEF); utilização do recurso gráfico aspas (“) (G-IFCH, G-FEF, PGD-FEF, PGD-IFGW e D-IFGW); palavra-chave em inglês coincidente com a palavra-chave elencada em português (D-IFCH, PGM-FCA e PGD-BAE); rascunho para anotação da quantidade de recursos informacionais considerada relevante (D-FCM); acentuação de termo (G-FEF); possibilidade de busca avançada na ferramenta (D-FCM); caixa de pesquisa (G-FCA); ordenação dos resultados de busca (D-FCA); busca booleana (D-FCA); digitação do termo em maiúsculo ou minúsculo (G-IFGW, PGD-FEA, PGM-IMECC e PGM-FEA); utilização de caracteres especiais contidos em um determinado termo e diferença nos resultados utilizando-se ou não o recurso gráfico aspas (“) (G-IMECC); acesso aos recursos

informativos do repositório na íntegra (D-FCA e G-IMECC); numeração dos resultados da busca (D-FCA); filtro de 50 resultados por página (G-FEA); e retorno de resultados em português e em inglês (PGD-FEA).

[busca por FORMAÇÃO INICIAL DE PROFESSORES] É ... anotar quantos, não precisa ser quais. (Não, só a quantidade). Tá, entendi (D-FE)

[busca por SEGURANÇA PÚBLICA] Dois. Aí eu pesquiso entre aspas em português também? (G-IFCH)

[busca por RURAL] Aqui eu acho que vai dar um problema porque ... é a mesma palavra em português e em inglês. ... (Ah, aí não vamos repetir, só isso). Aí em inglês vai aparecer a mesma coisa (D-IFCH)

[busca por VACINAS, RESPOSTA] É ... quando eu encontro, só pra mim não perder ... tem um rascunho que eu possa ir anotando? (D-FCM)

[busca por ATIVIDADE FÍSICA] Quando é duas palavras tem que colocar entre aspas? (...) [busca por SÍNDROME METABÓLICA] SÍNDROME tem acento no 'í'? (G-FEF)

[busca por ESPORTE ADAPTADO NA FRANÇA] Não precisa ser por tipo, eu posso pôr geral? (Não não, geral) (PGM-FEF)

[busca por "TEACHING/LEARNING"] Bom, e agora, ESCOLA DE CIRCO coloca entre aspas? (PGD-FEF)

[busca por HIV] É ... eu pretendia fazer uma busca ... é ... booleana ... que tivesse 'and' e 'or', ele tem algum ... mecanismo aqui? (D-FCM)

[busca por TECNOLOGIA SOCIAL] Aí eu coloco quantos eu achei? (...) Oito (...) [busca por "TECNOLOGIA SOCIAL"] É só colocar aqui, né (G-FCA)

[busca por BURKHOLDERIA] Aqui eu duplico? (PGM-FCA)

[busca por SECAGEM] E eu sigo as demais em português ou já passo pro inglês? (...) [busca por BAGASSE] Não tenho que te dizer qual que foi, né? (D-FCA)

[busca por SECAGEM] Assunto? (Isso). Deixa por relevância mesmo? (Isso). (...) Eu não posso abrir? (Pode, pode ficar à vontade). Posso ler? Por exemplo, eu não sei o que é OKARA, né, "secagem e caracterização das partículas de okara", daí como eu não sei o que é OKARA eu abriria pra ver. ... (Sim, pode avaliar do jeito que você achar melhor, tá. Se você conseguir avaliar pelo título ... ou se precisar avaliar mais a fundo ...). O resumo eu não consigo ver aqui. (É, nesse caso específico não tem o resumo). ... Eu posso abrir pra ver? (Claro). ...~ Você abre aqui pra ...? (Isso). ...~ É pesado ... 20 megas. Você deve ter que escanear pra ... 2004 ... É uma tese da FEA inclusive. (...) ...~ O que que eu faria agora? Iria no Google enquanto ele abre, o que que é OKARA ((RI)). ((RI)). Em 2003 o mestrado e em 2007 o doutorado, 2003 a minha dissertação tá toda escaneada, então eu acho que ela tá pesada no sistema porque na época ... não ia direto, né, digitalizado, era escaneado. Tanto que eu tenho a minha versão em PDF mas se eu jogar a palavra-chave não adianta, né, porque ela tá toda escaneada ... não vai localizar, né. (Não localiza). ...~ Bom, eu desistiria desse ((RI)). ((RI)). ...~ É muito lento ... 11 minutos. ...~ Deixa eu ver aqui ... vou procurar, tá ((RI)). ... (Fica à vontade ((RI))). ...~ É um alimento. (...) Resíduo de extrato de soja, então eu já nem pegaria esse. ...~ Ele não vem numerado pra gente, né, 1, 2,3,4 (...) (Porque o que dá pra fazer hoje é naquela caixinha inicial que a gente tem, colocar duas palavras-chave separadas, e aí ele acaba varrendo). Ele faz a busca. ... (Faz mas não é o ideal). Ah, entendi, aí colocaria as duas entre aspas ou não, sem aspas, sozinha. (É, se for sozinha assim não tem necessidade). Vamos supor que eu quisesse colocar SECAGEM e BAMBU, uma ao lado da outra naquela caixinha, teria que colocar entre aspas cada uma delas ou não? (Não, só se você quiser juntar duas palavras, por exemplo, Engenharia de Produção. Se você colocar entre aspas ele vai buscar exatamente Engenharia de Produção, aquele conjunto, um do lado do outro. Agora se você colocar Engenharia de Produção sozinho o que que ele vai fazer, buscar Engenharia num lugar, Produção no outro, não necessariamente juntos) (D-FCA)

[busca por ISOLANTES TOPOLÓGICOS] eu queria com letra maiúscula, que é o padrão ... vai maiúscula? (G-IFGW)

[busca por FÉRMIONS PESADOS] (Agora a gente vai buscar entre aspas). *Tá. Coloco cada uma delas entre aspas, é isso?* (Não. Coloca uma aspa aqui e uma aspa no final. Porque agora, o que que ele vai fazer? Ele vai buscar exatamente esse conjuntinho de palavras um do lado do outro). *Ah, tá.* (Quando a gente busca, pelo menos no repositório, né. Quando a gente busca sem aspas, o que que ele faz? Ele pode buscar o FÉRMIONS num lugar e PESADOS no outro). *Ah, entendi.* (Entendeu? Aí ele dá uma diferenczinha aí no ...). *Ele busca de maneira diferente.* (É. ...~~ se forem os mesmos você pode contar de novo, tá). *Ah, tá.* (Como se fosse uma nova busca) (...) [busca por SUPERCONDUTIVIDADE] Nesse caso não precisa, né. (Isso, esse não) (PGD-IFGW)

[busca por "FATOR DE FORMA"] Aspas no início e depois no final sem espaço? (...) ...~~ *Caiu bastante, né. ...~~ Esse aqui, seria, mais ou menos, que poderia ... vai ter essa palavra que eu to procurando mas ...* (...) (Agora é o inglês). ... Supostamente teria que pegar os mesmos? ... (D-IFGW)

[busca por PROBLEMA DE CORTE DE ESTOQUE] (Aqui como são palavras compostas que você colocou, a gente vai fazer ... buscar assim depois buscar entre aspas, tá). *Ah, tá.* (Pra ver se dá diferença). ...~~ E porque dá diferença? (Então, porque no repositório quando a gente coloca assim entre aspas ele traz exatamente esse conjunto de palavras um seguido do outro). *Ah, tá.* (Se você coloca separado ele pode buscar PROBLEMA num lugar, ESTOQUE no outro, entendeu?) *Tá, entendi.* (Então dá alteração, né). ...~~ Então primeiro eu procuro separado? (É) (...) [busca por "OTIMIZAÇÃO MULTIOBJETIVO"] A gente pode ter acesso a todas as coisas que estão aqui? ... (Ah, então ... tem uma parte que tá aberta, né, as teses e as dissertações, alguns artigos estão abertos e aí tem uma parte que está fechado o acesso, embora você consiga ver, você não consegue ter acesso ao documento, por conta de direito autoral, essas coisas). Mas dá pra ter acesso ao resumo pelo menos ou não? (Sim, sim. ...~~ Olha, aqui você consegue ver quando é aberto ou fechado. Aberto é quando o documento tá na íntegra). *Tá* (...) [busca por MÉTODO DO ϵ -RESTRITO] Agora o único problema é que eu não sei se ele aparece 'e', como E-RESTRITO, porque na verdade é um Épsilon ... é uma letra grega. (Ah, tá ... nos artigos costuma aparecer a letra grega mesmo?) (...) ...~~ (Pode ser dos dois jeitos?) *Então, normalmente a gente fala E-RESTRITO mas sempre escreve com o Épsilon.* (Tá ... nenhum documento, nenhum trabalho você viu com o 'e'?) ... *Não, que eu lembre não, sempre com Épsilon.* Eu deixo com o 'e' mesmo ou tento colocar o Épsilon? ...~~ *E tem o Épsilon maiúsculo e o Épsilon minúsculo.* (Qual é mais usado?) *Acho que é o minúsculo.* (Mas e quando você costuma fazer a pesquisa nas bases de dados, como você coloca? Com e ou com assim mesmo, com épsilon?) ...~~ *É uma boa pergunta. Eu não lembro ((RI)). (Não?) Não ((RI)). ...~~ Tipo, ele entende que é o Épsilon, to vendo aqui e aparece o Épsilon.* (...) ...~~ (Bom, deixa assim que aí depois a gente consulta na linguagem). ... *Assim como?* (Copia o Épsilon e coloca do jeito que você costuma). ...~~ *Agora resta saber se ele vai entender ((RI)). (É ((RI))). Nós saberemos ((RI)). ...~~ Ele não entendeu muito bem não.* (Nada? ...~~ Nenhum?) *Não. Só se eu escrever ÉPSILON mesmo, que nem estava naquele outro negócio. ... Posso ver se aparece alguma coisa?* (Pode) (G-IMECC)

[busca por PREDIÇÃO ESPACIAL] Tudo maiúsculo? (Tanto faz). *Ok. Aqui, né* (PGM-IMECC)

[busca por SISTEMAS DINÂMICOS] Eu tenho que anotar ou somente quantidade? (Só a quantidade). *Ok. ...~~* (E aí a gente vai fazer a pesquisa pra cada uma das palavras, tá). *Entendi* (PGD-IMECC)

[busca por ACELERÔMETROS] Eu acho que essa é a primeira. ... Tem que anotar o nome? (Não, só a quantidade mesmo). *Ah, tá bom* (G-BAE)

[busca por GASEIFICAÇÃO] Fica maiúscula? (...) [busca por GÁS DE SÍNTESE] Era pra gente anotar quantos resultados tinha dado? (Não, não, não tem problema porque como está gravando as telas, depois eu vou verificar). *Tá. ...~~* (Aqui, olha, quando você clica em 'retornar valores' ele vai trazer a palavra anterior que você buscou, tá vendo, olha, ele mudou, sempre clica aqui no 'ir'). *Tá. ...~~* (Não é muito usual, né, esse botão mas a ferramenta é assim). ...~~ *Será que foi?* (Foi. Aí o GÁS DE SÍNTESE como é uma palavra composta, eu vou pedir pra você buscar entre aspas também, tá, porque dá diferença) (PGD-BAE)

[busca por CASCA DE MANGA] Como é que eu sei que são 50 ... aparece aqui? (G-FEA)

[busca por SOBERANIA ALIMENTAR] Tem diferença escrever maiúscula ou minúscula? (PGM-FEA)

[busca por BIOFILME] Tem diferença maiúscula? (...) Uma por uma? Ou todas juntas? (...) Mas ele traz em ... português, em inglês, tudo misturado, né? (...) Eu ando procurando mais no Research Gate pra procurar artigo, sabe? ((RI)). Tipo ... ((RI)) eu pego mais no Google Acadêmico do que no repositório. ... (Não consegue encontrar?) É, não traz tanta variedade, agora tem um aplicativo também, né, chama Researcher, não sei se você já viu (...) eu só pego nele agora. ...~~~ (...) Você consegue buscar por palavras-chave ... é bem melhor (PGD-FEA)

[busca por "MICROENCAPSULATION OF OIL"] Agora apareceu os artigos ... por isso que aumentou. ...~~~ Nove. Tinha que anotar o total também? (D-FEA)

Os usuários expressam dúvidas particularmente quanto à avaliação da relevância: indicação de recurso informacional relevante que já apareceu em busca anterior (G-IFCH e PGM-IFCH); quantidade de recursos informacionais a serem apontados como relevantes (PGM-IFCH); e nível de detalhamento das buscas para avaliação da relevância, se por título, resumo ou o texto na íntegra (PGM-FCA, PGD-FCA, G-IFGW, PGM-IFGW, PGD-BAE e G-FEA). Particularmente, um usuário expõe que em sua área de pesquisa que resultados costumam ser forjados facilmente e considera que a avaliação apenas pelo título ou mesmo pelo resumo é superficial, sendo o ideal avaliar o texto completo, porém, em alguns casos o acesso ao mesmo está fechado (PGM-FCA).

Os usuários questionam a ausência de resumo (PGM-FCA e PGD-FCA). Em outras ocasiões, indagam sobre: confirmação do propósito da pesquisa (G-FCM); apontamento de quantos e não quais recursos informacionais são relevantes (G-FCM e D-IMECC); definição da relevância, direcionada para seu tema de pesquisa (G-FCM, D-FCM, G-IFGW, PGM-IMECC, D-IMECC, PGM-BAE, PGD-BAE e PGM-FEA); questionamento sobre ausência de resumo em português (PGD-FCA); e necessidade de indicar novamente quando um recurso informacional já havia sido considerado como relevante em busca anterior (PDG-FCM).

[busca por "SEGURANÇA PÚBLICA"] Pode ser o mesmo? (Pode, pode sim). ...~~~ Seis (G-IFCH)

[busca por ARQUITETURA BRASILEIRA] Eu tenho que escolher só uma ou ... (Não, não, todos que estão aí, você vai elencar todos que te interessam, desses 50, aí você vai contar) [busca por "ARTE PAULISTA"] E se apareceu esse nas outras? (Pode contar). ...~~~ Pode pôr zero? ((RI)). (Pode) (PGM-IFCH)

[busca por SÍNDROME DE TURNER] Então você quer que eu anote quantos que eu achei relevantes. (É). Não quais, só a quantidade mesmo? (Só a quantidade). (...) Se eu quiser eu posso clicar ou é só pelo título mesmo? ... (Você é que sabe ((RI))). Tá ((RI)). ...~~~ É especificamente para o meu trabalho, né. (É. ...~~~ pra esse conjunto de palavras, pro seu tema mesmo de pesquisa). Certo (G-FCM)

[busca por SILICOSIS] Conta repetido, quando tem dois iguais? (PGD-FCM)

[busca por HIV] Para a pergunta, né, porque essas minhas palavras-chave elas foram ... é ... ligadas a uma pergunta, né. (Sim) **(D-FCM)**

[busca por LIPASE] isso que eu ia perguntar pra você, é ... com base em que? Só olhando o título ... e o autor, a data ... (...) Meu assunto é meio abrangente, então como eu trabalho com ... coisa biológica ... é ... às vezes o ... geralmente nossos títulos são menores e ... o título às vezes não fala o que realmente a pessoa trabalhou. (...) O que é interessante pra mim são as análises que eles fazem, não do título, então às vezes não tem uma informação muito boa no título, por isso eu to perguntando, a sua avaliação é só no título ou não, em tudo? (...) ...~ É, tem até uns que, por exemplo, esse aqui é fechado, ou seja, não tem acesso a ele, né. (É, aí ... não dá pra avaliar na íntegra). Isso é uma observação pra você ou não também? (Se você perceber que ... pelo resumo você consegue dar uma olhadinha ...). É, com o resumo ele me dá um resultado muito satisfatório, mas como a gente trabalha na área a gente sabe que às vezes esses resultados eles são forjados, assim. (...) Então quando olha o que a pessoa fez você sabe se a pessoa forjou ou não ... aqui tá interessante ... eu teria que olhar o método que a pessoa utilizou pra ver se está forjado ou não, mas esse aqui é interessante. ...~ É só contar, não precisa escrever nome nem nada ... (Não, não). (...) Pronto, vi os 50 primeiros. (...) Um, dois, três, quatro, cinco (...) [busca por IMOBILIZAÇÃO] é uma pena que não tem resumo. ...~ No meu ponto de vista quem não tem resumo ... a não ser que eu tivesse muito tempo pra ler todo o artigo, mas também acho que não ia dar muito pra ler um artigo. ...~ Esse não ...~ ele tem resumo ... só no documento, né? (Só no documento). Mas no registro não? (Não) **(PGM-FCA)**

[busca por LAW INNOVATION] “tecnologia e inovação no saneamento básico”, não ...~ “pesticidas” não. ...~ É só pelo título mesmo, né? (É ... você que sabe, se você conseguir avaliar pelo título). ...~ “Plantas transgênicas” ...~ então foram seis **(PGD-FCA)**

Só olhando no título, né? (Você que sabe ... é ... pode ficar no título ou entrar no registro, resumo, no próprio documento ... é como você faz a busca mesmo) **(PGM-IFGW)**

[busca por TOPOLOGICAL INSULATORS] É. ...~ Acho que ... quando tem ... igual. ...~ (Igual, você fala, o mesmo registro?) É, eu acho que é. (Aí você conta um só, se estiver duplicado. ...~ Às vezes acontece). Acho que é o mesmo. ...~ Tem só a ver com o tema ou seria tipo um artigo que eu leria pra ... (É o que você acha que daria pra utilizar na sua pesquisa, que você leria e de repente alguma coisa no artigo poderia te ajudar) **(G-IFGW)**

[busca por DESEMPENHO JUDICIAL] tem algumas coisas que tem a ver mas não é diretamente, sabe? Mas seria interessante, eu poderia ler esse daqui. ...~ “reestruturação dos processos” ...~ esse só o resumo em inglês, né. ...~ “inovação social”. ...~ Também só em inglês esse, né. Então vai dar um total de 4 **(PGD-FCA)**

[busca por PREDIÇÃO ESPACIAL] Entendi. Considerando todas essas palavras-chave, né? (É, cada palavra-chave vamos fazer a busca e apontar os documentos relevantes). Mas por exemplo, esse termo aqui tem que ser pertinente a todos esses, não somente esse, né? (Não, é ...). Somente esse, tá. (Sobre PREDIÇÃO ESPACIAL ... o que você encontrou?) (...) Relacionado ao meu tema? (Ao seu tema, tudo direcionado ao seu tema). Joia. (...) ...~ Bom, eu tenho que falar pra você, né? (É, dá uma lida). ... Eu abro ou arquivo ou só ... (Fica à vontade, como você fosse pesquisar normal). Ah, legal. (Porque você vai ter que decidir se é relevante ou não pra sua pesquisa) **(PGM-IMECC)**

[busca por PROBLEMA DE CORTE DE ESTOQUE] A pergunta é, se eles são relacionados, se eles são importantes ... como é que ... qual que é a pergunta? ... (Se seria interessante pra esse conjunto aqui de pesquisa ... pra essa pesquisa em andamento que você tem). Então, na verdade pelo título é difícil, né (...) [busca por “PROBLEMA DE DIMENSIONAMENTO DE LOTES”] Não preciso falar de quantos? (Quais, você fala?) É, de quantos, porque que nem aqui, olha, aqui é 40, 25 de 50. E aqui é o total. (Ah, não precisa) (...) **(D-IMECC)**

[busca por ARQUITETURA PAULISTA] Eu tenho que pensar, também eu to pensando só nessa palavra, né. (Isso). ... Porque por exemplo, aparece teses importantes sobre outros conceitos que eu coloquei aqui que eu uso no meu trabalho ... E aí, como é que funciona? (Exatamente juntos? É, exatamente ... esse daqui.

Pode colocar. Como se a gente não estivesse utilizando as outras palavras). *Tá bom. ...~ Aí é a mesma coisa, eu coloco aspas, né, e busco novamente ...* (Isso) **(PGM-BAE)**

[busca por GASEIFICAÇÃO] *Vai só pelo título ou vai no abstract também?* (Você que sabe, se você conseguir avaliar só pelo título). ...~ *Mas eu queria te perguntar, eu tenho que selecionar algo que eu acharia interessante ou que poderia ser usado pra minha tese especificamente?* (Que seria relevante pra sua pesquisa, pra essa pesquisa aqui que você está fazendo). *Entendi. Independente se eu achar interessante pra qualquer coisa ...* (É). *Ah, tá, focado na tese.* (Focado na tese) **(PGD-BAE)**

[busca por KEFIR] *Ah, mas só por ser ... KEFIR DE LEITE me interessa, já. ...~ AÇÁI também tem antioxidante. (...)* *Se interessou eu já posso ...* (Se quiser ir anotando na folhinha a quantidade. ...~ Aqui você não precisa voltar). ...~ *É pra ver só pelo título ou ...* (Do jeito que você achar melhor ...~ que às vezes precisa entrar no título pra ter certeza, né, mas ...). *É, mas de vez em quando nem o título bate* **(GFEA)**

[busca por EMERGIA] *é só o que me interessa pro meu projeto, né? Específico. ...* (É, pra sua pesquisa, né, se esses documentos são ... se você acha que poderia ajudar de alguma forma) **(PGM-FEA)**

Os usuários expressam a hipótese de que com um determinado termo de busca (G-FE, D-FE, D-IFCH, PGD-FEF, D-IFGW e PGD-BAE) ou tipo de recurso informacional (D-IFGW) haverá resultados considerados relevantes.

[busca por EDUCAÇÃO ESPECIAL] *Esse vai dar bastante, né, quando faz composto acho que vai mais, né. (...) Na minha mesmo eu não coloquei nenhum composto ... mas se tivesse que pensar em mais ... vira quase a junção, né* **(G-FE)**

[busca por "PRÁTICAS DE LEITURA E DE ESCRITA"] *Eu acho que não vai ter tanta diferença, porque o que tinha de PRÁTICAS, né. (...)* [busca por READING AND WRITING PRACTICES] *Agora vai ser um monte. ...~ 24. É, esse é um item ... por isso que eu troquei, é um item muito rico* **(D-FE)**

[busca por GENDER] *Acho que o inglês vai dar mais ... porque aí GENDER é mais ... não vai ter essa ... esse triplo sentido que em português tem, é ... de GÊNERO LITERÁRIO, GÊNERO na Biologia* ((RI)) **(D-IFCH)**

[busca por STREET'S CIRCUS] *com certeza essa vai ser ...* ((RI)) **(PGD-FEF)**

[busca por PHYSICAL EDUCATION] *Aqui eu acho que tem mais, né? ... Mais termos* **(PGM-FEF)**

[busca por STRANGENESS] *Agora tem que aparecer um monte, porque é artigo, vai ter ...~ de grande colaboração também. ...~ Então ... 15, 16, 17* **(D-IFGW)**

[busca por GÁS DE SÍNTESE] *Eu acho que vai dar ... bastante. Só pelos primeiros resultados já. ...~ (Deu quantos aqui?) Deu só um* **(PGD-BAE)**

Os usuários expressam os recursos informacionais considerados relevantes ou potencialmente relevantes (G-FE, D-FE, G-IFCH, PGM-IFCH, PGD-IFCH, D-IFCH, G-FCM, D-FCM, G-FEF, PGM-FEF, PGD-FEF, PGD-FCA, G-IFGW, D-IFGW, PGM-IMECC, D-IMECC, PGD-BAE, PGM-FEA e D-FEA); expõem que não utilizavam o repositório para realizar buscas (PGD-FCA); consideram mais fácil a busca no *Google Acadêmico* do que no Repositório Institucional da

UNICAMP, já que no repositório há recuperação de recursos informacionais da *Área de Biomédicas* (G-BAE); indagam sobre a ordem de apresentação dos recursos informacionais (D-IFGW); apontam a ausência de resumo nos recursos informacionais (PGM-IMECC); indagam sobre o critério de ordenação de relevância da ferramenta (PGD-IFCH, PGD-FEF e D-IFGW); sugerem um filtro por área do conhecimento (PGD-IFCH); e expõem não compreender por que alguns temas foram recuperados na busca com o termo utilizado (G-IFGW).

Adicionalmente, observam que o sistema realiza a busca dos termos separados, quando não se utiliza o recurso gráfico aspas (“) (PGD-FCA e D-IFGW); indagam a existência no repositório da possibilidade de busca combinada como em outros sistemas de recuperação, já que os termos buscados são muito genéricos (PGD-IFCH, D-IFGW, PGD-BAE, PFGM-FEA e D-FEA); mostram-se hesitantes quanto à definição da avaliação da relevância (PGM-IMECC); consideram que seria necessário uma leitura mais atenta do recurso informacional para determinar com certeza a relevância (PGD-FEF); preveem que com o termo utilizado haverá muitos recursos informacionais recuperados/relevantes, devido à abrangência do mesmo (D-FEA); observam que a ferramenta desconsidera duplo sentido de um determinado termo (PDG-FCA); e que com a utilização do recurso gráfico aspas (“), a busca ficou mais precisa (PGM-IFCH, G-FEF e D-IFGW).

Os usuários consideram ainda ter sido muito específicos, pois com a utilização do recurso gráfico aspas (“) poucos recursos informacionais foram retornados (D-FEA); observam que com a inserção de outro termo na composição da palavra-chave houve um maior refinamento nas buscas (embora nesse caso não tenha sido utilizado o recurso gráfico aspas (“) (D-FE); que os documentos recuperados em português são teses e que em inglês são artigos (D-IFGW); que embora tenham buscado o termo em inglês, recursos informacionais em português também foram recuperados ou que já haviam sido recuperados e indicados como relevantes em busca anterior (PGD-IFCH); e que, embora a busca não tenha sido utilizada com o termo em português, houve retorno de vários registros em inglês (G-FEF). Em continuidade, mostram-se surpresos com o baixo número de recursos informacionais recuperados (D-IFGW), negativamente com os resultados de busca (PGD-FEF) ou com a quantidade de temas diferentes produzidos pela UNICAMP (PGD-IFCH); consideram que seu tema de pesquisa é bem específico, motivo pelo qual houve poucos recursos informacionais relevantes recuperados (G-IFCH, PGD-FCM e PGM-FEF); expressam que os pesquisadores da universidade

deveriam ter uma cultura de citar trabalhos da própria instituição, o que ocorre nos EUA e China (PGD-IFCH); indagam se teria sido interessante escolher termos mais específicos (D-IFGW); e manifestam que posteriormente consultarão os recursos informacionais relevantes de forma mais detalhada (PGM-IFCH, PGD-FCM, PGD-FCA e PGM-IMECC). Curiosamente, realizam a contagem dos recursos informacionais não relevantes em detrimento aos recursos informacionais relevantes (D-IMECC), comportamento comum dos demais usuários.

[busca por PRÁTICAS DISCURSIVAS E FORMAÇÃO] Nossa, 8. Imagina se eu não tivesse colocado o FORMAÇÃO. (É, então ((RI)) ia ser menos ainda ((RI))). Ia ser ... porque PRÁTICAS ia ficar aberto demais (D-FE)

[busca por violência] Tem muita coisa de VIOLÊNCIA DE GÊNERO, mas ... o meu é .. RACIAL. ...~~ Tem um que realmente me interessa. (...) ...~~ Um só (G-IFCH)

[busca por "ARQUITETURA BRASILEIRA"] É, bem menos. (...) ... Acho que quase nenhum aqui, só ...~~ só um (...) [busca por ICONOGRAFIA PAULISTA] (Esse tinha bastante). Eu vou procurar isso aqui depois ((RI)). (...) [busca por BRAZILIAN ARCHITECTURE] O mesmo trabalho ((RI)). ... (Conta também, tá, como se fosse uma outra busca). Tá (PGM-IFCH)

[busca por CONSUMO] Nossa, (...) eu escolhi uns termos muito, muito amplos. Quando eu faço pesquisa na base de dados eu sempre faço pesquisa combinada, então normalmente eu pesquiso CONSUMO e JUVENTUDE. E aí ele ... Porque CONSUMO aparece assim, desde ... fralda de neném até iogurte ((RI)). ((RI)). (...) ...~~ Vai muito pra galera da Economia, né. (...) [busca por CONSUMPTION] Como que eles ... colocam esse tipo de ... como que eles sabem, assim, como que eles elencam a relevância que eles me ... sugerem? (...) Porque ... muita coisa da Engenharia agora, né? (...) ... Engraçado, eles me dão coisas diferentes ((RI)). ...~~ Nossa, super da Engenharia ... eu achei que ia ser uma coisa mais ou menos ... Assim como o termo em português, né? (Exatamente). Eu achei que ia ter mais ou menos o mesmo tipo de ... de artigo. Mas até a área mudou ((RI)). Porque antes tinha bastante coisa de Economia mas bastante coisa de Ciências Humanas e aqui é praticamente tudo de ... tem alguma coisa de Economia mas praticamente tudo de ... de Engenharia, Engenharia de Alimentos e Engenharia Mecânica, what? ((RI)) Ah, tem uns em português ... (...) (Às vezes alguma parte do resumo, do abstract ou da palavra-chave tem essa palavra-chave que você pesquisou) ... Entendi. ... Um. ...~~ Esse aqui é o mesmo, aí eu marco de novo? (Não, marca um só). (...) Esse também de novo é o mesmo. ...~~ Seis. (...) [busca por JUVENTUDE] Nossa, é porque ... JUVENTUDE é uma coisa muito muito muito ampla. (...) ...~~ Meu Deus do céu ... (...) ... Tem de tudo, assim ((RI)). Eu to impressionada com a quantidade de ... temas diferentes ... que ... que a UNICAMP produz, sabe? ... Eu acho que a gente devia ter tipo ... uma ... uma cultura de citar trabalhos da própria universidade, sabe? Uma coisa que eu vi nos EUA e na China eles tem bastante, assim, tipo, você deve citar as pessoas do seu grupo de estudo, sabe assim? ...~~ Nossa, os JOVENS foi um mal assunto pra procurar. ...~~ Porque eu sempre faço nas bases de dados com busca conjunta, porque ... de todos aqui tem um que ... que ajudaria na minha tese, assim (...) [busca por BRICS] Nossa, esse não tem nem 50. Tem problema? (Não, não, é até 50). ...~~ É porque tem assim, desde a parte que me interessa, da parte da Sociologia até parasitas da malária, sabe? E ... isso é meio confuso ((RI)). E ajudaria nas pesquisas ... (Ficaria mais específica, né?) (...) Mas se a gente pudesse colocar só Ciências Humanas acho que já ... ajudaria muito. ...~~ É que quando eu procuro o mesmo termo dentro dos bancos de dados ... de vez em quando aparece, assim, uma planta, muito fora. ... (São as exceções, né?) É, é raro ter tipo "resistência humana de ... eritrócito" ((RI)). (...) Eu acho que faz todo sentido porque a UNICAMP tem muita pesquisa na área de Biológicas, sabe? (...) Por isso pra gente não se perder dentro do banco seria bom. ...~~ Quatro (PGD-IFCH)

[busca por TEORIA SOCIAL] 14 (D-IFCH)

[busca por SÍNDROME DE TURNER] 13 (G-FCM)

[busca por SAÚDE DO TRABALHADOR] (...) Aqui em Betim-MG eu sei que tem qarimpo então talvez seja uma fonte que eu possa dar uma. ... É um grupo bem específico, sabe? (...) Aí acha saúde do trabalhador, acha qualquer coisa menos do trabalhador ... da mineração nem se fala, mas trabalhador braçal, assim, é mais difícil, encontra talvez cortador de cana aqui na região ... ou trabalhador da área da saúde ... ou outros, nesse sentido, mas nunca em trabalhadores mais, de trabalhos mais remotos. ... Então esse seria um. ... Aqui tem outro que talvez eu consiga utilizar alguma coisa. ... Daqui eu conseguiria, talvez, usar umas três referências (...) [busca por FADIGA] (Tem até fadiga de materiais, né). Ah, a maioria. ... Vou até buscar porque achei dois interessantes ((RI)) **(PGD-FCM)**

[busca por "VACCINE, RESPONSE"] Ele reduziu muito ... só 14 entre aspas. ... Uma ... (É o mesmo registro?) É. ... Duas entre aspas **(D-FCM)**

[busca por EPILEPSIA] É, eu leria essa tese aqui, daria pra eu dar uma olhada. ... E mesmo não colocando em inglês apareceram vários em inglês já. ... 5 (...) [busca por "SÍNDROME METABÓLICA"] Nossa, quando coloca as aspas faz muito mais sentido. (Faz?) (...) Porque SÍNDROME separada aparece várias outras síndromes que não tem nada a ver. (...) 3 também **(G-FEF)**

[busca por ESPORTE ADAPTADO NA FRANÇA] "jogos na escola". ... Ah, esse aqui eu acho que sim ... esse aqui talvez ... é, teria que ler melhor, tá? ... Quatro (...) [busca por HANDSPORT IN FRANCE] O que entendem por HANDSPORT é outras coisas completamente diferentes. (...) ... Nossa, "constipação intestinal" ((RI)). ((RI)). ... "estrutura de proteínas" ... "mulheres com menopausa" ... ((RI)), o que que é isso ((RI)) ... "estudo de 8 casos" ... do que? ((RI)) ... "aromas e fragrâncias", nossa ((RI)), ... "colesterol" ... "corticoide" ... "influência do exercício" ... não. ... (Muita coisa da área da saúde, né). Na Biologia mesmo, olha, RNA ... proteína ... "síndrome", "um caso de síndrome" ... acho que não, né. ... Foi zero (...) [busca por EDUCAÇÃO FÍSICA] Essa coisa do ESPORTE NA FRANÇA é muito específico (...), às vezes eu procuro ESPORTE como palavra na FRANÇA mesmo e eu não consigo achar. (...) Ele ... mistura tudo ... então eu to pegando livros, coisas assim que eu trouxe de lá pra achar, porque ... porque aqui não tem, né. É, não tem muita coisa. E lá eles estudam bem essa parte de ... é bem específico nessa Federação, por isso que não sai muita coisa assim, entendeu? (...) ... E é muito prático também né. (...) ... É tanta coisa ... (É muito geral, né?) É. ... Só um. (...) [busca por "EDUCAÇÃO FÍSICA"] Só dois **(PGM-FEF)**

[busca por CIRCO] ((RI)) (...) (Você ainda não conhecia esse?) (...) ... "Veganismo e libertação animal: um estudo etnográfico" ((RI)). Até "Grupo Ginástico UNICAMP: 22 anos de ginástica geral" eu fico ... é ... dá ... mas ((RI)). ... Nossa, tá começando a ficar estranho, espera aí ((RI)) ... "trama poética entre literatura, narrativa, bordados e dança" ... (Foi CIRCO a primeira palavra?) Foi. ((RI)). (Então ... ((RI)), veja bem ((RI))). ((RI)). ... (...) Tá ficando engraçado ((RI)), "o papel educativo da fotografia de vida selvagem". ... Eu gostaria de saber ... qual a conexão que o sistema usou. ... (Então, eu também ((RI))). ((RI)). ... 7 **(PGD-FEF)**

[busca por INOVAÇÃO NAS CORTES] "corte em grafos" não ... esse aqui pode ser ... "alimentos processados" não (...) ele pega a palavra corte de cortar ((RI)). (...) Mas no inglês não tem esse problema. (...) ... Foi duas. (...) [busca por DESEMPENHO JUDICIAL] "judicialização" não ... (...) "o SUS no banco dos réus" não ... "desempenho psicométrico" ... não. ... Ah, esse daqui é interessante. ... Ok ... "cultura político-jurídico no sistema" seria interessante, dois. ... "Os modelos conciliatórios de solução de conflitos e a violência doméstica", não ... "justice in time and the time of justice", ah, deve ter a ver com a justiça no tempo, né, três. (...) ... Depois eu vou fazer isso de novo em casa ((RI)) pra procurar de novo ((RI)). ((RI)). Com certeza eu vou ler vários trabalhos aqui ... ((RI)). (Que bom que está te ajudando mesmo de alguma forma). É, porque eu não fazia essa busca no repositório da UNICAMP. ... (...) Então, três. Então é com aspas? [busca por LEGAL TECH] Nossa, "ensaio da odontologia legal na América Latina" ... "mudanças climáticas" ... esse aqui eu já contei. ... Não, um só. (...) É porque ele pesquisa separado as palavras, né? (...) Realmente um só (...) [busca por JUSTICE INNOVATION] "Terceirização" talvez **(PGD-FCA)**

[busca por ISOLANTES TOPOLÓGICOS] Tem umas coisas que eu não to entendendo porque tá aparecendo aqui. ((RI)). (Não tem nada a ver, né?) Não tem nada a ver. ... 1 ((RI)) **(G-IFGW)**

[busca por QCD] É, você vê que não é no título que ele vai, né, ele vai pegando coisas dentro, né. (...) ... 14 (...) [busca por SUPERSYMETRY] é que esse artigo tem vários autores ... é, então, tinha que aparecer de

grande colaboração, mesmo. ... É, esse daqui todos vão ser relevantes, assim, que é uma palavra meio que chavão, assim, entendeu? (...) *Aí SUPERSIMETRIA aqui, ó, tá no título. ... É que também não é uma palavra que busca ... uma busca mais ... restrita, né. (...)* *Talvez seria interessante colocar palavras que sejam mais restritas pra isso, né. ... (Qual seria mais restrito que esse, mais específico?) ... Uma coisa mais ... mais ... pode trocar uma aqui ou não? (Qual você trocaria?) Então, aí eu colocaria alguma coisa mais, é ... mais específico. (SUPERSIMETRIA com alguma coisa, é nesse sentido que você fala?) ... Especificar, por exemplo SUPERSIMETRIA mais, por exemplo, LHC. (Ah ... colocar duas palavras?) Duas. Aí eu acho que não, né. (Não). Porque a busca não tá em si errada ou ... “SUPERSIMETRIA, olha, SUPERSIMETRIA”, porque assim, a princípio esses artigos todos teriam uma relevância. ... (Mas todos os 50?) ... Ah, esse daqui não. Então vamos lá, vai. Esse sim. ... (..) É que é grande colaboração. (Sim, sim). (...) É. ... E o que ele encontra em português são teses. (...) ... Artigos sempre sai em inglês. ... Um, dois, três, quatro, vinte ... vinte e três (D-IFGW)*

[busca por SUPERSIMETRIA] *Só isso? (Só). (...)* [busca por NEUTRINO] *2014, 2006, qual é a ordem que ele coloca? (...)* *Quer dizer, que padrão doido, né? ... Esse daqui vai ser o mesmo pra português e inglês. (...)* *... Um, dois, três, quatro, cinco ... vinte e cinco, vinte e oito. (...)* *Ah, tem bastante, né. ... Agora esse daqui (...)* [busca por FORM FACTOR] *Você vê que não tem padrão, né, não é por ano. ... Foram três* [busca por “FORM FACTOR”] *Ele pega coisas diferentes do outro. ... Aqui são seis. ... Ele arruma o que, por ordem ... de ano? É aleatório? (D-IFGW)*

[busca por PREDIÇÃO ESPACIAL] *Aí ele tá mostrando o registro completo da tese. Eu queria o abstract, mas acho que daqui ... eu posso ir até o resumo? (Pode). ... Aqui, o resumo tem aqui. ... Maravilha, isso aqui inclusive acho que vai me ajudar de verdade ((RI)). ((RI)). Vai me ajudar, legal. ... Ah, aqui não tem resumo. (Essa não). ... Eu acho que eu vou fazer o seguinte ... pelo título, as palavras ... é ... eu não vou, eu não vou abrir isso aqui porque eu não sei se outras pessoas no seu caso abriram. (Não tem problema, não, de forma alguma, você tem que analisar se o documento é interessante ou não pra você, aí até onde você vai não tem ... é ... não tem problema nenhum. Às vezes como a tese é antiga, pode ser que ela não tenha o abstract). É, então ... bom, dois certamente me ajudam mas ... esse terceiro. ... Não tem resumo ... também não to a fim de ler ((RI)). ... Na verdade esse aqui (...) eu não sei se me ajuda ... pode ser que sim, pode ser que não, provavelmente não. Então eu considero como ... sim, baseado na minha percepção, né. Então vou considerar como um sim, porque ... eu tenho interesse em pelo menos ir mais a fundo, mas não sei se me ajuda, então ... três* [busca por “SPATIAL CLASSIFICATION”] *Agora eu quero com aspas aqui ... opa, achou. ... (Bom, é isso). É isso? É isso aí, mas eu não tenho certeza se todos realmente vão ajudar mas pelo menos são os que eu acho que ... poderiam né? É, talvez, uns certamente mas outros talvez (PGM-IMECC)*

[busca por PROBLEMA DE CORTE DE ESTOQUE] *Só tem um que não é, que não tá relacionado. (Mas é ... os demais, seria interessante pra sua pesquisa?) Sim, só esse aqui que não tem nada a ver. É “política de mudança climática com ênfase no termodinâmico” ... isso não tem nada a ver com PROBLEMA DE CORTE. O resto sim, tem coisa minha, coisa de colegas ... (..) (São documentos que você olharia, pra fazer um levantamento bibliográfico? É nesse sentido?) ... É que é assim, o problema é muito estudado, é enorme, eu não consigo, se eu tentar ler tudo o que tem ... não dá, então eu seleciono pra ver alguns. Mas todos são ... exceto aquele um. ... (Então vamos pensar assim, se você fosse selecionar quais desses você selecionaria ... quantos desses você selecionaria ... pra ler?) ... Aqui acabaram os 50? Assim, eu elimino dez. Os outros quarenta tem tudo a ver (...)* [busca por “PROGRAMAÇÃO LINEAR”] *5 ... então 45 (...)* [busca por PROGRAMAÇÃO INTEIRA] *Pouca diferença ... é que tem coisa, que assim, eles até usam porque isso é um método. ... (..) Mas não me é interessante. ... (Sim). 36 (...)* [busca por “CUTTING STOCK PROBLEM”] *38 também, porque tem dois que são do assunto mas que assim, não me interessa mesmo (...)* [busca por “COLUMN GENERATION”] *50 menos 14 ... 36, né, é isso? ... (É curioso porque você em vez de contar os relevantes você está contando os que não interessam, é isso? ((RI))). É, é ((RI)), é que como a maioria tem a ver eu achei melhor contar os ... ((RI)). Senão tem que ficar, um, dois ... então o que eu conto é menor ((RI)). (...) Aí é só contar o que tá fora (D-IMECC)*

[busca por ANÁLISE MODAL] (Quando você busca em conjunto aí você consegue especificar a sua área ... não vem nada da Medicina, geralmente, nas bases?) *Não, é porque geralmente eu só coloco no Google daí tem aquele Google Acadêmico, né, no Google Acadêmico é mais fácil. ... 7 (G-BAE)*

[busca por CATALISADOR] *talvez eu tenha colocado algumas palavras-chave meio genéricas, assim. ... ~ É porque geralmente o que eu faço? Eu ... pesquisei palavras-chave juntas, entendeu? Em conjunto. (...) Então, por exemplo, CATALISADOR é uma palavra-chave bem genérica ... mas se eu colocar junto de GASEIFICAÇÃO e FISHER-TROPCH ela já não é mais tão genérica e eu consigo filtrar melhor. (...) Nas bases de dados isso é mais comum, né. Aqui não dá pra colocar? Ai junto? (Então, até dá mas assim, teria que ser na mesma caixa de busca, porque ele ainda não tem essa opção de buscar dois assuntos diferentes, né, e trazer o mesmo registro que tenha os dois assuntos). ... ~ É bem genérico, né. (...) Só achei um ((RI))* **(PGD-BAE)**

[busca por SOBERANIA ALIMENTAR] *Oito* **(PGM-FEA)**

[busca por BIOFILM] *Eu tenho um problema quando eu vou buscar BIOFILME que ... existe BIOFILME em várias indústrias, né, e aí às vezes eu começo com BIOFILME e aí eu mudo às vezes BIOFILME eu coloco BIOFILM FOOD, sabe? (...) BIOFILME DE ALIMENTOS, porque aqui tá vindo tudo junto. É, pra refinar melhor, né. (...) ... ~ (Agora em português ...). Ah, eu já busquei o inglês (...) Vamos ver se dá a mesma quantidade ... português eu nem procuro ((RI)). (Tá ((RI))). Acho que quase ninguém tem costume de procurar em português, né* **(PGD-FEA)**

[busca por "PROTEÍNAS VEGETAIS"] *acho que eu fui muito específica* ((RI)). 2 (...) [busca por PROTEIN HIDROLYSIS] *Foram 11* **(D-FEA)**

[busca por EMULSÃO] *Nossa, EMULSÃO vai aparecer muitos, certeza. ... ~ CIMENTO? Nossa que engraçado ... (Às vezes pode ser que outras áreas usou esse termo né). É, esse termo né. Em um outro sentido. ... ~ (Muito genérico?) Muito genérico. ... ~ (...) O problema é a quantidade de resultados, né. ... É, eu fico pensando ... como que poderia ... essa é uma palavra bem difícil quando eu busco mesmo. ... ~ Normalmente quando eu faço busca ... é, igual, a Scopus permite você jogar uma palavra tipo EMULSÃO e aí tá 'buscar dentro' ... aí eu busco assuntos diferentes, EMULSÃO + alguma outra coisa. (Associa? Ele associa dois assuntos no mesmo registro?) É, porque agora por exemplo aqui dentro todos esses duzentos e pouco aí eu busco dentro desses daí ... aqui não tem essa ferramenta né? (...) ... ~ Que engraçado ... aparece sempre esses MÉTODOS QUIMIOMÉTRICOS, acho que é a terceira coisa, eu acho que eu ia clicar nele só pra ver qual que é a relação dele com EMULSÃO. ... (Às vezes eles aplicam esse método pra análise, né). (...) Olha, de novo. ... ~ Vai chegando pro final, parece que vai aparecendo umas coisas muito estranhas. (Já perde o foco, né?) É. ... ~ 35* **(D-FEA)**

Já outros usuários manifestam a hipótese de que não serão encontrados recursos informacionais relevantes com o termo de busca utilizado (G-FE, D-FE, PGM-IFCH, G-FEF, PGD-FCA, G-FEA, PGM-FEA e PGD-FEA).

[busca por INCLUSION] *Nossa, acho que só vai ser um ... é minha impressão, né. (...) Ainda tá no começo mas vira ... pura Biológicas, né. (...) [busca por ESCOLA] Esse não vai sobrar nenhum. (...) (Por que você acha?) Porque é muito geral. ... Mas é legal ver que os primeiros que aparecem são de professores da FE. (...) [busca por DIFERENÇA] É no singular mesmo, né. (Isso). (...) Esse eu acho que vai ter mais, né, porque é termo filosófico, aí vai bastante direto pro autor, né. (...) Ou não ... já foi diminuindo, mudando os caminhos da palavra* **(G-FE)**

[busca por "AUTHORSHIP AND TEACHER DEVELOPMENT"] *To achando que não vai encontrar. "Não há resultados"* **(D-FE)**

[busca por MIGUEL DUTRA] *Eu acho que ... não vai ter nada. (Não?) Impressionante como ... pega isso, ninguém chamava ele assim. (O MIGUEL DUTRA?) É ... (...) Miquelzinho, chamavam ... impressionante ... zero* **(PGM-IFCH)**

[busca por COMORBIDADE] *Eu acho difícil aparecer algo procurando só COMORBIDADE, porque nunca fala só de uma nos estudos, sempre tá falando mais ... e o que eu estudo também, são várias, (...) são*

várias COMORBIDADES que ... definem a SÍNDROME METABÓLICA. Então tipo, pra mim se aparecer só uma não seria interessante. (...) ...~~ E eu também acho que as pesquisas no geral elas não vão querer avaliar uma coisa só, quando elas podem fazer mais de uma pra pesquisa se tornar relevante, entendeu? (G-FEF)

[busca por "INOVAÇÃO NAS CORTES"] acho que não tem (PGD-FCA)

[busca por FERMENTAÇÃO] Nossa, esse vai vir muito ... (Vai?) FERMENTAÇÃO, minha nossa. ...~~ Não, até que não ... tinha mais em ANTIOXIDANTE. ...~~ "Camas de frango", o que seria isso? (...) ...~~ Já esperava mesmo, porque é muito geral (G-FEA)

[busca por "SEGURANÇA ALIMENTAR"] O próximo eu acho que não vai ter resultado ((RI)). (Não? ((RI))). Não sei, vamos ver ((RI)). ((RI)) (PGM-FEA)

[busca por BACILLUS SPOROTHERMODURANS] Essa bactéria quase não tem ... documento (PGD-FEA)

Os usuários expressam que os recursos informacionais recuperados não estão relacionados com seu tema de pesquisa (G-FE, PGM-IFCH, PGD-IFCH, D-FCM, PGD-FCM, D-FCM, G-FEF, PGM-FEF, PGD-FEF, G-FCA, PGM-FCA, PGD-FCA, D-FCA, PGM-IFGW, D-IFGW, G-IMECC, PGM-IMECC, G-BAE, PGM-BAE, D-BAE e G-FEA); observação que nos recursos informacionais recuperados em inglês houve um menor número da *Área de Artes e Humanidades* (G-FE); termo muito genérico, o que aumenta muito a imprecisão dos resultados (PGM-IFCH, D-FCM, PGD-FEF e D-FCA) ou muito específico (PGM-FEF); necessidade de busca booleana para refinamento das buscas (D-FCM, D-FCA e G-BAE); ausência de resultados de busca utilizando-se ou não recurso gráfico aspas (") (D-FCM, PGM-FEF, PGD-FEF, PGD-FCA, D-FCA, PGM-IFGW, G-IMECC, PGM-IMECC e PGM-BAE); ausência de relação direta com o tema de pesquisa (G-FEF, PGM-FEF, PGD-FEF e D-BAE) ou em outras áreas (G-FCA, G-BAE e G-FEA); possibilidade de dois termos para o mesmo conceito (G-IMECC); recursos informacionais muito antigos (G-FEA); ferramenta de busca que varre em todos os registros, sem refinamento de assunto (D-IFGW); recursos informacionais de tipologias muito diferentes (apresentações) (PGD-FEF); e observação de que as buscas não estão sendo realizadas especificamente por assunto, sugerindo que a ferramenta informe aos usuários quando estão realizando busca por assuntos (PGM-BAE).

[busca por DIFFERENCE] É curioso ver que as coisas ... traduzidas ... quase não são das Humanas, né ... tem pouquíssimos. Tem bastante da Psicologia. ...~~ Mas nas outras também, menos na escola, né, mas na INCLUSÃO quando eu vi em ... inglês começou a vir uns ... beta ((RI)). Da área de Biológicas, né, cromossomo (G-FE)

[busca por YOUTH] Mas é porque ... a produção é tanta, tanta, assim. ...~~ (...) É um termo tão amplo que tem desde coisas lá da Medicina, da Educação Física, da Educação. ...~~ Não tem nenhum ((RI)). (Tudo bem, é só colocar o zero) (PGD-IFCH)

[busca por MINING] (Não teve nenhum relevante porque?) É que eu preciso focar no que eu quero, né? Eu preciso buscar a mineração que trabalha ou que envolve mineral característico que causa silicose, porque quando se fala em mineração, principalmente no Brasil, é muito abrangente o termo, porque tem desde a mineração ... industrial, que é enorme, enfim, ... por exemplo da Vale mas o que eu busco é aquela mineração ainda bem manual, que é aquela que vai causar a doença que eu procuro usar (PGD-FCM)

[busca por CRIANÇA] Agora por exemplo, ... CRIANÇA é um termo muito vasto, né. (...) Eu faria ele como booleano, como eu utilizo por exemplo o MESH ... imagino que aqui vai ter muita coisa, né. (...)...~ Nessa vieram 5.800 registros. ...~ Sem o booleano fica muito difícil, né, porque é um tema muito vasto, né, essas palavras-chave que eu escolhi. (...) Mas realmente, pelo termo CRIANÇA, que é um termo muito amplo, né, os primeiros 50, zero (...) [busca por ADOLESCENTE] ADOLESCENTE. ...~ O grande problema é não cruzar, né. (É). ... Aí fica quase como uma questão aleatória acabar encontrando. ...~ Esse também, ADOLESCENTE deu zero (...) [busca por “TERAPIA ANTIRRETROVIRAL”] Se eu pudesse associar o VACCINE RESPONSE ... ‘and’ HIV, né (...) aí ... provavelmente teria ... (...) Tá bem parecido. ...~ Ah, tá, fica zero (...) [busca por “VACCINAS, RESPOSTA”] Entre aspas inclusive ele não localizou (...) [busca por HIV] Então especificamente para minha pergunta, nessa primeira página nenhum. (Nenhum?) Nenhum. (...) Para a minha pergunta, né? (Da sua pesquisa) Isso, que a minha pesquisa é RESPOSTA A VACINAS EM CRIANÇAS E ADOLESCENTES INFECTADOS POR HIV EM TERAPIA ANTIRRETROVIRAL (D-FCM)

[busca por ELT] (Teve algum?) (...) Não, né. ... Tinha vários estudos sobre EPILEPSIA DE LOBO TEMPORAL, mas nenhum é da minha área, assim (G-FEF)

[busca por REABILITAÇÃO] Isso daqui tá ... muito específico (...) “do ângulo” ... nada a ver ... “lombalgia”. ...~ Essa REABILITAÇÃO entra na parte de aparelhos. ...~ (Nada?) Nada (PGM-FEF)

[busca por “TEACHING/LEARNING”] A relação ENSINO/APRENDIZAGEM me interessa mas aí vem no EQUILÍBRIO QUÍMICO. ... (Ah, sim). ((RI)). ((RI)) (...) [busca por “STREET’S CIRCUS”] Não há resultados (PGD-FEF)

[busca por REAPPLICABILITY] (Nenhum desses tem a ver com ...) Não porque tem umas coisas de ... Biologia, entendeu? ((RI)) (...) Eu fiz um recorte na minha pesquisa sobre experiências de tecnologia social na modalidade reciclagem que gera renda (G-FCA)

[busca por ADSORÇÃO] Zero. (...) [busca por ADSORPTION] ...~ Zero (...) [busca por BIOLUBRIFICANTE] Dois artigos, só tem dois trabalhos no repositório e os dois não em interessam ((RI)). (...) [busca por BIOHUBRICANT] Os mesmos dois? Um é, o outro ... esse aqui eu acho que não ... nenhum (PGM-FCA)

[busca por “LEGAL TECH”] Zero (PGD-FCA)

[busca por SECAGEM] É, secagem ela acaba sendo uma palavra-chave muito geral ... no momento que ele tiver a combinação vai ser bacana, porque aí a gente consegue pôr ... mais específico, né. ... não sei se é um passo futuro ((RI)). (Espero que sim ((RI))). ...~ Nenhum. Mesmo tendo Engenharia Agrícola não saiu nada de ... MADEIRA, mesmo que saísse MADEIRA, BAMBU eu poderia dar uma olhada. (...) Bom, essa palavra não (D-FCA)

[busca por NANORRESONADORES] Não tem nada. (Pode pôr zero) (PGM-IFGW)

[busca por FATOR DE FORMA] Pois é, em artigos as palavras-chave ficam em inglês ou em português, as palavras-chave são traduzidas quando elas são colocadas no repositório? ...~ Então ... tem que colocar aspas, você disse, né? (...) ...~ Mas aí é que tá, FATORES E FORMA ele tá buscando, aqui. (Sim ((RI))). Ele não tá pegando a palavra composta. (...) (Ele pega, pelo menos no repositório, ele pega a composta quando a gente põe entre aspas). ...~ Ele pega separado ... não tem nenhum. (É só colocar o zero na frente e agora vamos buscar entre aspas) (D-IFGW)

[busca por MÉTODO DAS SOMAS PONDERADAS] O problema desse aqui é que esse tem dois nomes diferentes ... (Você fala como ...) MÉTODOS DAS SOMAS PONDERADAS ou ESCALARIZAÇÃO. (Deixa assim que aí depois a gente vai consultar na linguagem pra ver qual eles indicam como padronizado. ... É só pôr zero) (...) [busca por “ε-CONSTRAINT METHOD”] Nada (G-IMECC)

[busca por CLASSIFICAÇÃO] (Quando o termo é ... sozinho você não precisa colocar entre aspas). *Hum, beleza.* (Porque as aspas significam que você quer pesquisar um conjunto de termos ... aquele bloquinho, entendeu?) É, porque essa palavra é muito ampla ... Então, realmente, ela não faz sentido nenhum sozinha, vai ver tudo o que é área do conhecimento, mas quando eu junto essas duas isso aqui [referindo-se à CLASSIFICAÇÃO ESPACIAL] faz sentido, esse aqui não faz, então, (...) É que o problema é meu mesmo, não é nem da ferramenta ... se eu tivesse percebido a tempo ... (Não tem problema, buscas é assim mesmo, a gente vai ... se aperfeiçoando). É, redefinindo, né. (...) ... É que aqui ... cai de tudo o que é área. ... É. aqui é zero também (PGM-IMECC)

[busca por SMARTPHONES] Tem muita coisa da área da Saúde, né, nessa base ... pelo menos com as três que eu já digitei. Nenhum (...) [busca por “DYNAMIC EVALUATION”] (Continua muita coisa da Medicina?) *Muita. (...)* Quase tudo praticamente mas é engraçado porque mesmo esse também deu muita coisa da Medicina. ... Não, sei, talvez precisasse procurar eles em conjunto, quando eu fazia revisão bibliográfica era sempre em conjunto. (...) Muita coisa da Medicina, muita (...) (G-BAE)

[busca por “ARQUITETURA RESIDENCIAL”] É, quatro. Mesmo assim é engraçado, algumas teses que eu vi que eu tenho certeza que tem esse termo elas não estão aparecendo aqui ((RI)). (Então ((RI))). É, porque, olha ... aqui não tem nada ((RI)). Não tem nada ((RI)), então, eu coloco ... põe zero? (Põe zero) (...) [busca por DECIO TOZZI] Ele procura tudo, né ... É o que você falou, né, olha eu digitei DECIO TOZZI mas aqui foi só no ... o que eu digitei só apareceu na verdade nos autores, quer dizer, até agora ele não é assunto de nada do que tá aparecendo aqui, entendeu? ... (Apesar de estar assunto, né). Então ... olha, tá vendo, nossa, imagina, não tem nada a ver ((RI)). (...) Porque nossa, coisa de Medicina, coisa de Química, coisa de ... ((RI)) não tem nada a ver. Teria sentido buscar nos autores, né. (É, aí tudo bem né. Mas o certo seria não aparecer nada). ... (Mesma coisa). Olha ... DECIO ... DECIO ... mas é, não mudou. Se bem que aqui, engraçado né, eu acho que ele tinha que mostrar também que tá por assunto ... ah, se bem que não, né, não sei, ele tinha que mostrar em algum lugar que tá por assunto, sabe? Que né, mostra, realmente, a gente só tem certeza que tá por assunto se a gente voltar toda vez na página inicial. (...) Então, bom ... então eu vou pôr zero (PGM-BAE)

[busca por COMPÓSITOS DE MADEIRA] Aqui é uma outra linha de compósitos mas não especificamente essa (D-BAE)

[busca por CASCA DE MANGA] Eu não sei o que é isso mas eu acho que não é manga. ... Ai, nossa, não dá pra pesquisar ((RI)). (Datilografado, até ((RI))). Ah, eu nem uso, minha orientadora falou que eu não uso coisa muito antiga, então a gente nem usa. ... (Antigo quanto assim?) De 1952 não dá né ((RI)). (Ah, sim ((RI))). Ah, quando eu fiz o projeto tinha uma referência que era de 1986, dos anos 1990 ele já falou que não é bom, não é pra colocar muito antigo. É bom colocar mais ... atualizado. ... Esse aqui pode ser ... CONCRETO! O que tem a ver? ((RI)). (...) 2 (G-FEA)

Um dos usuários confirma novamente os recursos informacionais julgados por ele como relevantes na página de visualização dos resultados (D-FE).

[busca por “PRÁTICAS DISCURSIVAS E FORMAÇÃO”] “Não há resultados de busca”, é isso mesmo né, PRÁTICAS DISCURSIVAS E FORMAÇÃO. (...) [busca por DISCURSIVE PRACTICES AND FORMATION] (...) Não veio ((RI)). DISCURSIVE PRACTICES AND FORMATION. Nossa! Tá, zero. É, zero (D-FE)

Embora não localizem exatamente os recursos informacionais que poderiam ser relevantes para suas pesquisas, os usuários explicitam interesse em outros recursos informacionais recuperados (D-FCM, G-FEF e PGD-FEF).

[busca por CRIANÇA] Olha, uma coisa interessante, esse aqui, “Rousseau e a primeira infância”, que bacana (...) [busca por CHILD] CHILD veio um pouco menos. ... No final das contas ... o que a gente

acaba vendo é que a vida é curta pra tanto assunto interessante, né ((RI)). (Sim, ((RI))). ...~~
Especificamente aqui também ...~~ seria zero, de novo **(D-FCM)**

[busca por COMORBIDADES] Tem até coisa aparecendo que não tem nada a ver com a minha pesquisa mas que eu to com vontade de ler ((RI)). ((RI)). ... Mas isso não conta, né? (Não). ... Porque não tem nada a ver. (Tenta direcionar pra sua pesquisa). ...~~ Ah, que legal isso, “semente e do óleo de chia na prevenção da obesidade” ((RI)) **(G-FEF)**

[busca por STREET] Esse não tem nada a ver com a minha pesquisa mas eu leria só por causa do título ((RI)). (São criativos, né ((RI))). ((RI)) **(PGD-FEF)**

Os usuários comparam os recursos informacionais recuperados com um determinado termo de busca com recursos informacionais recuperados em buscas anteriores (G-FE, D-FE, PGD-IFCH, PGD-FCM, D-FCM, G-FEF, PGD-FEF, PGM-FCA, PGD-IFGW, G-IMECC, D-IMECC, PGM-BAE, D-BAE e D-FEA). De forma específica, expressam resultados diferentes com o termo em inglês do que em português (D-FE, D-IFCH, PGD-FCM, D-FCM, PGM-FCA, PGM-IMECC, PGM-BAE e D-FEA); quantidade menor de recursos informacionais utilizando-se o recurso gráfico aspas (“) (G-FE, D-FE, PGD-IFCH, G-FEF, PGD-FEA, D-IMECC e D-FEA), até mesmo conferindo se a digitação foi correta (D-FE) ou que o número de recursos informacionais relevantes aumentou em comparação à busca realizada com o termo sem utilização do recurso gráfico aspas (“) (D-FE) ou se manteve bem parecido (D-IFCH); comparação entre dois termos (G-FE, D-FE, PGD-FCM e D-FEA); e observação de que houve associação da ferramenta utilizando-se dois termos em inglês, isto é, nas duas buscas houve registros que foram recuperados com os dois termos (D-IFCH) e da importância da realização das buscas utilizando-se o recurso gráfico aspas (“) (G-FEF).

[busca por “EDUCAÇÃO INCLUSIVA”] Esse apareceu na DIFERENÇA também. (...) Esse apareceu quando eu pus a palavra DIFERENÇA (...) [busca por “SPECIAL EDUCATION”] Acho que eu vou voltar, acho que perdi a conta. (...) Porque esse foi um monte seguido. ...~~ Aqui apareceu muitos que não tinham aparecido **(G-FE)**

[busca por TEACHER DEVELOPMENT] Então o primeiro é TEACHER DEVELOPMENT. Vai dar, quer dizer, não é problema, é que a maioria é TEACHER TRAINING mas daí tem a ver exatamente com APPROACH. Vamos ver o que vai rolar, até to curiosa ((RI)), porque é TEACHER TRAINING. (...) Ah, não, veio bastante ((RI)). ...~~ Nossa (...) [busca por “TEACHER DEVELOPMENT”] É, daí tem a ver com a ... interessante! ... que daí quando fizeram pro inglês, né (...) o que que rolou, legal. ...~~ Que pouquinho ... é a mesma, então conta uma vez, né. (Isso). Conta uma vez (...) [busca por “PRÁTICAS DE LEITURA E DE ESCRITA”] (Mesmo se forem registros que já apareceram na sua pesquisa ...) ... Ah, tiveram vários, vários. (Você conta). Sim, eu to contando a mesma coisa, sim, isso tá acontecendo direto. (Isso ((RI))). To contando sim, eu não to descartando. Redundância, tá acontecendo. ...~~ É, eu pensei nisso agora ((RI)). ((RI)). ... Assim como aparece coisas, quer dizer, quando a gente vai vendo ... aparece coisa nova também, que nem esse aqui não tinha ...~~ “PRÁTICAS DE LEITURA E DE ESCRITA” ... (É ... pode ser, por conta do ‘de’). ... Ah, por causa do ‘de’? É uma coisa técnica, então olha. ... (É, porque ele busca exatamente esse bloquinho de palavras). Então ... 1, é, 1. Nossa, que estranho, entendi. ... Ou seja, claro que eu uso (...), o repositório eu não uso (...) mas só pra saber ... esse ... esse de, quer dizer, ele ... reduz ... , é isso? (É, reduz) (...) [busca por

“READING AND WRITING PRACTICES”] *E o quanto não é sinônimo mesmo, LETRAMENTO, quer dizer, embora teoricamente seja. ... Uh!! É isso mesmo, né, tá certinho. ... Olha como cai no negócio da palavra-chave, né (...)* [busca por ACADEMIC LITERACY] *Pra mim o inglês foi maior que em português. ... Uh! ... Todas* **(D-FE)**

[busca por “MEIO AMBIENTE”] *tem umas coisas diferentes, né? ... Tem umas coisas similares assim também* **(PGD-IFCH)**

[busca por “PENSAMENTO SOCIAL”] *Tá bem parecido (...)* [busca por SOCIAL THOUGHT] *Esse veio ... é o mesmo, eu acho (...)* [busca por GÊNERO] *9. É engraçado, né, GÊNERO aparece coisas da Biologia neste mesmo ((RI)). Agora com as aspas, né? (Não, esse não precisa, porque como é uma palavra só). ... Ah, tá, entendi. (Porque não ia fazer diferença a gente colocar com aspas ou não). É, exatamente ((RI)) (...)* [busca por SOCIAL THEORY] *Veio muita coisa da Medicina, da Saúde Coletiva, sabe? ... 12. E é engraçado, esse combinou em inglês, com as aspas combinou com algumas coisas que apareceram no GENDER, algumas coisas apareceram iguais. (...)* (E alguns registros foram os mesmos?) *Sim ... eles associaram os dois termos em inglês, nos termos em português não. Não teve associação* **(D-IFCH)**

[busca por “OCCUPATIONAL HEALTH”] (Foi o mesmo documento?) *Não foi, aqui apareceu um do professor Bagatin, que não tinha aparecido quando eu busquei em português (...)* [busca por SILICOSIS] *Agora apareceu 13 e mais ou menos os mesmos (...)* [busca por PNEUMOCONIOSIS] *Dez ... teve alguma diferença, apareceu, é, gripe, também, apareceu ... pneumonia, que não tinha aparecido antes* **(PGD-FCM)**

[busca por ADOLESCENT] *Alguns ... apareceram nas primeiras 50 em português, né, e outros agora em inglês também, então tem alguns que estão ... nas mesmas ... entre os primeiros 50, né. (...)* ... *Mas ... especificamente pra esse não* **(D-FCM)**

[busca por “FORMAÇÃO INICIAL DE PROFESSORES”] *18. (...)* *Aumentou, né* **(D-FE)**

[busca por “ATIVIDADE FÍSICA”] *O mesmo. (...)* *Quer dizer que não teve diferença, né? ... Vieram menos registros ainda do que ...* (Porque assim, quando você coloca entre aspas, o que que ele vai fazer? Pelo menos nessa ferramenta, né. Ele vai trazer ... ATIVIDADE FÍSICA junto. ... Ele vai buscar esse bloquinho de palavras. Agora se você põe separado ele pode buscar ATIVIDADE num lugar, FÍSICA no outro). ... *Ah, entendi. Então é importante, né, colocar entre aspas (...)* [busca por “PHYSICAL ACTIVITY”] *Apareceu a mesma anterior e uma nova que tinha a ver com EPILEPSIA, mas não ... não é interessante. (Tudo bem, é só colocar zero)* **(G-FEF)**

[busca por ENSINO/APRENDIZAGEM] *Vamos ver como isso vai sair ((RI)). ... (A barra acho que tá invertida, não tá?) É, eu acho melhor mudar porque tá vindo barra, tá vindo traço ... vamos arrumar então. ... Que engraçado, parece que ... ((RI)) ... mudou ... pra pior ((RI)). ((RI)). Pelo menos assim os dois primeiros que eu olhei naquela tinha a ver, agora ((RI)). ... Que louco, né* [busca por “ENSINO/APRENDIZAGEM”] *Esse já tinha vindo no outro ... é ... basicamente ... basicamente os mesmos. (...)* *Em outra ordem mas ... basicamente a mesma coisa. ... Tinha algumas diferenças mas ... vieram coisas que não me interessaram de qualquer forma. (Deu o mesmo então?) É ... os mesmos trabalhos inclusive ((RI)). ... Agora o inglês ... (Acho que ele busca os dois juntos mesmo (...)* [busca por CIRCUS SCHOOL] *Tipo ... zero? ((RI)). (Não sei ((RI))). ... Olha, até que veio mais até ((RI)). ... (É o mesmo?) É, esse daqui já tinha aparecido em outro mas também não me interessa* **(PGD-FEF)**

[busca por IMMOBILIZATION] *Praticamente são os mesmos artigos, parece. (...)* *De inglês pro português. O que não me interessou aqui é que a maioria dos artigos é voltado pra usar na cromatografia, que é uma outra coisa que usa pra imobilização* **(PGM-FCA)**

[busca por “JUDICIAL PERFORMANCE”] *“redução de cesáreas desnecessárias” não ... (accountability) ... ah, esse aqui é, então um. (Esse já tinha aparecido em outras buscas?) Já, esse já. Olha, “análise da competência do Conselho Nacional de Justiça”, esse eu já tinha visto antes* **(PGD-FCA)**

[busca por “ISOLANTES TOPOLÓGICOS”] *É o mesmo* **(G-IFGW)**

[busca por HEAVY FERMION SUPERCONDUCTORS] *apareceu bem mais em inglês do que em português. (É?) Apareceu. ...* (Na sua área as publicações são mais voltadas pra inglês, não é?) *A maioria é, a gente acha pouca coisa em português (PGD-IFGW)*

[busca por “PROBLEMA DE CORTE DE ESTOQUE”] *Fora que se aparecer o mesmo da outra vez eu tenho que ...* (Conta de novo. É, como se fosse uma outra pesquisa). *... Agora só apareceu 22. (Não tem problema ... é que só tem 22 mesmo). Ah, tá, deu 7 também (G-IMECC)*

[busca por SPATIAL PREDICTION] *saiu o mesmo, mesmo eu tendo colocado em inglês, eu conto? (Conta). ... Esse aqui também foi a mesma que o outro, só que agora com o título em inglês, eu também conto, né. (...)* *Mesmo sendo as mesmas, né. Isso. ... É, são dois, são os mesmos outros dois (PGM-IMECC)*

[busca por “PROBLEMA DE CORTE DE ESTOQUE”] *Agora vieram só 22. ... É, são os 22 que eu já tinha indicado (D-IMECC)*

[busca por “RESIDENTIAL ARCHITECTURE”] *Ah, então, é engraçado porque é uma tese ... que é em português, que foi uma orientanda da minha professora, né, que foi publicada agora, uma dissertação e ela só apareceu na busca por inglês, mas ela tá em português. Tem esse termo no título? (No título não). Ah, tá. (Tem dentro da obra, entendeu?) Mas eu achei interessante, engraçado isso porque na pesquisa em português ela não apareceu, entendeu? Mas na pesquisa em inglês ela apareceu. ... E é recente, aí eu não sei se é por causa disso, às vezes, né. (Aí a gente tem que dar uma olhadinha nas palavras-chave que foram colocadas). ... Entendi (...)* [busca por GRAPHIC ANALYSIS] *Engraçado que são 5 diferentes, deu 5 mas não são as mesmas. (...)* *Os documentos, alguns são iguais e alguns são outros documentos. (Você usando com ou sem aspas?) Sem aspas. (Em inglês e português?) Em inglês e em português. Um deu alguns documentos e outro deu alguns, dois similares e três que eram diferentes (PGM-BAE)*

[busca por WOOD WITH STEEL BARS] (Os documentos tem se repetido?) *... É ... basicamente o meu trabalho e mais uma ... dissertação também. ... É ... você diz em uma mesma página ou ... (Nas buscas diferentes) ... É, basicamente sim, basicamente. É ... esse meu artigo praticamente aparece em todos ... ele por exemplo apareceu em todos. (...)* *... (Ele é bem específico nesse tema mesmo, né). Exato, exato (D-BAE)*

[busca por “MANGO PEEL”] *É que são os mesmos anteriores. (Não tem problema, é como se fosse uma outra pesquisa). ... Um (G-FEA)*

[busca por “BACILLUS SPOROTHERMODURANS”] (...) *Nossa, parece que veio menos ... (Por que, o que que faz? Quando você busca entre aspas ele vai trazer exatamente aquele bloquinho de palavras, um seguido do outro (...), agora ... quando você coloca sem aspas, mais de uma palavra é, ele pode buscar um termo em um lugar, outro termo em outro ...). Por isso que veio vários de BACILLUS CEREBUS, outros BACILLUS ... Entendi. Já to tendo uma aula também ((RI)). (É ... ((RI))) (PGD-FEA)*

[busca por HIDRÓLISE DE PROTEÍNA] *Esse parece que aparece mais, né. ... 14 (...)* [busca por PROTEIN HYDROLYSIS] *Em inglês aparece bem mais (...)* [busca por VEGETABLE PROTEINS] *Parece que foi os mesmos 56. ... (Os mesmos, você acha?) Parece (D-FEA)*

Os usuários expressam o conhecimento dos autores dos recursos informacionais recuperados (G-FE, D-FE, PGM-IFCH, D-IFCH, G-FCM, D-FCM, G-FEF, PGM-FCA, PGD-FCA, D-FCA, D-BAE, G-FEA, PGM-FEA e PGD-FEA), orientadores (G-FE, G-FCM, PGM-FEA, D-BAE e PGM-FEA), orientandos (D-FE) ou mesmo autoria própria (D-IFCH, D-FCM, PGD-FCA e D-BAE). De forma específica, um usuário observa que alguns recursos informacionais são de autoria de sua orientadora, embora tenha suposto que haveria poucos resultados satisfatórios na

busca por se tratar de busca simples e não combinada utilizando-se o termo BURKHOLDERIA (PGM-FCA). Outro usuário manifesta conhecer o trabalho do qual participou em banca (D-IFCH), enquanto outro reconhece trabalhos dos quais participou em seu grupo de pesquisa. Salienta que o termo TERAPIA ANTIRRETROVIRAL pode ser encontrado sem o hífen (moderna grafia) ou com o hífen (antiga grafia) (TERAPIA ANTI-RETROVIRAL). Além disso, indaga sobre os critérios de ordenação da relevância pelo sistema e acredita que sem a opção de cruzamento entre os termos, ou seja, sem a busca combinada, a recuperação se torna muito difícil, já que é o cruzamento entre os termos que propicia o sentido (D-FCM). Um usuário percebe a opção de *busca facetada* por *Assunto* assim que a página com os resultados da primeira busca é carregada, porém, não utiliza este recurso para o refinamento das buscas (G-FEF).

[busca por EDUCAÇÃO ESPECIAL] *Nossa, eu nunca li, apareceu a tese da minha orientadora ((RI)). (...) Velha, já ((RI)). ...~~ Eu não usaria agora. (É antiga?) É. Ela já é aposentada, já. É que na Faculdade ela está como professora colaboradora (G-FE)*

[busca por FORMAÇÃO INICIAL DE PROFESSORES] *Olha, um que eu orientei aqui tá na primeira. ...~~ Acabou? (Acabou). 14. (...) Tenho uma orientanda minha que ela tá aparecendo direto em vários trabalhos ((RI)), é bom, né ((RI)). (...) [busca por "PRÁTICA DE LETRAMENTO"] *Olha eu aqui ((RI))* (D-FE)*

[busca por ARQUITETURA BRASILEIRA] *Apareceu a tese do meu orientador, a dissertação dele ((RI)) ... aqui, né, a tese do Chico que eu não conheço, que é interessante ... a tese do Professor Roberto que eu também já conheço, muito boa, também interessaria ... (...) A tese do André, também já conheço ... é importante ... a livre docência do meu orientador, super importante também, tem muito a ver ... ah não, é um artigo, né, esse é um artigo, tem a ver com a livre docência ... esse aqui eu já conheço também. ... Foram três. ...~~ Isso aqui também é muito importante. ...~~ É ... seriam quatro, cinco só (PGM-IFCH)*

[busca por PENSAMENTO SOCIAL] *Engraçado, parece coisas que eu fui banca ((RI)). (É? ((RI))). ...~~ (...) 23* (...) [busca por "PENSAMENTO SOCIAL"] *Encontrei o meu próprio trabalho ((RI)). (...) 21* (D-IFCH)

[busca por SÍNDROME DE TURNER] *É minha orientadora essa, inclusive ((RI))* (G-FCM)

[busca por HIV] *Já apareceu um que eu to ((RI)). (...) Fico feliz, é gostoso ver ((RI)). (...) Olha mais um meu aqui ((RI)), que bom. Meu não, né, que eu faço parte. (...) [busca por TERAPIA ANTIRRETROVIRAL] *Aqui às vezes tá com hífen ... a moderna grafia seria sem hífen. ...~~ Vieram 2.700. ...~~ Olha, o primeiro é uma tese que eu orientei ((RI)). ((RI)). ...~~ Interessante ter aparecido, não sei qual que é a hierarquia ... Como ele ordena, né? (...) (A ordem? Também não sei ((RI)), é um dos pontos da pesquisa, to tentando descobrir ((RI))). ... Mais um! ((RI)). ((RI)). Mais um! ((RI)). Três, e chegaram rápido aqui! ((RI)). Quatro! ((RI)). Os primeiros dez ... cinco! ((RI)). ... Seis! ((RI)). Tudo artigo que eu orientei ((RI)). ...~~ Olha, curioso aqui, só por curiosidade, dos primeiros dez ... sete são ... é ... teses que eu orientei. (...) ...~~ Mais um! ((RI)). ...~~ Outro! ((RI)). ...~~ Pra mim tá sendo bom, porque sair entre os primeiros ((RI)). ((RI)). ... É que é a minha linha de pesquisa, né ((RI)). (...) Outro! ((RI)). ...~~ Outro ... outro ... mais um. Temos 12 aqui dos 50. ...~~ Mas realmente para esta pesquisa específica ... nenhum. ... Fica difícil sem o cruzamento. ...~~ Fica difícil sem o cruzamento, né, porque é o cruzamento que dá os sentidos, né. (...) O 'and', o 'or'. (...) [busca por ANTIRRETROVIRAL THERAPY] *8.000 ... mais alguns ... que eu participei.* ...~~ É, pra mim está sendo interessante ver que tem uma amostra bem considerável de ... de trabalhos do nosso grupo aqui. (...) [busca por VACINAS, RESPOSTA] *Tem um que eu participei mas esse não era com HIV. ...~~ Minha tese de mestrado. ... (Foi sobre o mesmo tema?) Foi com VACINAS, mas não com HIV. (...) Três. ...~~ Em alguns já vem o link pro próprio arquivo do artigo. (Isso) (D-FCM)***

[busca por EPILEPSIA] *E aqui já apareceu, né, alguns ... resultados relacionados. (...) É, eu conheço uma professora aqui. (Conhece? ...) É a sua orientadora?) Não ... mas ela estuda a parte de neuro também (G-FEF)*

[busca por BURKHOLDERIA] *Minha orientadora. (...) Esse trabalho também é dela. (Você já conhecia?) Já, os dois trabalhos já, esse aqui não me interessa, esse aqui me interessa. (...) ...~~ Minha orientadora de novo. (...) Então ... não vai ter muita coisa ((RI)), imaginei que não ia ter muita coisa como busca, só uma palavra ... Esse trabalho aqui é interessante, que é da minha orientadora, eu conheço (PGM-FCA)*

[busca por DESEMPENHO JUDICIAL] *Sem aspas, ir. (...) Olha, adorei esse trabalho. (Já leu?) Li, mas não é tema que eu to procurando. ... Olha, minha dissertação! Ah, que legal, que felicidade ((RI)) (PGD-FCA)*

[busca por SECAGEM] *A próxima também eu já vi que é uma ex colega de mestrado. (...) [busca por BAGASSE] (Melhorou um pouquinho?) Melhorou. ...~~ Apareceu uma professora da Alimentos que já faleceu, que eu nem sabia que trabalhava com isso. (...) Mil e novecentos e pouco a tese dela. ...~~ Esse é o meu, vale? ((RI)). ((RI)) Se te interessar ((RI)))). ((RI)) Eu olhei, que título perfeito, esse eu quero ler ((RI)), mas é claro, né ((RI)). (É, exatamente esse ((RI))) (D-FCA)*

[busca por "GLUED LAMINATED TIMBER"] (Você já conhece esses trabalhos?) *É ... nem todos ... é ... mas a maioria sim ... aliás tem meu trabalho ((RI)). (Tá ((RI))). Quando ... Esse aqui foi meu orientador Nilson, né, da Civil ... tem ... diversos ... a maioria são conhecidos ... que já li, participei de banca, de alguma maneira. (...) [busca por MADEIRA LAMINADA REFORÇADA] Apareceram mais dois ... ((RI)) (D-BAE)*

[busca por KEFIR] *Engraçado, eu encontrei um professor meu que também trabalha com KEFIR ((RI))* (G-FEA)

[busca por EMERGIA] *Tudo é do meu orientador, todos ((RI)). (Vieram quantos?) 68. Todos dele ((RI)). (...) [busca por EMERGY] Todos são do meu orientador aqui ((RI)). (Bastante ((RI))). É ... é que todos são dele, assim, podem me ajudar com alguma coisa assim ... praticamente todos (PGM-FEA)*

[busca por FORMAÇÃO DE BIOFILMES] *Olha que nome cumprido ((RI)). Olha ... É, às vezes tem o título em português e em inglês. Ah, tá junto. Aparece, acaba aparecendo junto. ...~~ Legal porque aparece os trabalhos que eu to usando ((RI)). (É? ((RI))). Já até sei ((RI))* (PGD-FEA)

Os usuários anotam (D-FE) ou fotografam os registros recuperados considerados relevantes (D-FCM e PGM-IMECC) ou mesmo o *Roteiro de busca por assuntos no Repositório Institucional da Unicamp - Linguagem natural* (PDM-IMECC).

[busca por "READING AND WRITING PRACTICES"] *Eu to anotando porque isso é uma opção pra uma orientanda ((RI))* (...) [busca por AUTHORSHIP AND TEACHER DEVELOPMENT] *Vou pegar um outro pra minha orientanda, pra uma outra, tá ((RI))* (D-FE)

[busca por BURKHOLDERIA] *Você se importa se eu tirar uma foto disso daqui, porque ((RI)) ... (Fique à vontade ((RI))), porque realmente isso daqui é uma coisa que ... (que vai te ajudar? ((RI))) Que vai me ajudar ((RI)). (Que bom ((RI))). Isso aqui é interessante porque eu to ... começando esse trabalho. (...) ...~~ Deixa eu anotar o nome do título aqui super rápido (PGM-FCA)*

[busca por CLASSIFICAÇÃO ESPACIAL] *Eu posso tirar uma foto disso aqui? (Tudo bem). Ah, tá ... isso aqui é interessante* (PGM-IMECC)

6. Definição da linguagem controlada

A categoria “Definição da linguagem controlada” refere-se à consulta às linguagens de indexação ou tesouros em cada área do conhecimento, isto é, à compatibilização entre as palavras-chave elencadas pelos usuários (linguagem natural) para as buscas por assuntos no repositório institucional com uma linguagem controlada e aplicação do índice de precisão. A categoria está estruturada em duas principais partes: a) descrição das linguagens de indexação utilizadas para a tradução/compatibilização da linguagem natural; e b) comentários dos usuários sobre a linguagem de indexação utilizada. Ressalta-se que a definição desses termos foi realizada por meio de uma negociação entre a pesquisadora e o usuário.

Referente às linguagens controladas utilizadas, considerando-se a inexistência de uma linguagem própria e única que norteie a catalogação de assunto e a indexação no catálogo coletivo online e no repositório institucional no contexto da UNICAMP, durante as buscas por assuntos com os participantes foram utilizadas as seguintes linguagens controladas, de acordo com cada área do conhecimento das Bibliotecas, detalhadas a seguir (Quadro 10):

Quadro 10. Relação entre as áreas do conhecimento das Bibliotecas e as linguagens de indexação utilizadas

Área do conhecimento	Linguagem de indexação
Artes e Humanidades	<i>Terminologia de Assuntos da Biblioteca Nacional</i> <i>Library of Congress Subject Headings (LCSH)</i> <i>Descritores em Ciências da Saúde (DeCS)</i>
Biomédicas	<i>Descritores em Ciências da Saúde (DeCS)</i> <i>Library of Congress Subject Headings (LCSH)</i> <i>Terminologia de Assuntos da Biblioteca Nacional</i>
Ciências Aplicadas	<i>Ei Thesaurus</i> <i>Library of Congress Subject Headings (LCSH)</i> <i>Descritores em Ciências da Saúde (DeCS)</i> <i>FSTA Thesaurus</i>
Exatas	<i>Library of Congress Subject Headings (LCSH)</i> <i>Ei Thesaurus</i> <i>Descritores em Ciências da Saúde (DeCS)</i> <i>Terminologia de Assuntos da Biblioteca Nacional</i>
Tecnológicas	<i>Ei Thesaurus</i> <i>FSTA Thesaurus</i> <i>Library of Congress Subject Headings (LCSH)</i> <i>Terminologia de Assuntos da Biblioteca Nacional</i> <i>Descritores em Ciências da Saúde (DeCS)</i>

Fonte: Elaboração própria.

A *Library of Congress Subject Headings* (LCSH)⁴³ (Figura 38) é um vocabulário controlado multidisciplinar desenvolvido pela *Library of Congress* contendo 42 milhões de registros de autoridade e 9,3 milhões de registros de autoridade de nome pessoal, além de visualização de referências associadas e notas de escopo. Os resultados da consulta são apresentados como uma lista de cabeçalhos de assunto constantes no Catálogo online da *Library of Congress*, que se classificam alfabeticamente mais próximos das primeiras palavras da consulta. Os resultados da pesquisa incluem os cabeçalhos que possuem registros de autoridade associados a eles e cabeçalhos sem registros de autoridade associados. Embora seja atualizado regularmente, na LCSH os registros de autoridade ainda não suportam a consulta por palavra-chave de campos MARC, além de campos de títulos e acesso às subdivisões de assunto individualmente. O vocabulário controlado é atualizado regularmente. Todas as adições, alterações e exclusões a esses registros são propostas pelos catalogadores da *Library of Congress* da SACO (*Subject Authority Cooperative Program*)⁴⁴ e revisadas por especialistas em políticas da *Library of Congress* (LIBRARY ... 2019).

⁴³ Disponível em: <https://authorities.loc.gov>. Acesso gratuito.

⁴⁴ O SACO tem como principal objetivo a “criação e manutenção oportuna de registros de autoridade confiáveis e com boa relação custo-benefício por catalogadores experientes em instituições em todo o mundo”, permitindo que as instituições participantes submetam propostas de assuntos para inclusão na *Library of Congress Subject Headings* (LCSH), *LC Genre/Form Terms* (LCGFT), *LC Demographic Group Terms* (LCDGT), *LC Medium of Performance Thesaurus for Music* (LCMPT) e *LC Classification* (LCC) (SACO, 2019). Disponível em: <https://www.loc.gov/aba/pcc/saco/about.html>.

Figura 38. Página inicial da *Library of Congress Subject Headings* (LCSH)

The Library of Congress >> Go to Library of Congress Online Catalog

LIBRARY OF CONGRESS AUTHORITIES

Database Name: Library of Congress Online Catalog

Simple Search

Find this: Find Results in:

10 records per page

Authorities Search Tips

Search Type	Brief Help (select a search type for detailed Help)
Subject	<ul style="list-style-type: none"> - Searches are "left-anchored" (enter search starting with leftmost word) - Truncation is automatic - Searches LC Subject Headings and Annotated Card Program (AC) Headings - Searches LC Genre/Form Thesaurus, Thesaurus for Graphic Materials, etc. - Searches other subject systems (e.g., Medical subject headings) <p>Examples: united states history civil war 1861 1865 regimental histories bible criticism interpretation etc persian gulf war 1991</p>
Name	<ul style="list-style-type: none"> - Searches are "left-anchored" (enter search starting with leftmost word) - Truncation is automatic - Searches personal, family, corporate, meeting, and jurisdiction headings - For <u>personal names</u>, enter surname first - For <u>other names</u>, enter in direct order

Fonte: Dados da pesquisa.

A Terminologia de Assuntos da Biblioteca Nacional⁴⁵ (Figura 39) é um vocabulário multidisciplinar que segue a *Library of Congress Subject Headings* (LCSH). Está estruturada em forma de tesouros, onde para cada assunto são apresentados os termos gerais (TG), os termos específicos (TE) e os termos relacionados (TR), além da versão do termo em inglês na *Library of Congress* em alguns casos. Inclui tópicos, remissivas ver, remissivas ver também, subdivisões gerais, cronológicas e geográficas (BIBLIOTECA NACIONAL, 2019). O acréscimo dos cabeçalhos de assunto da *Library of Congress Subject Headings* (LCSH) torna-o bilíngue, característica importante para a interoperabilidade linguística, com cerca de 39.040 entradas de assuntos tópicos, 21.185 entradas de nomes geográficos e 2.789 subdivisões gerais, cronológicas e geográficas (BETTENCOURT, 2014, p. 130).

⁴⁵ Disponível em: http://acervo.bn.br/sophia_web. Acesso gratuito.

Figura 39. Página inicial da Terminologia de Assuntos da Biblioteca Nacional



Fonte: Dados da pesquisa.

Os *Descritores em Ciências da Saúde (DeCS)*⁴⁶ (Figura 40) é um vocabulário estruturado e trilingue (português, inglês e espanhol) desenvolvido pela BIREME a partir do *Medical Subject Headings (MeSH)*, da U.S. *National Library of Medicine (NLM)* para cumprir a função de linguagem única tanto na indexação de recursos informacionais quanto na busca e recuperação por assuntos da literatura científica nas fontes de informação disponíveis na Biblioteca Virtual em Saúde (BVS), tais como LILACS, MEDLINE e outras. O DeCS integra a metodologia LILACS e é um componente integrador da BVS, além de participar no projeto de *Unified Medical Language System (UMLS) da NLM* visando o desenvolvimento de uma terminologia única e rede semântica em Saúde, com a responsabilidade da atualização e envio dos termos MeSH em português e espanhol. Além dos termos médicos originais do MeSH, foram desenvolvidas as áreas específicas de *Saúde Pública, Homeopatia, Ciência e Saúde e Vigilância Sanitária*. Em relação aos assuntos, estes são organizados em uma estrutura hierárquica, permitindo buscas com termos mais amplos ou mais específicos ou todos os termos que pertençam a uma mesma estrutura hierárquica. Possui atualmente 33.966 descritores e qualificadores, sendo destes 29.431 do MeSH e 4.535 exclusivamente

⁴⁶ Disponível em: <http://www.decs.bvs.br>. Acesso gratuito.

do DeCS. Em constante atualização, registra a cada ano um mínimo de 1000 interações dentre alterações, substituições e criações de novos termos ou áreas (DESCRITORES ..., 2019).

Figura 40. Página inicial do Descritores em Ciências da Saúde (DeCS)

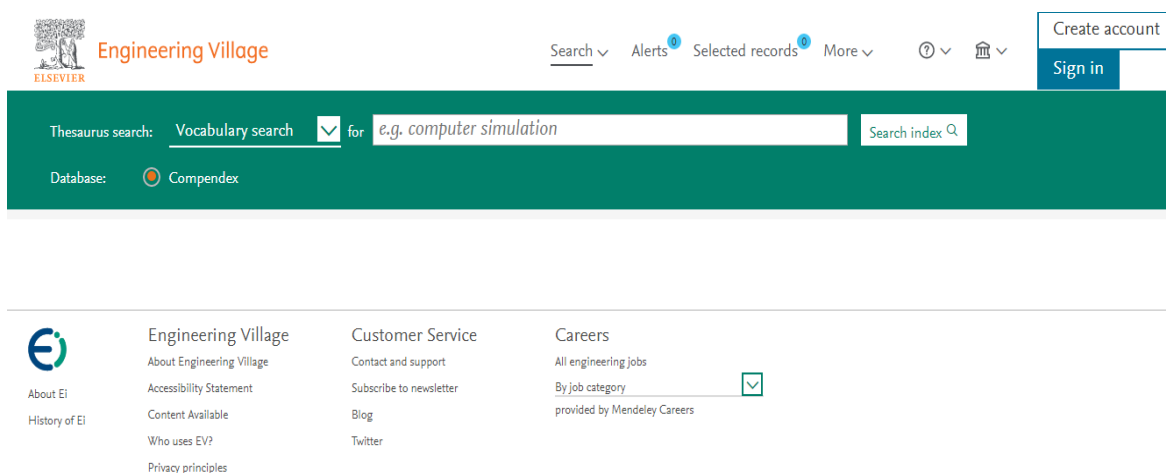
Fonte: Dados da pesquisa.

O vocabulário controlado *Ei Thesaurus*⁴⁷ ou *Engineering Index Thesaurus* (Figura 41) é utilizado por indexadores para a determinação de termos de uma lista de assuntos predeterminados para descrever o conteúdo dos artigos, permitindo as buscas por termos descritivos e sinônimos atribuídos a cada registro, com o objetivo de padronizar a indexação dos artigos e proporcionar resultados consistentes e precisos. Criado em 1884, juntamente com o *Engineering Index/Compendex (Compendium)* de Engenharia, nos EUA, é atualizado anualmente por especialistas, cobrindo mais de 190 disciplinas da Área de Engenharia e Ciências Físicas, com 22 mil termos/expressões. Inicialmente comercializado em formato impresso, foi convertido para o formato digital, disponível atualmente apenas através da plataforma EV Engineering Village, base Compendex, da Elsevier. Nesta pesquisa, a consulta ao *Ei Thesaurus* foi realizada na base de dados *Ei Compendex*, on-line desde 1970, sendo disponível ainda em outras cinco bases de dados da *Engineering Village: Inspec, GeoRef,*

⁴⁷ Disponível em: <http://www.periodicos.capes.gov.br>. Acesso restrito mediante assinatura da base de dados Engineering Village, da Elsevier.

GEOBASE, *EnCompassPAT* e *EnCompassLIT*, indexadas com seus próprios termos controlados do dicionário de sinônimos. Cada dicionário de sinônimos é organizado de maneira hierárquica, com palavras e sinônimos relacionados uns aos outros com termos amplos, restritos, equivalentes ou relacionados. Cabe salientar que a utilização da função *Thesaurus Search*, apenas a(s) base(s) de dados que a instituição assina será/serão exibida(s), permitindo ainda a pesquisa de apenas uma base de dados por vez ao usar tal função (Ei COMPENDEX, 2019).

Figura 41. Página inicial do *Ei Thesaurus*



Fonte: Dados da pesquisa.

O *FSTA Thesaurus*⁴⁸ (Figura 42), disponível na base de dados *Food Science and Technology Abstracts* (FSTA), é desenhado para facilitar e aprimorar a pesquisa do banco de dados bibliográfico da FSTA. Todos os termos de indexação são selecionados com base na Ciência dos alimentos, Tecnologia de alimentos e Nutrição humana e animal relacionados com alimentos. A seleção é baseada na frequência de ocorrência no FSTA (FSTA, 2019).

⁴⁸ Disponível em: <http://www.periodicos.capes.gov.br>. Acesso restrito mediante assinatura da base de dados *Food Science and Technology Abstracts*, da EBSCO.

Figura 42. Página inicial do *FSTA Thesaurus*

Fonte: Dados da pesquisa.

Referente aos resultados da compatibilização da linguagem natural com a linguagem controlada, cabe esclarecer que, no caso de termos localizados nas linguagens: *Ei Thesaurus* e *FSTA Thesaurus*, e da *Library of Congress Subject Headings* (LCSH), quando não localizada a versão em português na *Terminologia de Assuntos da Biblioteca Nacional*, a versão dos termos em português foi traduzida pelos próprios autores ou pela pesquisadora. Já no caso da *Terminologia de Assuntos da Biblioteca Nacional* e do *Descritores em Ciências da Saúde* (DeCS), não houve necessidade, visto que possibilitam a visualização dos termos em português e em inglês. Para a análise dos resultados, optou-se pela sistematização por área do conhecimento, conforme segue.

Na *Área de Artes e Humanidades*, do total de 80 palavras-chave elencadas pelos usuários (linguagem natural), 16 (20%) foram localizadas em linguagem controlada, 46 (57,5%) localizadas com alguma diferença, a maior parte na *Terminologia de Assuntos da Biblioteca Nacional*, seguida da *Library of Congress Subject Headings* (LCSH) e dos *Descritores em Ciências da Saúde* (DeCS); e 18 (12,5%) não foram localizadas em nenhuma das referidas linguagens (Quadro 11):

Quadro 11. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Artes e Humanidades

Usuário	Termo em linguagem natural	Termo em linguagem controlada	Linguagem controlada
G-FE	Inclusão	Inclusão educacional	BN

	Inclusion	Mainstreaming (Education)	<i>LCSH</i>
	Escola School	Escolas Schools	<i>BN</i> <i>LCSH</i>
	Diferença Difference	Diferença (Filosofia) Difference (Philosophy)	<i>DeCS</i>
PMG-FE	Ensino superior Higher education	Educação superior Higher education	<i>DeCS</i> <i>LCSH</i>
	Cooperação Cooperation	Cooperação universitária University cooperation	<i>BN</i> <i>LCSH</i>
	Privatização Privatization	Privatização na educação Privatization in education	<i>BN</i> <i>LCSH</i>
PGD-FE	Autorregulação da aprendizagem Auto-regulação da aprendizagem Self-regulation learning	Autodomínio Auto-domínio Self-control	<i>BN</i> <i>LCSH</i>
	Alunos Students	Estudantes do ensino médio High school students	<i>BN</i> <i>LCSH</i>
D-FE	Formação inicial de professores Teacher development	Professores - Formação Training of teachers	<i>BN</i>
	Práticas discursivas e formação Discursive practices and formation	Educação continuada Educação permanente Continuing education	<i>DeCS</i> <i>BN</i> <i>LCSH</i>
	Autoria e formação de professores Authorship and teacher development	Autoria Authorship	<i>BN e DeCS</i> <i>LCSH</i>
G-IFCH	Violência Violence	Violência policial Police brutality	<i>BN</i> <i>LCSH</i>
	Racismo Racism	Má-conduta policial Police misconduct	<i>BN</i> <i>LCSH</i>
	Polícia Police	Denúncia contra policiais Police - Complaints against	<i>BN</i> <i>LCSH</i>
	Justiça criminal Criminal justice	Violência contra os adolescentes Teenagers - Violence against	<i>BN</i>
PGM-IFCH	Arquitetura brasileira Brazilian architecture	Arquitetura - História	<i>DeCS</i>
	Arte paulista Paulist art	Arte - Brasil - História	<i>BN</i>
	Iconografia paulista Paulist iconography	Iconografia Iconography	<i>BN</i>
	Miguel Dutra* ⁴⁹ Miguel Dutra*	Dutra, Miguel Arcanjo Benicodé, 1812-1875	<i>BN</i> <i>LCSH</i>
PGD-IFCH	Consumo Consumption	Consumo alimentar Food consumption	<i>DeCS</i> <i>DeCS e LCSH</i>
	Juventude Youth	Adulto jovem Young adult	<i>DeCS</i>
D-IFCH	Gênero Gender	Estudos de gênero Gender studies	<i>DeCS</i> <i>DeCS e LCSH</i>
	Rural* Rural*	População rural Rural population	<i>DeCS</i> <i>DeCS e LCSH</i>

Fonte: Elaboração própria.

⁴⁹ A utilização do asterisco (*) visa sinalizar os termos equivalentes em português e em inglês.

Na *Área de Biomédicas*, do total de 80 palavras-chave elencadas pelos usuários (linguagem natural), 20 (25%) foram localizadas em linguagem controlada, 48 (60%) localizadas com alguma diferença, sendo a maior parte nos *Descritores em Ciências da Saúde* (DeCS), seguida da *Library of Congress Subject Headings* (LCSH) e da *Terminologia de Assuntos da Biblioteca Nacional*; e 12 (15%) não foram localizadas em nenhuma das referidas linguagens (Quadro 12):

Quadro 12. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Biomédicas

Usuário	Termo em linguagem natural	Termo em linguagem controlada	Linguagem controlada
G-FCM	Síndrome de Turner Turner Syndrome	Síndrome de Turner - Genética Turner Syndrome - Genetics	DeCS
	Cardiopatias congênitas Congenital cardiopathy	Cardiopatias congênitas - Complicações Congenital heart defects - Complications	DeCS
	Valva aórtica bicúspide Bicuspid aortic valve	Valva aórtica - Anormalidades Aortic valve - Abnormalities	DeCS
	Coarctação de aorta Aortic coarctation*	Coarctação aórtica Aortic coarctation*	DeCS
	Dismorfismos Dysmorphisms	Caracteres sexuais Sex characteristics	DeCS
PGM-FCM	Gestão Management	Public administration Administração pública	LCSH e BN BN e DeCS
	Administração Administration	Higher education - Administration Ensino superior - Administração	DeCS e BN BN
	Ensino médico Medical education*	Educação médica Medical education*	DeCS e BN
	Arquivo Archive	Archives Arquivos	DeCS, LCSH e BN DeCS e BN
PGD-FCM	Pneumoconiose Pneumoconiosis	Pneumoconiose - Prevenção e controle Pneumoconiosis - Prevention and control	DeCS
	Fadiga Fatigue	Fadiga muscular Muscle fatigue	DeCS
D-FCM	Terapia antirretroviral Antiretroviral therapy	Terapia antirretroviral de alta atividade Highly active antiretrovilar therapy	DeCS DeCS e LCSH
G-FEF	Epilepsia Epilepsy	Epilepsia do lobo temporal Temporal lobe epilepsy	DeCS DeCS e LCSH
	Comorbidades Comorbidities	Comorbidade Comorbidity	DeCS DeCS e LCSH
	Atividade física Physical activity	Atividade motora Motor activity	DeCS
PGM-FEF	Esporte adaptado na França Handsport in France	Sports for people with disabilities Esporte para pessoas com deficiência física	LCSH
	Paralimpíadas* Paralympics	Paralympic games Paralimpíadas*	LCSH
	Reabilitação Rehabilitation	People with disabilities - Rehabilitation Pessoas com deficiência física - Reabilitação	LCSH BN

PGD-FEF	Circo Circus*	Círcos Circus*	BN LCSH
	Escola de circo Circus school	Artistas circenses Circus performers	BN LCSH e BN
	Circo de rua Street's circus	Arte de rua Street art	BN LCSH e BN
D-FEF	História do esporte History of sport	Sports - History Esportes - História	LCSH
	Esporte e natureza Sport and nature	Nature Natureza	LCSH
	Corpo e natureza Body and nature	Human body (Philosophy) Corpo humano (Filosofia)	LCSH

Fonte: Elaboração própria.

Na *Área de Ciências Aplicadas*, do total de 40 palavras-chave elencadas pelos usuários (linguagem natural), 4 (10%) foram localizadas em linguagem controlada, 22 (40%) localizadas com alguma diferença, sendo a maior parte no *Ei Thesaurus*, seguido da *Library of Congress Subject Headings* (LCSH), *Descritores em Ciências da Saúde* (DeCS) e *FSTA Thesaurus*; e 14 (17,5%) não foram localizadas em nenhuma das linguagens controladas consultadas (Quadro 13):

Quadro 13. Compatibilização entre o termo em linguagem natural e linguagem controlada na *Área de Ciências Aplicadas*

Usuário	Termo em linguagem natural	Termo em linguagem controlada	Linguagem controlada
G-FCA	Tecnologia social Social technology	Tecnologia culturalmente apropriada Culturally appropriate technology	DeCS
	Redes Network	Redes sociais Social networking	LCSH
	Geração de renda Income*	Renda Income*	LCSH e DeCS
PGM-FCA	Lipase* Lipase*	Lipases* Lipases*	<i>Ei Thesaurus</i>
	Imobilização Immobilization	Imobilização enzimática Enzyme immobilization	<i>Ei Thesaurus</i>
PGD-FCA	Inovação nas cortes Court innovation	Tribunais - Administração Court administration	LCSH
	Desempenho judicial Judicial performance	Administração judiciária Administration of justice	LCSH
D-FCA	Secagem Drying	Secagem a ar Air drying Secagem solar Solar drying	<i>FSTA Thesaurus</i>
	Bagaço Bagasse	Etanol celulósico Cellulosic ethanol	<i>Ei Thesaurus</i>
	Etanol Ethanol	Bioetanol Bioethanol	<i>Ei Thesaurus</i>

	Cana-de-açúcar Sugarcane	Moagem de cana-de-açúcar Sugar cane Sugar cane milling	<i>Ei Thesaurus</i>
--	-----------------------------	--	---------------------

Fonte: Elaboração própria.

Na *Área de Exatas*, do total de 40 palavras-chave elencadas pelos usuários (linguagem natural), 16 (20%) foram localizadas em linguagem controlada, 42 (52,5%) localizadas com alguma diferença, sendo a maior parte na *Library of Congress Subject Headings* (LCSH), seguida do *Ei Thesaurus*, *Descritores em Ciências da Saúde* (DeCS) e *Terminologia de Assuntos da Biblioteca Nacional*; e 22 (27,5%) não foram localizadas em nenhuma das linguagens controladas consultadas (Quadro 14):

Quadro 14. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Exatas

Usuário	Termo em linguagem natural	Termo em linguagem controlada	Linguagem controlada
G-IFGW	Propriedades eletrônicas Electronic properties	Microscopia eletrônica Microscopia eletrônica de varredura Microscopia eletrônica de transmissão Electron microscopy Scanning electron microscopy Transmission electron microscopy	<i>DeCS</i>
	Técnicas de caracterização Characterization techniques	Characterization (Materials science) Caracterização (Ciência dos materiais)	<i>Ei Thesaurus</i>
PGM-IFGW	Bits quânticos Qubit	Qubits* Qubits*	<i>Ei Thesaurus</i>
	Nanorressonadores Nanoresonators	Superconducting resonators Ressonadores supercondutores	<i>Ei Thesaurus</i>
	Circuitos supercondutores Superconducting circuits	Josephson junction devices Superconducting devices - Josephson junction Dispositivos de junções Josephson Dispositivos supercondutores - Junções Josephson	<i>Ei Thesaurus</i>
	Computação quântica Quantum computing	Quantum computers Computadores quânticos	<i>Ei Thesaurus</i>
PGD-IFGW	Férmions pesados Heavy fermions	Heavy fermion superconductors Férmions pesados supercondutores Fermi liquids Líquidos de fermi	<i>Ei Thesaurus</i> <i>Ei Thesaurus e LCSH</i>
D-IFGW	QCD* QCD*	Quantum chromodynamics Cromodinâmica quântica	<i>LCSH</i>
	Fator de forma Form factor	Form factor (Nuclear physics) Fator de forma (Física nuclear)	<i>LCSH</i>
	Neutrino* Neutrino*	Neutrino astrophysics Astrofísica de neutrinos	<i>LCSH</i>
	Estranheza Strangeness	Strange particles Partículas estranhas	<i>LCSH</i>

G-IMECC	Otimização multiobjetivo Multicriteria optimization	Processo decisório por critério múltiplo Multiple criteria decision making	<i>BN</i> <i>LCSH</i>
	Programação linear inteira Integer linear programming	Integer programming Programação inteira Linear programming Programação linear	<i>LCSH</i> <i>DeCS,</i> <i>Ei Thesaurus e</i> <i>LCSH</i>
	Método das somas ponderadas Weighted sum method	Operations research Pesquisa operacional	<i>Ei Thesaurus</i>
PGM-IMECC	Predição espacial Spatial prediction	Digital soil mapping Mapeamento digital de solos	<i>LCSH</i>
	Abordagem bayesiana Bayesian approach	Bayesian field theory Teoria de decisão bayesiana	<i>LCSH</i>
	MCMC* MCMC*	Monte Carlo Method Método de Monte Carlo	<i>LCSH</i>
PGD-IMECC	Sistemas dinâmicos Dynamical system	Dynamics Sistemas dinâmicos	<i>LCSH</i>
	Processo estocástico Stochastic process	Stochastic processes Processos estocásticos	<i>LCSH</i>
	Espaço de Sobolev Sobolev space	Sobolev spaces Espaços de Sobolev	<i>LCSH</i>
D-IMECC	Problema de dimensionamento de lotes Lot sizing problem	Economic lot size Dimensionamento de lotes	<i>LCSH</i>

Fonte: Elaboração própria.

Por fim, na *Área de Tecnológicas*, do total de 80 palavras-chave elencadas pelos usuários (linguagem natural), 22 (27,5%) foram localizadas em linguagem controlada, 44 (55%) localizadas com alguma diferença, sendo a maior parte no *Ei Thesaurus*, seguido do *FSTA Thesaurus*, *Library of Congress Subject Headings* (LCSH), *Terminologia de Assuntos da Biblioteca Nacional* e *Descritores em Ciências da Saúde* (DeCS); e 14 (17,5%) não foram localizadas em nenhuma das linguagens controladas consultadas (Quadro 15):

Quadro 15. Compatibilização entre o termo em linguagem natural e linguagem controlada na Área de Tecnológicas

Usuário	Termo em linguagem natural	Termo em linguagem controlada	Linguagem controlada
G-BAE	Avaliação dinâmica Dynamic evaluation*	Análise dinâmica Dynamic evaluation*	<i>Ei Thesaurus</i>
	Modos de vibração* Vibration modes	Vibration measurement Modos de vibração*	<i>Ei Thesaurus</i>
PGM-BAE	Estratégias projetuais Design strategies	Arquitetura - Projetos e plantas Architecture - Design and plans	<i>BN</i> <i>LCSH</i>
	Análise gráfica Graphic analysis	Schematic diagrams Diagramas esquemáticos	<i>Ei Thesaurus</i>
PGD-BAE	Gás de síntese* Syngas	Synthesis gas Gás de síntese*	<i>LCSH e</i> <i>Ei Thesaurus</i>

	Fischer-tropsch* Fischer-tropsch*	Fischer-tropsch process Processo de Fisher-tropsch Fischer-tropsch synthesis Síntese de Fisher-tropsch	<i>LCSH e Ei Thesaurus</i>
	Catalisador Catalyst	Catalysts Catalisadores	<i>LCSH e Ei Thesaurus</i>
D-BAE	Compósitos de madeira Wood composite	Wood laminates Madeira laminada	<i>Ei Thesaurus</i>
	Madeira laminada colada Glued laminated timber	Wooden construction Construções em madeira	<i>Ei Thesaurus</i>
	Madeira laminada reforçada Reinforced glulam	Laminated composites Compósitos laminados	<i>Ei Thesaurus</i>
	Madeira com barras de aço Wood with steel bars	Reinforcement Reforço	<i>Ei Thesaurus</i>
G-FEA	Antioxidantes Antioxidants	Antioxidative activity Atividade antioxidante	<i>FSTA Thesaurus</i>
	Fermentação Fermentation	Fermented milk Leite fermentado	<i>FSTA Thesaurus e LSCH</i>
PGM-FEA	Energia Energy	Energy (Sustainability) Energia (Sustentabilidade)	<i>FSTA Thesaurus</i>
	Mudanças climáticas Climate change	Climate change - Social aspects Mudanças climáticas - Aspectos sociais	<i>FSTA Thesaurus</i>
	Agroecologia Agroecology	Agricultural ecology Agricultura ecológica	<i>FSTA Thesaurus e LSCH</i>
PGD-FEA	Biofilme Biofilm	Biofilms Biofilmes	<i>FSTA Thesaurus e LSCH</i>
	Dispersão de biofilme Biofilm dispersal	Dispersion Dispersão	<i>FSTA Thesaurus e LSCH</i>
	Microscopia eletrônica Microscopy electronic	Microscopia eletrônica de transmissão Electron transmission microscopy Microscopia eletrônica de varredura Scanning electron microscopy	<i>DeCS</i>
D-FEA	Microencapsulação de óleo Microencapsulation of oil	Microencapsulation Microencapsulação	<i>FSTA Thesaurus e LSCH</i>
	Hidrólise de proteína Protein hydrolysis	Protein hydrolysates Hidrolisado protéico	<i>FSTA Thesaurus e LSCH</i>
	Emulsão Oil emulsion	Emulsions Emulsões	<i>FSTA Thesaurus, LSCH e BN</i>

Fonte: Elaboração própria.

Em síntese, de um total de 360 palavras-chave elencadas pelos usuários, 78 (21,6%) foram localizadas em linguagem controlada, 202 (56%) localizadas com alguma diferença, sendo a maior parte na *Library of Congress Subject Headings* (LCSH), seguida dos *Descritores em Ciências da Saúde* (DeCS), *Terminologia de Assuntos da Biblioteca Nacional*, *Ei Thesaurus* e *FSTA Thesaurus*. Além disso, 80 (22,2%) não foram localizadas em nenhuma das linguagens controladas consultadas.

7. Avaliação da relevância - Linguagem controlada

A categoria “Avaliação da relevância - Linguagem controlada” refere-se aos comentários dos usuários sobre os recursos informacionais recuperados com seus temas de pesquisa, utilizando-se termos controlados já verificados nos tesouros de suas respectivas áreas do conhecimento.

Os usuários manifestam dúvidas sobre a atividade realizada ou sobre os resultados de buscas: não familiaridade utilizando-se hífen com espaço ou parênteses na digitação do termo (D-FE e D-IFGW); indagação sobre a versão do termo em inglês (D-FE); observação da ordenação da relevância da ferramenta (D-IFGW) ou da forma de varredura da ferramenta nos registros (G-IMECC); diferença em realizar as buscas com as letras maiúsculas ou minúsculas (PGM-IFCH); digitação correta do termo, pois não houve nenhum resultado utilizando-se o recurso gráfico aspas (“) (G-IFCH); confirmação sobre os 50 primeiros resultados na página, que é o número para avaliação da relevância (PGM-FCM); questionamento sobre a atualização do repositório institucional em termos de recursos informacionais incorporados (G-FEF); digitação do termo utilizando-se o recurso gráfico aspas (“) (D-FCA e D-IFGW); questionamento sobre a necessidade de repetição da avaliação da relevância com os termos da linguagem controlada (PGM-IMECC e PGD-IMECC); e confirmação da busca dos termos em português e em inglês (PGD-BAE).

[busca por PROFESSORES - FORMAÇÃO] E eu ponho o hífen junto, é isso? Sem espaço? (É). Ah, espacinho? Eu não colocaria o espaço. (Porque a linguagem está desse jeito) (...) [buscando TRAINING OF TEACHERS] Não é TEACHER TRAINING? (Não). Não? TRAINING OF TEACHERS? (TRAINING OF TEACHERS) (D-FE)

[busca por “DENÚNCIA CONTRA POLICIAIS”] Será que eu coloquei alguma coisa errada? (Não, é que não tem mesmo) (G-IFCH)

[busca por ICONOGRAFIA] Não faz diferença se for maiúsculo ou minúsculo, né? (Não) (PGM-IFCH)

[busca por ENSINO MÉDICO] Os 50 primeiros são só até o final do cursor, né. (Isso, são só na primeira página) (PGM-FCM)

[busca por TEMPORAL LOBE EPILEPSY] O que será que é hipossexualidade? ((RI)). (Não faço ideia ((RI)), sou da Humanas ((RI))). (...) ...~~ Essa base de dados está atualizada? Porque eu não vi nenhum estudo de 2019, eu acho. ... (Tá sim, é que 2019 eu não sei te dizer. Das bases de dados eles fazem importação em lote. Você conhece a Web of Science, Scopus ...). Não. (Você vai conhecer ((RI)). São bases de dados, as mais conhecidas, né, aí é importado em lote pro repositório ... eu não se 2019 já fizeram algum). Eu acho que não porque eu vi um estudo novo esses dias e não apareceu ele aqui. (Você viu onde?) No PubMed. (Mas era da UNICAMP?) Não. (Ah, então, porque no repositório é só da UNICAMP). Ah, não sabia. (É só produção da UNICAMP, é, professor, aluno ...). Ah, entendi porque o meu orientador tá aí em quase todos ((RI)). (É, então ((RI)), porque é da UNICAMP) (G-FEF)

[busca por MOAGEM DE CANA-DE-AÇÚCAR] (Como são termos compostos, eu vou pedir pra você colocar entre aspas agora). *Tá, o CANA-DE-AÇÚCAR?* (Todo o conjunto). Ah, entendi, todo o conjunto (D-FCA)

[busca por FORM FACTOR (NUCLEAR PHYSICS)] Esse daqui como busca? (Primeiro sem aspas, do que jeito que tá aí). *E coloca entre parênteses, com espaço?* (...) [busca por "STRANGE PARTICLES"] Não tá aparecendo nenhum que tem a palavra ... (No título?) É. ...~~ Ele procura dentro ... do arquivo. (...) ...~~ Deve ser ... eles devem ordenar por número de repetições em que a palavra aparece no documento (D-IFGW)

[busca por PROCESSO DECISÓRIO POR CRITÉRIO MÚLTIPLO] Ele faz essa pesquisa tanto no título quanto no resumo? (G-IMECC)

[busca por DIGITAL SOIL MAPPING] ...~~ Ah, espera aí, agora eu tenho que olhar de novo todos eles? (É). *Eu digitei, agora tenho que olhar todos eles? Quantos me interessam?* (Quantos, quantos dessa página). ... *Ok. Um. ...~~. Três (...)* [busca por BAYESIAN FIELD THEORY] Bom, esse aqui eu não vou colocar com aspas, mas acho que faz mais sentido com aspas porque vai ser muito, muito amplo do que eu quero ...~~ é ... não tem nada a ver ... É, no caso é melhor colocar com aspas mesmo. Não, assim, tem a ver mas não com o meu tema, sabe. (...) Porque falta uma palavra lá no meio, é só por isso (PGM-IMECC)

[busca por "DYNAMICS SYSTEM"] (Agora a gente vai no repositório da UNICAMP e vai buscar ...~~ a gente vai fazer a busca agora usando essas palavras no repositório). ...~~ Mas ... e por que? Eu acho que já buscamos. ...~~ (É que aqui ... teve o 's', não teve? SYSTEMS). *Ah, ok, entendi.* (É diferente, por isso a gente tá buscando de novo, tá. É diferente, DYNAMICS SYSTEM). ...~~ (Agora dá uma olhadinha nos registros pra ver se algum interessaria pra você). ...~~ *Não.* (Nenhum? Tá bom. Agora é só colocar aqui, agora a gente passa pro próximo) (PGD-IMECC)

[busca por PROCESSO DE FISCHER-TROPSCH] E aí português e em inglês? (PGD-BAE)

Os usuários expressam a hipótese de que com um determinado termo (G-FE e D-FEA) ou tipo de recurso informacional (D-IFGW) haverá mais resultados considerados relevantes. Todavia, após a realização das buscas, existe a percepção de que na verdade não houve mudança significativa nos resultados (G-FE).

[busca por DIFERENÇA (FILOSOFIA)] Agora vai ser fácil ((RI)). ...~~ Não mudou muito, 6 só (G-FE)

[busca por HIDROLISADO PROTEICO] Eu acho que vai aparecer bastante também. ...~~ Dez (D-FEA)

[busca por STRANGE PARTICLES] (...) Só vai aparecer artigo. (...) *Essas colaborações tem pelo menos ... no mínimo uns 500 autores. (...)* *Eles colocam uma linha pra cada um* (D-IFGW)

Já outros usuários manifestam a hipótese de que não serão encontrados recursos informacionais com um determinado termo de busca (G-FE, D-IFGW, G-IMECC e G-FEA).

[busca por INCLUSÃO EDUCACIONAL] Acho que vai aparecer coisas ... de educação não formal, também (...) *porque o meu é escolar, né, então eu falo unicamente da parte educativa da escola. E aí quando coloca educacional vai pra além (...)* ... *em qualquer tipo de ... de educação não formal também. (...)* Mas acho que vai aparecer também quase a mesma coisa (...) ...~~ *sete ... acho que não mudou tanto, não.* (Não?) ...~~ *11. ...~~ Deu a mesma quantidade (...)* [busca por DIFFERENCE (PHILOSOPHY)] Acho que nem vai ter, vai? (G-FE)

[busca por CROMODINÂMICA QUÂNTICA] Ele vai pegar um monte de QUÂNTICA mas não vai pegar o CROMODINÂMICA (...) [busca por QUANTUM CRHOMODYNAMICS] Então vai aparecer milhares de QUANTUM [busca por “FORM FACTOR (NUCLEAR PHYSICS)”] Eu coloco aspás onde? No FORM FACTOR ou em tudo? (No comecinho ... e no final). ... *Aí significa que ele supostamente tem que procurar nessa sequência?* (É, ele vai procurar nessa sequência). *Não vai dar nada.* ...~~ (É, não porque não tem nenhum documento que tenha ...). *Essa sequência particular.* (Exatamente). (...) Aqui até apareceu um monte de coisa ... aparece mas não é, não são os mesmos artigos que eu tinha olhado da outra vez (D-IFGW)

[busca por OPERATIONS RESEARCH] Esse aqui vai pra coisas extremamente amplas. (É amplo, né). (...) ...~~ Tem muita coisa específica, assim (G-IMECC)

[busca por LEITE FERMENTADO] LEITE FERMENTADO só não sei se vai ter outra conotação. ...~~ Talvez eu não ache nada. ...~~ (Muito geral?) Não tem muito a ver com o KEFIR mas é bom que eu fiz uma cinética de fermentação e aí ... pra ver sobre probiótico. Esses reviews são bem poucos, daí tipo, eu uso na introdução (G-FEA)

Os usuários expressam os recursos informacionais considerados relevantes (G-FE, PGD-FE, D-FE, G-IFCH, PGM-IFCH, G-FCM, D-FCM, G-FEF, PGM-FEF, PGD-FEF, PGM-FCA, PGD-FCA, D-IFGW, PGM-FEA e D-FEA).

[busca por “DIFERENÇA (FILOSOFIA)”] um ... dois ... três (G-FE)

[busca por ESTUDANTES DE ENSINO MÉDIO] esse deu três, não, quatro. Quatro? Mas que estranho, esse ... não, três, porque esse daqui não é de ensino médio. Mas tá aqui. Tá. Olha, terceiro e quarto ano, olha. (...) [busca por SELF-CONTROL] Três, esses foram três, bem direcionado à minha pesquisa mesmo (PGD-FE)

[busca por AUTHORSHIP] Espera aí, eu escrevi qual número que eu estava? 3. 3. Eu fiz isso, 3. (...) 5 (D-FE)

[busca por VIOLÊNCIA CONTRA OS ADOLESCENTES] Um (G-IFCH)

[buscando pro ARQUITETURA - HISTÓRIA - SÃO PAULO] 15. (Foi bastante, né? ((RI))). Foi ((RI)) (PGM-IFCH)

[busca por CONGENITAL HEART DEFECTS - COMPLICATIONS] Seis (G-FCM)

[busca por “FORMAÇÃO DE ANTICORPOS”] Uma ... que tem sido sempre a mesma, que é do nosso grupo por sinal ... Ok, só esse (D-FCM)

[busca por EPILEPSIA DO LOBO TEMPORAL] É, tá bem misturado, né. ...~~ (...) Teve um (G-FEF)

[busca por PARALYMPIC GAMES] Deu 8 (PGM-FEF)

[busca por CIRCUS PERFORMERS] Três. De vez em quando aparece umas coisas ... BIRDS não sei o que ... tudo a ver, né? E não tinha aparecido nas outras, tem outras coisas que não tem nada a ver tipo VEGANISMO que apareceu várias vezes, mas esse ... foi novo ((RI)). (...) [busca por STREET ART] Olha! Não apareceu em nenhuma. (Nenhuma?) Tem tudo, tem CIRCO, tem RUA, tem ... todas as palavras que eu precisava, que eu coloquei bem aqui. E não apareceu em nenhum, olha aqui ... (Pois é, talvez ele até estava, mas ... depois dos 50 primeiros). Pois é, mas, não faz sentido ... (Está então relacionado?). ...~~ E apareceu VEGANISMO que não sei o que e não apareceu ... coisa doida. Quatro (PGD-FEF)

[busca por “ENZYME IMMOBILIZATION”] É pra fazer um comparativo? (...) (Em português a versão ficaria IMOBILIZAÇÃO ENZIMÁTICA, alguma coisa assim?) Isso. ...~~ Um pouquinho diferente. ...~~ Esses três artigos aqui são muito ricos ... (Vão te ajudar mesmo?) Esses aqui ajudariam. (...) Três mesmo (PGM-FCA)

[busca por ADMINISTRATION OF JUSTICE] *Um ... dois ... três ...~~ quatro. (...) [busca por "ADMINISTRATION OF JUSTICE"] Apareceu um ... "a cultura da justiça: cultura jurídica e administração de justiça", interessante. (Esse não tinha aparecido ainda?) Não tinha. Parece interessante, pelo título né. Legal (PGD-FCA)*

[busca por QUANTUM CRHOMODYNAMICS] *Então esse aqui foram ... g (D-IFGW)*

[busca por "EMERGY (SUSTAINABILITY)"] *5. (Todos?) Sim (PGM-FEA)*

[busca por EMULSIONS] *Nossa, esse aqui tem um monte, vem muita coisa (D-FEA)*

Um usuário confirma novamente os recursos informacionais julgados por eles como relevantes na página de visualização dos resultados (D-FE), enquanto outro expressa desapontamento com o número de recursos informacionais recuperados (G-FE).

[busca por EDUCAÇÃO CONTINUADA] *Deixa eu só conferir, tá na página? Tá na página. Tá, deixa eu só conferir o último (D-FE)*

[busca por "MAINSTREAMING (EDUCATION)"] *Nossa, só apareceu três (G-FE)*

Outros usuários expressam não ter localizado nenhum recurso informacional relevante em suas buscas utilizando-se os termos em linguagem controlada (G-FE, G-IFCH, PGM-IFCH, G-FCM, PGD-FCM, D-FCM, G-FEF, G-FCA, PGM-FCA, PGD-FCA, PGM-IMECC e PGD-FEA).

[busca por ESCOLAS] *Nenhum mesmo (G-FE)*

[busca por "VIOLÊNCIA CONTRA OS ADOLESCENTES"] *Nada (G-IFCH)*

[busca por DUTRA, MIGUEL ARCHANJO BENICIO DE ASSUNÇÃO] *Vem de tudo menos o MIGUEL ((RI)) (PGM-IFCH)*

[busca por "SEX CHARACTERISTICS"] *Nenhum. (Nenhum também? Não). (Muito amplo?) É, fala tudo sobre ... é, parte psicológica do sexo ... (...) nada a ver (G-FCM)*

[busca por "PNEUMOCONIOSE - PREVENÇÃO E CONTROLE"] *Com aspas não apareceu nenhum (PGD-FCM)*

[busca por "TERAPIA ANTIRRETROVIRAL DE ALTA ATIVIDADE"] *É, especificamente não (...) [busca por HIGHLY ACTIVE ANTIRETROVIRAL THERAPY] Então nessa aqui sem aspas, zero (...) [busca por "HIGHLY ACTIVE ANTIRETROVIRAL THERAPY"] Com aspas sempre reduz muito. ...~~ Ok, então aí foi zero também (D-FCM)*

[busca por ATIVIDADE MOTORA] *Aqui não apareceu nenhum (G-FEF)*

[busca por REDE Social] *Acho que ele tá entendendo REDE SOCIAL tipo Facebook, sabe? (...) É ... tinha um talvez que parecia um estudo de caso ... mas não necessariamente ele tem TECNOLOGIA SOCIAL (G-FCA)*

[busca por "IMOBILIZAÇÃO ENZIMÁTICA"] *Caiu pra 5. (...) (Nenhum desses?) Não, porque esses daqui são todos associados a ... cromatografia, que é um outro tipo de trabalho. (...) Que também usa isso, então pra mim não interessa (PGM-FCA)*

[busca por ADMINISTRAÇÃO JUDICIÁRIA] *Esse daqui já foi, o “brinquedo e a administração” ... nada a ver ... “antropologia judiciária”. ...~~ Então dois ... três ... (PGD-FCA)*

[busca por MONTE CARLO METHOD] *Nada (PGM-IMECC)*

[busca por “ELECTRON TRANSMISSION MICROSCOPY”] *Zero (PGD-FEA)*

Os usuários expressam que os recursos informacionais recuperados não estão relacionados com seu tema de pesquisa (PGM-FE, PGD-FE, G-IFCH, PGD-IFCH, D-IFCH, G-FCA, PGD-FCA, D-IFGW, G-BAE e D-BAE). De forma mais específica, verifica-se: percepção de que com a linguagem controlada houve uma precisão menor, recuperando recursos informacionais não apenas relacionados ao seu tema de pesquisa mas com temas mais amplos (PGM-FE); redigitação do termo com outra grafia, para confirmação dos resultados (PGD-FE); exposição de que o termo da linguagem controlada não costuma ser utilizado em sua área de pesquisa (PGD-IFCH); indagação sobre a ausência de precisão voltada para seu tema de pesquisa, pois a maioria dos resultados tratam de outra temática (G-IFCH); e observação de que houve mais recuperação de recursos informacionais de outras áreas, o que o leva a se questionar se no repositório não há mais recursos informacionais da *Área de Biomédicas*, concepção reafirmada na categoria “Comentários adicionais sobre o repositório institucional” (D-IFCH) e questionamento sobre a linguagem controlada utilizada (D-IFGW).

[busca por PRIVATIZAÇÃO NA EDUCAÇÃO] *Esse aqui já apareceu mais ... termos ligados ao ensino fundamental e médio. (Ah, tá, mas no ensino superior ...). E no ensino superior também. (Ah, tá, tudo misturado). Tudo misturado. Esse foi o primeiro que apareceu misturado (PGM-FE)*

[busca por AUTODOMÍNIO] *deixa eu ver essa daqui ... (E você não costuma ver nos artigos que você costuma ler AUTODOMÍNIO?) Não, quase nada. (...) Olha que estranho, não é nem o que a gente estuda, tipo “o autocontrole foi avaliado a partir da percepção dos participantes em contexto distintos, pessoais, familiares” ... não, muito estranho. Então não tem nada no AUTODOMÍNIO. Posso tentar separado? ... (Pode). Porque antes era com hífen, né. (...) [busca por AUTO-DOMÍNIO] Aqui até que ... tem umas coisas que ...~~ deixa eu ver aqui ... ele colocou, não foi nem auto-domínio aqui ... mas AUTOCONCEITO. Mas AUTOCONCEITO é da teoria dele. ...~~ Ele buscou no resumo. É. Eu acho que, pode colocar zero, porque não é o que eu estudo (PGD-FE)*

[busca por TEENAGERS - VIOLENCE AGAINST] *São sempre os maus-tratos em casa (...) [busca por MÁ-CONDUTA POLICIAL] Não sei porque de MÁ-CONDUTA POLICIAL vai pra questão econômica ... (É?) ...~~ Vem de tudo ... se eu não entendi não é da minha pesquisa ((RI)). (É, então. ((RI))). ...~~ Dois (G-IFCH)*

[busca por “CONSUMO ALIMENTAR”] *Não sei, todas as vezes que a gente colocou entre aspas parece que vem coisas mais da Biológicas, assim (...) [busca por ADULTO JOVEM] Meu Deus do céu ((RI)) “aspectos regenerativos em duas espécies arbóreas em um fragmento florestal do Sudeste ... brasileiro”. ...~~ Eu escrevi certo? ((RI)). (Sim. ((RI))). ...~~ Veio essas coisas dos ratos aqui de novo (...) [busca por “ADULTO JOVEM”] (...) quando você descobrir como ele escolhe isso você ... me avise ((RI)). ...~~ Eu não faço ideia de como eles escolhem essas coisas, mas assim, esse, tipo, JOVEM ADULTO, tipo, tudo é da medicina, não tem um das Ciências Humanas. ... Bom, eu também acho que nas Ciências Humanas a gente usa o JOVEM como um marcador. (...) No tesouro que a gente consultou, juventude ele utiliza ADOLESCENTE ...*

JUVENTUDE nas ciências Humanas a gente não ... de ADOLESCENTES não tem delimitador de idade (...) [busca por YOUNG ADULT] Nos artigos que eu trabalho eu nunca vi marcador como YOUNG ADULT. (Não é utilizado na área, né). (...) É, dá pra perceber que a maioria, acho que quase todos ((RI)). E a gente não tem um DeCS das Ciências Humanas? (Não. Na verdade o DeCS contempla alguns termos que são interdisciplinares, né, mas a maioria é da área da Saúde). (...) **Nada (PGD-IFCH)**

[busca por ESTUDOS DE GÊNERO] Tá vindo coisas de outras áreas. (Ainda vem?) (...) GÊNERO MUSICAL ... ESTUDOS QUIMIOTAX ... de tudo ((RI)) (...) [busca por GENDER STUDIES] Que engraçado, veio mais coisas de outras áreas (...) [busca por RURAL POPULATION] Muita coisa de saúde, sabe, impacto de doenças no meio rural, na população rural. ... ~ 6. (...) [busca por "RURAL POPULATION"] Eu tenho impressão que no repositório tem mais coisas da área de saúde, não tem? **(D-IFCH)**

[busca por FORMAÇÃO DE ANTICORPOS] Veio uma busca, um resultado bem rico mas ... aqui fica muito aberto também, né. Pra ir cruzar com HIV, né, tem que ter. ... ~ ~ ~ Aqui entra muito a questão do uso do ... de anticorpos no diagnóstico das doenças **(D-FCM)**

[busca por TECNOLOGIA - ASPECTOS SOCIAIS] tem umas coisas aqui ... ((RI)) "análise de sentimento em tweets" ((RI)). (...) Nossa tem umas coisas que não tem nada a ver, né **(G-FCA)**

[busca por TRIBUNAIS - ADMINISTRAÇÃO] "o brinquedo e a administração no contexto escolar" ((RI)). É porque ele busca as palavras né. (É, cabe pensar como JUSTIÇA entraria ali, né ((RI))) **(PGD-FCA)**

[busca por FORM FACTOR (NUCLEAR PHYSICS)] NUCLEAR PHYSICS acabou com tudo ... (É?) **(D-IFGW)**

[busca por "MONTE CARLO METHOD"] Não, ele tá indo mais pro lado da Física. (Na Física esse termo também é usado?) É usado. Aqui tem umas coisas de termodinâmica que não tem nada a ver, muita coisa vai pra outra área **(PGM-IMECC)**

[busca por VIBRATION MEASUREMENT] Pacientes com derrame, isso aqui já apareceu muito ((RI)). (Apareceu? ((RI))). ... ~ ~ ~ Esse aqui por exemplo, tem o ACELERÔMETRO, VIBRAÇÃO só que é de ... MÁQUINAS. (...) Então MÁQUINAS pra mim ... seria de alguém da Mecânica daí, né (...) [busca por DYNAMIC ANALYSIS] Esse apareceu menos ainda, uma só ((RI)) (...) [busca por "DYNAMIC ANALYSIS"] Eu achei engraçado aparecer coisas de ... cadeiras de rodas, de doença, de plantação, tem coisas da engenharia civil mesmo que é mais .. tipo, tem uma monte de coisa de Química e de Mecânica aqui que pra mim **(G-BAE)**

[busca por COMPÓSITOS LAMINADOS] (Esse termo ... COMPÓSITOS LAMINADOS ... nada?) Cai muito em outras linhas de COMPÓSITOS. (...) AEROSPACIAL **(D-BAE)**

Um dos usuários aponta a baixa quantidade de recursos informacionais recuperados devido ao tema ou à área do conhecimento de suas pesquisas (PGD-FEA).

[busca por SCANNING ELECTRON MICROSCOPY] (Teve muita coisa de outra área?) Muito. Acho que o pessoal da minha área não tá usando, sabe, isso como palavra-chave, é bom saber. (É que como vieram só 28.000 registros ((RI))). ((RI)) Então ... não dá pra gente saber muito bem né ((RI)) é, tem que puxar né, usar 'and', né, FOOD alguma coisa assim, né. (Às vezes não é porque eles não estão usando, às vezes porque aqui não recupera mesmo). (...) Muitos registros ... Mas é que tem poucos, tem trabalhos mas não tem muitos, né, daqui. ... ~ ~ ~ E aí como eu acho que é fraco na nossa área vai ficando mais pro fundo ... não sei qual é a ordem do algoritmo. (É, então ... é um mistério também ((RI))). ((RI)). (...) ... ~ ~ ~ (Colocaria que sim?) É um trabalho que eu colocaria, mas não sei se eu vou usar. (...) [busca por "SCANNING ELECTRON MICROSCOPY"] Um. Mesmo assim não teve muito retorno ... mas tem poucos mesmos trabalhos, sabe, eu acho que ... diante de todos os outros institutos o nosso deve ser o que menos tem **(PGD-FEA)**

Os usuários expressam não localizar exatamente os recursos informacionais que poderiam ser relevantes para sua pesquisa, porém, explicitam interesse em outros recursos informacionais recuperados (G-FE e PGD-FEF):

[busca por “DIFFERENCE (PHILOSOPHY)”] Às vezes não acha nada que vai colaborar, mas tem umas coisas muito legais ((RI)). (...) Dá curiosidade ((RI)) (G-FE)

[busca por STREET ART] Sabe o que é louco? É encontrar trabalhos que se eu não estivesse no doutorado eu com certeza leria mas que nesse momento ((RI)) eu tenho que ... (Focar). Focar. E eu acho isso academicamente ... absurdo. É ... como a gente se restringe com o nosso ... ao invés de ampliar o espectro do conhecimento a gente é obrigado a restringir esse espectro porque tem que fazer uma tese em quatro anos. (...) É muito absurdo ... me deixa estudar gente! ((RI)) (...) [busca por “STREET ART”] To me divertindo ((RI)). ((RI)) Depois você pode continuar se divertindo ((RI)). ((RI)). (Usando o repositório ((RI))). ((RI)). Seria tão legal se a gente pudesse fazer isso mas não, tem que ler ((RI)). (É, ((RI))). Ah, eu começo a fazer isso e começa a me dar um certo desespero ((RI)). ((RI)) Meu Deus, quando que eu vou ler isso ((RI)) (PGD-FEF)

Os usuários comparam os recursos informacionais recuperados com um determinado termo de busca com recursos informacionais recuperados em buscas anteriores, entre o termo em português e em inglês utilizando-se ou não o recurso gráfico aspas (“) (G-FE, PGM-FE, D-FE, PGD-IFCH, D-IFCH, PGD-FCM, PGM-FCA, PGD-FCA, D-FCA, D-IFGW e G-BAE).

[busca por SCHOOLS] É o mesmo, esse aqui apareceu da outra vez (G-FE)

[busca por EDUCAÇÃO SUPERIOR] Agora já apareceu outros ... outros trabalhos, alguns repetidos mas surgiram outros que não estavam (...) [busca por “PRIVATIZAÇÃO NA EDUCAÇÃO”] Esse tem menos (PGM-FE)

[busca por “EDUCAÇÃO CONTINUADA”] Nossa, 23 [a mesma quantidade do anterior] 23. Mas foi diferente a busca. (Diferente?) ... Muitas repetições mas diferente. Muitas repetições mas teve um ou outro diferente. (...) E deu pra perceber que a saúde usa muito, foi fácil ... saúde usa bastante esse termo. Tá, justamente porque ... daí foi fácil, é, e deu pra perceber também que nós de humanas que fizeram, usaram a palavra-chave é ... como ... sinônimo, no título era FORMAÇÃO CONTINUADA (...) [busca por AUTHORSHIP] Olha! 5. [percebe que foi a mesma quantidade em português para AUTORIA] mas não são os mesmos, tem repetição (D-FE)

[busca por CONSUMO ALIMENTAR] (Vieram mais resultados?) Vieram mais resultados dessa vez (PGD-IFCH)

[busca por “POPULAÇÃO RURAL”] (São os mesmos mais ou menos?) Mais ou menos, é, são diferentes, não são os mesmos (...) [busca por CLASSE SOCIAL] É ... recuperou trabalhos ... recuperou coisas diferentes mas recuperou trabalhos que estavam só com CLASSE também. (...) Mas tem coisas diferentes. ...~ 14 (D-IFCH)

[busca por FADIGA MUSCULAR] (Esses documentos que vieram ...). Então, só FADIGA apareceu, “percepção de fadiga”, que é uma coisa que eu to procurando. E agora não apareceu (...) [busca por PNEUMOCONIOSE - PREVENÇÃO E CONTROLE] Aqui apareceu novo (PGD-FCM)

[busca por LIPASES] Acho que são os mesmos, porque nenhum tá aparecendo no plural aqui, a palavra LIPASE nos títulos estão todos no singular. ... (Praticamente não teve muita diferença?) Não teve diferença, eu acho ((RI)). Marca a quantidade? (Isso). Ficou bem próximo ao outro (...) [busca por IMOBILIZAÇÃO

ENZIMÁTICA] Um artigo que eu vi anteriormente também conta? (Conta) ... Por que é um outro tipo de pesquisa? (Isso, outra busca) **(PGM-FCA)**

[busca por COURT ADMINISTRATION] Esse apareceu o mesmo ... Então ... Três **(PGD-FCA)**

[busca por SUGAR CANE MILLING] Eles meio que se repetem, né, eu vou pôr, eu vi antes mas com a nova palavra-chave também aparece **(D-FCA)**

[busca por FATOR DE FORMA (FÍSICA NUCLEAR)] Aqui, por exemplo, apareceu esses dois que não tinha aparecido lá (...) [busca por “CROMODINÂMICA QUÂNTICA”] (Foram registros diferentes ou os mesmos?) Os mesmos, não foram diferentes **(D-IFGW)**

[busca por “VIBRATION MEASUREMENT”] (É ... foram os mesmos registros mais ou menos?) Basicamente sim. (É?) Esse aqui acho que apareceu um pouquinho mais **(G-BAE)**

Outros usuários expressam o conhecimento dos autores dos recursos informacionais recuperados (D-FE e G-IFCH), incluindo orientadores (G-FEF e PGM-FEA) ou mesmo autoria própria (D-FCM).

[busca por PROFESSORES - FORMAÇÃO] (...) O título tem tudo a ver mas eu conheço quem orientou ... eu não pegaria esse estudo pra minha pesquisa. (...) Faço isso também, faço essa limpa **(D-FE)**

[busca por POLICE BRUTALITY] Ah, é do meu professor esse ... não sabia que ele pesquisava isso também **(G-IFCH)**

[busca por TERAPIA ANTIRRETROVIRAL DE ALTA ATIVIDADE] Já apareceu o meu trabalho ((RI)). ((RI)). ...~~ Mais um ((RI)). ...~~ Mais um ((RI)) **(D-FCM)**

[busca por EPILEPSIA DO LOBO TEMPORAL] Apareceu um estudo aqui do meu orientador e da minha coorientadora ((RI)). (É, fizeram juntos?) Sim. (Já conhecia?) Já, mas eu acho que não apareceu aquela hora, mas é que aqui também tá DE LOBO TEMPORAL. (...) Ou eu não vi também, né, o olho fica cansado ((RI)) **(G-FEF)**

[busca por EMERGY (SUSTAINABILITY)] (Do jeito que está aí). (...) 38. É que a maioria, quase todos são do meu orientador **(PGM-FEA)**

Os usuários anotam (D-FE) ou fotografam os registros recuperados considerados relevantes (D-FCM e PGD-FEA) ou mesmo o roteiro de busca por assuntos (PGM-IMECC).

[busca por AUTHORSHIP] Vou anotar esse aqui. ...~~ Esse eu vou copiar mesmo ... Interessante. É, nossa, super a ver com o que eu to atualmente ... estudando **(D-FE)**

[busca por “FORMAÇÃO DE ANTICORPOS”] Você me permite que eu fotografe só um ... (Claro ((RI))). Tem esse aqui, eu achei interessante, esse aqui eu não conheço e é daqui. ...~~ Depois eu dou uma pesquisada nele. ...~~ (...) Não é especificamente sobre a questão das vacinas **(D-FCM)**

[busca por “MONTE CARLO METHOD”] Posso tirar uma foto? **(PGM-IMECC)**

[busca por SCANNING ELECTRON MICROSCOPY] Posso tirar uma foto? ((RI)). Porque eu acho que eu não sei se eu usaria mas eu quero ler certinho, porque eles estão fazendo uma outra técnica igual à que eu estou fazendo ((RI)) **(PGD-FEA)**

8. Determinação do índice de precisão

A categoria “Determinação do índice de precisão” apresenta a análise estatística da precisão da recuperação da informação dos termos utilizados nas estratégias de busca, visando determinar a exatidão de cada linguagem de indexação na recuperação por assuntos entre a linguagem natural e a linguagem controlada das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP (BOCCATO; FUJITA; GIL LEIVA, 2011, p. 61; 57). A linguagem natural foi caracterizada pelas palavras-chave elencadas pelos usuários para a primeira rodada de busca por assuntos, enquanto que a linguagem controlada foi caracterizada pelos termos utilizados para a segunda rodada de buscas pelos usuários, após consulta e compatibilização das palavras-chave elencadas com termos/descriptores em linguagens controladas.

De acordo com a literatura, a medida de precisão - “capacidade de evitar recursos informacionais inúteis” - está diretamente ligada à medida de revocação - “capacidade de recuperar recursos informacionais úteis”. Embora existam outras medidas para calcular o desempenho para buscas realizadas em repositórios institucionais e fatores que influenciam na eficácia dos resultados de busca, a revocação e a precisão são medidas úteis para expressar os resultados de qualquer busca que divida um repositório institucional em duas partes: recursos informacionais recuperados e não recuperados (LANCASTER, 2004, p. 4). Especificamente para as entrevistas de busca por assuntos com os usuários no Repositório Institucional da UNICAMP, houve a necessidade de realização das buscas utilizando-se ou não o recurso gráfico aspas (“), devido ao número expressivo de recursos informacionais recuperados sem a utilização deste recurso, solicitando-se também aos participantes que limitassem a avaliação da relevância nos 50 primeiros resultados.

As respectivas palavras-chave da linguagem natural e os termos localizados nas linguagens controladas foram categorizados por área de conhecimento e por Biblioteca, sendo cada cor na linguagem natural correspondente na linguagem controlada (APÊNDICES I, J, K, L, M, N, O, P e Q). Houve apenas uma coincidência de atribuição de mesma palavra-chave: “Problema de corte de estoque” e sua correspondência “Cutting stock problem” na Biblioteca do IMECC, com os participantes G-IMECC e D-IMECC, sendo orientanda e orientadora

respectivamente. Além disso, algumas palavras-chave coincidiram entre o português e o inglês, sendo: 3 na *Área de Artes e Humanidades*, 2 na *Área de Biomédicas*, 1 na *Área de Exatas* e 2 na *Área de Tecnológicas*. Na *Área de Ciências Aplicadas* não houve ocorrência. Neste caso, não foi necessário realizar novamente a busca da respectiva palavra-chave em inglês, bastando a repetição do resultado da busca realizada com a palavra-chave em português.

Considerando-se que cada um dos 36 participantes elencou 5 palavras-chave em português e 5 palavras-chave em inglês (isto é, a versão em inglês das palavras-chave em português), no total foram elencadas 360 palavras-chave pelos participantes, sendo 180 em português e 180 em inglês. Em linguagem natural, foram realizadas 352 buscas de termos sem utilização do recurso gráfico aspas ("), com uma variação de 0 a 143.034 recursos informacionais recuperados e 214 buscas de termos com aspas (60,8% do total de termos buscados), com uma variação de 0 a 4.341 recursos informacionais recuperados.

Quanto à relevância com os termos utilizando-se o recurso gráfico aspas ("), 2.342 recursos informacionais foram indicados pelos participantes como relevantes, perfazendo uma média de 6,50 recursos informacionais relevantes por busca, com uma variação de 0 a 40 recursos informacionais relevantes. Já em relação à relevância com os termos com aspas, 1.159 recursos informacionais foram indicados pelos usuários como relevantes, perfazendo uma média de 5,41 recursos informacionais relevantes por busca, com uma variação de 0 a 45 recursos informacionais relevantes, pouco maior que os recursos informacionais buscados sem a utilização do recurso gráfico aspas (").

No que se refere à linguagem controlada, foram realizadas 208 buscas sem a utilização do recurso gráfico aspas ("), com a seguinte distribuição: *Área de Artes e Humanidades* (43), *Área de Biomédicas* (44), *Área de Ciências Aplicadas* (23), *Área de Exatas* (51) e *Área de Tecnológicas* (47), com uma variação de 5 a 143.048 recursos informacionais recuperados e 178 buscas de termos utilizando-se o recurso gráfico aspas (") (85,58% do total de termos elencados), com uma variação de 0 a 13.030 recursos informacionais recuperados.

Quanto à relevância com os termos sem a utilização do recurso gráfico aspas ("), 1.207 recursos informacionais foram indicados pelos participantes como relevantes, perfazendo uma média de 5,77 recursos informacionais relevantes por busca, com uma variação de 0 a 38 recursos informacionais relevantes. Já em relação à relevância com os termos com aspas, 644 recursos informacionais foram indicados pelos usuários como relevantes, perfazendo uma

média de 3,64 recursos informacionais relevantes por busca, com uma variação de 0 a 34 recursos informacionais relevantes.

Em relação ao número de termos buscados utilizando-se o recurso gráfico aspas (“) em linguagem natural mas sem nenhum resultado, ou seja, resultado zero ou nulo, a *Área de Tecnológicas* aparece com 11 buscas, seguida da *Área de Exatas*, com 10 buscas, da *Área de Artes e Humanidades*, com 8 buscas, da *Área de Ciências Aplicadas*, com 6 buscas e por fim a *Área de Biomédicas*, com 5 buscas, totalizando 40 buscas, com 22,47% de todas as buscas realizadas nesta categoria (178). Já em relação aos termos buscados com a utilização do recurso gráfico aspas (“) em linguagem controlada, a *Área de Biomédicas* aparece com 14 buscas, seguida da *Área de Exatas*, com 9 buscas, da *Área de Ciências Aplicadas*, com 7 buscas, da *Área de Artes e Humanidades*, com 6 buscas e da *Área de Tecnológicas*, com 4 buscas, totalizando também 40 buscas, com 18,69% de todas as buscas realizadas nesta categoria (214).

Visando à determinação do índice de precisão, foram elaborados dois quadros de registros em *Excel*: *Quadro A - Linguagem natural* e *Quadro B - Linguagem controlada*, contendo cada um as seguintes informações: assuntos (palavras-chave); assuntos recuperados sem a utilização do recurso gráfico aspas (“); assuntos relevantes sem a utilização do recurso gráfico aspas (“); *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“); *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“); assuntos recuperados utilizando-se o recurso gráfico aspas (“); assuntos relevantes utilizando-se o recurso gráfico aspas (“); *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“); e *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“). Nesse sentido, foram levantados um total de 8 índices de precisão, que possibilitaram a análise comparada dos índices de precisão dos resultados de busca entre a linguagem natural e a linguagem controlada.

A determinação dos índices de precisão foi pautada na relação existente entre o número de recursos informacionais relevantes recuperados e o número total de recursos informacionais recuperados no Repositório Institucional da UNICAMP, conforme a seguinte fórmula matemática, amplamente utilizada na literatura, onde a) *recursos informacionais*

relevantes recuperados com base na avaliação de relevância do usuário; e b) recursos informacionais recuperados pelo repositório institucional (Figura 43):

Figura 43. Fórmula para a determinação do índice de precisão no Repositório Institucional da UNICAMP

$$\text{Precisão} = \frac{\text{a) número de recursos informacionais relevantes}}{\text{b) número total de recursos informacionais recuperados}}$$

Fonte: Gil Leiva (2008, p. 392).

Aplicando-se a fórmula da precisão ou da *pertinência* (MOREIRO, 2002, p. 59) nas buscas por assuntos em **linguagem natural** foram obtidos os seguintes resultados (Tabela 3):

Tabela 3. Índices de precisão em linguagem natural no Repositório Institucional da UNICAMP

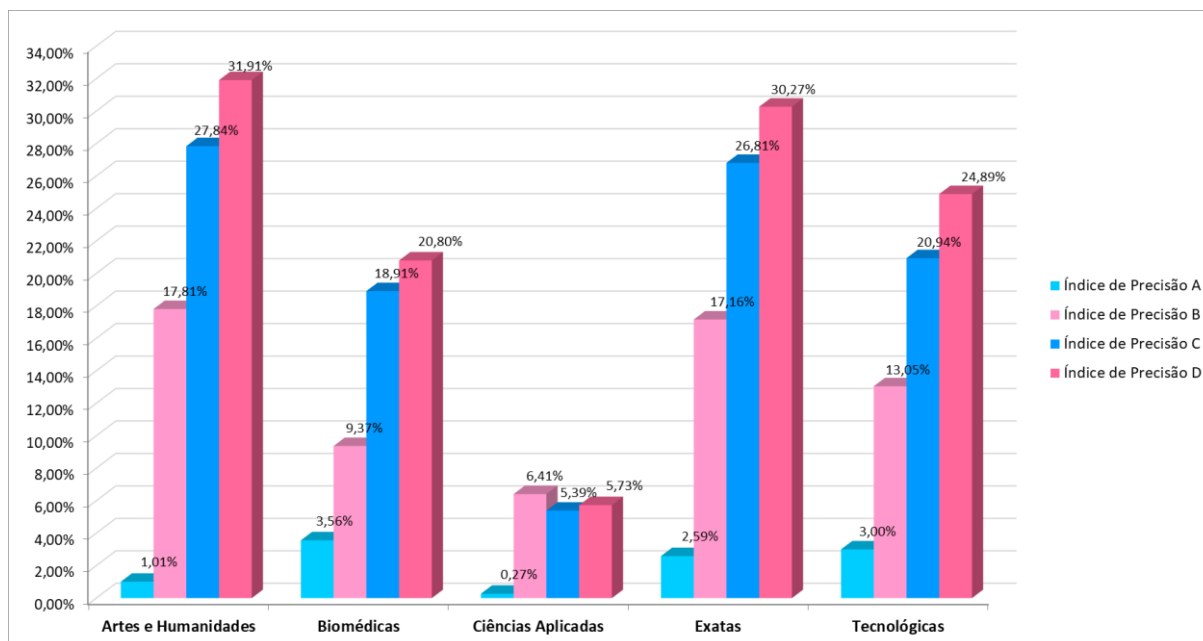
Área do conhecimento	Índice de precisão A	Índice de precisão B	Índice de precisão C	Índice de precisão D	Média por área do conhecimento
Artes e Humanidades	1,01%	17,81%	27,84%	31,91%	19,64%
Biomédicas	3,56%	9,37%	18,91%	20,80%	13,16%
Ciências Aplicadas	0,27%	6,41%	5,39%	5,73%	4,45%
Exatas	2,59%	17,16%	26,81%	30,27%	19,21%
Tecnológicas	3,00%	14,48%	20,94%	24,89%	15,82%
Média por índice	2,08%	7,09%	19,98%	22,72%	-

Fonte: Elaboração própria.

Calculando-se a média dos índices de precisão na recuperação por assuntos no Repositório Institucional da UNICAMP em linguagem natural, verifica-se que o maior nível foi obtido com o *índice de precisão D*, com o termo de busca utilizando-se o recurso gráfico aspas (“) e limitando-se no máximo em 50 recursos informacionais recuperados (22,72%). Em seguida, aparece o *índice de precisão C*, com o termo de busca utilizando-se o recurso gráfico aspas (“) sem limitação de recursos informacionais recuperados (19,98%), mantendo-se certo equilíbrio entre os índices encontrados. Por outro lado, aparece o *índice de precisão B* - sem a utilização do recurso gráfico aspas (“) e limitando-se no máximo em 50 recursos

informativos recuperados, com 7,09% e por fim o *índice de precisão A* - sem a utilização do recurso gráfico aspas e sem limitação de recursos informativos recuperados, com apenas 2,08% de precisão na recuperação (Figura 44):

Figura 44. Índices de precisão em linguagem natural por área do conhecimento



Fonte: Elaboração própria.

Em relação aos índices de precisão em linguagem natural por área do conhecimento, verifica-se que a média dos índices de precisão é maior na *Área de Artes e Humanidades*, com 19,64%, seguida da *Área de Exatas*, com 19,21%, índices bem semelhantes. Em seguida, aparece a *Área de Tecnológicas*, com 15,82% de índice de precisão, a *Área de Biomédicas*, com 13,16% e por fim a *Área de Ciências Aplicadas*, com apenas 4,45% de índice de precisão na recuperação com a linguagem natural. Já em relação à linguagem controlada, aplicando-se a fórmula da precisão nas buscas por assuntos foram obtidos os seguintes resultados (Tabela 4):

Tabela 4. Índices de precisão em linguagem controlada no Repositório Institucional da UNICAMP

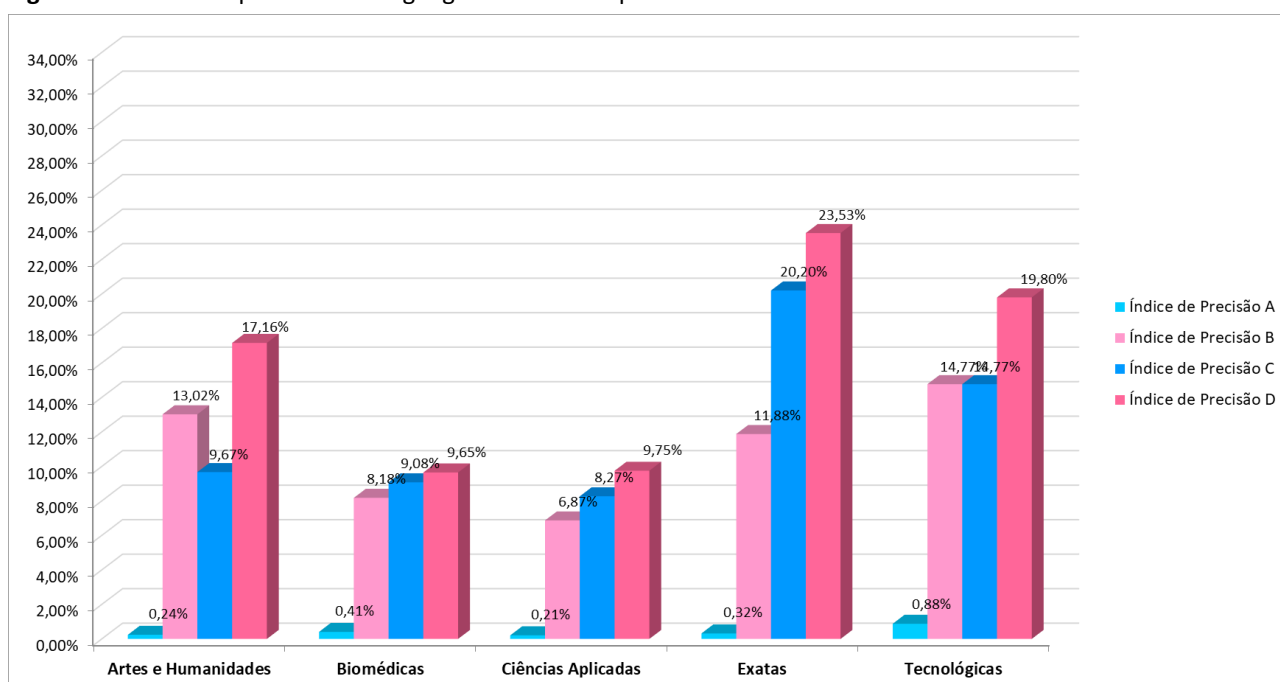
Área do conhecimento	Índice de precisão A	Índice de precisão B	Índice de precisão C	Índice de precisão D	Média por área do conhecimento
Artes e Humanidades	0,24%	13,02%	9,67%	17,16%	10,02%
Biomédicas	0,41%	8,18%	9,08%	9,65%	6,83%

Ciências Aplicadas	0,21%	6,87%	8,27%	9,75%	6,28%
Exatas	0,32%	11,88%	20,20%	23,53%	13,98%
Tecnológicas	0,88%	14,77%	14,77%	19,80%	12,55%
Média por índice	0,41%	10,95%	12,40%	15,98%	-

Fonte: Elaboração própria.

Calculando-se a média dos índices de precisão na recuperação por assuntos no Repositório Institucional da UNICAMP em linguagem controlada, verifica-se que o maior nível foi obtido com o *índice de precisão D*, com o termo de busca utilizando-se o recurso gráfico aspas (“) e limitando-se no máximo em 50 recursos informacionais recuperados (15,98%). Em seguida, aparece o *índice de precisão C*, com o termo de busca utilizando-se o recurso gráfico aspas (“) sem limitação de recursos informacionais recuperados (12,40%). Por outro lado, aparece o *índice de precisão B* - sem a utilização do recurso gráfico aspas (“) e limitando-se no máximo em 50 recursos informacionais recuperados, com 10,95% e por fim o *índice de precisão A* - sem a utilização do recurso gráfico aspas (“) e sem limitação de recursos informacionais recuperados, com apenas 0,41% de precisão na recuperação em linguagem controlada (Figura 45):

Figura 45. Índices de precisão em linguagem controlada por área do conhecimento

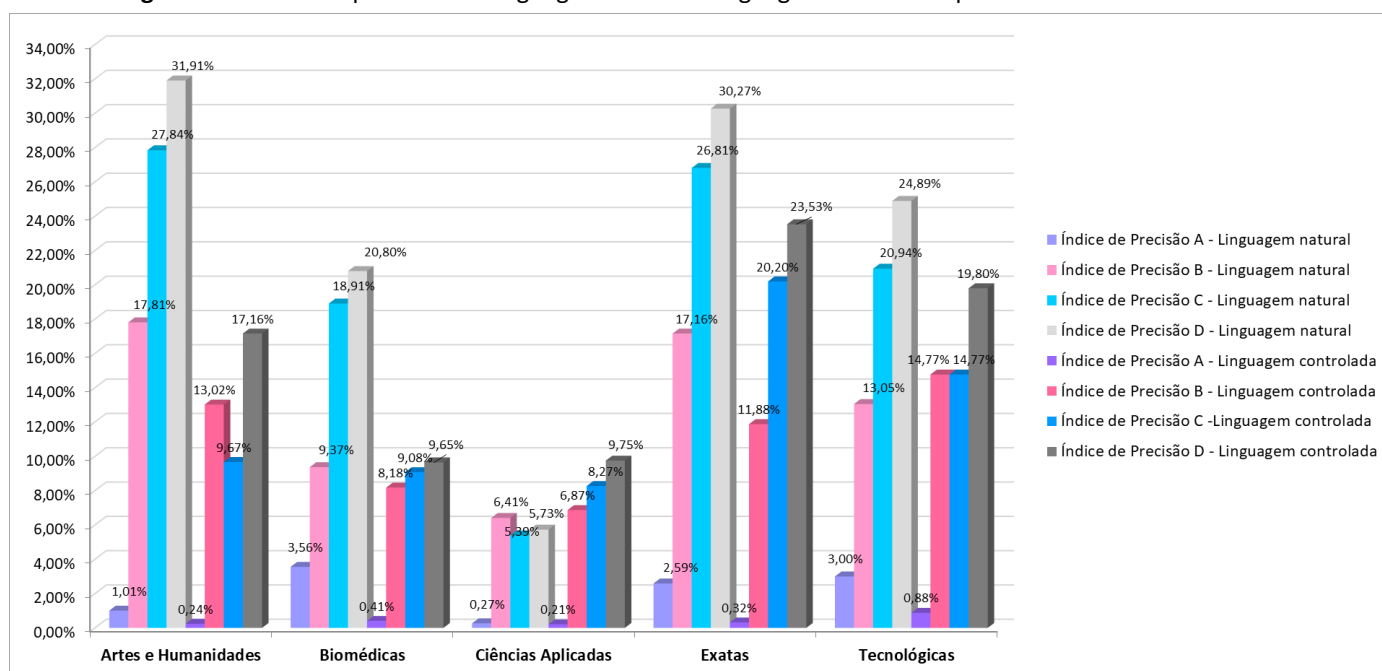


Fonte: Elaboração própria.

Em relação aos índices de precisão em linguagem controlada por área do conhecimento, verifica-se que a média é maior na *Área de Exatas*, com 13,98%, seguida da *Área de Tecnológicas*, com 12,55%. Em seguida, aparece a *Área de Artes e Humanidades*, com 10,02% de índice de precisão, a *Área de Biomédicas*, com 6,83% e por fim a *Área de Ciências Aplicadas*, com apenas 6,28% de índice de precisão na recuperação com a linguagem controlada.

Em uma análise sistematizada por área do conhecimento, ao realizar-se a comparação dos índices de precisão entre a linguagem natural e a linguagem controlada na recuperação por assuntos no Repositório Institucional da UNICAMP foram obtidos os seguintes resultados (Figura 46):

Figura 46. Índices de precisão em linguagem natural e linguagem controlada por área do conhecimento



Fonte: Elaboração própria.

Na *Área de Artes e Humanidades* houve um melhor desempenho na precisão da recuperação por assuntos com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (") em linguagem natural (31,91%) - maior índice dos 8 índices levantados, seguido do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (") em linguagem natural (27,84%), do *índice de precisão B* - Delimitação de 50 recursos informacionais

recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (17,81%) e do *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (17,16%). Com menores índices de precisão estão o *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (13,02%), o *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (9,67%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (1,01%) e em linguagem controlada (0,24%), respectivamente.

Na *Área de Biomédicas* houve um maior índice de precisão com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (20,80%) e com o *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (18,91%). Com menores níveis de precisão apareceram o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (9,65%), seguido do *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (9,37%) e do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (9,08%). Em continuidade, aparecem o *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada (8,18%), o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (3,56%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada (0,41%).

Na *Área de Ciências Aplicadas* os maiores índices de precisão foram alcançados com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (9,75%), seguido do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (8,27%), do *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada (6,87%) e do *índice de precisão B* - Delimitação de 50 recursos informacionais

recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (6,41%). Com índices semelhantes, apareceram o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (5,73%), o *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (5,39%), o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (0,27%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada, com apenas 0,21%, menor índice dos 8 índices levantados.

Na *Área de Exatas* os maiores índices de precisão foram obtidos com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (30,27%), seguido do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (26,81%), do *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (23,53%) e do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (20,20%). Um pouco mais abaixo aparece o *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural e em linguagem controlada com 17,16% e 11,88%, respectivamente. Com índices ainda menores, aparecem o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (2,59%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada (0,32%).

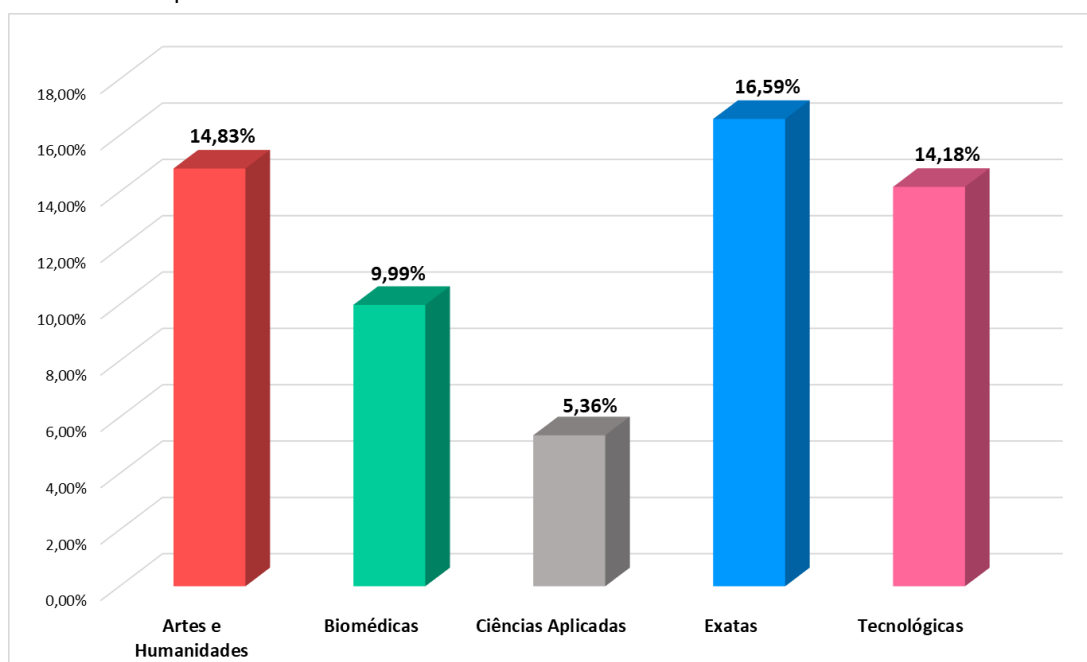
Já na *Área de Tecnológicas*, os maiores índices de precisão foram alcançados com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (24,89%), seguido do *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem natural (20,94%), do *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“) em linguagem controlada (19,80%), do *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural e do *índice de precisão C* -

Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspás (“), ambos com 14,77% e do *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“) em linguagem natural (13,05%). Com índices de precisão menores aparecem o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“) em linguagem natural (3,0%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“) em linguagem controlada (0,88%).

Em síntese, houve um maior desempenho na recuperação por assuntos no Repositório Institucional da UNICAMP com o *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspás (“) em linguagem natural na *Área de Artes e Humanidades* (31,91%), *Área de Exatas* (30,27%), *Área de Tecnológicas* (24,89%) e *Área de Biomédicas* (20,80%), com exceção da *Área de Ciências Aplicadas*, para quem o melhor índice de precisão foi alcançado com o mesmo índice, porém, em linguagem controlada (9,75%). Por outro lado, o menor desempenho ocorreu com o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“) em linguagem controlada, nas cinco áreas do conhecimento: *Área de Tecnológicas* (0,88%), *Área de Biomédicas* (0,41%), *Área de Exatas* (0,32%), *Área de Artes e Humanidades* (0,24%) e *Área de Ciências Aplicadas* (0,21%).

Calculando-se a média dos índices de precisão na recuperação por assuntos em linguagem natural e linguagem controlada no Repositório Institucional da UNICAMP por área do conhecimento, ou seja, por meio da somatória dos resultados dos 8 oito índices (*índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“); *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspás (“); *índice de precisão C* - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspás (“); e *índice de precisão D* - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspás (“), dividindo-se pelo total dos 8 índices analisados, foram obtidos os seguintes resultados (Figura 47):

Figura 47. Média dos índices de precisão entre linguagem natural e linguagem controlada por área do conhecimento no Repositório Institucional da UNICAMP



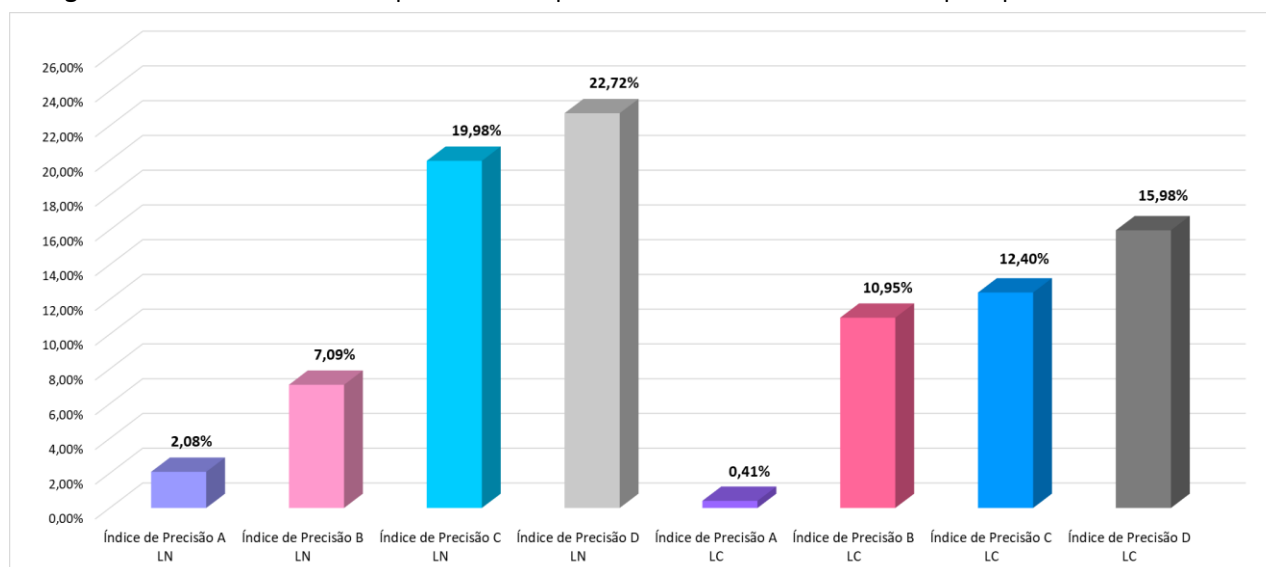
Fonte: Elaboração própria.

O maior índice de precisão foi obtido na *Área de Exatas*, com 16,59% de precisão na recuperação por assuntos, seguida da *Área de Artes e Humanidades*, com 14,83% e da *Área de Tecnológicas*, um pouco mais abaixo, com 14,18%. Em continuidade, os menores índices de precisão ocorreram com a *Área de Biomédicas*, com 9,99% e, por fim, com a *Área de Ciências Aplicadas*, com apenas 5,36% de precisão na recuperação por assuntos em linguagem natural e linguagem controlada.

Ao realizar-se a avaliação comparada da média dos índices de precisão entre a linguagem natural e a linguagem controlada na recuperação por assuntos no Repositório Institucional da UNICAMP por tipo de índice, verifica-se que o maior nível foi obtido com o *índice de precisão D - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“)* em linguagem natural (22,72%), seguido do *índice de precisão C - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“)* em linguagem natural (19,98%), do *índice de precisão D - Delimitação de 50 recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“)* em linguagem controlada (15,98%) e do *índice de precisão C - Todos os recursos informacionais recuperados utilizando-se o recurso gráfico aspas (“)* em linguagem controlada (12,40%).

Por outro lado, verifica-se que um menor nível de precisão foi obtido com *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada (10,95%) e do *índice de precisão B* - Delimitação de 50 recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (7,09%). Por fim, aparecem o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem natural (2,08%) e o *índice de precisão A* - Todos os recursos informacionais recuperados sem a utilização do recurso gráfico aspas (“) em linguagem controlada, com apenas 0,41% de precisão na recuperação por assuntos (Figura 48):

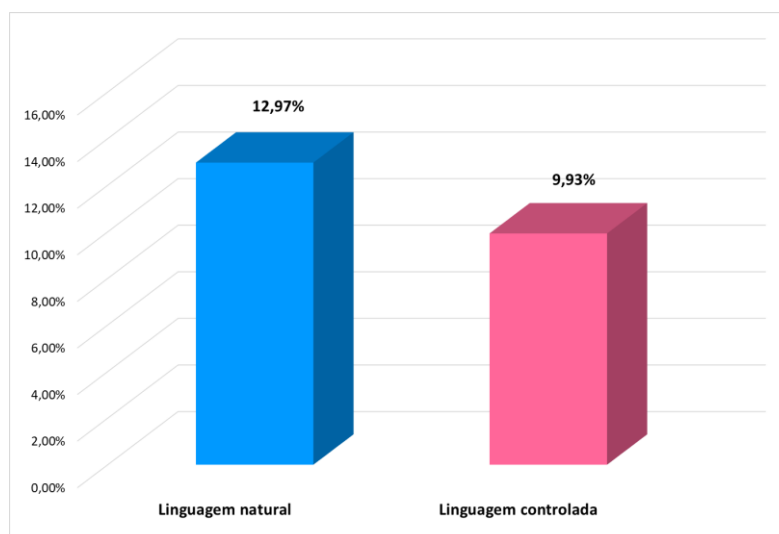
Figura 48. Média dos índices de precisão no Repositório Institucional da UNICAMP por tipo de índice



Fonte: Elaboração própria.

No geral, a linguagem controlada utilizada no Repositório Institucional da UNICAMP alcançou um índice de precisão na recuperação por assuntos menor em relação à linguagem natural, perfazendo 9,93% em relação a 12,97% alcançado pela linguagem natural (Figura 49):

Figura 49. Índice total de precisão das linguagens de indexação na recuperação por assuntos no Repositório Institucional da UNICAMP



Fonte: Elaboração própria.

Em estudo realizado por Boccato, Fujita e Gil Leiva (2011, p. 62) na comparação dos índices de precisão de recuperação por assuntos no catálogo online Athena da UNESP da *Lista de Cabeçalhos de Assunto da Rede BIBLIODATA* (LCCARB) - amplamente utilizada na Rede UNESP com outras duas linguagens: o VOCAUSP e a linguagem natural, isto é, a linguagem utilizada pelos usuários nas buscas por assuntos, os índices de precisão alcançados foram de 6,5% para a LCCARB, 7,8% para o VOCAUSP e 7,2% para a linguagem natural. Para os autores, a pouca diferença nos resultados está ligada possivelmente “à realização de estratégias de busca em modo simples e não avançado no catálogo online” e consideram a necessidade de continuidade da referida pesquisa.

Considerando-se que os índices de precisão podem variar entre 0% a 100% entre os recursos informacionais recuperados e os recursos informacionais relevantes recuperados, os resultados apresentam baixos índice de precisão na recuperação por assuntos tanto em linguagem natural (12,97%) quanto em linguagem controlada (9,93%) no Repositório Institucional da UNICAMP, no universo de recursos informacionais e de participantes analisado.

9. Representação descritiva

A categoria “Representação descritiva” refere-se à forma de apresentação dos resultados das buscas por assuntos no Repositório Institucional da UNICAMP, verificada em quatro principais aspectos: *forma de apresentação geral, data de publicação, título e autoria*, tanto pela pesquisadora quanto pelos próprios usuários durante as entrevistas.

No tocante à apresentação geral dos resultados, não há padronização na apresentação das pré-visualizações dos recursos informacionais, verificada apenas no caso de patentes. Nos demais recursos informacionais, não há a visualização (Figura 50):

Figura 50. Exemplo de pré-visualização de registros no Repositório Institucional da UNICAMP

Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

	31-Jan-1989	Sistema Multicanal De Deteccao Espectral Usando Fibra Optica	Marotta Aruy	-	Patente
	29-Nov-2013	Coating Compositions With Anticorrosion Properties	Foscante Raymond E [us]; Galembek Fernando [br]; Braga Melissa [br]	-	Patente
	27-Jun-2014	Coating Compositions With Anticorrosion Properties	Galembek Fernando[br]; Foscante Raymond E[us]; Braga Melissa[br]	-	Patente

[Artigo de evento](#)
[DISSERTAÇÃO](#)
[TESE](#)
[Artigo](#)
[TESE DIGITAL](#)
[Capítulo de livro](#)
[DISSERTAÇÃO](#)
[Editorial](#)

Fonte: Dados da pesquisa.

Um ponto a ser destacado é a duplicação de registros, notadamente em recursos informacionais do tipo artigo e artigo de periódico (Figura 51):

Figura 51. Exemplo de duplicação de registros entre artigo e artigo de periódico no Repositório Institucional da UNICAMP

Conjunto de itens:					
Pré-visualização	Data do documento	Título	Autor(es)	Orientador	Tipo
	2016	Mechanical Performance Of Glued-laminated Timber Beams Symmetrically Reinforced With Steel Bars	Soriano, J; Pellis, BP; Mascia, NT	-	Artigo de Periódico
	2016	Mechanical Performance Of Glued-laminated Timber Beams Symmetrically Reinforced With Steel Bars	Soriano, Julio; Pellis; Bruno Piva; Mascia; Nilson Tadeu	-	Artigo
	2016	Estudo de vigas de madeira de reflorestamento laminadas coladas reforçadas por fibras	Donadon, Bruno Fazendeiro, 1985-	Mascia, Nilson Tadeu, 1956-	DISSERTAÇÃO DIGITAL

Fonte: Dados da pesquisa.

A questão da duplicação de registros é manifestado pelos usuários (D-FE, D-FCM, G-FEF, PGM-FCA, PGD-FCA, D-FCA e D-BAE).

[busca por LETRAMENTO ACADÊMICO] *É ... deixa eu te fazer uma pergunta, até anotei pra não perder a contagem ... como eu conheço, quer dizer, um monte que eu conheço desde a primeira vez, no caso específico aqui ... eu sei que esse é igual a esse e apareceu duas vezes, o que que eu faço?* (Aí você conta uma vez só). *Eu conto uma vez só? Eu sei que é igual. Porque que tá duas vezes, hein.* (...) (Tem algumas inconsistências (...) que estamos corrigindo). *Tá, é isso.* (...) (Acontece muito quando peamos em lote das bases de dados). *É esse aqui, olha, engraçado, tá* (D-FE)

[busca por "TERAPIA ANTIRRETROVIRAL"] *Também tem uma razoável repetição entre os primeiros 50.* (...) ...~~ (De documentos duplicados?) *É, isso* (D-FCM)

[busca por COMORBIDITY] *Quando dá duplicado é só um, né.* (...) (Pode contar um só). *Nossa, três duplicados.* (Tem três?) (G-FEF)

[busca por IMOBILIZAÇÃO ENZIMÁTICA] *Acho que esse artigo e esse aqui acho que é o mesmo artigo, ele só tá em ... inglês e português. Conta? (Ah, é?) Olha, é o mesmo artigo.* (...) (É, conta um só) (...) [busca por "IMOBILIZAÇÃO ENZIMÁTICA"] *É ... tem muito artigo duplicado ... querendo ou não você não lembra o nome de um artigo que você já utilizou, você abre ali, não é um trabalho a mais, a hora que você tá no meio você fala, nossa, eu acho que eu já vi isso em algum lugar, você volta no trabalho anterior e você vê que tá de novo na mesma coisa* (PGM-FCA)

[buscando por JUDICIAL PERFORMANCE] *Esse eu já vi, tem uns que aparecem repetidos, né?* (PGD-FCA)

[busca por BAGASSE] *Tenho impressão que esse é o mesmo que eu já pontuei como um, que no final parece que é o mesmo conceito de BIOREFINARIA, parece bem semelhante ...~~ olha, esse também é duplicado* (D-FCA)

[busca por GLUED LAMINATED TIMBER] *Tem duas vezes o mesmo trabalho.* (Duplicado?) *É ... artigo e artigo de periódico,* (...) *até pelas definições de ... das citações, né?* (...) *É, olha, o nome dos autores, um completo e outro não* (D-BAE)

Além da ocorrência entre recursos informacionais do tipo artigo e artigo de periódico, a duplicidade aparece entre os próprios artigos de periódico (Figura 52):

Figura 52. Exemplo de duplicação de registros entre artigos de periódico no Repositório Institucional da UNICAMP

The screenshot shows the UNICAMP Institutional Repository interface. At the top, there is a navigation bar with links for 'Página inicial', 'Sobre', and 'Contato', along with a search bar labeled 'Buscar no repositório'. Below the navigation bar, a table displays search results. Two records are shown, which are duplicates of each other. The first record is dated '1-Mar-2010' and the second is dated '2010'. Both records have the same title: 'Body mass index percentiles in adolescents of the city of São Paulo, Brazil, and their comparison with international parameters'. The authors listed are 'Passos, Maria Aparecida Zanetti; Cintra, Isa de Pádua; Branco, Lúcia Maria; Machado, Helymar da Costa; Fisberg, Mauro' for the first record and 'Passos, MAZ; Cintra, ID; Branco, LM; Machado, HD; Fisberg, M' for the second record. Both records are classified as 'Artigo de periódico'.

1-Mar-2010	Body mass index percentiles in adolescents of the city of São Paulo, Brazil, and their comparison with international parameters	Passos, Maria Aparecida Zanetti; Cintra, Isa de Pádua; Branco, Lúcia Maria; Machado, Helymar da Costa; Fisberg, Mauro	-	Artigo de periódico
2010	Body mass index percentiles in adolescents of the city of Sao Paulo, Brazil, and their comparison with international parameters	Passos, MAZ; Cintra, ID; Branco, LM; Machado, HD; Fisberg, M	-	Artigo de periódico

Fonte: Dados da pesquisa.

A questão é também manifestada pelos usuários (G-FEF, D-FCA e D-BAE).

[busca por "SÍNDROME METABÓLICA"] *Nossa, que estranho isso, o mesmo título mas de autores diferentes* ... (É ... na verdade é o mesmo autor, são os mesmos autores só que esse tá abreviado e esse não, é o mesmo registro, é uma duplicação) (G-FEF)

[busca por SUGAR CANE MILLING] *Eu só acho que ele repete muito, acho que tem muitos trabalhos colocados* ... (Em duplicidade?) *Em duplicidade, exatamente. Daí tu vê que autor num momento o sobrenome tá escrito inteiro no outro ele tá escrito abreviado só com a sigla, né, o 's'. né, depois [o nome completo] ... então vários aparecem em duplicidade. (...) E daí na contagem eles aparecem inclusive nas minhas em duplicidade, né, porque às vezes eu acho que eu vi lá encima, não voltei pra ver se era o que eu tinha visto, mas tenho quase certeza que era o mesmo. (...) Mas acho que o que senti mesmo foi mais a duplicidade* (D-FCA)

[busca por "WOODEN CONSTRUCTION"] (É o mesmo?) *É o mesmo. A diferença é que ele põe ponto no ... (É) ... A citação também, acho que tudo por conta da forma de citação.* (É, um pontinho já acaba duplicando) (D-BAE)

Em continuidade, observa-se a visualização tanto da apresentação do título em português e em inglês quanto da duplicação do registro, questão também expressa pelos usuários (D-FE, PGM-FCA e G-IMECC) (Figura 53):

Figura 53. Exemplo de duplicação de registros em artigo de periódico com títulos em inglês e português no Repositório Institucional da UNICAMP

The screenshot shows the website header with navigation links: 'Página inicial', 'Sobre', and 'Contato'. A search bar contains the text 'Buscar no repositório'. Below the header, a table displays search results. The first record is for the year 2001, with the title 'Low Vision Student In The Regular School: A Challenge To The Teacher? [o Aluno Portador De Visão Subnormal Na Escola Regular: Desafio Para O Professor?]' and authors 'Gasparetto M.E.R.F.; Temporini E.R.; Carvalho K.M.M.; Kara-Jose N.'. The second record is for the date '1-Fev-2001', with the title 'O aluno portador de visão subnormal na escola regular: desafio para o professor?' and authors 'Gasparetto, Maria Elisabete R. Freire; Temporini, Edméa Rita; Carvalho, Keila Miriam Monteiro; Kara-José, Newton'. Both records are categorized as 'Artigo de periódico'. To the right of the table, there are buttons for 'aberto' and 'fechado'.

Fonte: Dados da pesquisa.

[busca por AUTHORSHIP] *Esse aqui é gigante. É, acho que repete o título em português e em inglês. É, mas não tem quase informação ... (...) não, não vou pegar ... ele não tem informação ... não tem autoria ... não tem informação. Bom, não vai ... é que é bem próximo mas é ... mas não tem informação (D-FE)*

[busca por IMMOBILIZATION] *então parece que os títulos são bem parecidos ... ele não faz a busca pelo mesmo título? Não foi cadastrado dois títulos? ... (...) Porque às vezes o repositório pode ser cadastrado com o título em inglês e em português ... (...) se ele foi cadastrado com os dois tipos então fazendo a busca tanto em inglês como em português ele vai me trazer os mesmos artigos que eu já vi em português, então eu vou estar olhando de novo o que eu acabei de ver. ...~~ Repete o mesmo artigo ... o primeiro artigo aqui ... (...) E o artigo aqui embaixo olha ... (Ah, tá, está duplicado) (PGM-FCA)*

[busca por OTIMIZAÇÃO MULTI OBJETIVO] *Nossa, olha o tamanho desse título. ...~~ (É que tem a versão em inglês e em português no mesmo) (G-IMECC)*

Além da duplicação de registros, verifica-se a duplicação de resumos nos próprios registros, a exemplo da busca realizada por um dos usuários com o termo de busca "Mango peel" (G-FEA) (Figura 54):

Figura 54. Exemplo de duplicação de resumos em registro no Repositório Institucional da UNICAMP

The screenshot shows the website interface of the UNICAMP Institutional Repository. At the top, there is a navigation bar with links for 'Página inicial', 'Sobre', and 'Contato', along with a search bar and language options. Below the navigation bar, the authors listed are Durigon, A., Hubinger, M. D., and Laurindo, J. B. The main content area displays a duplicated abstract for a research paper. The abstract text is repeated twice, with the second instance being a slightly shorter version of the first. The text describes the production of mango powder using spray drying and cast-tape drying, comparing the two processes in terms of moisture, particle size, bulk density, porosity, morphology, and total carotenoids content. It concludes that cast-tape drying is a suitable procedure for producing mango powders from whole fruit pulp without the addition of maltodextrin.

Fonte: Dados da pesquisa.

Em relação à apresentação das datas de publicação dos recursos informacionais, a maior parte aparece no formato “ano” (2018). Entretanto, em outros no formato “1-Abr-2008”, como é o caso de patentes, ou mesmo de forma incorreta (Figura 55):

Figura 55. Exemplo de visualização de datas no Repositório Institucional da UNICAMP

The screenshot shows the search results page of the UNICAMP Institutional Repository. The page features a navigation bar at the top with the same links as Figure 54. Below the navigation bar, there is a table of search results. The table has five columns: a date column, a title column, an author column, a language column, and a document type column. The first row shows a date of '-1-Uns- -1', a title 'Association Between Oral Health-related Quality Of Life And Atraumatic Restorative Treatment In School Children: An Exploratory Study.', authors 'Paula, J S; Tôrres, L H N; Ambrosano, G M B; Mialhe, F L', a language of '-', and a document type of 'Artigo de periódico'. The second row shows a date of '2016', a title 'Sixty Years Of The National Food Program In Brazil', authors 'Nogueira; Rosana Maria; Barone; Bruna; de Barros; Thiera Teixeira; Lima de Queiroz; Guimaraes; Katia Regina Leoni Silva; Sabbiao Rodrigues; Nilo Sergio; Behrens; Jorge Herman', a language of '-', and a document type of 'Artigo'.

Fonte: Dados da pesquisa.

Outro ponto verificado foi a ausência de padronização nos títulos dos recursos informacionais, com um maior nível de padronização em dissertação, dissertação digital, tese e tese digital, neste caso, iniciando-se somente com a primeira palavra do título com a letra

maiúscula. Em outros, as palavras do título iniciam-se com letras maiúsculas, caso de artigos, artigos de periódico e patentes (Figura 56):

Figura 56. Exemplo de visualização de inconsistência em títulos no Repositório Institucional da UNICAMP



Repositório da Produção Científica e Intelectual da Unicamp

Página inicial Sobre Contato

Buscar no repositório

2008	Carga viral vaginal de HIV em mulheres brasileiras infectadas pelo HIV	Villarreal, Marina; Amaral, Eliana; Portugal, Priscila; Campos, Angela; Levi, José Eduardo; Moraes, Sirlei Siani; Silva, Marcos T. Nolasco da; Bezerra, Karina C.	-	Artigo
2002	Hiv Infection In Women: Impact On Contraception.	Magalhães, Jarbas; Amaral, Eliana; Giraldo, Paulo Cesar; Simoes, Jose Antonio	-	Artigo de periódico
2001	Avaliação do uso do condom feminino em mulheres vivendo com o HIV	Magalhães, Jarbas	Silva, Eliana Martorano Amaral Freitas da	TESE
2008	Hiv Vaginal Viral Load In Brazilian Hiv-infected Women [carga Viral Vaginal De Hiv Em Mulheres Brasileiras Infectadas Pelo Hiv]	Campos A.; Amaral E.; Levi J.E.; Portugal P.; Villarreal M.; Bezerra K.C.; Nolasco Da Silva M.T.; Moraes S.S.	-	Artigo de periódico

Tipo de Acesso
 fechado
 aberto
 embargo

Fonte: Dados da pesquisa.

Referente à autoria, em recursos informacionais do tipo dissertação, dissertação digital, tese e tese digital verifica-se a existência de um maior nível de padronização dos autores, tendo o seguinte formato: “Fernandes, Eduardo Manoel”, enquanto que em outros, tais como artigos de periódico, não há padronização, a exemplo de “Fernandes, M.; Pires J.C.” ou “Fernandes M.; Pires JC.”, ambos sem a utilização da vírgula na separação de sobrenome e nome dos autores (Figura 57):

Figura 57. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - A



Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

[Entrar em](#)
[Idioma](#)

2002	Investigation of nonmetallic inclusions in continuously cast carbon steel by dissolution of the ferritic matrix	Fernandes, M.; Cheung, N.; Garcia, A	-	Artigo de periódico
2003	Influence Of Refining Time On Nonmetallic Inclusions In A Low-carbon, Silicon-killed Steel	Fernandes M.; Pires J.C.; Cheung N.; Garcia A.	-	Artigo de periódico
2002	Investigation of the chemical composition of nonmetallic inclusions utilizing ternary phase diagrams	Fernandes, M.; Pires, J.C.; Cheung, N.; Garcia, A	-	Artigo de periódico
2005	Physico-chemical characterization of the inclusion complex between a 2-propen-1-amine derivative and beta-cyclodextrin	De Souza, AO; Alderete, JB; Fajoni-Ataró, A; Silva, CL; Duran, N	-	Artigo de periódico

Ministry of Education and Youth o... 30
 próximo >

Tipo de Documento

Artigo de periódico	7535
DISSERTAÇÃO	761
TESE	570
Artigo de evento	295
Artigo	229
TESE DIGITAL	197
DISSERTAÇÃO DIGITAL	103

Fonte: Dados da pesquisa.

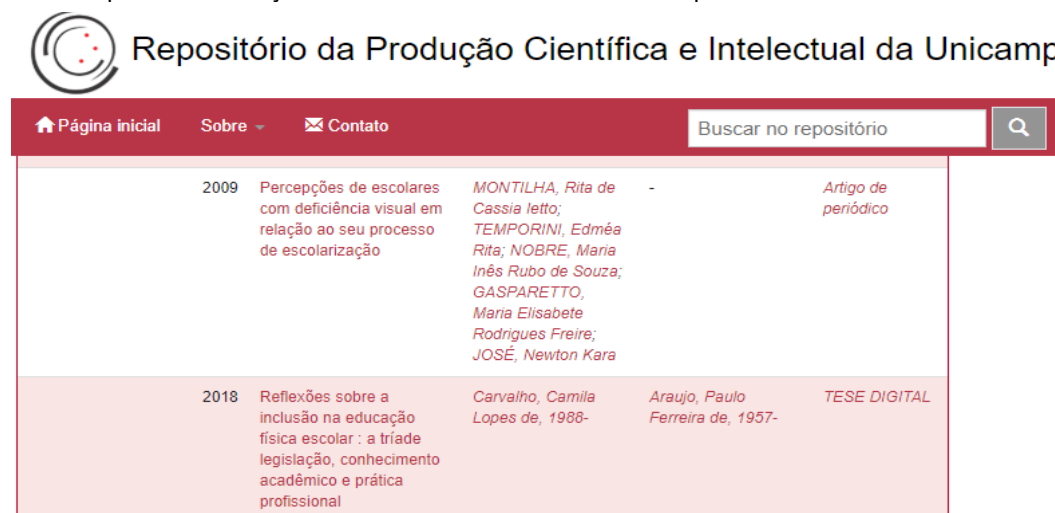
A questão da ausência de padronização no nome dos autores também é manifestada pelos usuários (D-IFGW e D-BAE):

[busca por QCD] ...~ Ah, meu nome tá escrito errado aqui. (É o seu?) É, tá com 'h', é com 'l' (D-IFGW)

[busca por "WOODEN CONSTRUCTION"] Mas é ... um problema relacionado a ... em todos eles aconteceram foi o fato de ... é ... da forma de citação. (...) (Nome dos autores ...). Dos autores. (...) Isso foi ... (Uma padronização, né). Mas acho que em termos de ... que é nosso mesmo tá bem ... legal (D-BAE)

Já em outros recursos informacionais verifica-se que somente o sobrenome do autor é visualizado em caixa alta (Figura 58):

Figura 58. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - B



Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

2009	Percepções de escolares com deficiência visual em relação ao seu processo de escolarização	MONTILHA, Rita de Cassia letto; TEMPORINI, Edmêa Rita; NOBRE, Maria Inês Rubo de Souza; GASPARETTO, Maria Elisabete Rodrigues Freire; JOSÉ, Newton Kara	-	Artigo de periódico
2018	Reflexões sobre a inclusão na educação física escolar : a tríade legislação, conhecimento acadêmico e prática profissional	Carvalho, Camila Lopes de, 1988-	Araujo, Paulo Ferreira de, 1957-	TESE DIGITAL

Fonte: Dados da pesquisa.

Em continuidade, o nome completo dos autores é visualizado por extenso (Figura 59):

Figura 59. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - C

Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

18-Nov-2010	He Conflict Of Difference At School: A Psychoanalytic View	Ana Archangelo; Daniela Kitawa Oyama; Maria Lígia Pompeu	-	Artigo de periódico
2016	School Trajectories Of Students With Autistic Spectrum Disorder	Teixeira da Silva Talarico; Mariana Valente; Frizman de Laplane; Adriana Lia	-	Artigo
2017	The Perception Of School Directors In The City Of Jaguariúna About Combat Sports	Rodrigues A.I.C.; Baião A.A.; Jr.; Antunes M.M.; De Almeida J.J.G.	-	Artigo

Tipo d
 Artigo d
 DISSEF
 TESE
 Artigo d
 Artigo

Fonte: Dados da pesquisa.

Em outros recursos informacionais, tanto os autores quanto os títulos aparecem em caixa alta (Figura 60):

Figura 60. Exemplo de visualização de inconsistência em autoria no Repositório Institucional da UNICAMP - D

Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

1992	PROPAGATION AND FLOWERING IN DESMODIUM-BARBATUM	SIQUEIRA, LOM; VALIO, IFM	-	Artigo de periódico
2001	Effects of 10% carbamide peroxide bleaching materials on enamel microhardness	Rodrigues, JA; Basting, RT; Serra, MC; Rodrigues, AL	-	Artigo de periódico

Fonte: Dados da pesquisa.

Verifica-se ainda em alguns recursos informacionais a ausência tanto do título quanto do tipo do recurso informacional (Figura 61):

Figura 61. Exemplo de visualização de inconsistência em título e tipo de recurso informacional no Repositório Institucional



Repositório da Produção Científica e Intelectual da Unicamp

[Página inicial](#)
[Sobre](#)
[Contato](#)

2012	Digital inclusion of official documents of the state of Parana, Brazil	Fuza, AF	-	-	Artigo de periódico
2015	Centrality Dependence Of Inclusive J/ψ Production In P-pb Collisions At (formula Presented.) Tev	-	-	-	-
2015	Rapidity And Transverse-momentum Dependence Of The Inclusive J/ψ Nuclear Modification Factor In P-pb Collisions At $\sqrt{s} = 5.02$ Tev	-	-	-	-
2015	Measurement Of Jet Quenching With Semi-inclusive Hadron-jet Distributions In Central Pb-pb Collisions At $\sqrt{s_{NN}} = 2.76$ Tev	-	-	-	-

Fonte: Dados da pesquisa.

Outro problema relatado pelos usuários refere-se à ausência de resumos nos registros (PGM-FCA e PGM-BAE):

[busca por LIPASE] *Essa pessoa aqui ... achou que resumo não é interessante ((RI)). (Não tem resumo?) Não. ...~ (Algumas teses mais antigas não têm mesmo). Ah é? (...) Ah, eu acho que ele traz ... às vezes ele tá insuficiente, como você já viu, né, não tem o resumo ali que ... dificulta um pouquinho o que a gente tá procurando (PGM-FCA)*

[busca por ARQUITETURA RESIDENCIAL] *Esse daqui eu tinha só aberto pra ver mas eu achei engraçado, ele não tem resumo! Tá vendo? (É ... às vezes é uma tese mais antiga ...). É, então. Não tem resumo. É só o PDF, né. (É). Eu acho que ... se bem que é 2003 não é tão antiga não, teve uma de 1994 aí ... bom, mas essa primeira já foi (PGM-BAE)*

No geral, nas entrevistas de busca por assuntos no Repositório Institucional da UNICAMP verificou-se a ausência de padronização na forma de visualização de alguns recursos informacionais nos resultados das buscas, que refletem inconsistências na representação descritiva e que precisam ser verificadas e corrigidas, sintetizadas pela seguinte fala:

É, vendo o que você trouxe, quer dizer, deu pra perceber que estão tendo muita falha ... (...) no próprio, sei lá, na própria forma de catalogar, de organizar os indicadores (D-FE)

As principais razões pelas quais tais inconsistências ocorrem nos registros dos recursos informacionais no Repositório Institucional da UNICAMP, principalmente em artigos e artigos de periódico, podem estar relacionadas à:

- importação em lote das bases de dados, que em muitos casos já não possuem uma padronização nos metadados e ao fato de um mesmo recurso informacional potencialmente estar publicado em duas ou mais bases de dados, corroborando para a duplicidade na importação para o repositório;
- grande quantidade de registros a serem verificadas pela Diretoria de Tratamento da Informação (DTRI), responsável pelo gerenciamento do repositório;
- limitação do tipo de ferramenta utilizada para busca e recuperação dos metadados (DSpace), duplicando registros facilmente mesmo com a menor diferença na descrição dos metadados em dois registros diferentes, ponto que poderia ser positivo caso possibilitasse um gerenciamento dos metadados mais dinâmico, sem a utilização do *Excel* pelos bibliotecários catalogadores-indexadores;
- ausência de mais bibliotecários que lidem diretamente com a representação descritiva dos metadados e especialmente com a representação temática no Repositório Institucional da UNICAMP.

10. Comentários adicionais sobre o repositório institucional

A categoria “Comentários adicionais sobre o repositório institucional” refere-se às sugestões ou comentários gerais dos usuários sobre o Repositório Institucional da UNICAMP.

O usuário G-FE manifesta a percepção de diferenças nas buscas considerando que em inglês os resultados de busca diminuem consideravelmente em sua área. Considera a utilização da ferramenta fácil e intuitiva e expõe que as fontes que mais utiliza para sua necessidade informacional é o *Google Acadêmico* ou o *Google* em geral, mas considera relevante o conhecimento sobre o Repositório Institucional da UNICAMP.

É curioso ((RI)), é interessante ficar vendo isso. ... Você deve se deparar com um monte de tipo de pesquisa, né. (...) Eu achei curioso ((RI)). E ... principalmente o negócio do inglês, quando vai pro inglês diminui muito o que é da área de Educação, né, vai virando Biológicas (...), vira muita Química também, quando põe INCLUSÃO, porque é INCLUSÃO de um ... um componente, uma molécula. (...) Eu achei legal e ... ah, eu achei curioso também de como aparece as pessoas ... que a gente conhece, que pesquisam a mesma coisa que a gente, que vai aparecendo. (...) (E a ferramenta em sim, assim, é fácil de pesquisar ...). É fácil. (Tem alguma sugestão de melhoria?) Não, é bem intuitivo. (...) Achei interessante, coisa que a gente não para pra ... pra ver. (Espero que essas pesquisas possam te ajudar). Ah, nossa, a gente acaba, eu acabo não

pesquisando muito pelas coisas da UNICAMP, vai mais pelo Google Acadêmico, ou direto no Google e vai vendo o que aparece. (...) É bom também conhecer (G-FE)

O usuário PGM-FE expressa que alguns dos itens que considerou relevante realmente no Repositório Institucional da UNICAMP poderão ajudá-lo em suas pesquisas. Quanto à ferramenta, manifesta o desconhecimento da importância das palavras-chave padronizadas para melhoria da recuperação da informação e expressa particular interesse em conhecer melhor os tesouros utilizados para a verificação das palavras-chave elencadas, anotando os tesouros utilizados.

Ah, tem uns três que eu quero ler ((RI)). ((RI)). Tem três que eu com certeza quero ver. (...) (E ... da ferramenta, você achou que é fácil buscar, você tem alguma sugestão de melhoria?) *Sim, na verdade eu não sabia disso, de procurar as palavras-chave, eu sempre tive essa dificuldade de ... que que é uma palavra-chave ou não ... (...) e como eu ainda não tinha pesquisado sobre ... foi interessante saber. Tem um ... quando você procura a palavra-chave ele te indica qual é a melhor. [referindo-se ao tesouro] (...). Porque eu mesmo iria olhar os trabalhos e ver as palavras-chave que o pessoal utiliza. (...) Mas agora eu procuro em vez de olhar os trabalhos já prontos. (...) Deixa eu só procurar o nome do site que você procurou, é Library ... o termo em inglês ... (É o Library of Congress Authorities). Ah, tá. (Mas tem o da Biblioteca Nacional, que é o ...). Esse Fundação Biblioteca Nacional e tem o DeCS? (Isso). Eu tinha anotado ... (O DeCS é da Área da Saúde mas muitas vezes eles tem termos que ... circulam em várias áreas). *Sim (PGM-FE)**

O usuário PGD-FE relata dificuldade em realizar as buscas com apenas uma palavra-chave na caixa de pesquisa, ou seja, busca simples e não combinada entre termos, apontando sobre o alto nível de revocação do Repositório Institucional da UNICAMP e a dificuldade em analisar tantos registros retornados nas buscas. Manifesta ter aprendido bastante com as buscas e como sugestão, aponta a implantação da busca combinada e a enumeração dos registros na visualização dos resultados.

[...] eu achei bem difícil fazer assim com uma palavra só (...) porque é bem aquilo mesmo, sai muito geral, né, se eu coloco PROFESSOR, PROFESSOR DE ENSINO MÉDIO, eu quase morro pra procurar, né ((RI)), imagina procurar nos ... 3000, 6000 que aparecem. Mas ... eu aprendi muita coisa que eu não sabia que tinha no Repositório da UNICAMP, então serviu também pra isso. (...) Bom, se pudesse ampliar, né, a busca, tipo ... é ... pra colocar mais palavra-chave e tal e aqui seria bom também enumerar, tipo, ah, esse é o primeiro, esse é o segundo, porque às vezes você tá aqui procurando, daí você se perde, ah, eu tava no primeiro, é mais fácil do que lembrar o nome, né, do artigo. Mas acho que é isso (PGD-FE)

O usuário D-FE expressa ter conhecimento do Repositório Institucional da UNICAMP mas por opção própria, não utiliza por certa resistência particular ao digital, defendendo sua concepção no Órgão Colegiado do qual faz parte como docente representante da Biblioteca. Entretanto, salienta que indica o repositório a seus orientandos, embora com menor frequência, preferindo bases mais consagradas. Observa que a questão do repositório

institucional na universidade tem se tornado um assunto em voga e cada vez mais presente e que a presente pesquisa é bastante relevante. Expressa a hipótese de que os usuários das demais áreas utilizam mais o repositório do que os usuários da *Área de Artes e Humanidades*. Novamente manifesta que repositórios é um campo da atualidade e do futuro e cada vez mais familiar devido às questões de produção e de indicadores da universidade. Reconhece a importância de elaboração de palavras-chave adequadas e estratégicas não apenas nos recursos informacionais como no próprio Currículo Lattes, possibilitando parcerias com outros autores que compartilham dos mesmos temas de pesquisa. Salaria que já reconhece a importância das palavras-chave há bastante tempo, desde a implementação do Currículo Lattes, no início dos anos 2000 e também às exigências de que as publicações contenham, além dos resumos, palavras-chave que representem adequadamente o conteúdo das mesmas.

Então, eu ia falar. Não, não conhecia, porque daí tem a ver até com aquilo que eu te falei, eu não ... eu sou deliberadamente resistente ao uso ... (...) tem coisa da minha pesquisa que eu como professora não aceito ... (...) então assim, eu sou da materialidade física e ... tanto é que e luto por isso, ah, até lá no SBU, pela manutenção de coisas ainda ligadas à materialidade. (...) É ... claro que os meus orientandos eu oriento mas em geral a minha orientação não é ... pro repositório. (...) É orientação pra fazer esse levantamento nas bases que ... não é referência, né. O repositório pra mim sempre foi referência porque eu sempre participei da ... eu sempre soube o que é, sempre soube, mas assim ... O que a gente tem decidido ... por ser coordenadora eu sempre participei, então tenho acompanhado que na universidade esse é um assunto, né. (Sim, sim. Cada vez mais presente). Então agora assim, acho que pelo o que eu pude perguntar pra você, do que você me deu de informação acho que é bem importante a sua pesquisa, e daí assim eu to curiosa ... você deve estar pesquisando assim também pesquisando gente de áreas diferentes ... (Sim). Porque assim, eu tenho um imaginário, as representações que eu tenho, os imaginários que eu tenho é que os caras da ... Exatas e Biológicas usam muito mais do que a gente da Humanas, não sei, é um imaginário que eu tenho. (...) Mas pelo o que eu tenho acompanhado eu acho que esse é um campo ... do momento. Não é nem do momento, do futuro. (...) Que ... tá cada vez mais presente (...) e acho que para daí os que, por exemplo, no meu caso, no meu assim, algumas coisas são deliberadas, é, quer dizer, eu sei, não quero saber ((RI)), mas e daí vou sempre caminhar na minha da resistência de ... manter, não que não considere importante. (...) Né, mas eu acho que esse é um campo que vai ser cada vez mais ... ele vai ser obrigatoriamente cada vez mais familiar (...) na área acadêmica. (...) Na área acadêmica eu acho que ele vai ser cada vez mais familiar, por todas essas questões de ... e até, até não, esse é um movimento que já tem se adiantado, né, é ... da questão das universidades, de produção, de números ... (Isso, de indicadores). De indicadores, dos indexadores. Então eu acho que é ... é do presente e do futuro. É super relevante, é super relevante a sua pesquisa. (...) É, então, eu ia dizer, voltando a falar das palavras-chave ... (Isso). Daqueles meus alunos, os pouquíssimos que me procuram. (...) Porque também, daí vem toda essa discussão que na prática ... material, também, leva pra não material, pra imaterial, que é ... a falta de cultura, a gente não tem uma cultura, os acadêmicos não tem uma cultura de se preocupar com palavras-chave (...) Então, não sei, daí é uma coisa mesmo de trabalhar, sensibilizar ... da importância, né. (...) Eu não vejo mais de forma ... há muitos anos eu não vejo mais de forma ... boba palavras-chave. (...) E acho que essa é uma coisa que tem sensibilizado mas ainda nós não temos essa ... nós não temos essa ... (essa visão). Essa visão. Quando eu falo nós é os acadêmicos em geral. Eu to bem ligada nisso. (...) Eu tenho ficado, eu tenho ficado ligada nisso aí, por várias questões ... (É, realmente, é super importante. Se uma palavra não tá adequada lá no registro você não vai conseguir recuperar, né, aquele documento). Então ... é, o próprio preenchimento do Lattes, né (...) ... bom, ou seja, essa palavra-chave eu já to prestando atenção faz tempo. É, é só prestar atenção nos movimentos, né. (...) Você vai entrar num congresso já tem, uns 15 anos pra você ... você tem que colocar palavras-chave. (Ah, sim,

qualquer artigo que você vai publicar precisa). É, então isso é uma prática ... mas se a pessoa não presta atenção ela pode pôr qualquer coisa. (É que intensificou a visibilidade com as tecnologias, né). Isso, daí tem isso também (D-FE)

O usuário G-IFCH expressa ter sido muito interessante a experiência de busca por assuntos no Repositório Institucional da UNICAMP e considera que os registros apontados como relevantes ajudarão na elaboração do relatório final de iniciação científica. Relata realizar pesquisas no *Google Acadêmico* mas que no repositório suas buscas ficarão bem mais específicas, por serem publicações da UNICAMP e o acesso facilitado. Manifesta o entendimento da importância das palavras-chave para aumentar o acesso de outros autores que trabalham com o mesmo tema de pesquisa.

Inclusive eu acho isso bem interessante, bem interessante. É bem interessante. (...) Eu achei muito interessante, viu. (...) Vai me ajudar muito pra entregar o relatório final e pra fazer o projeto de mestrado. (...) Vai ajudar muito. (...) O que eu fazia de busca era o seguinte, né, pra achar textos e tal, até pra achar livros pra procurar na Biblioteca ... eu vou no Google Acadêmico, se aparecer apareceu se não apareceu ((RI)). ((RI)). Não, e agora vai refinar bem mais a minha pesquisa, porque são coisas que tem na UNICAMP e eu consigo o acesso bem mais fácil. (...) Todo mundo falava do SophiA e tal, mas também eu não sabia mexer muito bem. Agora saber sobre o repositório foi muito importante, muito interessante, só de pesquisas feitas aqui dentro, né. (E também você pode até utilizar essas palavras que a gente colocou em outras bases). Sim, fica bem mais refinada, e aí acho que pelo o que eu entendi aumenta a facilidade de acesso de outras pessoas que tem mais ou menos o mesmo objetivo que o meu ... (...) achar a minha pesquisa, é isso, né? (Isso, exatamente). É importante. Muito obrigada viu, nossa, aprendi bastante. (...). Eu consegui dois textos pra eu fazer o projeto semana que vem (G-IFCH)

O usuário PGM-IFCH expressa contentamento pelo conhecimento de novos recursos informacionais e pelas reflexões em torno da importância das palavras-chave tanto na atribuição dos recursos informacionais quanto nas buscas. De forma geral, considera o Repositório Institucional da UNICAMP satisfatório, principalmente pela incorporação de artigos de periódico, visto ser uma forma mais rápida de leitura da pesquisa publicada em uma dissertação ou tese.

Eu achei legal porque tem coisas que nem eu tinha visto e às vezes a gente não para pra pensar na importância das palavras-chave, né, da pesquisa ... eu gostei. Eu até ... tem coisas que eu não conhecia ((RI)), então eu até vou ver depois ((RI)). (...) Gostei muito. (...) Não, eu acho que tá bom, o Repositório da UNICAMP é muito bom né. Bom, a gente lida com bancos de dados por aí e realmente esse do repositório de produção científica é legal e eu acho bom porque ele vai agregar também os periódicos, sabe? Uma coisa que faz toda diferença. (...) Porque às vezes você não consegue fazer uma leitura ... que nem várias vezes apareceu é... tese e aí eu artigo do mesmo autor, só que daí você não ... nem sempre você consegue, tem tempo de ler a tese, dá pra ler o artigo e tal. Gostei ((RI)) (PGM-IFCH)

O usuário PGD-IFCH sugere a categorização dos registros recuperados por grandes áreas do conhecimento e a opção de ordenação dos registros por data de publicação, sendo da mais

recente para a mais antiga. Cabe esclarecer que tal opção é possível, porém, considerando-se o foco da presente pesquisa, os registros recuperados foram mantidos na ordenação padrão por relevância. De forma geral, considera a ferramenta ótima, observa a importância da presente pesquisa para o diagnóstico e possibilidade de melhorias nas buscas por assuntos no Repositório Institucional da UNICAMP e expressa interesse em conhecer os resultados da mesma.

Bom, eu achei ... como a UNICAMP é uma universidade muito grande, com muita produção ... é aquilo que eu tinha te falado no começo, se ele pudesse organizar por grandes áreas eu acho que ajudaria muito nas pesquisas sabe, porque não tem as coisas do camundongo, sabe assim ((RI)). ((RI)). Das Ciências Humanas se aparece coisas da literatura, da economia, isso eu acho que é ótimo, assim, até pra gente ver que tipo de ... perspectiva que está sendo abordada, qual o recorte de pesquisa ... mas ... camundongos ((RI)) definitivamente não ia ser parte do meu aporte ((RI)). (...) Então se a gente pudesse só delimitar já pelo recorte de área eu acho que já ajudava muito. Se ele pudesse assim primeiro colocar o que foi publicado mais cedo. (...) Eu também acho que poderia ajudar. (...) Tem algumas coisas assim que foram muito, eu vou até dar uma olhada nas referências, mas que foram publicadas em 1980 ... então assim, muita coisa mudou então eu acho que seria bacana essas duas coisas. Bom, a ferramenta é maravilhosa, quando a gente pensar que a gente consegue fazer essas pesquisas assim, das referências é maravilhoso. (...) Muito legal a sua pesquisa ... Ah, obrigada. Quando você descobrir eu quero ver os resultados ((RI)). (Sim ((RI)), um dos objetivos é esse mesmo, tentar aprimorar a ferramenta pra melhorar as buscas). Isso é muito bom, aí ajudar outros doutorandos aí ((RI)) (PGD-IFCH)

O usuário D-IFCH expõe a percepção de que a maior parte dos recursos informacionais do Repositório Institucional da UNICAMP seja da *Área de Biomédicas*, considerando, por este fator, uma base não atrativa para a *Área de Artes e Humanidades*. Por outro lado, considera ser uma ótima ferramenta pela possibilidade de recuperação de publicações da universidade e de colegas autores e aponta a necessidade de incorporação de mais recursos informacionais da área.

Eu acho que é isso, assim, pra gente ele é muito ocupado pelas áreas médicas, assim, tem muito resultado, é ... de área médica, assim ... é ... então ele acaba sendo uma base que eu acho que ela ainda não ... ela não é atrativa o suficiente pra gente, assim, sabe? Eu indico pros meus alunos sempre pra fazer pesquisa de teses e de dissertações, pra área mesmo, de Ciências Sociais é isso, vai alimentando vai melhorando (...), mas ele é ótimo, porque ... ele recupera o que a gente faz, tipo, é, engraçado, por exemplo nessa busca aqui eu falei “olha, é verdade, eu tenho que lembrar do trabalho da Carla”, tipo, que às vezes é uma coisa que ... você conhece a pessoa, é ... mas você ainda não lembra direito o trabalho que ela fez, assim, que foi super legal, teve dois trabalhos que eu falei “nossa, é verdade, tem essa tese da Carla que eu não lembrava”, sabe, então assim, eu acho que tem esse ... isso que é bom, assim, da gente associar também um pouco ... é, se conhecer mais também, porque às vezes a gente convive com as pessoas, enfim, me lembrei de alguns trabalhos e associei a algumas pessoas (...), é, fazendo essa busca aqui. Mas acho que é isso, a gente precisa de um jeito de (...), incorporar um pouco mais os ... documentos da área de Humanas. (...) (Sim, um dos próximos passos é esse mesmo). Porque ele ainda é muito poluído pra gente, né, ainda é muito ... tem muita coisa que não é da área, não é nem que não interessa, da área médica, enfim. Tem muita coisa de saúde coletiva que fica muito ... que aparece muito pra gente aqui (D-IFCH)

O usuário G-FCM manifesta aprovação das buscas pela linguagem natural (palavras-chave), retornando vários recursos informacionais que poderia utilizar em sua pesquisa de iniciação científica. Considera que as buscas em linguagem controlada não se apresentaram muito diferentes, entretanto, observa que nas primeiras buscas utilizou uma linguagem mais genérica (Síndrome de Turner) e que com uma linguagem controlada (Síndrome de Turner - Genética) ficaria mais específica. Expressa não conhecer nenhum dos recursos informacionais recuperados, relatando ter maior familiaridade com o *PubMed*, mas considera importante conhecer o repositório e observa a questão da duplicação de registros no Repositório Institucional da UNICAMP.

Eu gostei da primeira que eu fiz, com as palavras-chave eu vi que tinha várias coisas que dava pra usar. (...) (E essa segunda rodada ...) Não achei que mudou, que acrescentou muito não ... (...) É, talvez ... é que eu tenho que dar uma estudada mesmo, né, pra ver, talvez na primeira busca com essa (...) ficaria mais amplo, aí com essa ficaria mais ... específico pra o que eu fizesse. Só a parte genética da SÍNDROME DE TURNER, pelo menos. (Você já conhecia os registros?) Eu uso mais ... PubMed ..., mas é bom saber que existe. (É porque no repositório ele vai reunir tudo o que é produção da UNICAMP). Sim (...) Eu achei que, assim, tem vários artigos repetidos, né, português, inglês, artigo de periódico e artigo só. Eu acho que poderia condensar, isso, não sei. Porque eu achava assim, eu via assim e falava “nossa, tem outro”, mas era o mesmo. (...) Uns escrito artigo de periódico, outros só artigo. (...) Vários, assim. Acho que quase todos estão assim (G-FCM)

O usuário PGM-FCM considera interessante a experiência de busca no Repositório Institucional da UNICAMP e expressa que alguns recursos informacionais recuperados realmente serão úteis para sua pesquisa. Expõe que, por falta de conhecimento em recuperação da informação, não há sugestões e que antes das buscas realizadas neste sistema de recuperação da informação, costumava utilizar apenas o catálogo online ou o *Google* como fontes de informação. Manifesta também o desconhecimento da ferramenta Busca Integrada, disponível no Portal do SBU.

Achei até legal, eu nunca tinha pesquisado, é interessante. (...) Tem uns que até eu to querendo ... tem um de EDUCAÇÃO MÉDICA que tem do início do ensino, depois eu vou pesquisar ele, não conhecia. Tem uns que eu usei. (Ah, você já conhecia). Já, já usei. (...). Ah, vou pegar, principalmente, a palavra mais interessante pra mim foi EDUCAÇÃO MÉDICA, é o que mais veio. Veio bastante sobre a história da FCM, se você jogar FCM-UNICAMP não vem, se você jogar EDUCAÇÃO MÉDICA ele vem. É, história do início da saúde coletiva, início não sei do quê, a melhor palavra pra procurar. (...) Eu vou pegar ela pra usar. ... (E você tem alguma sugestão de melhoria do repositório?) Como eu nunca usei, sou leigo no assunto, pra mim é ruim pra eu sugerir, eu era meio analfabeto de usar, eu pesquisava mais raiz, pelo Sophia ... e ia às vezes pelo Google. Aí às vezes eu achava outros artigos e ia atrás das referências, aí pegava aquelas referências. Era muito mais fácil. (É que aqui fica tudo reunido, mas aqui é só produção da UNICAMP). Vi, vi, mas já, já tá uma base ... principalmente pro meu, que estou estudando a história da faculdade, pra mim interessaria bastante. (É verdade). Eu não sabia disso. (Quando você quiser buscar em outras bases também, sem ser produção da UNICAMP, tem o Busca Integrada lá na página do SBU, é uma caixinha que parece um Google também. Ali você digita e ele vai varrer em todas as fontes, inclusive no repositório). Entendi. (Aí pesquisa em todos. É, é como se fosse um Google da Biblioteca). Legal, eu não sabia (PGM-FCM)

O usuário PGD-FCM expressa não utilizar com frequência o Repositório Institucional da UNICAMP pela possibilidade de críticas na utilização de recursos informacionais mais antigos, relatando que costuma utilizar outras fontes de busca. Expressa que algumas teses localizadas foram utilizadas em seu projeto de doutorado e que dois artigos são potencialmente utilizáveis em sua pesquisa. Quanto à usabilidade, considera uma ferramenta fácil, agradável e intuitiva, não expressando sugestões de melhoria.

O que eu encontro que eu poderia utilizar, as referências são antigas, então eu to procurando evitar justamente pra não ... levar essa crítica na tentativa de publicar os artigos ... então o que eu consigo encontrar normalmente é fora daqui. Mas o que tem aqui, que é principalmente do professor Bagatin, do professor é ... DECAPITANIA, que eles eram da UNICAMP, não sei se são ainda, são teses e tal que eu inclusive usei no projeto. (...) Então sim, com certeza eu poderia utilizar. (...) É ... naquela primeira busca da FADIGA, vou até voltar lá pra buscar que eu vi dois trabalhos interessantes que talvez eu possa utilizar o referencial que eles usaram na hora da pesquisa (...) (E em relação aos documentos que você colocou aqui que seria interessantes, eram mais teses ou artigos?). Mais teses. De artigo eu achei muito pouco, de 2006 ... e de 1998, que eu já não poderia usar, sabe? (...) Eu acho que sim, é agradável de trabalhar, é fácil, é bem intuitivo, então eu acho que nisso eu não tenho nada pra sugerir, pra ser mudado (PGD-FCM)

O usuário D-FCM considera o Repositório Institucional da UNICAMP uma ferramenta muito clássica e lógica e sugere a introdução de operadores booleanos nas buscas, considerando que aumentaria a especificidade das mesmas, focando mais nos temas de seu interesse e diminuindo a amplitude nos resultados. De forma geral, quanto à interface, considera ser habitual e de fácil navegação.

A sugestão seria ... eu acho que como é um mecanismo muito ... clássico, né, e ... filosoficamente muito ... lógico ... (...) seria a introdução dos indicadores booleanos. ... (...) Eu acho que isso ... aumentaria a especificidade das pesquisas. (E em relação à usabilidade, fácil de ...). Sim, tem uma interface bastante, é ... habitual, né, e de fácil navegação. (...) (É ... então nesse caso os operadores booleanos iriam trazer os resultados melhores). Eu imagino que sim, diminuiriam a amplitude mas focariam mais nos temas que eu to procurando (D-FCM)

O usuário da G-FEF expressa ter sido relevante a experiência de buscas no Repositório Institucional da UNICAMP e que os registros recuperados poderão ajudá-lo de alguma forma. Fotografou suas anotações no *Roteiro de busca por assuntos no Repositório Institucional da UNICAMP*. Considera as buscas no *PubMed* mais fáceis do que neste sistema de recuperação da informação, visto a possibilidade de busca combinada entre os termos de busca. Expõe seu descontentamento nas buscas realizadas pela falta de especificidade nas publicações de seu tema de interesse.

Eu gostei, não foi tão chato porque ((R1)) me ajudou de alguma forma ((R1)). Mas se fosse algo do tipo, "ah, faz aí ...", seria tipo "ai meu Deus do céu" ((R1)), seria muito cansativo ((R1)). ... Eu posso tirar foto da

folha? Pra mim olhar depois quantos que eu vi ... (Você já está acostumada com esse tipo de busca?) *Desse tipo não porque, por exemplo quando eu vou lá no PubMed e eu digito ATIVIDADE FÍSICA e EPILEPSIA, né, em inglês ou eu coloco ... SÍNDROME METABÓLICA e EPILEPSIA já aparece já ... esse assunto, só coisas desse assunto então é mais fácil. (Os dois assuntos reunidos no mesmo registro) Se aqui tivesse isso também seria melhor, tipo EPILEPSIA ADULTO só, só o pessoal adulto, porque vem muita coisa em CRIANÇA aqui ... é ... e eu não estudo né essa parte da infância (G-FEF)*

O usuário PGM-FEF considera que o Repositório Institucional da UNICAMP é de fácil utilização, não tendo nenhuma sugestão de melhoria. Observa que a entrevista de busca por assuntos contribuiu para suas pesquisas, possibilitando o conhecimento de mais recursos informacionais. Expõe utilizar este sistema de recuperação da informação em seu ambiente de trabalho em secretaria de pós-graduação, facilitando a confirmação de teses e dissertações de alunos para procedimentos internos.

Tranquilo, é bom pra mim porque ... eu vou olhando mais coisas, né ((RI)). (E teve algum documento que vai ser interessante pra você?) ... Sim, todos esses são interessantes pra mim, só que ... como a pesquisa é na França eu pego algumas coisas que eu encontro no Brasil, alguns artigos, algumas teses e dissertações que tem o assunto que fale sobre DEFICIÊNCIA no geral e ponho lá, lá é muito específico. (...) Ah, eu acho que tá bom, o que eu procuro às vezes ... falando do trabalho, eu sempre acho aí. (...) Principalmente quando eu preciso dar baixa em alguns alunos que fazem defesa, às vezes não está na DAC a tese, a dissertação, então eu vou no repositório e aí sempre ... (...). Pego a tese e ponho os dados que eu preciso lá no relatório, porque eu faço Sucupira, né, e aquilo lá é um horror ((RI)). (...) ... E aí não acha, não acha, e eu vou no repositório e eu sempre acho (PGM-FEF)

O usuário PGD-FEF relata ter localizado poucos registros relevantes para sua pesquisa no Repositório Institucional da UNICAMP, embora tenha o conhecimento de outras publicações que também estão relacionadas com seu tema. Expressa interesse em conhecer se tal resultado ocorreu em outras áreas do conhecimento. Acredita que este problema esteja relacionado com dois pontos: o nível de especificidade de seu tema de pesquisa, não estudado em outras áreas e por problemas no sistema de recuperação, já que tem o conhecimento de trabalhos da universidade que não foram recuperados. Manifesta ter trabalhos recuperados que poderão ser utilizados em sua tese e sugere o aumento no tamanho da fonte utilizada.

É, na verdade achei pouco, eu sei que tem mais coisas, tem mais, principalmente aqui na UNICAMP tem mais trabalhos que se relacionam com o meu ... e não apareceram, como esse ... que só agora na última, de trocentas buscas apareceu, tem outros que não apareceram. (...) E ... eu tinha curiosidade de saber se esse número ... tá mais ou menos equiparado ao número de retorno de outras áreas ... (...). Ah, me conta depois ... porque eu acredito que são pouquíssimos trabalhos ... você viu que ... tem coisa que nem apareceu nada, né. (...). Enfim, eu tinha curiosidade de saber isso. Eu acho que tem duas questões, né. Primeiro que o número de trabalhos realmente é menor que em outras áreas, não é uma área que é tão estudada, tem aumentado mas não pra chegar ... e segundo que tem algum problema aí ((RI)). ((RI)). Porque tem trabalhos que inclusive eu to estudando, to lendo que não apareceram. (...) Eu achei bem interessante aquela específica sobre o PALHAÇO NA RUA. ...~~ Custou ... ((RI)) a aparecer, quer dizer, e a gente cercou ela de todas as formas, né, a gente cercou ela com todas ... (...) Na verdade a única coisa que eu acho é o tamanho da letra ((RI)), mas isso não é uma questão do repositório, é uma questão geral, né.

(Espero que pelo menos algum desses registros que você buscou possa te ajudar na pesquisa). Com certeza, achei coisas bem legais ((RI)). E engraçado porque, enfim, teve coisas que apareceram pela primeira vez ... e inclusive uma palavra que eu costumo procurar e que eu não pus aqui porque ela é muito específica que é SALTIMBANCOS ... tem no título do trabalho e eu não tinha lido ainda ((RI)) (PGD-FEF)

O usuário D-FEF, embora não manifeste nenhuma sugestão específica de melhoria, reconhece que o Repositório Institucional da UNICAMP, assim como outros sistemas de recuperação da informação no contexto de bibliotecas, é fundamental e espera que continue desempenhando seu papel. Expõe descontentamento quando da elaboração da ficha catalográfica de teses e dissertações pelos alunos na compatibilidade entre a linguagem natural, isto é, das palavras-chave sugeridas pelos mesmos e a linguagem controlada - os tesouros da área ou pela não possibilidade de inclusão de novos termos de acordo com a especificidade das pesquisas. Verifica-se que tal problema é recorrente não apenas na *Área de Biomédicas*, mas em outras Bibliotecas do SBU, pela ausência de um vocabulário próprio e único.

Olha, eu acho que pra mim eu não teria nenhuma sugestão ... Eu não teria nenhuma crítica e nenhuma sugestão. Eu acho que esses instrumentos são ... muito fundamentais, importantes e espero que eles continuem existindo ((RI)). (...) (Então quando os alunos vão fazer a ficha catalográfica na Biblioteca, da dissertação e da tese, ocorre uma verificação das palavras). ... É, eu sei, e às vezes eu fico muito brava porque não tem as palavras. ... (Às vezes a linguagem não contempla o nível de especificidade dos assuntos). Exatamente, é. Eu lembro quando eu tinha CORPO lá nos anos 90, CORPO não ... não dava porque era sempre CORPO ... tinha que ter CORPO HUMANO. Não, mas não é CORPO HUMANO, é CORPO, CORPO é uma categoria da Sociologia, da Filosofia, da História (...) mas esse lugar não entendia isso, né (D-FEF)

O usuário G-FCA relata que os registros recuperados no Repositório Institucional da UNICAMP poderão auxiliar outros membros do grupo de pesquisa da qual faz parte, especialmente as referências bibliográficas. Sugere que, ao invés de utilização do recurso gráfico aspas (“”) como estratégia de busca, que o repositório possa ter a busca booleana entre os termos, por exemplo, utilizando-se o “and” ou “or”. Manifesta ter desconhecimento da estratégia de busca entre aspas. No geral, achou interessante, pois há vários recursos informacionais no repositório e considera não ter diferença nas buscas com a linguagem natural e a linguagem controlada, retornando registros já retornados anteriormente.

[...] porque assim, eu faço parte de um grupo que são 8 alunos (...) e o meu é geração de renda ... mas tem outros que são alimentação, meio ambiente e aí o eu acho que tem bastante casos sobre ... não bastante, uns três que tem a tecnologia social, redes, mas ligado mas especificamente à alimentação e meio ambiente, então eu acho que pode ajudar eles e a mim também pra escrever, a ter mais base, não sei, do que tiver no texto ... bibliográfico, sabe, se tiver um referencial bibliográfico sobre REDES ou TECNOLOGIA SOCIAL ... (Nas referências?) É. ... Eu acho que em vez de precisar colocar entre aspas podia ter aquele, sabe, colocar três palavras e você coloca ‘e’ e ‘ou’ pra buscar. (Os operadores booleanos, né, o ‘and’). Isso,

porque às vezes a gente não ... pelo menos eu não usaria se você não tivesse falado aspas, sabe? E aí teve alguns casos que eu vi textos diferentes usando aspas, então talvez eu não achasse nesse sentido, então pra filtrar seria ... ajudaria bastante. Mas eu achei legal, acho que tem bastante coisa. (Você achou que teve muita diferença entre as palavras?) Eu acho que não teve muita diferença, porque às vezes tipo, REDES SOCIAIS, eu acho que ele busca só REDES também, né? É, daí tinha alguns textos que eu achava que tinha sido repetido, então acho que não teve muita diferença não (G-FCA)

O usuário PGM-FCA manifesta descontentamento em relação ao Repositório Institucional da UNICAMP, considerando-o ruim e não atrativo, especialmente pela ausência de opção de combinação de termos durante as buscas. Relata a concepção em seu círculo de pesquisa da importância de citação de trabalhos considerados relevantes entre os pares ou bem citados. Por este e outros motivos, evita teses e dissertações e dá preferência a artigos, preferência compartilhada por seu círculo de pesquisa. Ressalta que um fator negativo deste sistema de recuperação da informação é a ausência da visualização de indicadores de citação, por exemplo, fator de impacto de autores e de periódicos juntamente com os recursos informacionais. Considera que indicadores de relevância poupam tempo do usuário e trazem maior segurança na citação de um determinado recurso informacional, pois percebe que em alguns casos há publicações falsificadas. Expressa o hábito de busca de recursos informacionais que já tenham certa relevância entre os pares, com receio de críticas na submissão em suas futuras publicações. Observa que os trabalhos desenvolvidos pela UNICAMP são mais relevantes para outras instituições do que na própria universidade, ponderando sobre a necessidade de uma mudança nessa cultura em sua Faculdade e indaga se esta é uma cultura geral da UNICAMP.

É ... não gosto da ferramenta de busca em si, acho que é uma ferramenta de busca ... ruim, não me atrai, não me agrada ... eu gosto muito de ter como classificar bastante ou usar a bastante palavra conjunta ... mas eu não ... não me atrai muito por esse sistema de busca, assim. (...) É ... a gente trabalha também muito com ... com muitos problemas, né, não é nem um problema, você também deve passar por isso, o orientador falar a mesma coisa assim como outras pessoas, é ... a importância do trabalho que você tá utilizando pra fazer uma citação, entendeu? Então ele não me traz se tem um artigo que é publicado, qual o conceito desse artigo, se ele é relevante, se ele não é ... você lê, às vezes você acha que é um artigo bom mas querendo ou não às vezes você tem artigos que são ... fake, né, que parece ser muito bom mas a hora que você coloca os assuntos na ponta da literatura você vê que a pessoa usou de artifícios ... forjou dados que a gente fala, entendeu, o que eu até tinha falado antes pra você, ah, tem que ver o que tá dizendo porque às vezes a pessoa coloca ali mas esse dado é meio mascarado ... então quando você tem ele ... você tem uma relevância, traz uma relevância encima daquele trabalho, você tem uma segurança encima do seu trabalho ... Se você não tem nenhuma relevância você vai ter que ler muito aí você vai ter que parar, você vai ter que analisar porque se você colocar isso no seu trabalho ... e vamos supor você manda pra uma revista com um conceito bom, o que que vai acontecer? Ele vai te mandar e falar assim, "olha, esse ponto de vista aqui ... não tá coerente com a literatura" ... então você tem que ficar voltando, buscando muito. Eu sempre procuro os artigos que me deem maior segurança no trabalho que eu to fazendo de forma rápida, claro que tem muito trabalho bom aqui mas entre parar e ficar analisando é mais fácil você vê que alguém já fez isso, que tem uma relevância e ele não traz isso aqui pra mim, entendeu, na busca, se você olha aqui você não tem um parâmetro de relevância. (...) No caso pra se rever muita coisa, só que

a gente tem que levar em consideração que a maioria tem tese ... umas coisas do tipo, então é difícil também, mas eu não uso ele por ser isso mesmo, eu prefiro pegar artigos ... a não ser que quando você não acha nada do que você tá procurando aí você acaba buscando nessas fontes de dados, porque ... não tem trabalho, você precisa saber se alguém já trabalhou, descobriu alguma coisa ... é igual aquele um que apareceu, é uma coisa que eu tava precisando muito e não tinha visto ainda, eu achei, vou pesquisar, vou dar uma lida ainda nele, que é uma coisa que eu vou ver o que que o cara trabalhou, quais foram os fundamentos que ele utilizou e se eu posso utilizar no meu também, mas o principal motivo é: relevância. Pra mim não ter essa insegurança e eu precisar de uma ... rapidez nas minhas respostas. (Aí a sua sugestão é que ele traga a relevância dos artigos mais ...). A relevância dos artigos em si já te traria uma visão melhor. (...) Os mais citados na área ... é, conceitos das revistas, aquele lance do JCR, fator de impacto, tudo isso ... Porque na UNICAMP o que eles cobram da gente é isso, "não me traga uma fonte que não tenha um fator de impacto relevante". E aí você fica tipo, ah, eu vou ter uma tese inteira e ... aí eu vou ter que pesquisar o autor ... depois de pesquisar o autor saber que ... quando você tem isso daqui aqui você já facilita a minha busca, aí eu tenho outras fontes que já me traz isso, então ao invés de eu procurar aqui não, vou procurar naquela lá que já tem essa resposta lá. (...) Às vezes até em tese, sei lá, se tem, é ... melhor ... descrito é, a relevância das citações, apesar que é difícil você colocar isso, uma tese tem ... 500 citações né mas ... em alguns outros documentos não sei, não sei como dizer pra melhorar nessa parte porque é bem difícil mas no meu ponto de vista eu não procuro, eu não gosto por causa disso. É, aquela coisa né, você começa a ter uma confiabilidade depois de um tempo que você tá utilizando, mas você precisa de algo atrativo pra isso que não precisar de um referencial pra ser confiável. Eu acho que inicialmente tem outras fontes de pesquisa e hoje você tem muito essa cobrança de ... é ... grau de confiança no que você está escrevendo, que vem da própria universidade, né, tipo, eles assustam a gente assim ((RI)), aqui eles falam tanto sobre isso que você fica até meio impactado quando você vai mostrar alguma coisa de uma revista que talvez não seja tão conhecida ou de uma tese de uma universidade, seu orientador já fica assim, mas tese, já abordam a gente assim. Tá, tem essa ... já que quer trazer isso na UNICAMP então eu acho que os professores também deveriam saber uma informação melhor de como aceitar e como passar essa informação, porque aqui você chega com uma tese e falam, mas isso aqui é uma tese. É uma tese, a pessoa dedicou a vida pra escrever aquilo ... não, não pode, é artigo. É um artigo, é de uma revista assim, entendeu? Então a gente já tá vindo com esse preconceito de algo que vocês estão querendo lançar, só que tá vindo não porque a gente criou e sim porque as pessoas da instituição criaram isso pra gente e a gente só tá levando pra frente. É uma questão cultural, né? É, aí já é uma questão que talvez eu até veria a resposta do que eu vi aqui com outros olhos se ... eu não tivesse tanto esse impacto, assim, então eu procuro o que, uma coisa bem específica que talvez ali essa pessoa pelo menos um referencial encima do que eu usaria ela tem. (...) Eu acho que aquele realmente vai me ajudar, porque os outros eu já conhecia, já tinha visto. Você ... tem uma relevância fora da UNICAMP um trabalho produzido na UNICAMP, só que infelizmente as pessoas dentro da UNICAMP não tem muita ... não acha tão relevante os trabalhos de dentro da instituição. Você pega uma universidade de outro estado e usa como referência uma tese da UNICAMP e é muito bem visto ((RI)). (...) Agora se você faz o mesmo dentro da UNICAMP utilizando as próprias fontes o professor fala, "não, esse aqui não é um conceito bom", entendeu? Aí é complicado ... então às vezes esse repositório vai ser muito utilizado por um aluno que não seja da UNICAMP, mais do que um aluno da UNICAMP hoje devido a essa limitação que a gente, que os professores trazem pra gente. (E será que isso é em todas as áreas ou você percebe isso mais na sua área?) ... Como eu não fiz graduação na UNICAMP, eu não sei como é na graduação UNICAMP, então dentro da minha área, dentro da FCA, posso dizer, dentro da FCA o que eu conheço é essa cobrança, "não, não, não e não". Como os nossos orientadores, a minha orientadora mesmo fez UNICAMP, como ela vem de um outro campus, com certeza ela já tinha essa mesma cobrança que ela passa pra mim, então eu acho que lá também não deve ser muito diferente, quando eu tava escrevendo a minha tese eu só buscava artigos relevantes, então eu acho que talvez seja uma coisa de Barão Geraldo (PGM-FCA)

O usuário PGD-FCA expressa seu contentamento nas buscas por assuntos no Repositório Institucional da UNICAMP pela possibilidade de maior entendimento de como as buscas funcionam, relatando ter sido uma experiência bastante construtiva. Considera que seu

campo de estudo é bem restrito e que precisa ser melhor consolidado no âmbito da universidade.

Ah, eu gostei muito porque eu tive uma noção muito mais ... vamos dizer, específica de como as buscas funcionam, eu acho que isso mostrou também que o campo que eu to trabalhando é muito restrito, então talvez eu deva trabalhar mais nisso. (...) Eu acho que realmente o campo precisa ser consolidado senão realmente eu teria achado mais coisas ... é que tem muita gente que começa um trabalho, chega no meio e aí que descobre ... mas acho que não é caso. Mas eu gostei bastante, foi uma experiência muito construtiva (PGD-FCA)

O usuário D-FCA manifesta sua percepção de que quanto os termos são buscados em inglês há uma maior quantidade de recursos informacionais recuperados em sua área de pesquisa no Repositório Institucional da UNICAMP. Entretanto, expressa admiração por não haver resultados com determinado termo pesquisado, mesmo com o curso de Engenharia Agrícola na universidade, que poderia ter pesquisas com o tema, ou o tema poderia estar sendo trabalhado na UNESP ou USP. Dos registros considerados relevantes, expressa o conhecimento de autores que costumam publicar sobre o tema. Sugere a enumeração dos itens recuperados visando facilitar a avaliação da relevância. Quanto à aparência, considera o tamanho de letra adequado, sendo, de forma geral, um bom sistema de recuperação.

[...] percebo que, pelo menos na minha área de pesquisa com BAGAÇO DE CANA, quando os termos são em inglês eu tenho mais material, deu pra perceber, quando eles estão em português e quando eles estão em inglês. (...) Então a gente acha mais coisas. Esse aqui realmente é uma dificuldade nossa, né, porque ela quer secar BAMBU mas eu ... pelo o que a gente tá vendo poucas pessoas secam o BAMBU mesmo em estufa e coisas do gênero, então pode ser por conta do que quando eu leio eu to procurando especificamente a secagem ou de madeira ou de bambu mas me admirou que mesmo sendo SECAGEM não apareceu nem de madeira, pensando que a gente tem, né, de Engenharia Agrícola, talvez tivesse trabalhos a respeito, mas ... (Sim, sim). Pode ser que eu esteja enganada, que não tenha mesmo muito. É ... seja mais a ... não sei se a USP ou a UNESP que tem Engenharia Florestal, que daí trabalhe mais com isso. (É, porque assim, aqui avaliamos os 50 primeiros. (...)) Pode ser que esses documentos estejam mais pra frente, né). É, isso é verdade. Aí ... não tem como saber exatamente agora, né. Mas assim, eu refinaria também, né, se pudesse refinar na busca acho que até dá né (...) A gente pode tentar. Como poderia fazer um refinamento ... acho realmente, dificilmente ... data de publicação. Mas acho realmente difícil que apareça específico ... mas acho que do bambu realmente foi zerado porque quando eu procuro SECAGEM mesmo em inglês eu to procurando trabalhos que tratem de ... eu vejo no título ou que fala, que pelo menos fale de madeira. Mas nenhum deles falou. Ou de pinus, ou de eucalipto, que são os mais conhecidos aqui pra nós, mais utilizados pra móveis. Mas em BAMBU realmente acho que não ia encontrar, é bem específico, é bem específico mesmo. Na verdade tem mas não assim, feita a secagem, você pega o BAMBU e seca. (...) Em geral é feita uma secagem solar, ao ar livre, que é dita solar por ser ao ar livre ... e depois é enviado pra secagem em estufa mas mesmo em SECAGEM SOLAR não tenha quem trabalha aqui. (E esses registros que foram recuperados a maioria você já conhecia ou ... eram novos pra você?) De ETANOL bastante conhecido, bastante autores conhecidos, mas não necessariamente os artigos em si porque não vou lembrar os títulos, né, mas são autores da área, são bem ... que realmente publicam ... quando a gente pega os trabalhos são bem esses autores mesmo. (...) (Além da duplicidade que você falou, de ter mais opções de talvez fazer uma busca combinada, né, entre palavras-chave tem mais alguma sugestão que você colocaria pro repositório ... pra melhoria?) (...) Eu sugeriria no conjunto de itens listar, 1,2,3,4,5, porque como às vezes o sistema tá lento ... você faz o mouse pra página correr, ela trava, mas daí ela corre, você faz a primeira vez ela não corre, você faz duas ou três daí ela corre as três. E daí eu esqueço qual eu vi antes, daí eu volto ... tive que voltar várias ((RI)). Se tivesse enumerado eu saberia por exemplo,

dos 50 primeiros o último título que eu li antes de correr o mouse foi o décimo sétimo. Daí no que ele travou, daí eu tenho que ficar lembrando um pouco do título. (...) Acho que isso pra mim foi uma coisa que eu achei ruim. Eu acho que em geral ... tirando isso o tamanho de letra tá bom, não achei ruim não, mas pra mim com óculos né ((RI)). Mas acho que aqui que dá um filtro de data é bom, ele já dá ali descendente. ... (...) Mas eu achei bom. Eu só listaria realmente pra facilitar por conta do sistema, às vezes ele carrega, né, sei lá, 14.000 documentos listados daí tu se perde, às vezes qual foi o último que você leu quando ele trava. E como os títulos são muito parecidos entre si às vezes, né, por conta do assunto (D-FCA)

O usuário G-IFGW manifesta nunca ter utilizado o Repositório Institucional da UNICAMP diretamente para buscas, geralmente utilizando o *Google* para buscas de teses mas considera interessante a possibilidade de encontrar os recursos informacionais diretamente no repositório. Sugere a categorização dos recursos informacionais recuperados por Instituto/Faculdade ou por área do conhecimento, considerando que as buscas seriam um pouco mais fáceis, possibilitando um refinamento nos resultados. Embora seja da *Área de Tecnológicas*, no caso específico de suas buscas manifesta o descontentamento com o alto índice de recuperação de registros da *Área de Biomédicas*.

Acho que tem bastante coisa que dá pra aproveitar mas talvez uma das coisas ... talvez definir de onde que foi, como é tudo produzido aqui ... de qual instituto, de qual faculdade foi ... não sei se tem essa opção nas ... (Ainda não tem). Seria interessante porque ... por mais que seja, ou colocar um campo de ... de áreas relacionadas , se for uma área interdisciplinar, ah, Física Química ... Física Engenharia ... que aí ficaria um pouco mais fácil ... (Aí você faria a pesquisa e usaria esse filtro pra refinar mais?) É, pra refinar mais. (...) Porque às vezes acaba escapando alguma coisa, assim (...), de tanta coisa, sei lá, da área médica que aparece ((RI)). (Sim, sim ((RI))). ... Eu nunca tinha ... geralmente quando você pesquisa no Google e aparece alguma tese e tá no repositório ele joga pro repositório direto (...) e aí sempre usei mas nunca tinha utilizado a busca em si. (...) É interessante (G-IFGW)

O usuário PGM-IFGW expressa já ter conhecimento de alguns recursos informacionais e de autores do Repositório Institucional da UNICAMP, incluindo seu orientador e considera em geral ser um sistema de recuperação da informação de fácil utilização. Não sugere nada específico, considerando que atende às suas necessidades.

[...] alguns documentos eu conheço os autores, tem meu orientador, tem meu orientador que é bem específico mas no geral é fácil de achar (...) uma ferramenta, ok, eu acho, boa ... é ... pra melhorar? Ah, eu não sei se eu tenho ... atende assim do jeito que tá. ... (Tranquilo pra fazer a busca?) ... Ah, acho que atende (PGM-IFGW)

O usuário PGD-IFGW manifesta o desconhecimento da possibilidade de utilização do recurso gráfico *aspas* (“) para refinamento da pesquisa e manifesta sua percepção de que as buscas com termos em inglês trazem mais resultados em recursos informacionais do tipo artigo e com termos em português, mais teses e dissertações. Expressa conhecer a maioria dos recursos informacionais apontados por ele como relevantes. Costuma utilizar o

Repositório Institucional da UNICAMP para localizar teses e dissertações para conhecer as pesquisas de outros autores. Quanto às sugestões, considera uma ferramenta clara e que não tem problemas maiores nas buscas. Já os artigos, costuma realizar buscas em outras fontes: plataforma ARCHIV e Portal da CAPES, usando raramente este sistema de recuperação da informação para este tipo de recurso informacional.

É, eu não sabia desse ... de colocar as aspas, na verdade, eu sempre procurava sem. (...) E ... mas eu percebi que às vezes no repositório eu procuro em português mas em inglês eu vi que aparece mais coisa, eu acho, aparecem mais os artigos, também, né. (...) Ele puxa mais os artigos do que ... o outro. (É porque em português tem mais teses e dissertações). Isso. Dava muitas teses mas aí em inglês ele traz bastante material. (...) Mas das aspas eu não sabia que ele fazia diferenciação, colocar junto ou não, pra mim ele entendia tudo igual. (E ... algum documento que você pesquisou, você acha que vai te ajudar na pesquisa?) No repositório? (É). Desses que eu vi agora? (...). Desses que eu vi agora eu já conhecia a maioria. (...) Bom, o que eu costumo buscar mais no repositório são teses, aí eu consigo pegar as teses, eu mais pego teses pra ver o que outras pessoas fizeram. (...) Tá legal assim. Eu acho o repositório da UNICAMP bom, assim, eu acho ele bem claro, assim, nunca tive muitos problemas com ele, assim. É que artigo eu não procuro pelo repositório ... é mais as teses, as teses eu nunca tive problema em pegá-las, não. (Os artigos você costuma pesquisar onde?) Os artigos ... em geral eu pesquiso nas revistas específicas e a gente acompanha muito o ARCHIV. (...) Que é uma outra plataforma, então a gente tá sempre olhando lá. E a gente usa muito o Portal da CAPES, também, a gente procura muito por lá, mas no repositório é mais difícil. (É que o repositório só tem produção da UNICAMP). Da UNICAMP, né (PGD-IFGW)

O usuário D-IFGW relata que a linguagem controlada não possibilitou a recuperação de recursos informacionais mais relevantes que a busca com a linguagem natural no Repositório Institucional da UNICAMP. Reconhece que a utilização do recurso gráfico aspas (“”) melhora as buscas, sugerindo que tal orientação esteja disponível aos usuários na página do repositório. Aponta a necessidade de, para recursos informacionais com significativo número de autores, ou seja, grande colaboração, que a visualização dos mesmos na página de resultados possa ser sintetizada, aparecendo apenas os primeiros autores, considerando-se o tempo significativo gasto para arrolar todos os autores na página de resultados. Além disso, sugere a categorização por área do conhecimento, tal como ocorre na base de dados *Web of Science*.

O design disso também deveria mudar (...) porque ... você tá fazendo uma busca, ficar rodando assim é uma coisa boba, eles deveriam colocar pra aparecer, sei lá, os dois primeiros aí se a pessoa tiver interessada no artigo ela clica e aparece o restante ... e vê o restante dos autores. (...) ...~ Fica ... fica poluído, né, fica muito poluído aqui. (...) É, na verdade eu acho que não ajudou muito não você pegar e colocar essas palavras aí no final ... ela não acaba selecionando muito mais do que já tinha pego na primeira rodada, né. Por exemplo, o primeiro foi exatamente o mesmo tipo de resposta, eu acho que saiu. O ... é, assim, quando você coloca as aspas você percebe que tem bastante diferença mesmo, de fato, né. (...) Eu acho que, por exemplo, coisas de grande colaboração mesmo, pegar e deixar o design de ... poucas, poucos autores aí ia ajudar, né. ...~ E ... aí não sei, aí busca realmente eu acho que talvez colocar a instrução que é melhor colocar com aspas, né. (...) Acho que é isso mesmo. ... (Filtrar por áreas?) É. Se você coloca por áreas fica mais fácil, né, por exemplo a Web of Science quando você tá procurando você coloca, é ... área de Saúde, área de Ciências Exatas aí dentro de Ciências Exatas você pode escolher qual que é a área, já vai limpando, né, porque ... (Filtrando mais?) É (D-IFGW)

O usuário G-IMECC relata não ter notado muita diferença nos resultados de busca entre a linguagem natural e a linguagem controlada utilizada no Repositório Institucional da UNICAMP, e que costuma encontrar mais recursos informacionais sobre os métodos específicos que utiliza em sua pesquisa de iniciação científica. Não manifesta nenhuma sugestão de melhoria por não ter muita experiência específica em buscas no repositório, estando mais familiarizado com buscas no *Google* e *Google Acadêmico*. No geral, considera um bom sistema de recuperação da informação, visto que possibilita o download na íntegra do recurso informacional em PDF.

(Você notou alguma diferença ... das palavras que você tinha colocado, das novas palavras). ...~~ Não muita. Acho que a OTIMIZAÇÃO MULTIOBJETIVO mudou um pouco. ...~~ (Deu pra refinar um pouquinho mais ...). Tipo, o que eu acho mais coisa é o nome do problema mesmo, né, e os métodos específicos que eu to usando. (...) Mas não deu pra procurar direito o ε-RESTRITO. ... (O que que você achou, é fácil pesquisar?) Ah, parece bom, parece que tem bastante coisa. (...) (Você tem alguma sugestão de melhoria?) ...~~ Boa pergunta. ...~~ É a primeira vez que eu usei ele, eu não sei se eu posso opinar muito ((R!)). (...) Então, por não ter muita experiência. (Por que assim, você costuma pesquisar em outros lugares, bases de dados e tal ...) Pesquisa no Google, no Google Scholar. (...) ...~~ Ah, parece bom, tem aqui o PDF pra fazer download (G-IMECC)

O usuário PGM-IMECC observa que existe a possibilidade de que os registros apontados por ele como relevantes contribuam para sua pesquisa e expressa ter ficado impressionado com a quantidade de respostas nas buscas no Repositório Institucional da UNICAMP, embora as publicações sejam apenas da universidade. Todavia, salienta a quantidade expressiva de assuntos não relacionados ao seu tema de pesquisa quando não é utilizado o recurso gráfico aspas (“) para termos compostos. Sugere que na página principal tenha informações/dicas para os usuários sobre as funcionalidades de busca, principalmente sobre o recurso gráfico aspas (“).

É isso aí, mas eu não tenho certeza se todos realmente vão ajudar mas pelo menos são os que eu acho que ... poderiam né? É, talvez, uns certamente, mas outros talvez. (...) A se considerar que o levantamento foi na base da UNICAMP, fiquei impressionado com a quantidade de respostas para a pesquisa. Porém, apesar da minha inexperiência com levantamento bibliográfico, também me surpreendeu a quantidade de assuntos não relacionados com o tema quando não foi utilizado a função aspas para tornar a busca específica e com a extrema especificidade ao usar a função (o que é bom e ruim). De forma geral achei razoavelmente útil pois a pesquisa é em âmbito global e não somente da UNICAMP, apesar de servir perfeitamente o propósito de ser uma pesquisa voltada a base só dá UNICAMP, destacar a existência das funções de busca (palavras chave, como por exemplo o uso das aspas para deixar a busca específica) seria bom, pois é raro as pessoas buscarem se informar sobre as funcionalidades da plataforma de busca antes de a usar (PGM-IMECC)

O usuário D-IMECC expressa nunca ter realizado buscas no Repositório Institucional da UNICAMP, tendo mais familiaridade com buscas no Portal de Periódicos da CAPES, mas

considera uma fonte que pode ser mais uma opção de buscas mais específicas. Não sugere melhorias, pela falta de utilização.

Não, eu nunca tinha buscado, porque na verdade eu faço busca geral, entendeu, vou no site da UNICAMP, busca geral, lá no Periódicos da CAPES dou busca geral e daí ele me traz tudo, me traz tese, me traz dissertação, traz ... mas é legal, mais um lugar pra eu procurar algo mais ... mais específico. (E você tem alguma sugestão de melhoria pro repositório ... pelo o que você ... buscou aqui ...). Ah, é que como eu não uso é difícil a gente, né (D-IMECC)

O usuário G-BAE manifesta já ter tido o conhecimento do Repositório Institucional da UNICAMP, mas nunca ter utilizado de fato. Considera ter publicações específicas, o que por um lado pode ser bom para alguns usuários, porém, manifesta seu descontentamento na recuperação do que se deseja. Embora seja da *Área de Tecnológicas*, relata que muitos recursos informacionais recuperados utilizando seus termos de busca são da *Área de Biomédicas*. Expressa que o repositório não é uma plataforma pobre, mas que a grande dificuldade é realmente recuperar os itens relevantes, e manifesta o entendimento da importância da presente pesquisa. Além disso, de forma geral, achou a ferramenta interessante pelo fato de os alunos de graduação terem pouca familiaridade com a mesma, geralmente utilizando-se de outras fontes como o *Google* e *Google Acadêmico*. Nesse caso, o repositório seria mais uma fonte de pesquisa. Sugere a categorização dos resultados por área do conhecimento e a implementação de buscas combinadas entre os termos.

Ah, eu já tinha visto essa plataforma mas eu acho que nunca usei ela de fato, eu não sei, eu acho que tem ... coisas muito muito muito específicas que pra algumas pessoas deve ser muito bom mas é muito difícil achar o que você quer, né (...) Tá parecendo, pelo menos. (...) Eu acho que filtrar pela área do conhecimento, eu acho que pelo menos daí, nossa, dos 50 eu tiraria 40 ((RI)). (...) Só da área da Medicina mas ... principalmente nesses primeiros, o primeiro que era ... ACELERÔMETROS tinha muita coisa de ... derrame, umas coisas da Medicina. (...) Eu achei muito interessante mas pra minha área do conhecimento já não ... não é uma plataforma ... pobre, tem muita coisa, mas é difícil ... (Recuperar). Recuperar, exato. É por isso que ... você tá fazendo a pesquisa, é, exato. E aí fica difícil você só conseguir pesquisar uma coisa, se você consegue mesclar várias coisas e ele desse um jeito de achar, assim ... (Uma busca avançada, que pudesse pôr dois termos...). É ... exato. ...~~ Mas é legal porque ajuda todo mundo, né. (...) (E tamanho de fonte, essas coisas, achou tranquilo?) Não, isso eu achei bom, só achei difícil de achar mesmo, de retornar. ... (...) Mas é bom, é uma boa plataforma. ...~~ Gostei, é porque é difícil pra estudante de graduação pesquisar essas coisas, né. A gente vai no Google, o Google Scholar então é bem mais fácil, né. Mas seria bom ... além do Google buscar em outras bases (G-BAE)

O usuário PGM-BAE considera o Repositório Institucional da UNICAMP de fácil utilização, embora não tivesse encontrado recursos informacionais específicos para sua necessidade informacional e sugere a categorização por área do conhecimento.

Achei fácil de realizar as buscas, apesar de não encontrar muitas publicações que me interessassem como em outros repositórios mas isso acredito que seja porque na UNICAMP a área que eu pesquisa ainda está se desenvolvendo. (...) Não sei, mas não vi se dava para selecionar a macro área de pesquisa, pois algumas

das palavras que buscamos apareciam muito resultados de áreas diferentes da minha como Medicina, Química (PGM-BAE)

O usuário PGD-BAE sugere a combinação de termos nas buscas no Repositório Institucional da UNICAMP, permitindo um maior refinamento e que esta possibilidade já lhe é familiar em outros sistemas de recuperação que utiliza.

Eu acho que realmente conseguir combinar as palavras-chave porque você consegue refinar bem mais a pesquisa. (...) Eu não consigo viver sem isso pelo menos ((RI)). Eu acho que ... é um caminho bom (PGD-BAE)

O usuário D-BAE percebe que as buscas das palavras-chave nos tesouros de sua área permitiram um maior refinamento e que este mesmo processo é realizado pelos bibliotecários nas Bibliotecas do SBU por ocasião da elaboração da ficha catalográfica da tese ou dissertação. Reforça que o Repositório Institucional da UNICAMP reúne a produção científica e intelectual da universidade. Não realiza sugestões, porém, indaga à pesquisadora a não localização de um trabalho de evento publicado pelo mesmo e que tem relação direta com seu tema de pesquisa. Após verificação pela pesquisadora, conclui-se que o referido trabalho não foi indexado pelas bases de dados que são utilizadas para importações em lote neste sistema de recuperação da informação e que provavelmente será incorporado pela Biblioteca da BAE.

É, o que foi feito foi um refinamento, né, pras palavras que são indexadas na verdade, né, partindo das que a gente acostuma, é o mesmo que se passa quando o aluno chega pra montar a ficha catalográfica, né. (Isso, exatamente). (...) É, o repositório tá bem ... condensado ao que é nosso. (Isso, é só produção da UNICAMP mesmo). Só produção da UNICAMP né (D-BAE)

O usuário G-FEA considera boa a experiência de busca por assuntos no Repositório Institucional da UNICAMP, sendo útil para a elaboração de seu relatório de iniciação científica e manifesta o desconhecimento da utilização das buscas utilizando-se o recurso gráfico aspas (“) nos termos de busca. Entretanto, aponta a necessidade de melhoria na precisão da recuperação na maioria das buscas, relatando que vários dos resultados trouxeram termos sem relação alguma com seu tema de pesquisa e sugere a busca combinada para facilitar as buscas.

Ah, eu achei bom, eu não sabia disso de aspas, vai ser bem útil agora pra mim fazer o relatório mas ... ah, eu achei que, assim, tem coisas que funciona bem mas tem outros termos que estão viajados assim, tipo, tinha CIMENTO no meio aí eu fiquei ((RI)). (...) E não era um, era tipo vários, assim, não sei se é algum bug que dá, mas de forma geral é bom. (Ou às vezes algum termo que caminha em várias áreas). É, é que ... não lembro qual era ... (ANTIOXIDANTE?) Talvez. Eu acho que deu alguma coisa assim. É, algum componente, vai saber ((RI)). (...) É ... fazer uma busca combinada talvez seria bom. Dá pra fazer

combinado? Porque aí seria mais fácil. Tipo que nem a da Biblioteca? (Colocar dois termos juntos?) É (G-FEA)

O usuário PGM-FEA considera o Repositório Institucional da UNICAMP muito bom, sendo interessante o conhecimento de que existe esta ferramenta na universidade.

Interessante. ... saber que tem, muito bom. (Vai te ajudar em alguma coisa ...). *Sim, é interessante pensar em outros nomes também (PGM-FEA)*

O usuário PGD-FEA manifesta ter o conhecimento de muitos trabalhos recuperados nas buscas e apontados como relevantes no Repositório Institucional da UNICAMP, sendo um deles provavelmente a ser utilizado em sua tese. Expressa o interesse pelos tesouros, indagando à pesquisadora se o ideal é a utilização de termos controlados. Como sugestão, menciona a possibilidade de categorização por área do conhecimento, visando uma maior delimitação e refinamento dos resultados de busca.

Então, muitos eu conhecia quando eu busquei algumas palavras aqui, né [palavras-chave]. (...) São trabalhos que eu uso, bastante. (...) E teve aquele que agora eu pequei que pode ser que me ajude ((RI)). ((RI)). Mas ah, é legal saber que tem mais coisas do que a gente imagina, né. (E você tem alguma sugestão pra melhorar as pesquisas no repositório). ...~~ Boa pergunta, eu acho que não. É porque ... esse negócio do tesouro, o certo é a gente ir se acostumando e tal, o certo é a gente usar a palavra-chave padronizada, né? ... Eu acho que talvez alguma coisa assim ... que eu possa selecionar o que que é de alimentos. (Ah, por área). É, entendeu? Porque aí ficaria mais fácil, por exemplo, se eu buscar MICROSCOPIA ELETRÔNICA em ... ou em bactérias, porque também tem gente da Biologia que faz, tem gente da Física que faz, se você coloca muito específico também você acaba não achando ... (É uma sugestão interessante colocar por área). Eu acho que talvez ... delimitar melhor, pra refinar mesmo, né (PGD-FEA)

O usuário D-FEA reforça a importância da melhoria no Repositório Institucional da UNICAMP em geral, principalmente no que se refere à duplicação de registros, especialmente aqueles que possuem tanto o título em português quanto em inglês.

Ah, eu acho que tem que ter uma melhorada, porque ... percebi que estava repetindo alguns itens (...) ... mas acho que é mais isso de estar repetindo os itens ... o que aparece em inglês às vezes aparece em português porque eu acho que tem um ... resumo em português e em inglês, ... mas acho que é mais isso (D-FEA)

Tendo em vista os resultados de todas as coletas de dados realizadas na pesquisa, são apresentados, a seguir, uma síntese geral dos mesmos.

6 SÍNTESE DOS RESULTADOS

Nesta seção busca-se realizar uma síntese geral dos resultados das entrevistas semiestruturadas com gestores e bibliotecários catalogadores-indexadores dos repositórios institucionais como metodologia qualitativa e da avaliação da indexação de assuntos como metodologia quantitativa em duas abordagens: pela perspectiva do sistema - *Avaliação intrínseca quantitativa mediante a interconsistência* e pela perspectiva do usuário - *Avaliação extrínseca mediante a recuperação*.

Referente aos resultados das **entrevistas semiestruturadas com gestores e bibliotecários catalogadores-indexadores dos repositórios institucionais** da USP, UNESP e UNICAMP, no geral demonstram a ausência de uniformidade entre as três universidades quanto às políticas de informação dos repositórios institucionais, sendo pouco mais uniforme entre a USP e a UNESP e mais notadamente menos uniforme entre estas e a UNICAMP, conforme apresentado no quadro-síntese a seguir (Quadro 16):

Quadro 16. Síntese dos principais resultados das entrevistas com gestores e bibliotecários catalogadores-indexadores de repositórios institucionais

Categoria	USP	UNESP	UNICAMP
Autoarquivamento de teses e dissertações	Realizado diretamente pelos alunos de pós-graduação na Biblioteca Digital de Teses e Dissertações (BDTD)	Realizado diretamente pelos alunos de pós-graduação no repositório institucional	Não há
Catálogo de teses e dissertações	Realizada diretamente pelos alunos de pós-graduação ou pelas Bibliotecas na Biblioteca Digital de Teses e Dissertações (BDTD) e importada para o catálogo online DEDALUS	Realizada diretamente pelos alunos de pós-graduação no repositório institucional e importada para o catálogo online Athena	Realizada no catálogo online Acervus pelos bibliotecários catalogadores-indexadores e importada para o repositório
Tipo de linguagem utilizada para padronização dos assuntos de teses e dissertações	Linguagem controlada, composta pelo Vocabulário Controlado da USP (VOCAUSP)	Linguagem controlada, composta pelo Vocabulário Controlado da UNESP	Linguagem controlada, composta por vocabulário misto, de acordo com cada área do conhecimento, sendo geralmente: LC, BN,

			DeCS, <i>Ei Compendex</i> , <i>FSTA Thesaurus</i> , dentre outros
Autoarquivamento de outros recursos informacionais	Somente pelos bibliotecários	Não há	Não há
Catálogo de outros recursos informacionais	Coleta automática das bases de dados <i>Web of Science</i> , <i>Scopus</i> , <i>SciELO</i> e <i>PubMed</i> pela equipe gestora do repositório. Após, os recursos informacionais são encaminhados às Bibliotecas para validação/verificação e inserção no repositório institucional	Coleta automática das bases de dados <i>Web of Science</i> , <i>Scopus</i> , <i>SciELO</i> e <i>PubMed</i> pela equipe gestora do repositório	Coleta automática das bases de dados <i>Web of Science</i> , <i>Scopus</i> , <i>SciELO</i> e <i>PubMed</i> . Após, os recursos informacionais são encaminhados às Bibliotecas para validação/verificação, que retornam os recursos informacionais validados à gestão, para inserção
Tipo de linguagem utilizada para padronização dos assuntos de outros recursos informacionais	Linguagem controlada, composta pelo Vocabulário Controlado da USP (VOCAUSP)	Linguagem natural no repositório institucional Linguagem controlada no catálogo online, composta pelo Vocabulário Controlado da UNESP	Linguagem controlada, composta por vocabulário misto, de acordo com cada área do conhecimento, sendo geralmente: LC, BN, DeCS, <i>Ei Compendex</i> , <i>FSTA Thesaurus</i> , dentre outros
Política de indexação formalizada	Sim	Sim	Não
Manual de catalogação	Sim	Sim	Sim
Manual de indexação	Sim	Sim	Não
Grupo gestor de linguagem controlada	Sim	Sim	Não
Adoção do ORCID para autoridade de autores	Em implementação	Em implementação	Em implementação

Fonte: Elaboração própria.

Na avaliação da indexação de assuntos pela perspectiva do sistema - ***Avaliação intrínseca quantitativa mediante a interconsistência***, considerando-se que os índices de interconsistência podem variar entre 0% a 100% de correspondência/consistência entre os assuntos atribuídos na indexação de assuntos, os resultados revelaram baixos índices de interconsistência na atribuição de assuntos entre autores e bibliotecários catalogadores-indexadores no Repositório Institucional da UNICAMP, no universo de recursos informacionais

analisado, com uma variação média de 2,38% a 21,05% nas cinco áreas do conhecimento analisadas. Cabe salientar que o caráter de originalidade da aplicabilidade da avaliação da indexação de assuntos entre autores e bibliotecários catalogadores-indexadores em repositórios institucionais inviabilizou a comparação dos resultados obtidos nesta pesquisa com outros índices já levantados na literatura, devido à lacuna de pesquisas nesta temática. Nesse sentido, torna-se importante o encaminhamento e continuação das pesquisas, para que seja possível obter um panorama mais amplo da avaliação da indexação de assuntos em repositórios institucionais.

Na **avaliação da indexação de assuntos pela perspectiva do usuário - Avaliação extrínseca mediante a recuperação**, considerando-se que os índices de precisão também podem variar entre 0% a 100% entre os recursos informacionais recuperados e os recursos informacionais relevantes recuperados, os resultados apresentaram baixos índice de precisão tanto com a linguagem natural (12,97%) quanto com a linguagem controlada (9,93%) no Repositório Institucional da UNICAMP, no universo de recursos informacionais e de participantes analisado. No geral, os usuários relataram suas dificuldades e sugeriram melhorias significativas no Repositório Institucional da UNICAMP. De acordo com a literatura, o desempenho dos mesmos durante as buscas por assuntos é determinado, dentre outros aspectos, pelo nível de apoio prestado pelos profissionais da informação, a quantidade de informações que possa auxiliá-lo para realizar uma busca por assuntos, a visualização das informações, a facilidade de interação com o repositório institucional e suas próprias habilidades para buscas em geral (MOREIRO, 2002, p. 54). Logo, as melhorias sugeridas precisam ser implementadas o quanto antes para que as buscas por assuntos sejam mais precisas, dinâmicas e relevantes a suas necessidades informacionais.

Em síntese, os resultados da pesquisa alertam para a qualidade da indexação de assuntos nos repositórios institucionais visto os baixos índices de consistência entre a indexação realizada pelos autores e a indexação/validação/verificação realizada pelos bibliotecários catalogadores-indexadores. Tal situação foi refletida na busca e recuperação por assuntos dos recursos informacionais no Repositório Institucional da UNICAMP, como apontado pelos baixos índices de precisão entre a linguagem natural e a linguagem controlada nas buscas por assuntos realizadas pelos usuários. Depreende-se que tais resultados estão relacionados a três principais fatores:

- ausência de recursos informacionais nas referidas temáticas de buscas;
- ausência de especificidade na atribuição dos recursos informacionais; ou
- ausência de descrição dos recursos informacionais por meio de uma linguagem controlada/padronizada no Repositório Institucional da UNICAMP.

Tais fatores são decorrentes de questões mais amplas:

- ausência de uma política de indexação no contexto da UNICAMP e formalização em uma Manual de Política de Indexação que norteie a indexação de assuntos no repositório Institucional;
- ausência de um vocabulário próprio e único que padronize a indexação e a recuperação por assuntos no repositório institucional;
- ausência de disponibilização da linguagem controlada utilizada na indexação dos recursos informacionais aos usuários, tendo em vista que a recuperação ocorre por meio de coincidências terminológicas entre a representação dos recursos informacionais e as buscas dos usuários. Assim, a disponibilização do vocabulário como recurso de busca melhoraria essas coincidências e conseqüentemente aumentaria a precisão.

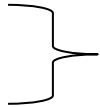
Ao recorrermos à literatura e à prática da Organização do Conhecimento sobre outros tipos de sistemas de recuperação da informação no contexto de bibliotecas universitárias, verifica-se que particularmente os modernos catálogos online permitem o acesso de forma individual aos itens das coleções, o que Buckland (1992) denomina de *acesso bibliográfico*. Para o autor, este processo permite aos usuários conectar-se a diversos tipos de recursos informacionais contidos em diferentes suportes e que inclui três pontos principais: a *identificação* dos recursos informacionais, sua *localização* e o *acesso físico* ao material.

É pertinente destacar que no decorrer da história da Biblioteconomia, algumas tentativas têm sido feitas para identificar os objetivos dos catálogos. Nos primórdios da Biblioteconomia moderna, Cutter (1904) apresentou estes objetivos na publicação *Rules for*

a *Dictionary Catalog* (Regras para um Catálogo Dicionário⁵⁰), que embora tenham sido concebidas inicialmente apenas para os catálogos de bibliotecas, são aplicáveis a arquivos, museus e coleções e tais regras ainda parecem representar o catálogo na atualidade (TAYLOR; JOUDARY, 2009, p. 44). Adaptando-se tais objetivos concebidos para catálogos no contexto de repositórios institucionais, percebe-se que um repositório institucional deve:

1. Permitir que um usuário encontre um recurso informacional por

(A) autor
(B) título
(C) **assunto**



que é conhecido.

2. Indicar o que o repositório institucional possui

(D) de um determinado autor

(E) **sobre um determinado assunto**

(F) em um determinado tipo de recurso informacional.

3. Facilitar a escolha de um recurso informacional

(G) quanto à **tipo de recurso informacional**.

(H) quanto ao seu caráter (**na especialidade do assunto**).

Conforme Martinho (2010, p. 122), na relação estabelecida entre os objetivos de Cutter e as funções de cada tipo de catálogo (autor, título, assunto e forma) Miksa (1973⁵¹) verificou que

Objetivos A e D são funções das entradas em um catálogo de autor;

Objetivos **C e E** são funções das entradas em um catálogo de **assunto**;

Objetivo B, função das entradas em um catálogo de título;

Objetivo F, função das entradas em um catálogo de forma;

Objetivo G e H, não foram relacionados a um tipo específico de catálogo, mas sim para a plenitude da informação dada por qualquer um dos tipos anteriores.

⁵⁰ CUTTER, C. A. *Rules for a Dictionary Catalog*. 4th ed. (Washington, D.C.: Government Printing Office, 1904; reprint, London: The Library Association, 1962).

⁵¹ MIKSA, F. L. The Making of the 1876 Special Report on Libraries. *Journal of Library History*, 1973.

Ainda para a autora, Miksa (1973) analisa a sistematicidade intrínseca aos objetivos idealizados por Cutter, salientando que a enumeração e agrupamento que Cutter deu aos objetivos estão pautados em três principais aspectos, que nesta pesquisa podem ser aplicados no contexto de repositórios institucionais: *Finding principle*: identificação individual de cada recurso informacional para possibilitar o acesso aos mesmos pelos usuários; *Gathering principle*: reunião dos recursos informacionais identificados individualmente em categorias definidas; e *Evaluating principle*: fornecimento de informações bibliográficas e outras fontes possibilitando a escolha do recurso informacional por si mesmos (MARTINHO, 2010, p. 122).

Nesse sentido, com base nos ideais propostos por Cutter, reitera-se o objetivo do repositório institucional quanto à representação temática: “permitir que *um usuário* encontre um recurso informacional por assunto” e “*indicar* o que o *repositório institucional* possui sobre um determinado assunto”. Logo, considerando-se todos os resultados da pesquisa, advoga-se a adequada indexação de assuntos em repositórios institucionais, permitindo que tais objetivos sejam alcançados, ou seja, a compatibilidade entre os recursos informacionais do repositório institucional e as necessidades informacionais dos usuários.

7 RECOMENDAÇÕES PARA A INDEXAÇÃO E RECUPERAÇÃO POR ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS

A fim de cumprir-se o quarto objetivo específico da pesquisa - sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais -, tomando-se como base os principais resultados da pesquisa, são elencadas algumas recomendações para a melhoria da representação e recuperação por assuntos no Repositório Institucional da UNICAMP, podendo ser replicadas em outros repositórios institucionais que compartilham das mesmas características:

- composição de um Grupo de Trabalho de representação temática: *GT de Indexação*, no âmbito da UNICAMP, composto por docentes especialistas e bibliotecários catalogadores-indexadores representantes das seis áreas do conhecimento do SBU: *Área de Artes e Humanidades, Área de Biomédicas, Área de Ciências Aplicadas, Área de Conhecimentos Gerais, Área de Exatas e Área de Tecnológicas* para composição de diretrizes e planos de trabalho. Para tanto, seriam estabelecidas parcerias com o *Grupo de Trabalho em Política de Indexação da Rede de Bibliotecas UNESP* e com o *Grupo Gestor do VOCAUSP*. Tanto a pesquisadora quanto a orientadora da presente pesquisa estão à disposição para colaborar no desenvolvimento da construção do referido vocabulário UNICAMP, assistindo os bibliotecários participantes nas ocorrências técnico-práticas que vierem a surgir no decorrer desse trabalho;
- composição de um Grupo de Trabalho de representação temática no âmbito do CRUESP: *GT de Catalogação/Indexação*, composto por gestores e bibliotecários catalogadores-indexadores representantes da USP, UNESP e UNICAMP;
- elaboração/formalização da política de indexação em um Manual de Política de Indexação que norteie a indexação de assuntos no Repositório Institucional da UNICAMP pelos profissionais que realizam o tratamento temático dos recursos informacionais;
- elaboração de um vocabulário próprio e único no âmbito da UNICAMP: *Vocabulário Controlado da UNICAMP*, que possa contemplar todas as áreas do conhecimento. Para tanto, sugere-se as seguintes principais etapas: levantamento dos termos já atribuídos às dissertações e teses incluídas no repositório; levantamento dos termos já atribuídos

aos demais recursos informacionais e validados pelos bibliotecários indexadores-catalogadores; padronização dos termos e verificação de inconsistências; compilação dos termos para formalização do *Vocabulário Controlado da UNICAMP*; reindexação dos demais recursos informacionais no Repositório Institucional da UNICAMP dos últimos 5 anos; e reindexação dos demais recursos informacionais no Repositório Institucional da UNICAMP dos demais anos, dentre outras atividades;

- disponibilização da linguagem controlada utilizada na indexação dos recursos informacionais aos usuários tanto na página do Repositório Institucional da UNICAMP quanto na página do catálogo online;
- padronização dos registros no Repositório Institucional da UNICAMP quanto à representação descritiva, com eliminação de recursos informacionais duplicados;
- melhoria da ferramenta para que execute buscas condizentes de acordo com as opções/estratégias de busca elencadas pelos usuários no Repositório Institucional da UNICAMP, já que, conforme detectado na pesquisa, tal filtro é inexistente, não fazendo diferença, no momento, a busca do mesmo termo utilizando-se a opção “Todos os campos”, “Título”, “Autor”, “Assunto” ou “Data de publicação”, isto é, trazendo exatamente os mesmos resultados. Especificamente referente às buscas com a opção “Assunto”, todas as palavras, tanto dos registros quanto dos próprios recursos informacionais são verificadas e retornadas pelo sistema e não apenas no campo “Palavras-chave”, isto é, no elemento 3 Subject (Assunto) e respectivo esquema de classificação <dc:subject> do Dublin Core. Todavia, mesmo que todos os recursos informacionais do Repositório Institucional da UNICAMP estivessem padronizados de acordo com um vocabulário controlado, ainda assim as buscas continuariam com baixíssimos índices de precisão, especialmente aquelas em que não são utilizadas o recurso gráfico aspas (“) nos termos de busca. Torna-se necessária uma remodelação tanto no tratamento quanto na recuperação do sistema;
- em termos gerais de sistema, sugere-se para correção ou desenvolvimento: implementação de busca avançada e utilização de operadores booleanos; disponibilização de um guia rápido aos usuários com informação de que buscas realizadas com termos utilizando-se o recurso gráfico aspas (“) possibilitam maior precisão nos resultados; especificação ao usuário de que tipo de busca está sendo

realizada, tendo em vista que, após a realização da primeira pesquisa, tal informação não aparece; implantação de recursos de acessibilidade; alteração do nome “base Acervus” para “catálogo Acervus”, considerando-se suas características próprias de catálogo e não de base de dados; melhoria da categorização dos tipos de recursos informacionais; supressão do botão “Retornar valores”, visto que confunde os usuários; e categorização dos recursos informacionais por Área do conhecimento, Unidade e Departamento, dentre outras;

- por fim, desenvolvimento de perfis de aplicação para geração de indicadores no Repositório Institucional da UNICAMP, atendendo à demanda da universidade de indicadores de produção confiáveis.

8 CONSIDERAÇÕES FINAIS

Ao fim deste percurso, são pertinentes algumas reflexões em torno da proposição e delineamento que possibilitaram a elaboração desta pesquisa. Reconhecendo que o objetivo do processo de indexação é permitir a adequada recuperação por assuntos em um sistema de recuperação, buscamos verificar suas duas faces em repositórios institucionais: o processo de indexação pela abordagem do tratamento/da representação temática e o processo de indexação pela abordagem da recuperação da informação. Desse modo, em sua vertente teórica foram investigados alguns pressupostos da Organização do Conhecimento e em sua vertente prática, se as necessidades informacionais dos usuários são supridas de forma adequada, pela abordagem da recuperação.

Partindo-se da relevância da recuperação da informação aos usuários e pesquisadores no âmbito acadêmico, visamos contribuir para a resolução de problemas relacionados à organização conceitual do conhecimento para indexação de assuntos da Organização do Conhecimento, o que levou ao delineamento do problema de pesquisa: a necessidade de verificação/avaliação da indexação de assuntos na representação e recuperação da informação científica especializada em repositórios institucionais como sistemas de recuperação da informação no contexto de bibliotecas universitárias.

Tendo em vista o aprimoramento do tratamento temático da informação documental e da recuperação por assuntos nestes sistemas de recuperação tanto no contexto de bibliotecas universitárias quanto no âmbito acadêmico, o objetivo geral da pesquisa foi contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento.

Para que o objetivo geral pudesse ser alcançado, foram realizados quatro objetivos específicos: estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento temático/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias;

investigação sobre o tratamento temático da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como *metodologia qualitativa* com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores, com coleta de documentação; verificação da recuperação documental de assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de bibliotecas universitárias, por meio da aplicação da avaliação da indexação de assuntos como *metodologia quantitativa*; e sistematização da análise comparada com delineamento de diretrizes de indexação de assuntos para repositórios institucionais, visando propor um modelo pertinente de avaliação da indexação de assuntos em repositórios institucionais.

A hipótese inicial da pesquisa foi confirmada, pautada na concepção de que investigações em torno da avaliação da recuperação por assuntos em repositórios institucionais demonstrariam a necessidade de adequação destes contemporâneos sistemas de recuperação da informação, tanto na vertente do tratamento temático da informação realizado pelo bibliotecário catalogador-indexador quanto na busca e recuperação da informação realizada pelos usuários. Diante da proposição inicial, que foi investigar/avaliar o processo de indexação de assuntos em repositórios institucionais e sua influência na recuperação da informação pelos usuários, a tese da pesquisa foi confirmada, ou seja, verificamos a existência de problemas na recuperação por assuntos, especificamente no Repositório Institucional da UNICAMP pela ausência de formulação e aplicabilidade de diretrizes de indexação nestes sistemas de recuperação da informação.

Na prática, o delineamento de procedimentos metodológicos de avaliação da indexação de assuntos em repositório institucional possibilitou o levantamento de um estudo diagnóstico da atual situação da indexação e da recuperação por assuntos de forma detalhada no Repositório Institucional da UNICAMP. Todavia, devido à complexidade dos elementos envolvidos, as principais limitações do estudo centram-se na escolha de um dentre os três repositórios institucionais que compõem o CRUESP para a aplicabilidade da metodologia de avaliação da indexação de assuntos. Todavia, considerando-se que a avaliação da indexação de assuntos é um instrumento importante para determinar a eficácia da biblioteca universitária em suprir as necessidades informacionais de seus usuários no que se refere a seus repositórios institucionais, torna-se relevante o encaminhamento de estudos detalhados

nos Repositórios da USP e UNESP. A análise comparada de tais resultados possibilitará um panorama mais amplo da indexação e recuperação por assuntos nestes repositórios e o delineamento de procedimentos padronizados nos três repositórios, como componentes integradores do Repositório CRUESP.

Verifica-se notadamente uma tendência e cada vez maior relevância dos repositórios institucionais como sistemas de recuperação da informação não apenas no contexto de bibliotecas universitárias e, em uma perspectiva mais ampla, no contexto acadêmico no qual estão inseridos, tornando-se ainda instrumentos de gestão universitária por meio de indicadores científicos. No contexto acadêmico, os repositórios institucionais propiciam visibilidade à coleção de recursos informacionais produzida pela universidade, permitindo conhecer quais títulos, autores ou determinados assuntos compõem a coleção e reunindo todas as publicações de um autor e todas as publicações sobre um mesmo assunto, além de outros pontos de acesso secundários, conforme a necessidade de seus usuários e da universidade.

Nesse sentido, tanto a indexação de assuntos (processo) quanto o repositório institucional (produto) realizam uma *mediação* entre os recursos informacionais de uma universidade e as necessidades informacionais dos usuários no âmbito acadêmico. Neste cenário, reitera-se a relevância da recuperação da informação condizente com as necessidades informacionais dos usuários, pautada em princípios/diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, em consonância com o corpus teórico-prático da abordagem da indexação de assuntos da Organização do Conhecimento e os atuais recursos tecnológicos de representação e recuperação da informação.

Em atenção à complexidade diversa e crescente que permeia a representação e a recuperação da informação por assuntos de recursos informacionais, a biblioteca universitária deve abarcar ferramentas, padrões e diretrizes de organização, tornando os demais atores envolvidos no processo parceiros ativos na crescente interconectividade dos fenômenos sociais, do conhecimento acadêmico e da inovação tecnológica, de modo a compor um panorama compatível com a realidade atual, visando a proposição de melhorias no *acesso*, *uso* e *reuso* dos recursos informacionais em repositórios institucionais no âmbito acadêmico. De fato, as questões técnicas sobre os repositórios institucionais precisam ser desenvolvidas,

gerenciadas e promovidas pelo bibliotecário, de modo a envolver não apenas a biblioteca universitária como toda a comunidade acadêmica. Neste cenário, as bibliotecas universitárias estão em boa posição para criar e gerenciar repositórios institucionais, pois além de estarem ‘mais próximas dos pesquisadores’ que produzem, são instituições marcadas historicamente pela preservação, tratamento e disseminação dos recursos informacionais acadêmicos, cada vez mais relevantes.

As reflexões delineadas nesta pesquisa visam reforçar o papel-chave desempenhado pelas bibliotecas universitárias em um cenário de profundas transformações de suas atividades, especialmente no que tange à indexação de assuntos de recursos informacionais em repositórios institucionais, tanto pela perspectiva da *representação* quanto da *recuperação por assuntos* por meio dos metadados temáticos. Neste cenário, destaca-se o protagonismo do bibliotecário na criação e gerenciamento de metadados temáticos, possibilitando a representação e a recuperação por assuntos em repositórios institucionais.

Entretanto, isto requer uma atuação profissional condizente com este (novo) papel na definição e criação de metadados em repositórios institucionais, pautada em uma melhor formação e educação continuada do profissional da informação tanto nos diversos aspectos que cercam a criação e gerenciamento de metadados de recursos informacionais como nos fundamentos teórico-metodológicos em indexação de assuntos da Organização do Conhecimento. Tais mudanças têm trazido novos desafios aos profissionais bibliotecários, reiterando a necessidade cada vez maior de se pensar seu *modus operandi*.

Certamente, algumas questões aqui discutidas levam a outras indagações e demonstram que a avaliação e reflexão crítica do processo de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias é uma temática imprescindível no escopo da Organização do Conhecimento. Espera-se que alguns desses caminhos tenham sido evidenciados nesta Tese, contribuindo para a efetiva recuperação por assuntos dos recursos informacionais nestes contemporâneos sistemas de recuperação da informação.

REFERÊNCIAS

- ABAD GARCÍA, M. F.; PÉREZ, I. A.; BENAVENT, A. Evaluación de la consistencia en la indización del repertorio documentación médica española. *Revista Española de Documentación Científica*, v. 21, n. 4, p. 389-401, 1998.
- ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The indexer*, v. 18, n. 4, p. 219-223, oct. 1993.
- ALMEIDA, M. C. B. de. *Planejamento de bibliotecas e serviços de informação*. 2. ed. rev. e ampl. Brasília: Briquet de Lemos, 2005.
- ALVES, R. C. V. *Metadados como elementos do processo de catalogação*. 2010. 132 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- ALVES, R. C. V.; SANTOS, P. L. V. A. da C. Metadados: organização e acesso à informação no domínio bibliográfico. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9., 2013, Rio de Janeiro; ENCONTRO NACIONAL DE CATALOGADORES, 2., 2013, Rio de Janeiro. *Anais [...]*. Rio de Janeiro: Fundação Biblioteca Nacional, 2013.
- ARAKAKI, F. A. *O padrão de metadados Dublin Core: simples e codificado* [livro eletrônico]. Marília: NEaD, 2015.
- ARRIOLA NAVARRETE, O.; TECUATL QUECHOL, M. G. M. (Orgs.). *Evaluación de bibliotecas: um compendio de experiências*. México: Escuela Nacional de Biblioteconomía y Archivonomía: Library Outsourcing Service, 2011.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 12676: *Métodos para análise de documentos* - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992. 4 p.
- BABIK, W. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, p. 327-331, v. 41, n. 4, 2014.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. *Recuperação de informação: conceitos e tecnologia das máquinas de busca*. 2. ed. Porto Alegre: Bookman, 2013.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (Org.). *Educação, universidade e pesquisa*. Marília: UNESP-Marília Publicações; São Paulo: FAPESP, 2001. p. 35-60.
- BARRY, C. User-defined relevance criteria: an exploratory study. *Journal of the American Society for Information Science*, v. 45, p. 149-159, 1994.

BETTENCOURT, A. M. *A representação da informação na Biblioteca Nacional: do documento tradicional ao digital*. Rio de Janeiro: Fundação Biblioteca Nacional, 2014.

BLAIR, D. C. Indeterminacy in the subject access to documents. *Information Processing & Management*, v. 22, n. 2, p. 229-241, 1986.

BOCCATO, V. R. C.; FUJITA, M. S. L.; GIL LEIVA, I. Avaliação comparada do uso de linguagens de indexação em catálogos de bibliotecas universitárias para recuperação por assunto. *Scire*, Zaragoza, v. 17, n. 1, p. 55-64, en.-jun. 2011.

BRÄSCHER, M. *Organização do conhecimento no Brasil: pós-graduação e grupos de pesquisa*. I EnReDo - Encontro de Representação Documental, 2017, Universidade Federal de São Carlos, São Carlos, SP, Brasil.

BRUCE, H. A cognitive view of the situational dynamism of user-centered relevance estimation. *Journal of the American Society for Information Science*, v. 45, p. 142-149, 1994.

BUCKLAND, M. K. Information as thing. *Journal of American Society for Information Science*. v. 42, n. 5, p. 351-360, 1991.

BUDD, J. M. *The academic library: its context, its purpose, and its operation*. Englewood: Libraries Unlimited, 1998.

CALHOUN, K. *Exploring digital libraries: foundations, practice, prospects*. London: Facet, 2014.

CAMARGO, L. S. de A. de; VIDOTTI, S. A. B. G. *Arquitetura da informação: uma abordagem prática para o tratamento de conteúdo e interface em ambientes informacionais digitais*. Rio de Janeiro: LTC, 2011.

CAMINITA, C. E-books and patron-driven acquisitions in academic libraries. In: BRIDGES, K. (Ed.). *Customer-based collection development: an overview*. London: Facet Publishing, 2014.

CASE, D. O.; GIVEN, L. M. *Looking for information: a survey of research of information seeking, needs, and behavior*. 4th ed. Howard House, Wagon Lane, Bingley: Emerald, 2016.

CAVALCANTI, M. do C. *Interação leitor-texto: aspectos de interpretação pragmática*. Campinas: Editora da UNICAMP, 1989.

CESARINO, M. A. N., PINTO, M. C. M. F. Análise de assunto. *Revista de Biblioteconomia de Brasília*, Brasília, v. 8, n. 1, p. 32-43, jan./jun., 1980.

CHAN, L. M. *Cataloging and classification: an introduction*. 2nd ed. New York: McGraw-Hill, 1994.

CHOWDHURY, G. G.; CHOWDHURY, S. *Organizing information: from the shelf to the web*. London: Facet Publishing, 2007.

CHU, C. M., O'BRIEN, A. Subject analysis: the critical first estage in indexing. *Journal of Information Science*, Amsterdan, v. 1, n. 19, p. 439-454, 1993.

CLEVELAND, D. B.; CLEVELAND, A. D. *Introduction to indexing and abstracts*. 2nd ed. Englewood: Libraries Unlimited, 1990.

COOPER, W. S. Is interindexer consistency a hobgoblin? *American Documentation*, v. 20, n. 3, p. 268-278, 1969.

CROW, R. *The case for institutional repositories: a SPARC position paper*. Washington: SPARC, 2002.

CRUESP. Portal. Disponível em: <http://www.repositorio.unicamp.br>. Acesso em: 01 abr. 2019.

DAHLBERG, I. Knowledge organization: its scope and possibilities. *Knowledge Organization*, v. 20, n. 4, p. 211-222, 1993.

DAHLBERG, I. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, p. 327-331, v. 41, n. 4, 2014.

DAL'EVEDOVE, P. R. *O tratamento temático da informação em abordagem sociocultural: diretrizes para definição de política de indexação em bibliotecas universitárias*. 2014. 268 p. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

DAL'EVEDOVE, P. R.; FUJITA, M. S. L.; TARTAROTTI, R. C. D. A produção científica periódica na temática indexação: análise bibliométrica no período de 2003 a 2012. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 14., 2013, Florianópolis. *Anais [...]*. Florianópolis: Universidade Federal de Santa Catarina (UFSC), 2013.

DAL'EVEDOVE, P. R.; TARTAROTTI, R. C. D.; FUJITA, M. S. L. Reflexões acerca do desenvolvimento de metodologias para análise de assunto. *In: GUIMARÃES, J. A. C.; DODEBEI, V. (Orgs.). Organização do conhecimento e diversidade cultural*. Marília: ISKO-Brasil: FUNDEPE, 2015. (Série: Estudos Avançados em Organização e Representação do Conhecimento, v. 3).

DAL'EVEDOVE, P.; TARTAROTTI, R. C. D.; FUJITA, M. S. L. Estudos sobre análise de assunto no Brasil: estado da arte e perspectivas futuras. *Scire*, Zaragoza, p. 35-44, v. 24, n. 1, en.-jun. 2018.

DAL'EVEDOVE, R. C. *Compreensão de leitura em análise de assunto para identificação e seleção de conceitos*. 2002. 154 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.

DAVID, A. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, p. 327-331, v. 41, n. 4, 2014.

DESCRITORES em Ciências da Saúde: DeCS. ed. rev. e ampl. São Paulo: BIREME/OPAS/OMS, 2019. Disponível em: <http://www.decs.bvs.br>. Acesso em: 16 ago. 2019.

DÍEZ CARRERA, C. *La biblioteca digital*. Somonte-Cenero: Trea, 2012.

DODEBEI, V. Repositórios institucionais: por uma memória criativa no ciberespaço. In: SAYÃO, L.; TOUTAIN, L. B.; ROSA, F. G.; MARCONDES, C. H. (Orgs.). *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. Salvador: EDUFBA, 2009.

DUBLIN Core Metadata Initiative. *About us*, 2014. Disponível em: <http://www.dublincore.org/about>.

FAIRTHORNE, R. A. Content analysis, specification, and control. *Annual Review of Information Science and Technology*, Medford, NJ, v. 4, p. 73-109, 1969.

FARROW, J. F. A cognitive process model of document indexing. *Journal of Documentation*, v. 47, n. 2, p. 149-66, 1991.

FERNEDA, E. *Ontologia como recurso de padronização terminológica em um sistema de recuperação da informação*. 2013. 98 f. Relatório de Pesquisa (Pós-doutorado em Ciência da Informação) - Universidade Federal da Paraíba. João Pessoa, 2013.

FIDEL, R. User-centered indexing. *Journal of the American Society for Information Science*, v. 48, n. 8, p. 572-576, 1994.

FROEHLICH, T. J. Relevance reconsidered: towards an agenda for the 21st century: introduction to special topic issue on relevance research. *Journal of the American Society for Information Science*, v. 45, p. 124-134, 1994.

FUGMANN, R. *Subject analysis and indexing: theoretical foundation and practical advice*. Frankfurt/Main: INDEKS Verlag, 1993.

FUJITA, M. S. L. *Linguagem documentária em Odontologia: uma aplicação do sistema de indexação PRECIS*. 1992. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo. 3 v.

FUJITA, M. S. L. *Leitura em análise documentária: uma contribuição à formação do indexador*. Marília: UNESP; CNPq, 2001. 190 p. (Relatório Parcial de Pesquisa).

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, dez. 2003.

FUJITA, M. S. L. Aspectos evolutivos das bibliotecas universitárias em ambiente digital na perspectiva da rede de bibliotecas da UNESP. *Informação & Sociedade: Estudos*, v. 15, n. 2, p. 97-112, 2005.

FUJITA, M. S. L. Representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *Ponto de Acesso*, Salvador, v. 7, n. 1, p. 42-66, abr. 2013.

FUJITA, M. S. L.; GIL LEIVA, I. Avaliação da indexação por meio da recuperação da Informação. *Ciência da Informação*, Brasília, v. 41 n. 1, p. 50-66, jan./abr., 2014.

FUJITA, M. S. L.; RUBI, M. P.; BOCCATO, V. R. C. As diferentes perspectivas teóricas e metodológicas sobre indexação e catalogação de assuntos. In: FUJITA, M. S. L. (Org.). *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. São Paulo: Cultura Acadêmica, 2009. p. 19-42.

FUMANI, M. R. F. Q. Inter-indexer consistency (IIC) in a Persian context. *The Indexer*, v. 28, n. 1, p. 12-17, mar. 2010.

FUNK, M.; REID, C. A.; MCGOOGAN, L. E. Indexing consistency in MEDLINE. *Bulletin of the Medical Library Association*, v. 71, p. 176-183, 1983.

GIASSON, J. *A compreensão na leitura*. Lisboa: Asa, 1993.

GIL LEIVA, I. *Manual de indización: teoría y práctica*. Gijón: Trea, 2008.

GIL LEIVA, I.; FUJITA, M. S. L. (Eds.). *Política de indexação*. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012.

GIL LEIVA, I.; RUBI, M. P.; FUJITA, M. S. L. Consistência na indexação em bibliotecas universitárias brasileiras. *Transinformação*, Campinas, v. 20, n. 3, p. 233-253, set./dez. 2008.

GILLILAND-SWETLAND, A. J. La definición de los metadatos. In: BACA, M. (Ed.). *Introducción a los metadatos vías a la información digital*. Traducido al español por Marisol Jacas-Santoll. Los Angeles, CA: J. Paul Getty Trust, 1998.

GNOLLI, C. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, p. 327-331, v. 41, n. 4, 2014.

GOMES, F. A. *Padronização de metadados na representação da informação em repositórios institucionais de universidades federais brasileiras*. 2015. 277 f. Dissertação (Mestrado em Ciência da Informação) - Instituto de Ciência da Informação, Universidade Federal da Bahia, Salvador, 2015.

GONÇALVES, M. Digital libraries. *In: BAEZA-YATES, R.; RIBEIRO-NETO, B. Modern information retrieval: the concepts and technology behind search.* 2nd ed. Harlow: Pearson, 2011.

GONZALO NAVARRO, M. Indexação e busca. *In: BAEZA-YATES, R.; RIBEIRO-NETO, B. Recuperação de informação: conceitos e tecnologia das máquinas de busca.* 2. ed. Porto Alegre: Bookman, 2013.

GUIMARÃES, E. *Texto, discurso e ensino.* São Paulo: Contexto, 2009.

GUIMARAES, J. A. C. Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. *Informação & Informação*, Londrina, v. 22, n. 2, p. 84-98, maio/ago., 2017.

HARTER, S. P. Psychological relevance and information Science. *Journal of the American Society for Information Science*, v. 43, p. 602-615, 1992.

HASKINS, C. H. *The rise of universities.* Cornell University Press, 2013.

HERNON, P.; MATTHEWS, J. R. *Reflecting on the future of academic and public libraries.* Chicago: Ala Editions, 2013.

HJØRLAND, B. *Information seeking and subject representation: an activity-theoretical approach to Information Science.* Westport: Greenwood Press, 1997.

HJØRLAND, B. Domain analysis in information science. *Journal of Documentation*, v. 58, n. 4, p. 422-462, 2002.

HJØRLAND, B. Fundamentals of knowledge organization. *International Society for Knowledge Organization*, v. 30, n. 2, p. 87-111, 2003.

HJØRLAND, B. What is Knowledge Organization (KO)? *Knowledge Organization*, v. 35, n. 2/3, p. 86-101, 2008.

HJØRLAND, B. Theories of Knowledge Organization: theories of knowledge. *Knowledge Organization*, Germany, v. 40, n. 3, p. 169-181, 2013.

HOOPER, R. S. *Indexer consistency tests: origin, measurement, results, and utilization.* Bethesda: IBM Corporation, 1965.

HOUAISS, A.; VILLAR, M. de S. *Dicionário Houaiss da língua portuguesa.* Rio de Janeiro: Objetiva, 2009.

HOWARD, D. L. Pertinence as reflected in personal constructs. *Journal of the American Society for Information Science*, v. 45, p. 172-185, 1994.

HUGHES, A. V.; RAFFERTY, P. M. Inter-indexer consistency in graphic materials indexing at the National Library of Wales. *Journal of Documentation*, v. 67, n. 1, p. 9-32, 2011.

IFLA. *Declaração dos Princípios Internacionais de Catalogação*. 2009.

INÁCIO, M. de O. *Avaliação da indexação em bibliotecas universitárias: uma aplicação em catálogos online*. 2012. 157 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.

INGWERSEN, P.; JÄRVELIN, K. *The turn: integration of information seeking and retrieval in context*. Dordrecht: Springer, 2005.

ISKO. *International Society for Knowledge Organization*. ISKO's mission. Disponível em: <http://www.isko.org/about.html>. Acesso em: 15 jun. 2018.

ISO 5963 *Documentation*. Methods for examining documents, determining their subjects, and selecting indexing terms. Geneve: International Organization for Standardization, 1985.

JOHNSON, P. *Fundamentals of collection development and management*. 3rd. ed. London: Facet Publishing, 2014.

JONES, C. *Institutional repositories: content and culture in an open access environment*. Oxford: Chandos Publishing, 2007.

KIESLER, S.; SPROULL, L. Managerial response to changing environments: perspectives on problem sensing from social cognition. *Administrative Science Quarterly*, v. 27, p. 548-570, 1982.

KURAMOTO, H. Sintagmas nominais: uma nova abordagem no processo de indexação. In: NAVES, M. M. L.; KURAMOTO, H. (Orgs.). *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. p. 117-137.

KURAMOTO, H. Prefácio. In: LEITE, F. C. L. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Brasília: IBICT, 2009. p. 7-10.

LANCASTER, F. W. *Avaliação de serviços de bibliotecas*. Brasília: Briquet de Lemos, 1996.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Tradução de: Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos, 2004.

LANGRIDGE, D. *Subject analysis: principles and practice*. London: Bowker-Sour, 1989.

LARA, M. L. G. *A representação documentária: em jogo a significação*. 1993. 133 f. Dissertação (Mestrado em Ciência da Comunicação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1993.

LEITE, F. C. L. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Brasília: IBICT, 2009.

LEONARD, L. E. *Inter-indexer consistency studies, 1954-1975: a review of the literature and summary of study results*. Graduate School of Library Science, Occasional Papers, n. 31, University of Illinois, Urbana, IL. 1977.

LYNCH, C. A. Institutional repositories: essential infrastructure for scholarship in the digital age. *The Association of Research Libraries*, n. 226, p. 1-7, feb. 2003.

MACIEL, A. C. *Planejamento de bibliotecas: o diagnóstico*. 2. ed. Niterói: EDUFF, 1997.

MAI, J.-E. The concept of subject: on problems in indexing. Knowledge Organization for Information Retrieval. *Proceedings of the 6th International Study Conference on Classification Research*, v. 6, p. 60-67, 1997a.

MAI, J.-E. The concept of subject in a semiotic light. Digital Collections: implications for users, funders, developers and maintainers. *Proceedings of the ASIS Annual Meeting*, v. 34, p. 54-64, 1997b.

MAI, J.-E. Deconstructing the indexing process. *Advances in Librarianship*, v. 23, p. 269-298, 2000.

MAI, J.-E. Semiotics and indexing: na analysis of the subject indexing process. *Journal of Documentation*, v. 57, n. 5, p. 591-622, 2001.

MAI, J.-E. Folksonomies and the new order: authority in the digital disorder. *Knowledge Organization*, v. 38, n. 2, p. 114-122, 2011.

MESQUITA, R. M. A.; STUMPF, I. R. C. Estudo de citações de documentos eletrônicos on-line em revistas da área de comunicação. *Em Questão*, Porto Alegre, v. 10, n. 2, jul./dez. 2004.

MIRANDA, M. L. C. de. A organização do conhecimento e seus paradigmas científicos: uma abordagem epistemológica. *Informare: Cadernos do Programa de Pós-Graduação em Ciência da Informação*, Rio de Janeiro, v. 5, n. 2, p. 64-77, jul.-dez. 1999.

MIZZARO, S. Relevance: the whole history. In: HAHN, T.; BUCKLAND, M. (Eds.). *Historical studies in Information Science*. Medford, NJ: Information Today, 1998. p. 221-243.

MOOERS, C. Zatacoding applied to mechanical organization of knowledge. *American Documentation*, v. 2, p. 20-30, 1951.

MORALES DEL CASTILLO, J. M. *Hacia la biblioteca digital semántica*. Ediciones Trea: Gijón, 2011.

MOREIRO, J. A. Criterios e indicadores para evaluar la calidad del análisis documental de conteúdo. *Ciência da Informação*, Brasília, v. 31, n. 1, p. 53-60, jan./abr. 2002.

NAVARRETE, O. A.; QUECHOL, M. G. M. T. (Orgs.). *Evaluación de bibliotecas: um compendio de experiências*. México: Escuela Nacional de Biblioteconomía y Archivonomía: Library Outsourcing Service, 2011.

NAVARRO, G.; ZIVIANI, N. Documentos: linguagens e propriedades. In: BAEZA-YATES, R.; RIBEIRO-NETO, B. *Recuperação de informação: conceitos e tecnologia das máquinas de busca*. 2. ed. Porto Alegre: Bookman, 2013. p. 187-247.

NAVES, M. M. L. *Fatores interferentes no processo de análise de assunto: estudo de caso de indexadores*. 2000. 273 f. Tese. (Doutorado em Ciência da Informação) - Universidade Federal de Minas Gerais, Belo Horizonte, 2000.

NAVES, M. M. L.; KURAMOTO, H. (Orgs.). *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006.

NEVES, D. A. B. Ciência da Informação e cognição humana: uma abordagem do processamento da informação. *Ciência da Informação*, Brasília, v. 35, n. 1, p. 39-44, jan./abr. 2006.

NEVES, D. A. de B.; DIAS, E. W.; PINHEIRO, Â. M. V. Uso de estratégias metacognitivas na leitura do indexador. *Ciência da Informação*, Brasília, v. 35, n. 3, p. 141-152, set./dez. 2006.

NISO (National Information Standards Organization). *Guidelines for indexes and related information retrieval devices* (TR-02-1997). Bethesda, MD: NISO, 1997.

OHLY, H. P. *et al.* ISKO and Knowledge Organization's 25th anniversary: the future of Knowledge Organization and ISKO Panel Discussion. Reported by Rebecca Green. *Knowledge Organization*, p. 327-331, v. 41, n. 4, 2014.

OLSON, H. How we construct subjects: a feminist analysis. *Library Trends*, v. 56, n. 2, p. 509-541, 2007.

OPEN RESEARCHER AND CONTRIBUTOR ID. Disponível em: <http://orcid.org>. Acesso em: 19 nov. 2019.

PICKARD, A. J. *Research methods in information*. 2nd ed. London: Facet Publishing, 2013.

PINHEIRO, L. V. R.; Medidas de consistência da indexação: interconsistência. *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 109-114, 1978.

PINTO MOLINA, M. *Análisis documental: fundamentos y procedimientos*. 2. ed. rev. y aum. Madrid: Eudema, 1993.

PIOVEZAN, L. B. *Avaliação da indexação em catálogos em bibliotecas universitárias por meio da recuperação da informação*. 2015. 104 p. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2015.

PIOVEZAN, L. B.; FUJITA, M. S. L. Análise de cocitação de autores: uma aplicação em estudos de indexação. *Em Questão: Revista da Faculdade de Biblioteconomia e Comunicação da UFRGS*, v. 21, n. 1, p. 110-129, jan./abr. 2015.

RAJU, J.; RAJU, R. *Descriptive and subject cataloguing: a workbook*. Oxford: Chandos Publishing, 2006.

REDIGOLO, F. M.; DAL'EVEDOVE, R. C.; FUJITA, M. S. L.; BOCCATO, V. R. C. Elementos de política de indexação em biblioteca universitária da área médica. *Scire*, Zaragoza, v. 18, n. 2, p. 75-86, jul.-dic. 2012.

RICE, R.; SOUTHALL, J. *The data librarian's handbook*. London: Facet Publishing, 2016.

ROLLING, L. Indexing consistency, quality and efficiency. *Information Processing & Management*, v. 17, p. 69-76, 1981.

RUBIN, R. E. *Foundations of Library and Information Science*. 4. ed. London: Facet Publishing, 2016.

SÁNCHEZ-AMBRIZ, G. El modelo EFQM: instrumentación teórico-práctica en procesos de evaluación de bibliotecas universitarias. In: NAVARRETE, O. A.; QUECHOL, M. G. M. T. (Orgs.). *Evaluación de bibliotecas: un compendio de experiencias*. México: Escuela Nacional de Biblioteconomía y Archivonomía: Library Outsourcing Service, 2011.

SANTOS, P. L. V. A. da C.; SANTANA, R. C. G. Dado e granularidade na perspectiva da Informação e tecnologia: uma interpretação pela Ciência da Informação. *Ciência da Informação*, Brasília, v. 42, n. 2, jan. 2013.

SMIRAGLIA, R. P. Bibliocentrism, cultural warrant, and the ethics of resource description: a case study. *Cataloging & Classification Quarterly*, v. 47, n. 7, p. 671-686, 2009.

SMIRAGLIA, R. P. Domain coherence within Knowledge Organization: people, interacting theoretically, across geopolitical and cultural boundaries. In: *Annual CAIS/ACSI Conference*, 39., June 2-4, 2011, Canada. Proceedings [...]. Canada: University of New Brunswick, 2011. p. 1-6.

SMIRAGLIA, R. P. *The elements of Knowledge Organization*. Cham: Springer, 2014.

SIQUEIRA, J. C. Biblioteconomia, Documentação e Ciência da informação: história, sociedade, tecnologia e pós-modernidade. *Perspectivas em Ciência da Informação*, v. 15, n. 3, p. 52-66, set./dez. 2010.

SISTEMA DE BIBLIOTECAS DA UNICAMP (SBU). Portal. Disponível em: <http://www.sbu.unicamp.br>. Acesso em: 01 abr. 2019.

SOUSA, B. P. de; FUJITA, M. S. L. Análise de assunto no processo de indexação: um percurso entre teoria e norma. *Informação & Sociedade*, João Pessoa, v. 24, n. 1, p. 19-34, jan./abr. 2014.

SPARC. *The case for institutional repositories: a SPARC position paper*. Prepared by Raym Crow. Washington, DC: SPARC, 2002.

TAMAYO, A. M. M.; VALDEZ, J. C. *Indización y clasificación en bibliotecas*. Buenos Aires: Alfagrama, 2008.

TARGINO, M. das F.; GARCIA, J. C. R.; PAIVA, M. J. R. Repositórios institucionais brasileiros: entre o sonho e a realidade. *Revista FSA*, Teresina, v. 11, n. 1, art. 6, p. 117-133, jan./mar. 2014.

TARTAROTTI, R. C. D. *Atuação bibliotecária no tratamento temático da informação em unidades informacionais: um estudo comparativo qualitativo-quantitativo*. 2014. 277 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) - Universidade Federal de São Carlos - UFSCar, São Carlos, 2014.

TARTAROTTI, R. C. D., BOCCATO, V. R. C. A abordagem teórica bakhtiniana e o processo de indexação: diálogos. *Versão Beta: sob o signo da palavra*, ano XI, v. 74, p. 33-47, 2013.

TARTAROTTI, R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Avaliação da consistência da indexação em bibliotecas universitárias federais da Região Nordeste do Brasil. *Anales de Documentación*, v. 20, n. 1. p. 1-19, 2017.

TARTAROTTI, R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Avaliação da indexação de documentos não-textuais: uma análise da literatura. *In: XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 19., 2018, Londrina. Anais [...]. Londrina: Universidade Estadual de Londrina (UEL), 2018. p. 547-563.

TARTAROTTI, R. C. D.; FUJITA, M. S. L. Produção e colaboração científica em Organização e Representação do Conhecimento: análise bibliométrica do GT2 do ENANCIB no período de 2009 a 2014. *Em Questão*, Porto Alegre, v. 22, n. 3, p. 136-160, set/dez. 2016.

TARTAROTTI, R. C. D.; FUJITA, M. S. L.; MOREIRA, W. Aspectos temáticos da aplicabilidade de ontologias em catálogos online. *In: MOURA, M. A.; SILVEIRA, F. J. N. (Orgs.). In: ENCONTRO DA ASSOCIAÇÃO DE EDUCAÇÃO E PESQUISA EM CIÊNCIA DA INFORMAÇÃO DA IBERO-AMÉRICA E CARIBE*, 10., 2016, Belo Horizonte. *Anais [...]*. Belo Horizonte: ECI/UFMG, 2017. p. 1704-1720.

TARVER, H.; PHILLIPS, M. Metadata enhancement through name authority in the UNT Digital Library. *In: HOFFMAN, S. Dynamic research support for academic libraries*. London: Facet Publishing, 2016. p. 133-147.

TAYLOR, A. G.; JOUDAREY, D. N. *The organization of information*. 3rd ed. Wesport: Libraries Unlimited, 2009.

TODD, R. J. Academic indexing: what's it all about? *The Indexer*, v. 18, n. 2, p. 101-104, 1992.

TONTA, Y. A study of indexing consistency between Library of Congress and British Library catalogers. *Library Resources & Technical Services*, v. 35, p. 177-185, 1991.

UNISIST: Princípios de indexação. *Revista Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 10, n. 1, p. 83-94, mar. 1981.

UNIVERSIDADE ESTADUAL DE CAMPINAS. Portal. Disponível em: <http://www.unicamp.br>. Acesso em: 01 abr. 2019.

VAN SLYPE, G. *Lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Trad. de Pedro Hípola e Félix de Moya. Madrid: Fundación Germán Sánchez Ruipérez, 1991.

WELLISH, H. H. *Glossary of terminology in Abstracting, Classification, Indexing and Thesaurus Construction*. 2nd ed. American Society of Indexers, 2000.

WILSON, P. Situational relevance. *Information storage and retrieval*, v. 9, p. 457-471, 1973.

WOLFRAM, D.; OLSON, H.A.; BLOOM, R. Measuring consistency for multiple taggers using vector space modeling. *Journal of The American Society for Information Science and Technology*, v. 60, n. 10, p. 1995-2003, 2009.

ZICK, G. Digital collections: history and perspectives. *Journal of Library Administration*, v. 49, p. 687-693, 2009.

ZUNDE, P.; DEXTER, M.E. Indexing consistency and quality. *American Documentation*, v. 20, n. 3, p. 259-267, 1969.

Apêndice A. Roteiro da Entrevista de Diagnóstico Organizacional com os gestores dos repositórios institucionais

Projeto de pesquisa: Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação.

Autoria: Roberta Cristina Dal'Evedove Tartarotti (doutoranda do Programa de Pós-graduação em Ciência da Informação da UNESP/Campus de Marília), sob a orientação da Profa. Dra. Mariângela Spotti Lopes Fujita.

Objetivo da aplicação da entrevista: Coleta de dados com os gestores dos repositórios institucionais de universidades públicas do Estado de São Paulo: USP, UNESP e UNICAMP, para caracterização geral do Repositório Institucional, visando delinear recomendações para a indexação de assuntos em repositórios institucionais.

PARTE A - CARACTERIZAÇÃO DO REPOSITÓRIO INSTITUCIONAL

1) Dados gerais sobre o Repositório Institucional

- 1.1 Qual a Unidade(s) Organizacional(is) responsável pela gestão do Repositório Institucional?
- 1.2 Qual sua função/cargo no Repositório Institucional?
- 1.3 Qual sua formação acadêmica?
- 1.4 Quantos profissionais atuam na gestão do Repositório Institucional e qual a formação destes profissionais?
- 1.5 Quantos profissionais atuam especificamente no tratamento da informação no Repositório Institucional e qual a formação destes profissionais?
- 1.6 Há quanto tempo o Repositório Institucional está em funcionamento? Como foi o percurso histórico da implementação do Repositório Institucional?
- 1.7 Os bibliotecários catalogadores-indexadores recebem formação continuada sobre indexação/catalogação de assuntos para a atuação no Repositório Institucional? Se sim, de que tipo?

2) Infraestrutura do Repositório Institucional

- 2.1 O Repositório Institucional possui uma política de funcionamento formalizada e divulgada no contexto da universidade?
- 2.2 O Repositório Institucional possui um manual/tutorial de funcionamento? Se sim, está divulgado?
- 2.3 Qual o software utilizado para implantação do Repositório Institucional?

3) Depósito/submissão de recursos informacionais no Repositório Institucional

- 3.1 O autoarquivamento/autodepósito está implementado?
- 3.2 Se sim, quem possui responsabilidade de realizar o depósito de recursos informacionais no Repositório Institucional?
- 3.3 O Repositório Institucional oferece treinamento para os responsáveis pelo depósito de recursos informacionais?

3.4 O Repositório Institucional realiza revisão/controle de metadados antes de validar o depósito de recursos informacionais? Se sim, quem realiza a revisão/controle de metadados?

Em sua opinião, qual ação ainda não realizada pelo Repositório Institucional poderia contribuir para a melhoria da padronização da representação e recuperação?

Outros comentários e/ou sugestões para a pesquisa:

Apêndice B. Roteiro da Entrevista de Diagnóstico Organizacional com os bibliotecários catalogadores-indexadores dos repositórios institucionais

Projeto de pesquisa: Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação.

Autoria: Roberta Cristina Dal'Evedove Tartarotti (doutoranda do Programa de Pós-graduação em Ciência da Informação da UNESP/Campus de Marília), sob a orientação da Profa. Dra. Mariângela Spotti Lopes Fujita.

Objetivo da aplicação do questionário: Coleta de dados com os bibliotecários catalogadores-indexadores dos repositórios institucionais de universidades públicas do Estado de São Paulo: USP, UNESP e UNICAMP, para caracterização do tratamento temático da informação, visando delinear recomendações para a indexação de assuntos em repositórios institucionais.

Parte B - REPRESENTAÇÃO DA INFORMAÇÃO NO REPOSITÓRIO INSTITUCIONAL

1) Representação descritiva da informação no Repositório Institucional

1.1 Existe algum instrumento de representação da informação para a padronização dos metadados descritivos dos recursos informacionais no Repositório Institucional?

1.2 É utilizada alguma base de dados para o controle de autoridades de autores no Repositório Institucional?

2) Representação temática da informação no Repositório Institucional

2.1 A instituição dispõe de uma política de indexação regulamentada em seu contexto de atuação profissional?

2.2 Existe uma política de indexação específica para o Repositório Institucional?

2.3 A política de indexação está regulamentada em um documento?

2.4 Você faz uso desta política de indexação para direcionar a sua prática cotidiana na atividade de indexação?

2.5 A política de indexação vigente corresponde com suas necessidades profissionais?

2.6 A indexação dos recursos informacionais no Repositório Institucional é realizada a partir de registros copiados de algum catálogo ou base de dados? Se sim, qual(is)?

2.7 Durante o processo de indexação é utilizado algum auxílio automático ou semi-automático para facilitar essa atividade?

3) Qualidade da indexação no Repositório Institucional

3.1 O nível de especificidade na indexação no Repositório Institucional está estabelecido?

3.2 Existe indicação sobre o número de termos/assuntos por recurso informacional no Repositório Institucional?

3.3 A instituição segue alguma norma nacional ou internacional para a indexação no Repositório Institucional?

4) Ferramentas para a indexação no Repositório Institucional

4.1 Qual o tipo de linguagem utilizada na indexação dos recursos informacionais no Repositório Institucional (linguagem controlada/natural)?

4.2 No caso de linguagem controlada, o vocabulário controlado/linguagem documental utilizado para a indexação atende satisfatoriamente às necessidades da realização deste processo?

4.3 A instituição participa de projetos de compatibilidade/interoperabilidade de vocabulários controlados?

4.4 No Repositório Institucional é utilizado algum sistema de validação/verificação automática de termos/assuntos para garantir a consistência?

5) Avaliação da indexação de assuntos no Repositório Institucional

5.1 A instituição realiza testes ou ensaios para a avaliação periódica da indexação no Repositório Institucional? Se sim, existem relatórios publicados ou públicos dessa avaliação?

5.2 A instituição realiza testes ou ensaios para a avaliação da recuperação da informação com usuários no Repositório Institucional? Se sim, existem relatórios publicados ou públicos dessa avaliação?

Em sua opinião, qual ação ainda não realizada pelo Repositório Institucional poderia contribuir para a melhoria da padronização da representação e recuperação?

Outros comentários e/ou sugestões para a pesquisa:

Apêndice C. Termo de Consentimento Livre e Esclarecido (TCLE) - Entrevista semiestruturada com gestores e bibliotecários catalogadores-indexadores dos repositórios institucionais

Termo de Consentimento Livre e Esclarecido (TCLE)

1. Você está sendo convidado(a) para participar da pesquisa “Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação”.
2. Você foi selecionado(a) devido à atuação profissional como gestor(a) ou catalogador(a)/indexador(a) do repositório institucional da universidade pública do Estado de São Paulo a qual você está vinculado(a). Sua participação não é obrigatória.
3. O objetivo geral da pesquisa é contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento, com os seguintes objetivos específicos: a) realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias; b) investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como *metodologia qualitativa* com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP, com coleta de documentação; c) verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária por meio da aplicação da avaliação da indexação de assuntos como *metodologia quantitativa*, delineada em: avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*, em 9 bibliotecas da UNICAMP, divididas em cinco áreas do conhecimento: Área de Artes e Humanidades: Faculdade de Educação (FE); Instituto de Filosofia e Ciências Humanas (IFCH); Área de Biomédicas: Faculdade de Ciências Médicas (FCM); Faculdade de Educação Física (FEF); Área de Ciências Aplicadas: Faculdade de Ciências Aplicadas (FCA); Área de Exatas: Instituto de Física Gleb Wataghin (IFGW); Instituto de Matemática, Estatística e Computação Científica (IMECC); Área de Tecnológicas: Biblioteca da Área de Engenharia e Arquitetura (BAE); Faculdade de Engenharia de Alimentos (FEA); e d) sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais, visando propor um modelo pertinente de avaliação da indexação de assuntos em repositórios institucionais.
4. Sua participação nesta pesquisa consistirá em entrevista semiestruturada sobre o repositório institucional de sua universidade.
5. Os riscos relacionados com sua participação são mínimos, tanto nos aspectos físicos, psíquicos, quanto morais, ou seja: constrangimento ao realizar a atividade da pesquisa, desgaste no raciocínio, redução da liberdade, alteração no comportamento social, extravio de informações ou algum tipo de desrespeito a sua privacidade. Os benefícios na realização da pesquisa estão fundamentados nos objetivos do projeto.

6. Conforme os riscos já citados, a qualquer momento você pode desistir de participar deste estudo sem sofrer qualquer prejuízo. Sua recusa não trará nenhum dano em sua relação com a pesquisadora ou com a instituição.

7. Você poderá solicitar informações da pesquisa a qualquer momento por meio do endereço eletrônico ou contato telefônico das pesquisadoras descritas neste termo.

8. Garantimos todo e qualquer tipo de esclarecimentos, antes e durante o curso da pesquisa, a respeito dos procedimentos quando for necessário.

9. Não haverá despesas ou compensação financeira em qualquer fase do estudo para o(a) entrevistado(a).

10. Garantimos que as informações obtidas por meio dessa pesquisa serão confidenciais e asseguramos o sigilo sobre sua participação.

11. Você receberá uma cópia deste termo onde consta o telefone e o endereço da pesquisadora principal, podendo tirar suas dúvidas sobre o projeto e sua participação, agora ou a qualquer momento.

Declaro que entendi os objetivos, riscos e benefícios de minha participação na pesquisa e concordo em participar. A pesquisadora me informou que o projeto foi aprovado pelo Comitê de Ética e Pesquisa em Seres Humanos da FFC que funciona na UNESP, Campus de Marília. Endereço eletrônico: sta@marilia.unesp.br.

_____, ____ de _____ de _____.

Participante da Pesquisa

Apêndice D. Modelo de Autorização para coleta de dados nas Bibliotecas participantes

Campinas, 22 de outubro de 2018.

Prezado(a) Diretor(a),

Sou Roberta Cristina Dal'Evedove Tartarotti, aluna regularmente matriculada em nível doutorado, junto ao Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP), sob a orientação da Profa. Dra. Mariângela Spotti Lopes Fujita. Estamos desenvolvendo a pesquisa intitulada "Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação". O objetivo geral da pesquisa é contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento, com os seguintes objetivos específicos: a) realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias; b) investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como *metodologia qualitativa* com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP, com coleta de documentação; c) verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária por meio da aplicação da avaliação da indexação de assuntos como *metodologia quantitativa*, delineada em: avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*, em 9 bibliotecas da UNICAMP, divididas em cinco áreas do conhecimento: Área de Artes e Humanidades: Faculdade de Educação (FE); Instituto de Filosofia e Ciências Humanas (IFCH); Área de Biomédicas: Faculdade de Ciências Médicas (FCM); Faculdade de Educação Física (FEF); Área de Ciências Aplicadas: Faculdade de Ciências Aplicadas (FCA); Área de Exatas: Instituto de Física Gleb Wataghin (IFGW); Instituto de Matemática, Estatística e Computação Científica (IMECC) e Área de Tecnológicas: Biblioteca da Área de Engenharia e Arquitetura (BAE); Faculdade de Engenharia de Alimentos (FEA); e d) sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais, visando propor um modelo pertinente de avaliação da indexação de assuntos em repositórios institucionais. Venho por meio desta solicitar a autorização de V.Sa. para proceder aplicação de instrumentos de coleta de dados para a referida pesquisa junto aos usuários da Biblioteca nas categorias: 1 discente de graduação, 1 discente de pós-graduação (mestrado), 1 discente de pós-graduação (doutorado) e 1 docente, em datas e horários a serem acordados com tais participantes e dentro das dependências da Biblioteca. Sendo o que temos a apresentar, solicitamos vossa autorização por escrito e informamos que a mesma será anexada ao projeto de pesquisa a ser submetido à avaliação do Comitê de Ética em Pesquisa da FFC/UNESP.

De acordo.

Diretor(a) da Biblioteca

Apêndice E. Termo de Consentimento Livre e Esclarecido (TCLE) - Entrevista semiestruturada de busca por assuntos em repositório institucional pela perspectiva do usuário

Termo de Consentimento Livre e Esclarecido (TCLE)

1. Você está sendo convidado(a) para participar da pesquisa “Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação”.

2. Você foi selecionado(a) devido à categoria de usuário que possui na Universidade Estadual de Campinas (UNICAMP) a qual você está vinculado. Sua participação não é obrigatória.

3. O objetivo geral da pesquisa é contribuir para a implementação/melhorias de diretrizes de indexação de assuntos em repositórios institucionais no contexto de bibliotecas universitárias, possibilitando a adequada representação e recuperação da informação por assuntos e, de forma mais ampla, com elementos teórico-metodológicos em torno da avaliação da indexação de assuntos no âmbito da Organização do Conhecimento, com os seguintes objetivos específicos: a) realizar estudo teórico sobre os repositórios institucionais como sistemas de recuperação da informação e do processo de indexação de assuntos no âmbito da Organização do Conhecimento em perspectiva dicotômica: do tratamento/da representação temática e da recuperação por assuntos no contexto de bibliotecas universitárias; b) investigar o tratamento/a representação temática da informação documental em repositórios institucionais no contexto de bibliotecas universitárias, por meio da aplicação de entrevista de diagnóstico organizacional como *metodologia qualitativa* com os gestores dos repositórios institucionais e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP, com coleta de documentação; c) verificar a recuperação documental por assuntos de forma comparada entre linguagem natural e linguagem controlada em repositório institucional no contexto de biblioteca universitária por meio da aplicação da avaliação da indexação de assuntos como *metodologia quantitativa*, delineada em: avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *sistema - Avaliação intrínseca quantitativa mediante a interconsistência* e avaliação da indexação de assuntos em repositórios institucionais pela perspectiva do *usuário - Avaliação extrínseca mediante a recuperação*, em 9 bibliotecas da UNICAMP, divididas em cinco áreas do conhecimento: Área de Artes e Humanidades: Faculdade de Educação (FE); Instituto de Filosofia e Ciências Humanas (IFCH); Área de Biomédicas: Faculdade de Ciências Médicas (FCM); Faculdade de Educação Física (FEF); Área de Ciências Aplicadas: Faculdade de Ciências Aplicadas (FCA); Área de Exatas: Instituto de Física Gleb Wataghin (IFGW); Instituto de Matemática, Estatística e Computação Científica (IMECC); Área de Tecnológicas: Biblioteca da Área de Engenharia e Arquitetura (BAE); Faculdade de Engenharia de Alimentos (FEA); e d) sistematizar a análise comparada e delinear recomendações para a indexação de assuntos em repositórios institucionais, visando propor um modelo pertinente de avaliação da indexação de assuntos em repositórios institucionais.

4. Sua participação nesta pesquisa consistirá em realizar busca por assuntos no repositório institucional da Universidade Estadual de Campinas (UNICAMP).

5. Os riscos relacionados com sua participação são mínimos, tanto nos aspectos físicos, psíquicos, quanto morais, ou seja: constrangimento ao responder o questionário da pesquisa, desgaste no raciocínio, redução da liberdade, alteração no comportamento social, extravio de informações ou algum tipo de desrespeito a sua privacidade. Os benefícios na realização da pesquisa estão fundamentados nos objetivos do projeto.

6. Conforme os riscos já citados, a qualquer momento você pode desistir de participar deste estudo sem sofrer qualquer prejuízo.
7. Sua recusa não trará nenhum dano em sua relação com a pesquisadora ou com a instituição.
8. Você poderá solicitar informações da pesquisa a qualquer momento por meio do endereço eletrônico ou contato telefônico das pesquisadoras descritas neste termo.
9. Garantimos todo e qualquer tipo de esclarecimentos, antes e durante o curso da pesquisa, a respeito dos procedimentos quando for necessário.
10. Não haverá despesas ou compensação financeira em qualquer fase do estudo para o(a) entrevistado(a).
11. Garantimos que as informações obtidas por meio dessa pesquisa serão confidenciais e asseguramos o sigilo sobre sua participação.
12. Você receberá uma cópia deste termo onde consta o telefone e o endereço da pesquisadora principal, podendo tirar suas dúvidas sobre o projeto e sua participação, agora ou a qualquer momento.

Declaro que entendi os objetivos, riscos e benefícios de minha participação na pesquisa e concordo em participar. A pesquisadora me informou que o projeto foi aprovado pelo Comitê de Ética e Pesquisa em Seres Humanos da FFC que funciona na UNESP, Campus de Marília. Endereço eletrônico: sta@marilia.unesp.br.

_____, ____ de _____ de _____.

Participante da pesquisa

Apêndice F. Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem natural

Escreva, na tabela abaixo, as principais palavras-chave do **tema de sua pesquisa** em português e sua versão em inglês.

Realize a busca de cada uma das palavras-chave no Repositório Institucional da UNICAMP e anote o número de recursos informacionais relevantes para sua pesquisa.

Palavras-chave em português		N. de recursos informacionais relevantes para sua pesquisa	Versão das palavras-chave em inglês		N. de recursos informacionais relevantes para sua pesquisa
1			1		
2			2		
3			3		
4			4		
5			5		

Apêndice G. Roteiro de busca por assuntos no Repositório Institucional da UNICAMP - Linguagem controlada

Com o auxílio da pesquisadora, pesquise cada uma das palavras-chave da etapa anterior no vocabulário controlado de sua Área. Caso sejam diferentes, anote a versão dessas palavras-chave em português e em inglês na tabela abaixo, de acordo com o número correspondente das palavra-chave em português e em inglês da tabela anterior. Realize a busca de cada uma das palavras-chave no Repositório Institucional da UNICAMP e anote o número de recursos informacionais relevantes para sua pesquisa.

Palavras-chave em português		N. de recursos informacionais relevantes para sua pesquisa	Versão das palavras-chave em inglês		N. de recursos informacionais relevantes para sua pesquisa
1			1		
2			2		
3			3		
4			4		
5			5		

Apêndice H. Tempo de duração das entrevistas de busca por assuntos no Repositório Institucional da UNICAMP

Área de Conhecimento	Biblioteca participante	Categoria	Tempo de duração da definição da Linguagem natural	Tempo de duração das buscas por assuntos - Linguagem natural	Tempo de duração da conversão da linguagem natural em linguagem controlada	Tempo de duração das buscas por assuntos - Linguagem controlada
Área de Artes e Humanidades	Faculdade de Educação (FE)	G-FE	5 minutos	53 minutos	11 minutos	22 minutos
		PGM-FE	7 minutos	43 minutos	15 minutos	20 minutos
		PGD-FE	3 minutos	85 minutos	27 minutos	18 minutos
		D-FE	5 minutos	82 minutos	12 minutos	48 minutos
	Instituto de Filosofia e Ciências Humanas (IFCH)	G-IFCH	2 minutos	41 minutos	20 minutos	29 minutos
		PGM-IFCH	2 minutos	24 minutos	19 minutos	15 minutos
		PGD-IFCH	2 minutos	48 minutos	13 minutos	26 minutos
Área de Biomédicas	Faculdade de Ciências Médicas (FCM)	G-FCM	8 minutos	50 minutos	11 minutos	24 minutos
		PGM-FCM	6 minutos	32 minutos	8 minutos	25 minutos
		PGD-FCM	2 minutos	39 minutos	7 minutos	13 minutos
		D-FCM	1 minuto	33 minutos	13 minutos	25 minutos
	Faculdade de Educação Física (FEF)	G-FEF	3 minutos	50 minutos	4 minutos	24 minutos
		PGM-FEF	4 minutos	31 minutos	13 minutos	18 minutos
		PGD-FEF	8 minutos	48 minutos	12 minutos	18 minutos
		D-FEF	8 minutos	48 minutos	12 minutos	18 minutos
Área de Ciências Aplicadas	Faculdade de Ciências Aplicadas (FCA)	G-FCA	3 minutos	41 minutos	9 minutos	29 minutos
		PGM-FCA	2 minutos	57 minutos	19 minutos	23 minutos
		PGD-FCA	4 minutos	53 minutos	16 minutos	14 minutos
		D-FCA	2 minutos	41 minutos	8 minutos	32 minutos
Área de Exatas	Instituto de Física Gleb Wataghin (IFGW)	G-IFGW	7 minutos	67 minutos	14 minutos	25 minutos
		PGM-IFGW	4 minutos	26 minutos	7 minutos	10 minutos

		PGD-IFGW	2 minutos	22 minutos	16 minutos	12 minutos
		D-IFGW	3 minutos	35 minutos	8 minutos	37 minutos
	Instituto de Matemática, Estatística e Computação Científica (IMECC)	G-IMECC	5 minutos	54 minutos	23 minutos	35 minutos
		PGM-IMECC	6 minutos	36 minutos	9 minutos	19 minutos
		PGD-IMECC	4 minutos	48 minutos	16 minutos	15 minutos
		D-IMECC	4 minutos	48 minutos	16 minutos	18 minutos
Área de Tecnológicas	Biblioteca da Área de Engenharia e Arquitetura (BAE)	G-BAE	3 minutos	48 minutos	7 minutos	17 minutos
		PGM-BAE	4 minutos	42 minutos	11 minutos	13 minutos
		PGD-BAE	1 minuto	39 minutos	14 minutos	29 minutos
		D-BAE	4 minutos	51 minutos	12 minutos	34 minutos
	Faculdade de Engenharia de Alimentos (FEA)	G-FEA	3 minutos	53 minutos	8 minutos	35 minutos
		PGM-FEA	2 minutos	39 minutos	16 minutos	19 minutos
		PGD-FEA	1 minuto	42 minutos	12 minutos	29 minutos
		D-FEA	3 minutos	58 minutos	9 minutos	28 minutos

Fonte: Elaboração própria.

Apêndice I. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Artes e Humanidades - Biblioteca da FE

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-FE	Inclusão	3.264	-	6	-	Escolas	8784	-	0	-
	Escola	8.798	-	0	-	Diferença (Filosofia)	19534	23	6	3
	Diferença	13.772	-	5	-	Inclusão educacional	6173	28	11	4
	Educação especial	33.432	271	11	13	Schools	16739	-	1	-
	Inclusão escolar	6.315	87	11	12	Difference (Philosophy)	64393	20	4	10
	Inclusion	9.878	-	3	-	Mainstreaming (Education)	13800	3	3	1
	School	16.745	-	1	-					
	Difference	64.376	-	0	-					
	Special education	26089	179	13	20					
School inclusion	24314	30	13	9						
PGM-FE	Ensino superior	25624	1337	11	6	Cooperação universitária	6204	5	8	1
	Cooperação	995	-	3	-	Educação superior	28638	325	10	11
	Mercantilização	97	-	3	-	Privatização na educação	71873	8	3	1
	Privatização	353	-	4	-	University cooperation	42452	9	7	0
	Ensino superior privado	26343	26	8	7	Privatization in education	16273	39	8	3
	Higher education	46061	1419	6	9					
	Cooperation	5447	-	8	-					
	Commodification	26	-	1	-					
	Privatization	3889	-	4	-					
Private higher education	47654	19	5	4						
PGD-FE	Estratégias de aprendizagem	142953	95	36	40	Autodomínio	5	-	0	-
	Autoregulação da aprendizagem	90836	18	25	15	Auto-domínio	8609	-	0	-

	Auto-regulação da aprendizagem	91710	16	30	15	Estudantes do ensino médio	100830	54	2	3
	Intervenção	3052	-	2	-	Self-control	49477	-	3	-
	Alunos	3778	-	1	-	High school students	60397	182	0	1
	Professores	10745	-	1	-					
	Learning strategies	17997	134	30	40					
	Self-regulation learning	23594	34	33	26					
	Intervention	4727	-	0	-					
	Students	9180	-	0	-					
	Teachers	3059	-	1	-					
D-FE	Formação inicial de professores	142843	64	14	18	Professores - Formação	18551	560	18	24
	Práticas de leitura e de escrita	142842	5	35	1	Educação continuada	11263	231	23	23
	Letramento acadêmico	2093	12	24	12	Educação permanente	12521	235	19	28
	Práticas discursivas e formação	124805	0	8	0	Autoria	735	-	5	-
	Autoria e formação de professores	142842	0	6	0	Training of teachers	9075	99	18	19
	Teacher development	51562	39	19	10	Continuing education	30937	328	24	28
	Reading and writing practices	1797	17	24	10	Authorship	281	-	5	-
	Academic literacy	11761	12	31	12					
	Discursive practices and formation	5522	0	7	0					
	Authorship and teacher development	44	0	13	0					

Fonte: Elaboração própria.

Apêndice J. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Artes e Humanidades - Biblioteca do IFCH

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-IFCH	Violência	1334	-	1	-	Denúncia contra policiais	6282	0	6	0
	Racismo	214	-	2	-	Violência contra os adolescentes	59943	0	1	0
	Segurança pública	12102	95	2	6	Violência policial	15739	18	6	3
	Polícia	398	-	5	-	Má-conduta policial	38503	0	2	0
	Justiça criminal	1308	20	6	2	Police - Complaints against	16348	1	2	0
	Violence	835	-	1	-	Teenagers - Violence against	15687	0	2	0
	Racism	116	-	1	-	Police brutality	782	2	6	0
	Public security	21870	23	1	1	Police misconduct	689	0	5	0
	Police	680	-	4	-					
Criminal justice	996	23	6	2						
PGM-IFCH	Arquitetura brasileira	22893	19	5	1	Arquitetura - História	12665	7	7	1
	Arte sacra	9850	11	6	6	Arte - Brasil - História	40848	4	8	1
	Arte paulista	16371	1	5	0	Iconografia	131	-	5	0
	Iconografia paulista	7727	1	10	1	Dutra, Miguel Arcanjo Benicodé, 1812-1875	3043	0	0	0
	Miguel Dutra	2986	1	0	1					
	Brazilian architecture	29453	0	6	2					
	Art sacred	9927	0	6	6					
	Paulist art	9817	0	2	0					
	Paulist iconography	96	0	10	0					
	Miguel Dutra	2986	1	0	1					
PGD-IFCH	Consumo	4765	-	12	-	Consumo alimentar	6398	204	16	10
	Juventude	555	-	1	-	Adulto jovem	4788	61	3	0

	Meio ambiente	27398	2697	11	10	Food consumption	17475	363	7	10
	China*	2995	-	8	-	Young adult	12707	3	0	0
	BRICS*	33	-	4	-					
	Consumption	6814	-	6	-					
	Youth	935	-	0	-					
	Environment	17075	-	0	-					
	China*	2995	-	8	-					
	BRICS*	33	-	4	-					
D-IFCH	Pensamento social	17342	123	23	21	Estudos de gênero	143074	272	6	15
	Gênero	5720	-	9	-	População rural	10875	111	9	9
	Rural*	3876	-	13	-	Gender studies	89761	72	7	10
	Classe	16410	-	10	-	Rural population	20383	119	8	6
	Teoria social	21850	186	15	14					
	Social thought	19387	52	15	18					
	Gender	4775	-	14	-					
	Rural*	3876	-	13	-					
	Class	16410	-	9	-					
	Social Theory	31144	144	4	12					

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Apêndice K. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Biomédicas - Biblioteca da FCM

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-FCM	Síndrome de Turner	142969	58	13	16	Síndrome de Turner - Genética	142969	0	11	0
	Cardiopatas congênicas	802	48	5	7	Cardiopatas congênicas - Complicações	2326	0	3	0
	Valva aórtica bicúspide	225	2	7	2	Valva aórtica - Anormalidades	864	0	11	0
	Coarctação de aorta	142969	1	3	0	Coarctação aórtica	87	2	6	0
	Dismorfismos	29	-	7	-	Caracteres sexuais	1616	28	1	4
	Turner Syndrome	7392	103	18	19	Turner Syndrome - Genetics	16370	3	14	0
	Congenital cardiopathy	1434	6	2	1	Congenital heart defects - Complications	15456	0	6	0
	Bicuspid aortic valve	2195	6	13	5	Aortic valve - Abnormalities	6970	1	11	0
	Aortic coarctation	708	16	13	4	Sex characteristics	32386	33	0	0
Dysmorphisms	137	-	9	-						
PGM-FCM	Gestão	3893	-	5	-	Public administration	25038	154	4	3
	Administração	3565	-	5	-	Archives	2731	-	1	-
	Ensino médico	11906	99	7	7	Higher education - Administration	49143	0	1	0
	FCM-UNICAMP*	142968	4341	1	1	Administração pública	12328	309	7	3
	Arquivo	1222	-	4	-	Ensino superior - Administração	27719	0	7	0
	Management	13649	-	2	-	Educação médica	26619	175	6	6
	Administration	6733	-	5	-	Arquivos	12222	-	3	-
	Medical education	25896	221	5	5					
	FCM-UNICAMP*	14968	4341	-	-					
Archive	2731	-	3	-						

PGD-FCM	Saúde do trabalhador	101688	275	3	1	Pneumoconiose - Prevenção e controle	127586	0	9	0
	Silicose	28	-	16	-	Fadiga muscular	3415	47	0	0
	Pneumoconiose	37	-	13	-	Pneumoconiosis - Prevention and control	49166	0	12	0
	Mineração	572	-	1	-	Muscle fatigue	7578	132	0	0
	Fadiga	658	-	2	-					
	Occupational health	19818	363	2	1					
	Silicosis	38	-	13	-					
	Pneumoconiosis	39	-	10	-					
	Mining	6386	-	0	-					
Fatigue	1388	-	1	-						
D-FCM	HIV*	1752	-	0	-	Terapia antirretroviral de alta atividade	142984	0	0	0
	Criança	5837	-	0	-	Highly active antiretrovilar therapy	67818	0	0	0
	Adolescente	2498	-	0	-					
	Terapia antirretroviral	2724	46	0	0					
	Vacinas, resposta	8902	0	3	0					
	HIV*	1752	-	0	-					
	Child	4728	-	0	-					
	Adolescent	3590	-	0	-					
	Antiretroviral therapy	8443	210	0	0					
Vaccine, response	30708	14	1	0						

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Apêndice L. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Biomédicas - Biblioteca da FEF

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-FEF	Epilepsia	742	-	5	-	Epilepsia do lobo temporal	101327	71	1	0
	Comorbidades	484	-	2	-	Comorbidade	484	-	2	-
	Atividade física	28343	1211	0	0	Atividade motora	14603	160	0	0
	Síndrome metabólica	3194	182	3	3	Comorbidity	924	-	5	-
	ELT	263	-	0	-	Motor activity	40132	241	0	0
	Epilepsy	1137	-	2	-	Temporal lobe epilepsy	7376	398	2	2
	Comorbities	15	-	7	-					
	Physical activity	54597	1470	0	0					
	Metabolic syndrome	13044	606	2	3					
TLE	432	-	0	-						
PGM-FEF	Esporte adaptado na França	70832	0	4	0	Esporte para pessoas com deficiência física	117952	1	8	1
	Paralimpíadas	5	-	1	-	Pessoas com deficiência física - Reabilitação	111321	0	8	0
	Educação física	25683	1865	1	2	Paralympic games	1314	23	8	5
	Reabilitação	1245	--	0	-	Sports for people with disabilities	10569	6	5	0
	Atividade motora	14604	160	5	4	People with disabilities - Rehabilitation	8927	0	3	0
	Handsport in France	6592	0	0	0					
	Paralympics	54	-	6	-					
	Physical education	37135	901	1	2					
	Rehabilitation	1739	-	0	-					
Motor activity	40133	241	6	2						

PGD-FEF	Circo	60	-	7	-	Circos	60	-	4	-
	Rua	9476	-	2	-	Artistas circenses	886	9	4	2
	Ensino/aprendizagem	9071	802	3	3	Arte de rua	142993	2	6	1
	Escola de circo	142993	3	6	1	Circus performers	48555	2	3	0
	Circo de rua	142993	0	6	0	Street art	11424	1	4	1
	Circus	48	-	3	-					
	Street	1923	-	6	-					
	Teaching/learning	8878	322	-	-					
	Circus school	16789	4	3	1					
	Street's circus	124867	0	5	0					
D-FEF	Natureza e educação	125373	7	10	7	Esportes - História	10	1	1	1
	Vida ao ar livre	64361	12	5	7	Sports - History	9	1	1	1
	História do esporte	101027	15	0	0	Nature	32394	-	1	-
	Esporte e natureza	124934	1	3	0	Natureza	5935	-	7	-
	Corpo e natureza	124933	2	7	2	Corpo humano (Filosofia)	16309	0	0	0
	Nature and education	42383	3	4	1	Human body (Philosophy)	35121	0	0	0
	Outdoor life	16419	3	3	2					
	History of sport	16339	8	1	1					
	Sport and nature	33997	3	3	2					
	Body and nature	40566	25	3	1					

Fonte: Elaboração própria.

Apêndice M. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Ciências Aplicadas - Biblioteca da FCA

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-FCA	Tecnologia social	25954	26	8	3	Tecnologia culturalmente apropriada	13292	0	4	4
	Redes	7533	-	3	-	Redes sociais	13645	299	7	5
	Reaplicabilidade	1	-	0	-	Renda	2235	-	6	-
	Geração de renda	143034	89	6	5	Culturally appropriate technology	40067	0	1	2
	Reciclagem	517	-	3	-	Social networking	25231	414	2	4
	Social technology	34843	22	5	5					
	Network	10763	-	5	-					
	Reapplicability	36	-	0	-					
	Income	3341	-	5	-					
Recycling	1456	-	2	-						
PGM-FCA	Lipase*	869	-	5	-	Lipases			5	-
	Imobilização	415	-	1	-	Enzyme immobilization			6	3
	Burkholderia*	99	-	2	-	Imobilização enzimática			3	0
	Adsorção	953	-	0	-					
	Biolubrificante	2	-	0	-					
	Lipase*	869	-	5	-					
	Immobilization	2524	-	4	-					
	Burkholderira*	99	-	2	-					
	Adsorption	3221	-	0	-					
Biohubricant	0	-	0	-						

PGD-FCA	Inovação na justiça	39765	0	5	0	Tribunais - Administração	3637	0	2	0
	Inovação nas cortes	31600	0	2	0	Administração judiciária	3589	0	3	0
	Desempenho judicial	7611	1	3	0	Court administration	7196	0	3	0
	Law tech	11666	0	0	0	Administration of justice	7130	1	4	1
	Legal tech	8672	0	1	0					
	Law innovation	9277	0	6	0					
	Court innovation	3850	0	4	0					
	Justice innovation	3871	31	4	0					
	Judicial performance	4	-	6	1					
D-FCA	Secagem	1157	-	0	-	Moagem de cana-de-açúcar	143034	0	1	0
	Bambu	102	-	0	-	Bioetanol	268	-	3	-
	Bagaço	328	-	8	-	Etanol celulósico	1590	18	3	1
	Cana-de-açúcar	143034	-	3	-	Secagem a ar	135036	0	0	0
	Etanol	1567	-	6	-	Secagem solar	3321	1	0	0
	Drying	12810	-	0	-	Sugar cane	4510	1191	5	-
	Bamboo	247	-	0	-	Sugar cane milling	6834	23	11	6
	Bagasse	725	-	5	-	Bioethanol	415	-	6	-
	Sugarcane	1757	-	8	-	Cellulosic ethanol	8759	78	4	10
	Ethanol	7270	-	8	-	Air drying	19389	981	0	0
						Solar drying	14548	10	0	0

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Apêndice N. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Exatas - Biblioteca do IFGW

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-IFGW	Isolantes topológicos	255	1	1	1	Microscopia eletrônica	4701	1417	4	5
	Técnica ARPES	16653	0	4	0	Microscopia eletrônica de varredura	142969	922	4	4
	Clivador de amostras	142969	0	0	0	Microscopia eletrônica de transmissão	142969	293	4	3
	Propriedades eletrônicas	10892	75	4	5	Electron microscopy	24637	5169	6	9
	Técnicas de caracterização	142969	100	5	1	Scanning electron microscopy	28200	3068	6	4
	Topological insulators	4554	14	6	2	Transmission electron microscopy	28212	1683	5	6
	ARPES technique	29940	2	6	1	Characterization (Materials science)	75576	2	3	0
	Sample cleaver	35828	0	1	0	Caracterização (Ciência dos materiais)	78849	0	2	0
	Electronic properties	40017	656	6	7					
	Characterization techniques	48310	188	5	2					
PGM-IFGW	Bits quânticos	2531	1	6	1	Qubits	152	-	17	-
	Nanorresonadores	0	-	0	0	Superconducting resonators	7778	1	2	1
	Circuitos supercondutores	1368	0	3	0	Josephson junction devices	11261	4	12	3
	Computação quântica	8950	14	2	5	Superconducting devices - Josephson junction	11799	0	9	0
	Metamateriais	20	-	4	-	Quantum computers	24240	182	10	11
	Qubit	137	-	16	-	Ressonadores supercondutores	139	0	3	0
	Nanoresonators	3	-	1	-	Computadores quânticos	3460	10	2	3
	Superconducting circuits	3885	6	6	6	Dispositivos de junções Josephson	143048	0	5	0
	Quantum computing	5853	168	7	16	Dispositivos supercondutores - Junções Josephson	3230	0	5	0

	Metamaterials	19	-	4	-					
PGD-IFGW	Férmions pesados	1619	14	12	14	Heavy fermion superconductors	4919	46	10	11
	Fase nemática	9266	4	1	1	Fermi liquids	14012	153	31	34
	Supercondutividade	68	-	14	-	Férmions pesados supercondutores	1713	2	15	17
	Pares de Cooper	142967	2	1	1	Líquidos de fermi	142967	8	2	6
	Teoria de Landau	142967	4	5	5					
	Heavy fermions	4548	190	17	2					
	Nematic phase	24139	20	1	1					
	Superconductivity	913	-	16	-					
	Copper pairs	14963	3	2	2					
	Landau theory	17866	39	5	7					
D-IFGW	QCD*	440	-	14	-	Cromodinâmica quântica	628	17	8	8
	Fator de forma	142969	34	0	4	Fator de forma (Física nuclear)	142969	0	0	0
	Supersimetria	4	-	2	-	Partículas estranhas	3149	5	2	2
	Neutrino*	511	-	28	-	Astrofísica de neutrinos	142969	5	12	4
	Estranheza	96	-	3	-					
	QCD*	440	-	14	-	Quantum chromodynamics	5854	149	8	16
	Form factor	56026	226	3	6	Form factor (Nuclear physics)	67633	0	0	0
	Supersymmetry	114	-	23	-	Strange particles	10813	53	26	18
	Neutrino*	511	-	28	-	Neutrino astrophysics	941	29	14	2
Strangeness	552	-	17	-						

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Apêndice O. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Exatas - Biblioteca do IMECC

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-IMECC	Problema de corte de estoque	143034	22	7	7	Processo decisório por critério múltiplo	58310	30	0	0
	Otimização multiobjetivo	3132	40	5	3	Programação inteira	2331	80	2	3
	Programação linear inteira	23453	68	3	4	Programação linear	23035	280	4	2
	Método das somas ponderadas	56486	0	0	0	Pesquisa operacional	42659	221	2	3
	Método do ϵ -restrito	101589	0	0	0	Multiple criteria decision making	40018	20	3	4
	Cutting stock problem	30734	40	9	8	Integer programming	20695	213	2	2
	Multicriteria optimization	15085	20	2	2	Linear programming	35500	644	3	2
	Integer linear programming	36035	166	4	2	Operations research	56600	501	1	2
	Weighted sum method	61146	3	1	0	Multiobjective optimization	15036	97	4	3
ϵ -constraint method	56950	0	0	0						
PGM-IMECC	Predição espacial	2911	3	3	0	Digital soil mapping	29005	11	0	3
	Abordagem bayesiana	6987	11	0	0	Bayesian field theory	34535	0	0	0
	Processo de Dirichlet	142959	0	0	0	Monte Carlo Method	67187	353	0	0
	MCMC*	132	-	1	-	Método de Monte Carlo	143048	109	0	0
	Classificação espacial	6041	3	2	0	Mapeamento digital de solos	143048	3	0	0
	Spatial prediction	20352	14	2	0	Teoria de decisão bayesiana	143048	0	0	0
	Bayesian approach	28195	153	1	0					
	MCMC*	132	-	1	-					
	Dirichlet process	49179	1	0	0					
Spatial classification	14471	3	2	0						
PGD-IMECC	Sistemas dinâmicos	30209	344	0	2	Dynamics	57334	1191	0	0
	Processo estocástico	26178	102	0	3	Stochastic processes	49368	342	1	1

	Transporte ótimo	11664	2	1	1	Sobolev spaces	17183	248	2	2
	Espaço de Sobolev	142969	17	3	2	Processos estocásticos	26208	102	5	4
	Medida de probabilidade	142969	5	2	1	Espaços de Sobolev	143041	17	4	3
	Dynamical system	57334	1191	0	4					
	Estochastic process	48948	0	0	0					
	Optimal transport	23358	17	3	2					
	Sobolev space	43646	0	2	2					
	Probability of measure	46718	24	3	3					
D-IMECC	Problema de corte de estoque	142984	22	40	22	Dimensionamento de lotes	143034	15	25	15
	Problema de dimensionamento de lotes	142984	9	25	9	Economic lot size	56587	15	19	14
	Geração de colunas	142984	14	11	13					
	Programação linear	23032	280	39	45					
	Programação inteira	2328	80	36	38					
	Cutting stock problem	30714	40	38	38					
	Lot sizing problem	60316	27	30	27					
	Column generation	23922	58	12	36					
	Linear programming	35495	644	31	32					
Integer programming	20691	213	24	29						

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português

Apêndice P. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Tecnológicas - Biblioteca da BAE

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-BAE	Acelerômetros	49	-	3	-	Análise dinâmica	36424	103	3	3
	Smartphones*	78	-	0	-	Avaliação dinâmica	21419	18	0	0
	Avaliação dinâmica	21419	18	1	0	Dynamic analysis	67254	270	1	3
	Análise modal	35716	101	7	6	Vibration measurement	41076	40	2	2
	Modos de vibração	143034	28	5	3					
	Accelerometers	144	-	5	-					
	Smartphones*	78	-	0	-					
	Dynamic evaluation	58670	36	0	0					
	Modal analysis	62050	227	4	5					
Vibration modes	14362	659	2	1						
PGM-BAE	Arquitetura residencial	2930	4	9	0	Arquitetura - Projetos e plantas	125034	4	10	2
	Arquitetura paulista	10122	3	11	3	Diagramas esquemáticos	898	42	0	0
	Decio Tozzi*	565	0	0	0	Architecture - Design and plans	30233	5	7	3
	Estratégias projetuais	7661	3	7	2	Schematic diagrams	6761	584	0	0
	Análise gráfica	34425	35	5	5					
	Residential architecture	4557	3	6	1					
	Achitecture of paulista	7630	0	6	0					
	Decio Tozzi*	565	0	0	0					
	Design strategies	31904	85	0	3					
Graphic analysis	62243	122	3	4						
	Gaseificação	70	-	12	-	Processo de Fisher-tropsch	143034	7	9	6
	Gás de síntese	143034	30	1	4	Síntese de Fisher-tropsch	143034	8	9	5

PGD-BAE	Fischer-tropsch*	2827	56	9	-	Catalisadores	846	-	1	-
	Termoquímica	84	-	2	-	Fischer-tropsch process	50483	9	12	3
	Catalisador	846	-	1	-	Fischer-tropsch synthesis	14536	13030	34	11
	Gaseification	4	-	13	-	Synthesis gas	19825	51	5	10
	Syngas	65	-	14	-	Catalysts	2636	-	1	-
	Fischer-tropsch*	2827	56	9	-					
	Thermochemical	459	-	3	-					
	Catalyst	2636	-	3	-					
D-BAE	Compósitos de madeira	143034	0	3	0	Madeira laminada	1650	13	6	5
	Madeira laminada colada	1695	13	6	5	Construções em madeira	84944	0	3	0
	Madeira laminada reforçada	2020	0	8	0	Compósitos laminados	756	12	0	0
	Propriedades mecânicas da madeira	91977	6	4	0	Reforço	548	-	3	-
	Madeira com barras de aço	143034	0	2	0	Wood laminates	4658	3	7	1
	Wood composite	22568	19	5	3	Wooden construction	15117	2	2	0
	Glued laminated timber	2343	8	9	5	Laminated composites	19882	27	3	0
	Reinforced glulam	4655	1	3	1	Reinforcement	4652	-	5	-
	Wood mechanical proprieties	32716	0	4	0					
Wood with steel bars	13983	0	2	0						

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Apêndice Q. Assuntos pesquisados e itens recuperados no Repositório Institucional da UNICAMP - Área de Tecnológicas - Biblioteca da FEA

Usuário	Assuntos pesquisados - Linguagem natural	Itens recuperados		Itens relevantes		Assuntos pesquisados - Linguagem controlada	Itens recuperados		Itens relevantes	
		Sem aspas	Com aspas	Sem aspas	Com aspas		Sem aspas	Com aspas	Sem aspas	Com aspas
G-FEA	Kefir*	22	-	5	-	Atividade antioxidante	14585	318	3	2
	Casca de manga	142994	2	2	1	Leite fermentado	6460	41	4	6
	Compostos bioativos	7739	192	1	3	Antioxidative activity	38501	1053	1	4
	Antioxidantes	1059	-	1	-	Fermented milk	4672	110	9	9
	Fermentação	846	-	0	-					
	Kefir*	22	-	5	-					
	Mango peel	1268	11	7	1					
	Bioactive compounds	15560	610	3	2					
	Antioxidants	3309	-	1	-					
Fermentation	2633	-	4	-						
PGM-FEA	Soberania alimentar	2743	6	8	4	Energia (Sustentabilidade)	933	0	18	0
	Segurança alimentar	5200	243	6	10	Mudanças climáticas - Aspectos sociais	22910	0	6	0
	Energia	68	-	13	-	Agricultura ecológica	4157	13	5	2
	Mudanças climáticas	9494	288	7	10	Emergy (Sustainability)	4998	5	38	5
	Agroecologia	75	-	6	-	Climate change - Social aspects	60016	0	7	0
	Food sovereignty	13550	2	3	1	Agricultural ecology	10709	14	4	4
	Food security	15377	197	8	7					
	Emergy	39	-	38	-					
	Climate change	41776	759	8	13					
Agroecology	89	-	9	-						
PGD-FEA	Biofilme	1150	-	8	-	Biofilmes	1150	-	7	-
	Bacillus Sporothermodurans*	1095	8	9	5	Dispersão	1794	-	0	-

	Formação de biofilme	143006	90	14	11	Microscopia eletrônica de transmissão	143006	294	2	0
	Dispersão de biofilme	143006	0	7	0	Microscopia eletrônica de varredura	143006	922	1	0
	Microscopia eletrônica	4705	1418	1	2	Biofilms	1150	-	5	-
	Biofilm	1150	-	12	-	Dispersion	8490	-	0	-
	Bacillus Sporothermodurans*	1095	8	9	5	Electron transmission microscopy	28216	90	1	0
	Biofilm formation	22202	339	11	12	Scanning electron microscopy	28203	3038	2	1
	Biofilm dispersal	9538	5	6	1					
	Microscopy electronic	24640	334	0	1					
D-FEA	Microencapsulação de óleo	143006	8	25	8	Microencapsulação	99	-	21	-
	Secagem por pulverização	52184	0	0	0	Hidrolisado protéico	1008	45	20	10
	Hidrólise de proteína	1715	5	14	4	Emulsões	203	-	13	-
	Emulsão	301	-	35	-	Microencapsulation	361	-	22	-
	Proteínas vegetais	5997	10	5	2	Protein hydrolysates	16399	143	22	25
	Microencapsulation of oil	7536	12	18	9	Emulsions	1235	-	15	-
	Spray drying	13559	344	27	34					
	Protein hydrolysis	16081	0	11	0					
	Oil emulsion	7948	111	18	16					
	Vegetable proteins	22966	57	5	8					

Fonte: Elaboração própria.

* Termo equivalente em inglês e em português.

Anexo A. Notações específicas para transcrição das entrevistas⁵²

itálico: vocalização do participante⁵³

(): questões ou comentários da pesquisadora

....: pausas curtas

...~~: pausas longas

(...): omissão de trecho não relevante na transcrição da entrevista

((RI)): riso do participante ou pesquisadora

((FR)): vocalização e riso ao mesmo tempo pelo participante ou pesquisadora

((RM)): tom de ironia do participante ou pesquisadora

“...”: paráfrase do participante ou pesquisadora

/ : auto interrupção de um pensamento pelo participante ou pesquisadora

(->->->): aceleração do ritmo da vocalização do participante ou pesquisadora

(~~~): desaceleração do ritmo da vocalização do participante ou pesquisadora

{...}: trecho do texto-base⁵⁴ vocalizado pelo participante

(<-): releitura a trecho do texto-base vocalizado pelo participante

(->): trecho do texto-base “saltado” (ignorado) pelo participante

[]: inclusão nas transcrições, de descrição de gestos significativos do participante ou de comentários analíticos da pesquisadora

MAIÚSCULO: termo em linguagem natural buscado no repositório institucional

NEGRITO MAIÚSCULO: termo em linguagem controlada localizado no vocabulário controlado

NEGRITO MAIÚSCULO SUBLINHADO: termo final em linguagem controlada buscado no repositório institucional

Sublinhado: trecho específico que demonstra o fenômeno estudado

⁵² Adaptadas de notações originais para transcrição de protocolos verbais de CAVALCANTI, M. C. *Interação leitor-texto*: aspectos de interação pragmática. Campinas: Editora da UNICAMP, 1989 e readaptadas por TARTAROTTI (2014, p. 269).

⁵³ Originalmente, utilizado o termo “sujeito”, sendo adaptado nesta pesquisa para “participante” (vide Nota de Rodapé 41).

⁵⁴ O texto-base refere-se às informações contidas nos registros do repositório institucional recuperados durante as buscas por assuntos e vocalizadas pelos participantes.

Anexo B. Resolução Nº. 6444 (USP), de 22 DE OUTUBRO DE 2012 (D.O.E. 23.10.2012)

Dispõe sobre diretrizes e procedimentos para promover e assegurar a coleta, tratamento e preservação da produção intelectual gerada nas Unidades USP e pelos Programas Conjuntos de Pós-Graduação, bem como sua disseminação e acessibilidade para a comunidade.

O Reitor da Universidade de São Paulo, usando de suas atribuições legais, tendo em vista o deliberado pelo Presidente da d. Comissão de Legislação e Recursos, “ad referendum” daquele Colegiado, e considerando a necessidade de:

- preservar a memória institucional;
- ampliar a visibilidade e acessibilidade da produção intelectual (científica, acadêmica, artística e técnica) da USP;
- potencializar o intercâmbio com outras instituições nacionais e internacionais;
- certificar o uso de indicadores confiáveis referentes à produção intelectual da USP;
- aperfeiçoar a gestão de investimentos em pesquisa, ensino e extensão nesta Instituição, baixa a seguinte.

RESOLUÇÃO:

Artigo 1º - A Biblioteca Digital da Produção Intelectual (doravante denominada BDPI) passa a ser o instrumento oficial incumbido de reunir a produção intelectual da USP, de modo a:

- I - aumentar a visibilidade, acessibilidade e difusão dos resultados da atividade acadêmica e de pesquisa da USP por meio da coleta, organização e preservação em longo prazo;
- II - facilitar a gestão e o acesso à informação sobre a produção intelectual da USP, por meio da oferta de indicadores confiáveis e validados;
- III - integrar-se a um conjunto de iniciativas nacionais e internacionais, por meio de padrões e protocolos de integração qualificados e normalizados.

Do Conselho Supervisor do SIBi

Artigo 2º - Fica o Conselho Supervisor do Sistema Integrado de Bibliotecas - SIBi incumbido de estabelecer e validar normas para coleta, tratamento e preservação da produção intelectual gerada na Universidade (atendendo às especificidades da produção impressa e digital), bem como definir os tipos de documentos para depósito, além das teses e dissertações defendidas nas Unidades USP.

Da constituição da memória documental

Artigo 3º - Para a formação e desenvolvimento da memória da produção intelectual da USP, os docentes, servidores técnicos e administrativos, alunos e pós-doutorandos deverão depositar na BDPI o conteúdo integral de produtos de sua autoria, à medida que forem publicados ou editados.

§ 1º - A inserção de conteúdos na BDPI poderá ser feita por auto-arquivamento (depósito feito diretamente pelo próprio autor do trabalho), pela equipe da biblioteca de sua Unidade funcional ou por importação de dados executada pela gerência da BDPI.

§ 2º - O depósito da produção intelectual deverá ser realizado de forma não exclusiva, mantendo os autores dos documentos todos os seus direitos.

§ 3º - Se de direito, o acesso aos documentos poderá ser aberto, embargado (por tempo limitado pelo contrato assinado pelo autor com a casa editorial), restrito para uso apenas pelos computadores da USP ou restrito completamente (neste caso, o arquivo digital depositado servirá apenas para gestão e governança da produção).

§ 4º - Quando produção intelectual não disponível em formato digital, os metadados deverão ser registrados na BDPI e um exemplar da produção deverá ser depositado na biblioteca de sua Unidade funcional.

Artigo 4º - As teses e dissertações seguem o padrão estabelecido pela Resolução CoPGr nº 6018, de 13.10.2011.

Artigo 5º - Recomenda-se a todos os membros da comunidade USP a publicação de seus resultados de pesquisa, preferencialmente, em fontes que se encontrem em livre acesso ou que façam constar em seus contratos de publicação a permissão para depósito na BDPI.

Das Bibliotecas do SIBi

Artigo 6º - Compete às Bibliotecas do SIBi, em relação à BDPI:

- I - efetuar o registro técnico de produção intelectual na BDPI, desde que solicitado por sua Unidade de vínculo ou por membros daquela comunidade;
- II - a edição, revisão, validação e disponibilização online da produção intelectual auto-arquivada pelos autores;
- III - a organização de ações periódicas de capacitação sobre procedimentos e esclarecimentos das funcionalidades existentes, dirigidas à comunidade USP;
- IV - o apoio aos autores USP na averiguação da situação de suas publicações perante entidades externas, a quem tenham eventualmente sido cedidos os direitos de autor;
- V - o fornecimento de dados, informações e estatísticas institucionais requeridas por suas Unidades de vínculo;
- VI - a garantia da atualização permanente dos registros da produção intelectual na BDPI, a partir de ações periódicas junto aos autores de sua Unidade de vínculo.

Da Coordenação pelo Departamento Técnico do SIBi

Artigo 7º - O Departamento Técnico do SIBi, em relação à BDPI, será responsável pela:

- I - gerência e atualização constante do sistema de gestão decorrente de evolução tecnológica;
- II - geração de dados e indicadores sobre a produção intelectual da USP para fins diversos, dentre eles o Anuário Estatístico ou outros que venham a ser requeridos pelos Órgãos da Universidade;
- III - garantia da disseminação de indicadores confiáveis e certificados sobre a produção intelectual gerada na Universidade;
- IV - preparação de diretrizes e mecanismos para garantir o controle e a preservação digital da produção intelectual gerada pela USP;
- V - formação das competências necessárias às equipes das Bibliotecas do SIBi, visando à plena realização das atividades relativas à BDPI;
- VI - criação de mecanismos de estímulo e ações de integração que possibilitem a interoperabilidade e racionalização de recursos com bancos de dados informacionais internos e externos à USP.

Artigo 8º - Para o pleno desenvolvimento das atividades da BDPI, deverá ser assegurada sua integração aos sistemas corporativos da Universidade.

Artigo 9º - Esta Resolução entrará em vigor na data de sua publicação, ficando revogadas a Resolução nº 4221, de 17.11.1995, e todas as disposições em contrário.

Reitoria da Universidade de São Paulo, 22 de outubro de 2012.

JOÃO GRANDINO RODAS

Reitor

Anexo C. Regulamento Interno do Repositório Institucional (UNESP)

Política de Gestão do Repositório Institucional da UNESP

Regulamento interno aprovado por unanimidade em reunião realizada em 08 de novembro de 2016.

Introdução

A missão do Repositório Institucional UNESP é promover a Universidade em âmbito nacional e internacional, por meio do acesso aberto, da disseminação, da preservação, da valorização e do reconhecimento de sua produção científica, acadêmica, artística, técnica e administrativa, bem como, prover à sociedade o acesso ao conhecimento resultante das ações realizadas pela Universidade.

Para esclarecimento, definimos os termos utilizados neste Regulamento da seguinte forma:

- I - documento: unidade constituída pela informação e seu suporte, podendo abarcar um ou mais arquivos, digitais ou não;
- II - registro: conjunto de dados (metadados) relacionados ao documento: título, criador, data de publicação, assunto, entre outros;
- III - item: conjunto formado pelo documento e seu registro.

CAPÍTULO I DOS RESPONSÁVEIS E DE SUAS ATRIBUIÇÕES

Artigo 1º - No âmbito do Repositório, compete ao Grupo Gestor do Repositório Institucional UNESP:

- I - criar normas para a gestão e o uso do Repositório Institucional UNESP;
- II - estabelecer políticas para o arquivamento e a preservação de itens;
- III - acompanhar o trabalho dos membros da Equipe Técnica do Repositório Institucional UNESP.

CAPÍTULO II DA ORGANIZAÇÃO

Artigo 2º - O Repositório Institucional UNESP está organizado em comunidades que representam as Unidades universitárias; os Câmpus experimentais, as Unidades complementares e a Reitoria, essas comunidades estão divididas em subcomunidades que representam as Unidades auxiliares; os Centros interdepartamentais, os Departamentos, os Programas de pós-graduação, e as Seções técnico-administrativas de cada unidade universitária.

Parágrafo único: Cada subcomunidade contém coleções para os diferentes tipos de documentos nela arquivados.

Artigo 3º - A criação e a alteração de comunidades e de subcomunidades ocorrerá em consonância com a criação ou alteração de denominação de Unidades universitárias, Unidades complementares; Unidades auxiliares, Centros interdepartamentais, departamentos, programas de pós-graduação, e seções técnicoadministrativas, de modo que o Repositório Institucional UNESP reflita a organização da Universidade.

Artigo 4º - A criação de coleções ocorrerá mediante a demanda pelo arquivamento de documentos dos tipos listados neste regulamento.

CAPÍTULO III DOS CRITÉRIOS PARA O ARQUIVAMENTO DE DOCUMENTOS

Artigo 5º - Para ser incluído no Repositório Institucional UNESP, o documento deverá atender aos seguintes critérios gerais:

- I - possuir entre seus criadores ao menos uma pessoa vinculada à Universidade;
- II - estar em formato digital;
- III - ser resultado de atividade científica, acadêmica, artística, técnica ou administrativa realizada nas dependências da Universidade ou em seu nome;
- IV - estar completo, pronto para publicação ou já ter sido publicado e com a cessão de direitos patrimoniais, quando for o caso;
- V - estar categorizado como:

- a) artigo publicado em periódico científico, em sua versão publicada ou aceita para a publicação;
- b) trabalho completo ou resumo publicado em anais de congresso, conferência, encontro, simpósio ou qualquer outra reunião de natureza técnicocientífica; c) livro ou capítulo de livro;
- d) tese de doutorado, tese de livre-docência, dissertação de mestrado ou trabalho de conclusão de curso de graduação ou de especialização;
- e) notação musical;
- f) imagem;
- g) material cartográfico;
- h) gravação de som ou gravação de vídeo;
- i) software;
- j) patente;
- k) outra categoria a ser estabelecida pelo Grupo Gestor do Repositório Institucional UNESP.

Parágrafo único - Além dos critérios gerais elencados nesse artigo, deverão também, ser cumpridos os critérios específicos para cada tipo de recurso informacional, estabelecidos nas políticas de arquivamento definidas pelo Grupo Gestor.

VI - não violar direitos autorais.

Artigo 6º - Documentos de tipos não listados no inciso V, do artigo 5º, deste Regulamento, poderão ser arquivados no Repositório Institucional UNESP, mediante encaminhamento de ofício para o Grupo Gestor do Repositório Institucional UNESP, juntamente com a lista de documentos a serem arquivados, que deverá ter sido previamente enviada ao gestor da comunidade a que pertence, que irá emitir o parecer.

Parágrafo único: O Grupo Gestor do Repositório Institucional UNESP emitirá o parecer final sobre o arquivamento dos documentos, somente após a emissão do parecer do gestor da comunidade.

Artigo 7º - Para o arquivamento no Repositório Institucional UNESP não serão impostas restrições quanto ao país de publicação, ao idioma ou a data de publicação do documento.

Artigo 8º - Na impossibilidade de arquivar um documento, seja devido às restrições contratuais acordadas entre o criador do documento e seu publicador ou, devido à necessidade de proteção de propriedade intelectual ou industrial, poder-se-á:

- I - incluir no Repositório Institucional UNESP, apenas o registro do documento e, se possível, o endereço para acesso ao documento na página de seu publicador;
- II - durante o período de embargo, poderá, como definido na política, ser arquivada versão com o conteúdo parcial para acesso público imediato, ficando o acesso integral restrito ao administrador.

Artigo 9º - Os formatos dos arquivos digitais aceitos no Repositório Institucional UNESP serão definidos em políticas específicas para o arquivamento de cada tipo de recurso informacional, devendo ser dada preferência aos padrões abertos (formatos não proprietários), de modo a facilitar as ações para a preservação digital.

Artigo 10º - O documento arquivado no Repositório Institucional UNESP, não poderá ser alterado por outro que apresente qualquer mudança em seu conteúdo.

Parágrafo único - Excepcionalmente, quando for necessária qualquer alteração no conteúdo de um documento arquivado, poderá ser acrescentado ao item, com a devida autorização dos responsáveis, um arquivo digital contendo uma errata, uma nova edição ou versão.

Artigo 11º - Os itens que não atenderem aos critérios elencados nesse documento e nas políticas específicas para arquivamento, serão removidos do Repositório Institucional UNESP.

CAPÍTULO IV DO ACESSO AOS DOCUMENTOS E AOS METADADOS

Artigo 12º - A acessibilidade aos metadados e documentos arquivados no Repositório Institucional UNESP será livre e gratuita a qualquer indivíduo, não sendo necessário qualquer cadastro ou pagamento, respeitando-se os embargados, quando for o caso, para acesso aos documentos.

CAPÍTULO V DAS FUNÇÕES DA COORDENAÇÃO EXECUTIVA

Artigo 13º - Cabe à Coordenação Executiva acompanhar as atividades da Equipe Técnica.

Artigo 14º - No âmbito do Repositório, compete à Equipe Técnica do Repositório Institucional UNESP a partir das políticas estabelecidas pelo Grupo Gestor:

- I - estabelecer procedimentos para o arquivamento e a gestão dos itens;
- II - prover orientações sobre o uso do Repositório Institucional UNESP e o arquivamento de itens;
- III - supervisionar o arquivamento de itens;
- IV - definir metadados para a representação e preservação dos documentos, bem como, estabelecer padrões para o preenchimento dos valores desses metadados, considerando a missão do Repositório Institucional UNESP e padrões internacionalmente utilizados;
- V - definir e realizar tarefas para a melhoria da qualidade dos metadados;
- VI - gerenciar as contas de usuário criadas, concedendo e/ou revogando suas permissões; VII - desenvolver ações visando à manutenção, à atualização e à customização dos componentes de hardware e de software utilizados no Repositório Institucional UNESP; VIII - realizar e apoiar as ações para a promoção do Repositório Institucional UNESP junto à comunidade da UNESP e fora dela.

Artigo 15º - Os casos omissos serão resolvidos pelo Grupo Gestor do Repositório Institucional UNESP e pela Equipe Técnica dentro de suas respectivas competências.

Grupo Gestor do Repositório Institucional UNESP.

Anexo D. Resolução GR-013/2015 (UNICAMP), de 06/07/2015

Reitor: José Tadeu Jorge

Dispõe sobre a criação do Repositório da Produção Científica e Intelectual da UNICAMP.

O Reitor da Universidade Estadual de Campinas, no uso de suas atribuições legais, considerando a necessidade de criação de um repositório institucional para armazenamento, em texto completo no formato digital, da produção científica e intelectual da Universidade, baixa a seguinte RESOLUÇÃO:

Artigo 1º - Fica criado o Repositório da Produção Científica e Intelectual da Universidade, qualificando-o como instrumento oficial incumbido de armazenar a produção científica e intelectual da Universidade dos docentes, pesquisadores, alunos de graduação e pós-graduação, e servidores técnicos administrativos vinculados à UNICAMP, de modo a:

- I - Aumentar a visibilidade, acessibilidade e difusão dos resultados das atividades acadêmicas e de pesquisa da UNICAMP, por meio da coleta, organização e preservação da produção científica em longo prazo;
- II - Facilitar a gestão e o acesso à informação sobre a produção científica e intelectual da UNICAMP, por meio da oferta de indicadores confiáveis e validados;
- III - Integrar-se a um conjunto de iniciativas nacionais e internacionais, por meio de padrões e protocolos de integração qualificados e normalizados.

Artigo 2º - Institui as bibliotecas do Sistema de Bibliotecas da UNICAMP (SBU) como depositária da produção científica e intelectual da Universidade, por meio do Repositório Institucional.

Artigo 3º - Designa membros dos órgãos abaixo para comporem o Grupo de Trabalho do Repositório da UNICAMP:

Coordenadoria do SBU
Pró-Reitoria de Pesquisa
Pró-Reitoria de Pós-Graduação
Pró-Reitoria de Graduação
DTRI/SBU
Biblioteca Coleções Especiais e Obras Raras
SIARQ
Bibliotecários representantes por áreas de conhecimento
COCEN

Artigo 4º - É de competência do Grupo de Trabalho:

- I- Estabelecer e validar normas para coleta, tratamento, disseminação e preservação da produção científica gerada na Universidade;
- II- Definir os tipos de documentos para depósito;
- III- Definir procedimentos para captação e disponibilização de documentos científicos, tecnológicos e culturais existentes no formato digital;
- IV- Definir a estrutura do sistema computacional para armazenamento dos conteúdos a serem disponibilizados no REPROCITEC-UNICAMP ou RPCTC-UNICAMP, bem como os aplicativos de gestão dos conteúdos e da interoperabilidade com outros sistemas informatizados;
- V- Estabelecer padrões e metodologia para digitalização de documentos que não exista no formato eletrônico;
- VI- Estabelecer políticas de preservação digital para o Repositório Institucional da UNICAMP.

Artigo 5º - As coleções retrospectivas em formato impresso, não existentes em formato digital, serão tratadas por processos de conversão para o formato digital, conforme procedimentos a serem estabelecidos pelo Grupo de Trabalho aqui designado.

Artigo 6º - O desenvolvimento do Repositório poderá contar com parcerias de outras áreas/órgãos nas definições de estrutura tecnológica e/ou no controle de captação da produção científica, tecnológica e cultural da Universidade.

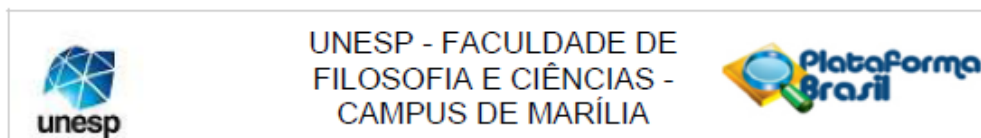
Artigo 7º - O Repositório da Produção Científica e Intelectual da UNICAMP deverá obedecer a Legislação vigente sobre Direitos Autorais, Lei 9.160 de 19/02/1998, bem como os embargos interpostos por editoras nas publicações de documentos científicos.

Artigo 8º - O Grupo de Trabalho do terá o prazo de 3 meses, a partir desta data, para estabelecer os procedimentos para promover e assegurar a coleta, tratamento, disseminação e preservação da produção científica, tecnológica e cultural gerada na UNICAMP

Artigo 9º - Esta Resolução entra em vigor na data da sua publicação, revogando a Portaria GR-085/2001, pela qual criou a Biblioteca Digital da UNICAMP, que passará a ter outras atribuições de acordo com a Resolução GR-014/2015.

Publicada no D.O.E. em 21/07/2015, p. 50.

Anexo E. Parecer Final do Comitê de Ética de Pesquisa (CEP) da Universidade Estadual Paulista (UNESP/Marília)



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: O PROCESSO DE INDEXAÇÃO DE ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS PELA ABORDAGEM DA RECUPERAÇÃO DA INFORMAÇÃO

Pesquisador: ROBERTA CRISTINA DAL EVEDOVE TARTAROTTI

Área Temática:

Versão: 2

CAAE: 03217318.0.0000.5406

Instituição Proponente: UNIVERSIDADE ESTADUAL PAULISTA JULIO DE MESQUITA FILHO

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.233.439

Apresentação do Projeto:

Apresentação bem elaborada.

Objetivo da Pesquisa:

Objetivos bem delimitados

Avaliação dos Riscos e Benefícios:

De acordo.

Comentários e Considerações sobre a Pesquisa:

Pesquisa relevante para a área da biblioteconomia e ciência da informação

Considerações sobre os Termos de apresentação obrigatória:

Termos de apresentação obrigatória estão corretos.

Conclusões ou Pendências e Lista de Inadequações:

Aprovado.

Considerações Finais a critério do CEP:

O CEP da FFC da UNESP de MARÍLIA, em reunião ordinária de 27/03/2019, após acatar o parecer do membro relator previamente aprovado para o presente estudo e atendendo a todos os dispositivos das resoluções 466/2012, 510/2016 e complementares, bem como ter aprovado o Termo de Consentimento Livre e Esclarecido como também todos os anexos incluídos na pesquisa,

Endereço: Av. Hygino Muzzi Filho, 737

Bairro: Campus Universitário

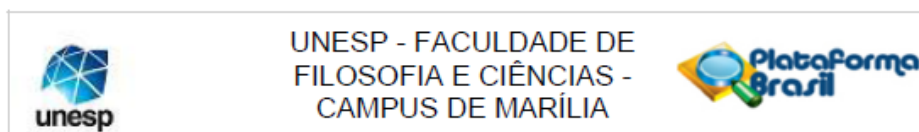
UF: SP

Município: MARILIA

CEP: 17.525-900

Telefone: (14)3402-1346

E-mail: cep.marilia@unesp.br



UNESP - FACULDADE DE
FILOSOFIA E CIÊNCIAS -
CAMPUS DE MARÍLIA

Continuação do Parecer: 3.233.439

resolve APROVAR o projeto de pesquisa O PROCESSO DE INDEXAÇÃO DE ASSUNTOS EM REPOSITÓRIOS INSTITUCIONAIS PELA ABORDAGEM DA RECUPERAÇÃO DA INFORMAÇÃO

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1194277.pdf	07/03/2019 18:16:37		Aceito
Declaração de Instituição e Infraestrutura	FCM.pdf	07/03/2019 18:12:43	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Projeto Detalhado / Brochura Investigador	Novo.pdf	07/03/2019 15:38:04	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	FCA.pdf	07/03/2019 15:34:15	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	BAE.pdf	07/03/2019 15:33:49	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	IMECC.pdf	09/11/2018 13:53:11	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	IFGW.pdf	09/11/2018 13:53:01	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	IFCH.pdf	09/11/2018 13:52:55	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	FEF.pdf	09/11/2018 13:52:48	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	FEA.pdf	09/11/2018 13:52:37	ROBERTA CRISTINA DAL EVEDOVE	Aceito
Declaração de Instituição e Infraestrutura	FE.pdf	09/11/2018 13:52:09	ROBERTA CRISTINA DAL EVEDOVE	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_Usuário.pdf	08/11/2018 17:57:19	ROBERTA CRISTINA DAL EVEDOVE TARTAROTTI	Aceito
TCLE / Termos de Assentimento / Justificativa de	TCLE_Gestor.pdf	08/11/2018 17:57:04	ROBERTA CRISTINA DAL EVEDOVE	Aceito

Endereço: Av. Hygino Muzzi Filho, 737
Bairro: Campus Universitário CEP: 17.525-900
UF: SP Município: MARILIA
Telefone: (14)3402-1346 E-mail: cep.marilia@unesp.br



UNESP - FACULDADE DE
FILOSOFIA E CIÊNCIAS -
CAMPUS DE MARÍLIA



Continuação do Parecer: 3.233.439

Ausência	TCLE_Gestor.pdf	08/11/2018 17:57:04	ROBERTA CRISTINA DAL EVEDOVE	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_Catalogador.pdf	08/11/2018 17:56:47	ROBERTA CRISTINA DAL EVEDOVE TARTAROTTI	Aceito
Folha de Rosto	Roberta.pdf	14/08/2018 10:09:41	ROBERTA CRISTINA DAL EVEDOVE	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

MARILIA, 29 de Março de 2019

Assinado por:
SIMONE APARECIDA CAPELLINI
(Coordenador(a))

Endereço: Av. Hygino Muzzi Filho, 737
Bairro: Campus Universitário
UF: SP Município: MARILIA

CEP: 17.525-900

Telefone: (14)3402-1348

E-mail: cep.marilia@unesp.br