

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

MARINA PENTEADO DE FREITAS

**AUTOARQUIVAMENTO E REPRESENTAÇÃO DE ASSUNTO: estudo analítico de
teses e dissertações do Repositório Institucional da UFSCar**

São Carlos, SP
2019

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

MARINA PENTEADO DE FREITAS

**AUTOARQUIVAMENTO E REPRESENTAÇÃO DE ASSUNTO: estudo analítico de
teses e dissertações do Repositório Institucional da UFSCar**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos, como requisito para a obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Conhecimento, Tecnologia e Inovação.

Linha de pesquisa: Tecnologia, Informação e Representação.

Orientador(a): Profa. Dra. Paula Regina Dal'Evedove

São Carlos, SP
2019

Freitas, Marina Pentead de

Autoarquivamento e representação de assunto: estudo analítico de teses edissertações do Repositório Institucional da UFSCar / Marina Pentead de Freitas. -- 2019.

89 f. : 30 cm.

Dissertação (mestrado)-Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador: Paula Regina Dal'Evedove

Banca examinadora: Mariângela Spotti Lopes Fujita, Ariadne Chloe Mary Furnival

Bibliografia

1. Consistência da indexação. 2. Autoarquivamento. 3. Teses e Dissertações. I. Orientador. II. Universidade Federal de São Carlos. III. Título.

Ficha catalográfica elaborada pelo Programa de Geração Automática da Secretaria Geral de Informática (SIn).

DADOS FORNECIDOS PELO(A) AUTOR(A)

Bibliotecário(a) Responsável: Ronildo Santos Prado – CRB/8 7325

MARINA PENTEADO DE FREITAS

AUTOARQUIVAMENTO E REPRESENTAÇÃO DE ASSUNTO: estudo Analítico de Teses e Dissertações do Repositório Institucional da UFSCar

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos como parte das exigências para a obtenção do Título de Mestre em Ciência da Informação.

Linha de Pesquisa: Tecnologia, Informação e Representação.

Data da Defesa: 22 de agosto de 2019

BANCA EXAMINADORA

Profa. Dra. Paula Regina Dal'Evedove (Presidente)
Universidade Federal de São Carlos
PPGCI/UFSCar

Profa. Dra. Mariângela Spotti Lopes Fujita (Membro externo)
Universidade Estadual Paulista "Júlio de Mesquita Filho"
PPGCI/UNESP/Marília

Profa. Dra. Ariadne Chloe Mary Furnival (Membro interno)
Universidade Federal de São Carlos
PPGCI/UFSCar.

Prof. Dr. Roniberto Morato do Amaral (Suplente)
Universidade Federal de São Carlos
PPGCI/UFSCar

Profa. Dra. Giovana Deliberali Maimone (Suplente)
Universidade de São Paulo
PPGCI/USP

AGRADECIMENTOS

Agradeço a todos que contribuíram direta ou indiretamente na construção e realização desta dissertação de mestrado. Agradeço em especial:

À Universidade Federal de São Carlos pela oportunidade do estudo. Aos meus chefes pela compreensão nas ausências e incentivo. A orientadora Profa. Dra. Paula Regina Dal'Evedove. A todos os professores do Programa de Pós-graduação do Departamento de Ciência da Informação da UFSCar. Aos Bibliotecários Ronildo e Milena. Ao Prof. Dr. Roberto A. Martins e a Profa. Dra. Maria das Graças Volpe Nunes, amigos e companheiros. Não poderia deixar de agradecer ao “Bond do Capurro” meus colegas de mestrado, pela amizade e pelas mensagens de incentivo nos momentos desanimadores deste percurso (sobrevivemos afinal). A minha família e aos meus filhos incluindo os de quatro patas.

RESUMO

Pesquisas dedicadas à questão do autoarquivamento em repositórios institucionais são escassas na literatura nacional e internacional de Biblioteconomia e Ciência da Informação, especificamente acerca da qualidade da representação temática da informação e, por consequência, da posterior recuperação por assuntos nestes ambientes digitais. Considerando-se o exposto, o problema de pesquisa decorre da carência de estudos que investiguem as implicações do autoarquivamento na recuperação por assuntos em repositórios institucionais de acesso aberto. Sendo assim, propõe-se investigar a consistência da indexação de assunto em repositórios institucionais com autoarquivamento tendo a coleção de teses e dissertações como recorte de pesquisa. Para tanto, analisa-se a consistência na indexação livre realizada pelo autor no processo do autoarquivamento em comparação com a de um bibliotecário especialista no processo de indexação de assuntos, com uma abordagem investigativa centrada na avaliação da indexação de teses e dissertações do Repositório Institucional da Universidade Federal de São Carlos defendidas nos Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade e Engenharia de Produção. Neste sentido, busca-se contribuir com perspectivas teórico-metodológicas que propiciem consonância entre a prática do autoarquivamento e a consistência da indexação das publicações científicas presentes nos repositórios institucionais brasileiros, tendo como objetivos específicos: a) contextualizar os repositórios institucionais na perspectiva da Ciência da Informação, com atenção às práticas contemporâneas de representação temática da informação no contexto digital; b) caracterizar o autoarquivamento em repositórios institucionais nacionais e internacionais; e c) investigar a consistência da indexação de Teses e Dissertações do RI/UFSCar por meio da aplicação da avaliação da indexação na abordagem da avaliação intrínseca qualitativa. Trata-se de uma pesquisa exploratória e descritiva, tendo como método a pesquisa bibliográfica visando a realização de uma análise qualitativa. Os resultados demonstram concordância entre as indexações avaliadas, sendo a indexação do usuário mais exaustiva, porém sem a mesma precisão da indexação realizada pelo bibliotecário especialista. Como alternativa para a falta de precisão dos termos, tem-se a disponibilização de linguagens documentários para acesso e uso pelos autores como referência para a atribuição de termos, a validação do metadado assunto por bibliotecários, bem como a elaboração de políticas de indexação para a melhoria da representação e da recuperação por assuntos em repositórios institucionais.

Palavras-chave: Repositórios Institucionais; Autoarquivamento; Indexação por extração. Consistência da indexação; Teses e Dissertações.

ABSTRACT

There is a lack of research in the Information Science literature on self-archiving in institutional repositories, regarding the quality of representation of the documentary information and, consequently, the information retrieval capabilities. Investigations on the indexing practices of scientific production using self-archiving procedures are relevant in order to contribute to the improvement of quality both of information representation and of information retrieval by subject in institutional repositories. The research focus of this dissertation is on the lack of empirical investigation on the consequences of information retrieval in open access institutional repositories. Therefore, the aim is to investigate the self-archiving effects on the consistency in indexing by subject in institutional repositories focusing on the theses and dissertation collection. The empirical investigation concentrated on analyzing the results of unmediated self-archiving by authors/researchers compared to the same process executed by a librarian specialized in self-archiving. The evaluation centered on the dissertations and theses of the Science, Technology, and Society Postgraduate Programme available in the Institutional Repository of the Federal University of Sao Carlos (UFSCar). The aim is to contribute from a theoretical-methodological perspective to harmonize self-archiving practices and the consistency in indexing scientific production of the Brazilian institutional repositories. The secondary objectives are the following: a) contextualize the institutional repositories from the Information Science perspective emphasizing the modern practices of documentary information representation in a digital context; b) characterize self-archiving practices in both national and international institutional repositories; c) investigate the indexing consistency of theses and dissertations in the RI/UFSCar applying the intrinsic qualitative evaluation method. The investigation is exploratory and descriptive, applying the bibliographic research method to perform a qualitative analysis. The results show a similarity between the indexes evaluated, but the user's indexation took more time and was less precise than the expert librarian's indexation. A possible solution to deal with the lack of terms precision is to provide documentary languages to users/researchers as references for terms assignment, subject metadata validation by librarians, and indexing policies elaboration to improve representation and retrieval by subjects in institutional repositories.

Keywords: Consistency in indexing. Institutional repositories. Self-archiving. Institutional Repository of the Universidade Federal de São Carlos.

RESUMEN

La investigación dedicada al autoarchivo en repositorios institucionales es escasa en la literatura nacional e internacional de bibliotecas y ciencias de la información, específicamente sobre la calidad de la representación temática de la información y, en consecuencia, la posterior recuperación por materias en estos entornos digitales. Teniendo en cuenta lo anterior, el problema de la investigación se debe a la falta de estudios que investiguen las implicaciones del autoarchivo en la recuperación de sujetos en repositorios institucionales de acceso abierto. Por lo tanto, se propone investigar el impacto del autoarchivo en la consistencia de la indización de sujetos en repositorios institucionales con la recopilación de tesis y disertaciones como un recorte de investigación. Con este fin, la consistencia de la indexación libre realizada por el autor / investigador en el proceso de autoarchivo se compara con la de un bibliotecario que se especializa en el proceso de indexación de temas, con un enfoque de investigación centrado en la evaluación de la tesis del índice y la tesis. Institución institucional de la Universidad Federal de São Carlos defendida en el Programa de Postgrado en Ciencia, Tecnología y Sociedad e Ingeniería de Producción. En este sentido, buscamos contribuir con perspectivas teórico-metodológicas que proporcionen una consonancia entre la práctica del autoarchivo y la consistencia de la indexación de publicaciones científicas presentes en los repositorios institucionales brasileños, teniendo como objetivos específicos: a) contextualizar los repositorios institucionales desde la perspectiva de Información, con atención a las prácticas contemporáneas de representación temática de la información en el contexto digital; b) caracterizar el autoarchivo en repositorios institucionales nacionales e internacionales; y c) investigar la consistencia de la indexación de Tesis y Disertaciones de RI/UFSCar aplicando la evaluación de indexación en el enfoque de evaluación intrínseca cualitativa. Es una investigación exploratoria y descriptiva, que tiene como método la investigación bibliográfica con el objetivo de realizar un análisis cualitativo. Los resultados muestran un acuerdo entre los índices evaluados, siendo la indexación del usuario más exhaustiva, pero sin la misma precisión que la indexación realizada por el bibliotecario experto. Como alternativa a la falta de precisión de los términos, existe la disponibilidad de lenguajes documentales para el acceso y uso por parte de los autores / investigadores como referencia para la atribución de términos, la validación de los metadatos de los sujetos por parte de los bibliotecarios y la elaboración de políticas de indización. Para mejorar la representación y recuperación por sujetos en repositorios institucionales.

Palabras clave: Repositorios institucionales; Autoarchivo; Indización; Consistencia en la indización; Teses y Disertaciones.

LISTA DE ILUSTRAÇÕES

Figura 1 – Fórmula para obter o índice de consistência entre duas indexações	59
Figura 2 – Página inicial do Repositório Institucional da UFSCar	67

LISTA DE QUADROS

Quadro 1 –	Etapas do processo de indexação	29
Quadro 2 –	Instituições brasileiras com política de autoarquivamento registradas no ROARMAP	52
Quadro 3 –	Fórmula para obter o índice de consistência entre duas indexações	
Quadro 4 –	Resultados da indexação por extração de teses e dissertações do PPGCTS	61
Quadro 5 –	Resultados da indexação por extração de teses e dissertações do PPGEF	63
Quadro 6 –	Índices de consistência entre os termos atribuídos às teses e dissertações do PPGCTS	70
Quadro 7 –	Índices de consistência entre os termos atribuídos às teses e dissertações do PPGEF	72
Quadro 8 –	Correspondência entre os termos das indexações a partir do catálogo de assuntos da Biblioteca Nacional	74

LISTA DE ABREVIATURAS E SIGLAS

BDTD	Biblioteca Digital Brasileira de Teses e Dissertações
BOAI	<i>Budapest Open Access Initiative</i>
BRAPCI	Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação
FAPESP	Fundação de Apoio à Pesquisa do Estado de São Paulo
HP	<i>Hewlett-Packard</i>
IBICT	Instituto Brasileiro de Informação Ciência e Tecnologia
ISBD	<i>International Standard Bibliographic Description</i>
ISKO	<i>International Society for Knowledge Organization</i>
MIT	<i>Massachusetts Institute of Technology</i>
OAI	<i>Open Archives Initiative</i>
OAM	<i>Open Access Movement</i>
PPGCTS	Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade
PPGEP	Programa de Pós-Graduação em Engenharia de Produção
RI/UFSCar	Repositório institucional da Universidade Federal de São Carlos
ROARMAP	<i>Registry of Open Access Repository Mandates and Policies</i>
SIBi	Sistema Integrado de Bibliotecas
SISA	Sistema de Indización Semi-Automático
TBCI	Tesouro Brasileiro da Ciência da Informação
TIC	Tecnologias de Informação e Comunicação
TTI	Tratamento Temático da Informação
UFSCar	Universidade Federal de São Carlos
UNESP	Universidade Estadual Paulista

SUMÁRIO

1 INTRODUÇÃO	13
1.1 Objetivos	18
1.2 Justificativas	18
1.3 Estrutura da Pesquisa	19
2 REFERENCIAL TEÓRICO	21
2.1 REPRESENTAÇÃO TEMÁTICA EM CONTEXTOS INFORMACIONAIS CONTEMPORÂNEAS	21
2.1.1 O Processo de Indexação	26
2.1.2 Práticas Contemporâneas de Indexação	32
2.1.3 Avaliação da Indexação	36
2.2 REPOSITÓRIOS INSTITUCIONAIS	41
2.2.1 Autoarquivamento em Repositórios Institucionais	46
3 PROCEDIMENTOS METODOLÓGICOS	56
3.1 Universo da Pesquisa	65
4 ANÁLISE DA CONSISTÊNCIA DA INDEXAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFSCar	68
5 CONSIDERAÇÕES FINAIS	78
REFERÊNCIAS	80

1 INTRODUÇÃO

A implantação do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos (UFSCar) objetiva fortalecer a Ciência da Informação enquanto campo científico dedicado à produção de conhecimentos e métodos para a inovação, tanto em sua perspectiva social como tecnológica. Assim, a presente pesquisa encontrou bases acadêmicas de realização no contexto da linha 2 – Tecnologia, Informação e Representação, com contribuições ao universo de estudos da Organização do Conhecimento.

O avanço de inovações científicas e tecnológicas culminou em uma rápida e crescente disponibilidade de conteúdos na Internet. Como resultado, o acesso à informação e ao conhecimento como ativos econômicos passou a ser instantâneo e ilimitado. Este fenômeno contemporâneo repercute diretamente na forma de publicação e acesso às pesquisas científicas e a exigência deste movimento pelas agências de fomento, instituições ligadas à ciência, tecnologia e inovação, assim como diversos setores da sociedade que fazem uso dos resultados provenientes destes estudos.

Nesta ótica, se uma publicação não puder ser encontrada eletronicamente, ela é pouco útil e, em termos práticos, também não existe. Esta colocação pode ser contextualizada em bibliotecas universitárias que ao longo dos anos deixaram de disponibilizar em seus acervos físicos teses e dissertações. Na UFSCar, por exemplo, desde o ano de 2016 está prática é implantada, sendo as publicações disponíveis para acesso apenas em seus formatos digitais.

Com as alterações dos sistemas de comunicação científica decorrentes da expansão tecnológica e o uso da Internet, o volume de publicações oriundas de pesquisas científicas aumentou exponencialmente nas últimas décadas. De modo geral, esse conhecimento científico se materializa em teses, dissertações, artigos de periódicos e outros documentos provenientes do contexto universitário, os quais passam a integrar ambientes especializados de armazenamento, acesso e recuperação desses recursos informacionais.

Diante da complexidade da vida contemporânea, a sociedade conclama o acesso à informação cada vez mais interativo e facilitado. Neste cenário, novos ambientes informacionais digitais são desenvolvidos e viabilizados pelas tecnologias de informação e comunicação (TIC). Um dos principais mecanismos utilizados para o acesso aberto às informações científicas são os repositórios digitais.

Os repositórios surgiram com a crise dos periódicos vivenciada inicialmente nos Estados Unidos em meados dos anos 1980, em que as bibliotecas não conseguiam mais manter as suas assinaturas e, conseqüentemente, atender as demandas de seus usuários. Esse fato desencadeou a busca por alternativas de fontes de informação.

A partir dos anos 1990, apareceram as primeiras publicações eletrônicas e a mudança no sistema tradicional de publicação científica. Uma iniciativa que teve destaque importante neste meio foi o movimento dos Arquivos Abertos, propulsor da bandeira para o acesso livre e gratuito às publicações científicas, assim como o Movimento de Acesso Livre, representado por instituições como a *American Research Libraries*, *Open Society* e *Max Planck Society*, líderes nesta militância (MUELLER, 2006).

O movimento de acesso aberto à informação científica tem como princípio que os resultados das pesquisas financiadas com dinheiro público devem ser de acesso aberto. Diante da importância do tema, o movimento ganhou adeptos em diversos países, sendo lançados manifestos como o de Bethesda, Budapeste, Berlim e o manifesto brasileiro, proposto pelo Instituto Brasileiro de Informação Ciência e Tecnologia - IBICT (KURAMOTO, 2006).

A *Open Archives Initiative* (OAI) é uma iniciativa para se disponibilizar o conteúdo de publicações científicas na *Web*, de forma gratuita. A ideia é diminuir os custos das publicações impressas e manter a qualidade exigida pela comunidade científica. Enquanto o acesso aberto gratuito se refere ao acesso *online* sem restrições, o acesso livre consiste no acesso *online* com alguns direitos adicionais de uso e reuso concedidos mediante o uso de licenças do autor. (MEDEIROS, 2010).

A iniciativa de arquivos abertos tem como proposta aspectos técnicos e de organização de uma estrutura para publicações científicas, podendo ser estabelecidas esferas livres e comerciais. A OAI, introduziu alguns conceitos e ideais a respeito, resumidos em uso de *software* livre (*open source*), autoarquivamento, criação de repositórios de acesso aberto (institucionais ou temáticos), uso de padrões de preservação de objetos digitais, além de mudanças na organização e na recuperação da informação (KURAMOTO, 2006).

Um dos principais conceitos da filosofia da OAI é o autoarquivamento, o qual consiste no depósito de um documento digital de acesso público, que esteja de acordo com os princípios estabelecidos pela OAI. Para Baggio (2007), o Brasil tem uma baixa adesão de

autoarquivamento nos repositórios institucionais e ainda há muita resistência por parte dos autores, mesmo que a instituição tenha uma política pré-estabelecida.

Em um cenário mais atual, Veiga (2015) analisou a questão a partir das instituições brasileiras e o envolvimento com o autoarquivamento, sendo constatado que a maioria dos repositórios institucionais, 64% do total, tem implementado o autoarquivamento, porém apenas 28% possibilitam a participação do autor no depósito de sua produção científica. Assim como Baggio (2007), Veiga (2015) concorda que não bastam políticas mandatórias para garantir a adesão ao autoarquivamento, é preciso empenho dos gestores dos repositórios institucionais e novos estudos que identifiquem as dificuldades pela adesão.

Suber (2004) aponta que o maior obstáculo pode ser a inercia ou a omissão dos próprios autores. Para Sawn (2005) o autoarquivamento não é uma alternativa para publicação, mas para preservação das publicações científicas em um repositório. Autores como Swan (2005) e Keefer (2007), veem a resistência ao autoarquivamento como um risco de baixar a qualidade das publicações científicas nos sistemas de informação de acesso aberto (ASSIS, 2013).

Com efeito, além da iniciativa de arquivos abertos e acesso aberto às pesquisas científicas, o autoarquivamento é resultante da nova configuração do ambiente *Web*, tonando-se uma rede de interação. Na *Web* social, usuários interagem e colaboram na descrição dos conteúdos disponíveis, produzindo, classificando e reformulando o que já está disponível (BLATTMANN; SILVA, 2007). Na literatura científica da Ciência da Informação, a participação dos usuários na representação por assuntos dos conteúdos e recursos informacionais no ambiente *Web* pode ser denominada de indexação social, etiquetagem colaborativa, classificação social ou folksonomia. De modo geral, essa prática consiste na atribuição livre de termos/palavras-chave (*tags*) por parte dos usuários aos documentos disponíveis em ambientes digitais colaborativos (GUEDES; DIAS, 2010).

Na prática, o autoarquivamento é um processo que permite maior rapidez na divulgação da pesquisa científica, mas por outro lado, isso não garante a qualidade da representação temática, situação que pode repercutir na recuperação da informação. Apesar da facilidade de envio e preenchimento dos campos obrigatórios nos sistemas de recuperação de informação, a liberdade deste procedimento pode não significar qualidade e precisão do preenchimento dos metadados fornecidos pelo autor, principalmente no metadado assunto

quando este não for padronizado, podendo impactar negativamente no processo de busca e recuperação da informação.

No cenário internacional, destacam-se os estudos de Rousidis et al. (2014), Palavitsinis, Manouselis e Sanchez-Alonso (2014) e Barton, Currier e Hey (2003) que trazem eminentes preocupações acerca da qualidade dos metadados utilizados na descrição de dados armazenados em repositórios. No escopo da Ciência da Informação brasileira, as pesquisas dedicadas aos repositórios institucionais de acesso aberto que trazem preocupações semelhantes privilegiam discussões acerca da utilização das linguagens natural e controlada como forma de melhoria da pesquisa por assunto (BACHA; ALMEIDA, 2013), bem como observam as práticas de indexação em repositório institucional (SANTOS; NEVES, 2018).

Neste contexto, destaca-se a pesquisa apresentada por Santos e Neves (2018), em que as autoras observaram as práticas de indexação no Repositório Institucional da Universidade Federal do Rio Grande do Norte (RI/UFRN). Na análise sobre as implicações da indexação por extração para representação de conteúdo no campo “assunto” atribuídos em linguagem natural, o referido estudo revelou a presença de algumas implicações, dentre elas o emprego de frases, termos com polissemia, ocorrência de erros ortográficos, presença de descritores abrangentes e com multiplicidade de sentidos, abreviações nos termos, uso de siglas, entre outras.

Cabe ressaltar que o objetivo do estudo não foi o de excluir o uso de termos em linguagem natural, considerando-se a importância de manter os descritores informados pelos autores. Contudo, as referidas autoras alertam para as implicações negativas que esse tipo de indexação pode causar na representação e recuperação da informação. A ocorrência de plurais, existência de polissemia, emprego de sinonímia, homógrafos, erros ortográficos, expressões regionais e outras são um alerta para que os administradores de repositórios e os profissionais indexadores elaborem políticas de indexação que visem o aperfeiçoamento das práticas utilizadas em repositórios institucionais, incluindo o auxílio de linguagens documentárias para elevar o nível da qualidade da descrição do conteúdo desses documentos e, conseqüentemente, a precisão no processo de busca e recuperação da informação nestes ambientes (SANTOS; NEVES, 2018).

Apesar do avanço das pesquisas nacionais, investigações dedicadas à questão do autoarquivamento em repositórios institucionais são escassas, especificamente acerca da

qualidade da representação temática da informação e, por consequência, da posterior recuperação por assuntos nestes ambientes digitais. Mesmo que a prática contribua para uma maior rapidez na divulgação da pesquisa científica, ainda são incertos os benefícios e as implicações do autoarquivamento para a qualidade do metadado assunto utilizado na descrição dos recursos informacionais.

Este cenário originou o seguinte questionamento: quais as implicações do autoarquivamento na representação e recuperação por assuntos em repositórios institucionais de acesso aberto? Soma-se ao exposto a incipiente presença de investigações dedicadas a analisar a qualidade da indexação livre realizada pelo autor em repositórios institucionais de bibliotecas universitárias em que o autoarquivamento é praticado.

Face ao problema apresentado, considera-se que a discussão sobre a prática da indexação seja pertinente e necessária, de modo a contribuir para a melhoria na representação e recuperação por assuntos em repositórios institucionais. Sendo assim, a proposta dessa pesquisa é investigar a consistência da indexação de assunto em repositórios institucionais que aderem à prática do autoarquivamento, tendo a coleção de teses e dissertações como recorte de pesquisa.

Para tanto, analisa-se a consistência da indexação livre realizada pelo autor no processo do autoarquivamento em comparação com a de um bibliotecário especialista no processo de indexação de assuntos, com uma abordagem investigativa centrada na avaliação da indexação de teses e dissertações do Repositório Institucional da Universidade Federal de São Carlos defendidas nos Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade e Engenharia de Produção.

Dessa forma, este estudo teórico-aplicado contribui com discussões importantes sobre os benefícios e implicações do autoarquivamento para o tratamento de recursos informacionais, em que os resultados aqui apresentados podem colaborar para a sistematização de diretrizes de política de indexação que satisfaçam as exigências atuais da representação da informação em ambientes digitais e, ao mesmo tempo, primam pela qualidade da indexação de assuntos praticada em repositórios institucionais.

1.1 Objetivos

O objetivo geral da pesquisa é contribuir com perspectivas teórico-metodológicas que propiciem consonância entre a prática do autoarquivamento e a qualidade da indexação das publicações científicas presentes nos repositórios institucionais brasileiros.

Como objetivos específicos, pretendemos:

- Contextualizar os repositórios institucionais na perspectiva da Ciência da Informação, com atenção às práticas contemporâneas de representação temática da informação no contexto digital;
- Caracterizar o autoarquivamento em repositórios institucionais nacionais e internacionais; e
- Investigar a qualidade da indexação de Teses e Dissertações do RI/UFSCar por meio do índice de consistência mediante aplicação da avaliação da indexação na abordagem da avaliação intrínseca qualitativa.

1.2 Justificativa

Apesar das diversas possibilidades oferecidas pela Internet para o acesso à informação digital, a agilidade no processo nem sempre é sinônimo de qualidade e relevância das informações recuperadas. Esta situação reforça a importância de iniciativas na área de Organização do Conhecimento que observem e discutam as exigências contemporâneas que se colocam à organização da informação, muito em razão da introdução e evolução das tecnologias de informação e comunicação no contexto da Biblioteconomia e Ciência da Informação.

Este cenário acentua a necessidade de se avaliar a qualidade da indexação de assuntos em repositórios institucionais que adotam o autoarquivamento na representação dos recursos informacionais. A discussão teórica e metodológica das práticas de indexação de assuntos em repositórios institucionais ainda é um tema escasso na literatura especializada. No Grupo de Pesquisa Representação e Humanidades Digitais da Universidade Federal de São Carlos, o tema é investigado no âmbito do Projeto de Pesquisa “Representação e recuperação

da informação por assuntos em repositórios institucionais”, sob coordenação da Profa. Da. Paula Regina Dal’Evedove – Professora Permanente do Programa de Pós-Graduação em Ciência da Informação desta mesma instituição. Portanto, a presente pesquisa corrobora com as atuais discussões e projeta esforços para a construção de referenciais teóricos e metodológicos no âmbito da representação e recuperação por assuntos em repositórios institucionais brasileiros.

Para além das contribuições acadêmicas e sociais, a motivação pela pesquisa decorre de uma preocupação enquanto bibliotecária atuante no RI/UFSCar com a inconsistência de termos no processo de busca por assuntos em repositórios institucionais, por entender que uma recuperação por assuntos eficaz decorre de uma representação de assuntos consistente. Almeja-se, portanto, contribuir com discussões na área e favorecer a melhoria dos repositórios institucionais brasileiros.

O *Evidence-based librarianship* é uma tendência que teve início na Medicina e hoje se estende para outras áreas, em que o objetivo é aperfeiçoar produtos e serviços da biblioteca utilizando as melhores evidências disponíveis combinadas com uma perspectiva pragmática desenvolvida a partir de experiências de trabalho em Biblioteconomia. Portanto, os resultados provenientes desta pesquisa contribuem para as discussões relativas ao tema em Organização do Conhecimento, além de favorecer a implementação de melhorias em repositórios institucionais, com viabilidade de aplicação dos conhecimentos construídos no decorrer do mestrado acadêmico no RI/UFSCar.

1.3 Estrutura do trabalho

Para atingir os objetivos propostos, além deste capítulo introdutório, esta pesquisa é composta por outros cinco capítulos.

Capítulo 2 – REFERENCIAL TEÓRICO – contextualiza a representação temática da informação em Organização do Conhecimento, sendo contemplada a indexação por assuntos e sua importância para a recuperação da informação. Apresenta a consistência na indexação como um importante recurso para avaliar a qualidade da indexação e, por consequência, a acessibilidade temática em sistemas de recuperação da informação. De modo complementar, descreve trabalhos correlacionados dedicados à avaliação da indexação em diferentes

cenários e propósitos. De modo a situar o universo da pesquisa, apresenta o Repositório Institucional como ambiente contemporâneo de armazenamento, acesso, busca, compartilhamento, uso e reuso da pesquisa científica. Discute-se, ainda, o autoarquivamento como prática contemporânea no ambiente dos repositórios institucionais, com destaque para as experiências nacionais e internacionais que adotam esta prática.

Capítulo 3 – PROCEDIMENTOS METODOLÓGICOS – em que são descritas as escolhas metodológicas adotadas, além de apresentar o universo da pesquisa e descrever os dados utilizados para a realização da etapa empírica.

Capítulo 4 – ANÁLISE DA CONSISTÊNCIA DA INDEXAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFSCar – apresenta e discute os valores de consistência obtidos entre as indexações avaliadas mediante aplicação da fórmula do cálculo de consistência.

Capítulo 5 – CONSIDERAÇÕES FINAIS – são contextualizados os principais resultados da pesquisa e sua pertinência para o avanço do tema em Organização do Conhecimento, com recomendações para pesquisas futuras. Finalizando, são apresentadas as referências utilizadas na realização desta Dissertação.

2 REFERENCIAL TEÓRICO

Esse capítulo visa contextualizar a organização e a representação da informação a partir das dimensões teóricas e aplicadas da Organização do Conhecimento, tendo como referência o processo de indexação e sua importância para a recuperação da informação em ambientes digitais. A partir deste entendimento, apresenta a avaliação da indexação e descreve algumas experiências brasileiras com a aplicação da fórmula para obtenção do índice de consistência na indexação. Essa contextualização favorece a caracterização do repositório institucional como ambiente contemporâneo de acesso e recuperação às publicações científicas, com atenção à prática do autoarquivamento e as contribuições no campo da informação dedicadas à temática.

2.1 REPRESENTAÇÃO TEMÁTICA EM CONTEXTOS INFORMACIONAIS CONTEMPORÂNEAS

Desde a origem da humanidade existem tentativas de organizar e representar o conhecimento, repercutindo diretamente nas formas de sociabilidade e suas relações. Logo, essa não é uma necessidade atual – organizar e representar o conhecimento é uma preocupação que nasce com a evolução da sociedade, que deseja compartilhar e usar o conhecimento produzido e registrado como parte das ações humanas.

Os tempos mudam, mas a necessidade de organizar o conhecimento para se ter acesso mais facilmente permanece. Desde quando o homem busca o conhecimento, ele também estuda meios para organizá-lo e socializá-lo. Já na Pré-História havia a necessidade de se marcar o tempo, trocar experiências, transmitir mensagens e sentimentos, registrando o conhecimento com as pinturas nas paredes das cavernas, sendo essa a forma utilizada para representar a realidade cotidiana.

Representar é colocar “algo em lugar de”, assim define Alvarenga (2003). Na Antiguidade a preocupação recaía em como o conhecimento seria organizado com fins de acesso. Temos como exemplo a Biblioteca de Alexandria, que reunia obras de várias partes do mundo, e isso inspirou a construção de catálogos e bibliografias com a finalidade de controlar o que havia sido publicado, bem como a criação de instrumentos para classificar, identificar e hierarquizar o conhecimento produzido.

A organização e a representação do conhecimento foram influenciadas por filósofos e cientistas de várias áreas que, preocupados com a divisão do conhecimento, deram contribuições teóricas ao seu desenvolvimento (PINHO, 2009). As primeiras classificações bibliográficas foram fundamentadas a partir das classificações construídas para sistematizar o conhecimento desenvolvido pelo homem (PANDO, 2018).

Nesse sentido, a Filosofia teve um papel fundamental para a construção de uma teoria da classificação. Ela fornece a fundamentação para a ciência. Historicamente, foi a filosofia que iniciou a discussão nessa área, da investigação e da apresentação dos sistemas de organização e representação do conhecimento. Todo esse processo de representação tem a ver com a produção do pensamento do ser humano, que gera o conhecimento. “Conhecer significa dividir e classificar para depois poder determinar relações sistemáticas entre o que se separou” (SANTOS, 1988, p. 50).

Pando (2018) faz uma retrospectiva histórica em sua tese que o leva a Grécia antiga, a qual foi o berço da filosofia clássica, em que os primeiros filósofos gregos estudaram o conhecimento contemporâneo, e a Filosofia se distinguiu da religião, tornando-se fundamental para a compreensão da realidade. Um dos pontos principais discutidos e sistematizados pelos filósofos gregos foi o conhecimento, surgindo então as construções e contribuições da noção de conhecimento, sendo estabelecidas as diferenças entre conhecimento sensível e conhecimento intelectual; aparência e essência; opinião e saber as regras da lógica para se chegar à verdade (CHAUI, 202).

Com este breve histórico a respeito do pensamento sobre o conhecimento, percebe-se uma preocupação inata do homem em entender como se dá o conhecimento e como organizá-lo, registrá-lo. O conhecimento não é algo simples como pode parecer, tampouco a sua organização.

Os processos de organização do conhecimento transformaram-se em disciplina científica. Hoje, a Organização do Conhecimento configura-se como área base que norteia boa parte das pesquisas teóricas e instrumentais da Ciência da Informação (MARTINS, 2014). Originalmente descrita como “Organização do Conhecimento e Documentação”, proposta por Soergel em sua dissertação em 1971, expressão anteriormente adotado pelo bibliotecário norte americano Henry Evelyn Bliss (1870-1955) e designada como “Organização do

Conhecimento”, dentro da Biblioteconomia, sendo amplamente difundida por Ingetraut Dahlberg (1927 – 2017).

Neste contexto, Dahlberg é considerada a responsável por atribuir à Organização do Conhecimento o *status* de disciplina científica. A partir de seus estudos sobre teoria do conceito a pesquisadora constrói uma base teórica para a organização do conhecimento, sendo institucionalizada com a fundação da *International Society for Knowledge Organization* (ISKO) em 1989. A ISKO constitui-se em uma sociedade internacional dedicada aos aspectos teóricos-metodológicos e interdisciplinares da organização do conhecimento e do tratamento da informação (ISKO, 2019).

Assim, no processo de evolução da área, as práticas científicas ou pragmáticas receberam definições até se certificarem como Organização do Conhecimento, em um cenário internacional, assim como Organização e Representação do Conhecimento, expressão atualmente adotada pelo Capítulo Brasileiro da ISKO. Considerada uma ciência tridimensional por estudar os princípios, os métodos e os instrumentos, denominados Representação, Organização e Comunicação do Conhecimento (GUIMARÃES, 2014). Dessa forma, a Organização do Conhecimento conquistou espaço, não só como uma atividade pragmática, mas também como campo de estudos teóricos e aplicados e as suas contribuições deram origem a uma disciplina legítima (MARTINS, 2014).

Além da dimensão epistemológica e aplicada da área, a inclinação e contribuição da Organização do Conhecimento ao social são evidenciadas na conceituação de Barité (2001, p. 41), para quem:

O objeto de estudo da organização do conhecimento é – a nosso juízo – o conhecimento socializado, e como disciplina dá conta do desenvolvimento de técnicas para a construção, gestão, uso e avaliação de classificações científicas, taxonomias, nomenclaturas e linguagens documentais. De outra parte, traz metodologias de uso e recuperação por linguagem natural. É esta visão integral do conhecimento, em que se associam as classificações filosóficas ou científicas do saber com as classificações destinadas à organização de documentos em bibliotecas, arquivos e outras unidades de informação que abre maiores perspectivas para um importante desenvolvimento disciplinar e interdisciplinar no âmbito da Biblioteconomia e Documentação.

A Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB) é outra importante associação de representação científica e política para o campo da

informação. Ela se organiza em 11 Grupos de Trabalho, cada grupo se destina a assuntos específicos. Composta por pesquisadores, docentes e pós-graduandos, tem a finalidade de estimular e avaliar o ensino e a pesquisa na pós-graduação em Ciência da Informação no Brasil e nas Américas. O grupo de trabalho GT-2 intitulado “Organização e Representação do Conhecimento” é outro espaço científico importante em que as reflexões e discussões relativas ao desenvolvimento de teorias e metodologias em relação aos processos, produtos e instrumentos dedicados à organização e representação da informação e do conhecimento são conduzidos, conforme consta a seguir:

Teorias, metodologias e práticas relacionadas à organização e preservação de documentos e da informação, enquanto conhecimento registrado e socializado, em ambiências informacionais tais como: arquivos, museus, bibliotecas e congêneres. Compreende, também, os estudos relacionados aos processos, produtos e instrumentos de representação do conhecimento (aqui incluindo o uso das tecnologias da informação) e as relações inter e transdisciplinares neles verificadas, além de aspectos relacionados às políticas de organização e preservação da memória institucional. (ANCIB – GT2, 2019).

A organização e a representação do conhecimento foram incorporando as sistematizações filosóficas e científicas, e suas técnicas surgem em atenção às diversidades da organização de documentos, fracionando-se em representação descritiva e representação temática para um melhor entendimento e condução de discussões. A representação descritiva dedica-se a organização dos aspectos físicos dos documentos, enquanto a representação temática considera a organização dos documentos a partir do conteúdo, isto é, do “conhecimento neles contido” (GUIMARÃES, 2008). Ambas, porém, compõem a representação da informação, sendo tal diferenciação importante para o avanço das pesquisas na área e a verticalização das abordagens.

A representação temática da informação, objeto de estudo nesta pesquisa, decorre da necessidade pragmática do tratamento dos documentos (GUIMARÃES; MORAES; GUARIDO, 2007). Bräscher e Guimarães (2018, p. 242) reforçam o exposto ao sinalizarem que “as questões relativas à análise, descrição e representação do conteúdo dos documentos inserem-se no contexto dos estudos da organização do conhecimento e constituem um desafio, em função da complexidade dos aspectos envolvidos nesses processos”.

Neste bojo, tem-se o Tratamento Temático da Informação (TTI) como núcleo dedicado às questões de análise e representação temática, ou seja, terminologia adotada por parte da

comunidade científica em Organização do Conhecimento para compreender e propor soluções para os problemas relacionados ao conteúdo dos documentos. Guimarães (2008) assinala, ainda, que o TTI tem um espaço primordial na Biblioteconomia e Ciência da Informação ao revelar a mediação entre a produção e o uso da informação, sendo considerado uma ponte entre documento e usuário, o que resulta no acesso à informação.

A acessibilidade temática da informação é oportunizada por processos, produtos e instrumentos dedicados à organização e representação do assunto/conteúdo dos documentos com fins de recuperação em sistemas de informação. As abordagens teóricas¹ do TTI são constituídas pela catalogação de assunto (matriz norte-americana), indexação (matriz inglesa) e análise documentária (matriz francesa) enquanto operações que visam, conforme a sua especificidade e instrumentos para a elaboração de produtos documentários, a análise, descrição e representação do conteúdo informacional (GUIMARÃES, 2008).

Os estudos em indexação manifestam uma preocupação mais efetiva com a valorização do usuário e suas expectativas diante do sistema de recuperação da informação, entendendo-o como elemento central das práticas de organização e representação do conteúdo do documento. No âmbito da indexação, a primazia consiste em determinar o assunto do documento para efetivar a sua posterior recuperação. Para delimitar o teor do documento, tem-se a análise conceitual ou análise de assunto como etapa inicial, seguida da tradução desse conteúdo em termos representativos provenientes da linguagem documental. As linguagens documentais ou Sistemas de Organização do Conhecimento (SOC) têm o objetivo de padronizar os termos, sendo construídos artificialmente para auxiliar as atividades de gestão e recuperação da informação por meio de instrumentos de classificações, ontologias, taxonomias, tesouros, dicionários, normas, etc. (SOUZA; ALBUQUERQUE, 2016).

A lógica dos SOC é tornar familiar o que ainda não conhecemos, agrupando em categorias os objetos, os pensamentos, os indivíduos, os animais, conforme suas características e semelhanças por experiências vivenciadas que nos são transmitidas culturalmente. Assim, essa organização tem como base as representações sociais vivenciadas e compartilhadas (MOSCOVICI, 2004 apud MARTINS; MORAES, 2015).

¹ Para um panorama mais detalhado das abordagens teóricas do Tratamento Temático da Informação e suas matrizes, recomenda-se a leitura do trabalho de Guimarães (2008).

Ao longo da história, os sistemas de organização do conhecimento vêm sendo aprimoradas conforme as necessidades dos usuários e as tecnologias disponíveis no ambiente digital. A primeira contribuição para a construção de SOC foi a do bibliógrafo Konrad von Gesner (1516-1565) com a *“Bibliotheca universalis”* (1545-1549), considerada a primeira lista com todos os livros conhecidos impressos em Latim, Grego e Hebraico, além de publicar um índice de assunto de classificação por temas (EISENSTEIN, 1980). Atualmente, os tesouros são os instrumentos mais empregados na indexação de documentos, por oferecerem um repertório alfabético de termos que descrevem sem ambiguidade os conceitos a ele atinentes. Por abarcarem as diferentes áreas e especialidades do conhecimento, constituem-se como valiosos instrumentos de representação da informação.

No escopo da Organização do Conhecimento, a indexação ganha fundamentação teórica a partir das contribuições das várias iniciativas de construção de SOC. Atualmente, a indexação está no âmbito da ISKO como uma área científica e profissional (FUJITA, 2013).

2.1.1 O Processo de Indexação

A indexação é um processo essencial para as atividades informativas, ao passo que visa identificar o conteúdo do documento e descrevê-lo mediante termos provenientes de uma linguagem de indexação que o represente, isto é, que consiga favorecer a busca e o entendimento do usuário sem a necessidade de leitura do seu conteúdo. Seu objetivo é, portanto, atender as necessidades dos usuários na busca por informação.

O conceito de indexação surge por volta do VI milênio A.C, nas tarefas executadas pelos antigos escribas da Mesopotâmia, quando estes começam a fazer cópias de tábuas de argila, com a confecção de colofão (nota com informações sobre a obra, indicação de responsabilidade, impressão etc.) e rótulos para recuperação das placas. Essas cópias eram armazenadas em prateleiras, nichos e cestos, e para saber o que havia em seus recipientes, era anexado um pequeno rótulo de argila no qual constava o conteúdo dos documentos depositados. Nestas ações básicas, vemos os primeiros passos do que chamamos hoje de indexação (GIL LEIVA; RODRIGUEZ MUÑOZ, 1996 apud GIL LEIVA, 2008).

Na Idade Antiga a indexação se manifesta a partir das bibliotecas com a criação de índices ou listas de documentos gerados como forma de organizar, tratar e recuperar esses

documentos. No final do Século XIX surge a preocupação com a recuperação de documentos que tratem de um mesmo assunto. Sendo assim, a partir do momento em que esses índices precisaram de uma organização por assunto, deixaram de ser um ato mecânico de construção de listas e passaram a fazer parte de um processo de análise de conteúdo dos documentos. Atualmente, a indexação está relacionada ao conceito de análise de assunto (SILVA; FUJITA, 2004).

Com o crescente volume de informações e de publicações científicas periódicas, se faz necessária a recuperação da informação cada vez mais rápida, precisa e especializada. Com efeito, a indexação passa a contar com instrumentos metodológicos cada vez mais diversificados e mais voltados para o contexto de cada documento (SILVA; FUJITA, 2004).

Silva e Fujita (2004) ressaltam a popularidade que a Documentação como área científica, e os serviços de informação em áreas especializadas, ganham a partir dos anos 1960 com a indexação e a produção de resumos usados nos serviços bibliográficos para recuperação de artigos de periódicos científicos. O termo documentação passa a ter um significado mais abrangente quando Bradford (1961) usa em seu livro “Documentação”, o termo Indexação para análise documentária (SILVA; FUJITA, 2004).

No campo da análise documental, de acordo com a linha teórica de Gardin (1981) a indexação é vista como um processo de representação de documentos com a finalidade de recuperar a informação. A partir da perspectiva de outros teóricos, principalmente do Reino Unido e dos norte-americanos, a análise documental e a indexação são compostas pelos mesmos passos operativos com o objetivo de representar o conteúdo de um documento para o desenvolvimento de índices (SILVA; FUJITA, 2004).

As linguagens documentárias foram criadas para minimizar a subjetividade visando à padronização dos termos. A subjetividade decorre de uma série de fatores, como por exemplo, a formação do bibliotecário, o nível de conhecimento do assunto, as experiências vivenciadas, fatores linguísticos, culturais, cognitivos e lógicos.

Mesmo assim, as pessoas podem fazer escolhas diferentes a respeito do conceito de um documento que trata do mesmo assunto, sobre os quais serão indexados. Esta situação é possível de ser contextualizada quando se trata da indexação praticada por pessoas diferentes, as quais, mesmo indexando um mesmo documento, tendem a ter alguma divergência na atribuição dos termos. Além disso, uma pessoa pode tomar decisões diferentes

em momentos diferentes durante a indexação do mesmo documento. No entanto, o bibliotecário deve se manter neutro mesmo sendo a subjetividade uma característica da atividade de indexação (LANCASTER, 1993).

Lucas (2000), diverge deste pensamento de Lancaster ao entender que o bibliotecário não é totalmente neutro em sua leitura, uma vez que carrega consigo uma história de experiências vivenciadas, uma cultural social, memórias que influenciam totalmente sua leitura. Teoricamente, a indexação deve ser objetiva. Contudo, esse elemento é impossível de ser atendido, visto que o ser humano adquire conhecimento, valores e atitudes a partir da convivência em sociedade, na qual as linguagens são construídas e trazem por trás a vivência do sujeito que realiza o processo de indexação. Sendo assim, a representação de assunto não é um fazer imparcial, mas um processo subjetivo de grande complexidade técnica (FOSKETT, 1973).

Para Gil Leiva (2008), o processo técnico dos documentos é mais amplo, engloba várias etapas que vão desde a seleção, aquisição, registro e tratamento dos documentos, a fim de armazenar, recuperar e divulgar a informação. Nesta ótica, o processo de indexação faz parte do processo do tratamento de documentos como um todo, se divide em duas fases, a descrição bibliográfica ou catalogação e a fase de análise de conteúdo ou análise de assunto do documento, que abrange a classificação, o resumo e a indexação. Por sua vez, Guimarães (2008) e outros teóricos da matriz francesa consideram a indexação como fase final da análise documental, na qual as linguagens documentais são usadas na produção de elementos documentários como índices. Para esta pesquisa vamos nos ater à segunda fase, análise de assunto para indexação.

De modo geral, as definições conceituais sobre indexação têm relação direta com o processo – atividade informativa. Em Mai (2005, p. 599) a operação é descrita como ato de “determinar o assunto temático dos documentos e expressá-lo em índices (por exemplo, descritores, cabeçalhos de assuntos, números de chamada, códigos de classificações ou índices) para que seja possível a representação temática”. Esta definição tem consonância com a apresentada pela Unesco (1975, p. 2) ao sinalizar “Descrever e caracterizar um documento com a ajuda de representações dos conceitos em um documento para permitir uma busca eficaz das informações contidas no documento” e com a indicação da ABNT NBR 12676 (1992) delimitada como “ato de identificar e descrever o conteúdo de um documento

com termos representativos dos seus assuntos e que constituem uma linguagem de indexação”.

Em Gil Leiva (2008) o processo de indexação tem o objetivo geral armazenar a informação e atender as necessidades de informação, cujas etapas compreendem a análise dos documentos, responde questões para extrair conceitos explícitos e implícitos e define as palavras-chave em linguagem natural ou padronizadas. No entanto, alguns autores divergem no número das etapas do processo de indexação, como se vê no quadro a baixo.

Quadro 1 - Etapas do processo de indexação.

Duas etapas	a) Reconhecer e extrair os conceitos, e b) Traduzir os conceitos em linguagem documental.	Chaumier, (1988, p. 23)
	a) Análise do conteúdo que resulta na seleção dos conceitos para representar o documento, e b) Tradução dos conceitos selecionados em linguagem de indexação utilizada pelo sistema de informação ou base de dados.	Fidel (1994, p. 573)
	a) Determinação do assunto, e b) Tradução do assunto pela linguagem de indexação.	Mai (2000)
	a) Análise dos documentos e das perguntas para a seleção dos conceitos explícitos ou implícitos, e b) Armazenamento dessas palavras chave tal como estão ou sua normalização por meio de um vocabulário controlado.	Gil Leiva (2008, p. 245)
	a) Análise conceitual, e b) Representação ou formalização.	Tamayo & Valdez (2008, p. 38)
Três etapas	a) Examinar o documento e estabelecer os conceitos, b) Identificar os principais conceitos do assunto do documento, e c) Traduzir os conceitos ou termos selecionados em termos de linguagem de indexação.	Amat (1989, p. 177)
	a) Exame do documento e determinação de seu conteúdo, b) Identificação, e c) Seleção dos conceitos principais de seu conteúdo e seleção dos termos de indexação	ABNT (1992)
	a) Análise de conteúdo, b) seleção e generalização, e c) tradução do conteúdo para termos de indexação.	Mones (2000, p. 56)
Quatro etapas	a) Conhecimento do conteúdo conceitual do documento, b) Extração dos conceitos em linguagem natural, c) Tradução desses conceitos com a linguagem documental, e d) Busca de outros conceitos pertinentes não expressados pelo autor.	Dijk e Slype (1972, p. 105)
	a) Contato com o documento, b) Identificação dos conceitos explícitos e implícitos do documento, c) Tradução dos conceitos expressados em linguagem natural por descritores, e d) Estabelecimento de ligações sintáticas entre os descritores.	Slype (1991, p. 116)

Cinco ou mais etapas	a) Revisão dos objetivos desta operação, se necessário, b) Conhecimento prévio do documento, c) Determinação de seu tema principal, d) Identificação dos elementos do conteúdo que se devem descrever e a extração dos termos correspondentes, e) Verificação da pertinência dos termos selecionados, f) Tradução dos termos da linguagem natural pelos termos correspondentes da linguagem documental, se procede, g) Verificação da pertinência desta descrição, e h) Formalização da descrição quando o sistema prevê regras especiais de apresentação ou de redação.	Guinchat e Menou (1983, p. 179)
	a) Registro dos dados bibliográficos, b) Análise do conteúdo dos documentos a partir do título, resumo e texto completo, c) Determinação dos assuntos, d) conversão dos conceitos extraído em linguagem natural para linguagem de indexação, e e) Revisão da indexação.	Cleveland e Cleveland (1990, p. 104)

Fonte: Adaptado de Gil Leiva (2008) e Fujita (2013).

Nota-se que há uma divergência quanto ao número das etapas de indexação. De modo geral, são três as etapas que confluem entre os autores – análise do documento e conhecimento do texto; escolha de conceitos; e representação dos conceitos por meio de linguagens documentais.

Hjorland (2001) considera que a indexação deve ser feita para atender as necessidades de determinados usuários, ação que vai além da escolha de uma linguagem de indexação adequada, devendo ser analisada por uma perspectiva do profissional que trabalha no processo de indexação. Autores como Rafferty e Hilderley (2007, p. 399) destacam a existência de três tipos de indexação, quais sejam:

Indexação orientada por especialistas (bibliotecários, indexadores, editores voluntários) essa é uma indexação dispendiosa e cara;

Indexação orientada pelo autor que irá utilizar termos que são comumente compreendidos e geralmente aceitos. Porém, o problema deste tipo de indexação é o fato de o autor não ser necessariamente um gestor de informação com os conhecimentos de um especialista; e

Indexação orientada pelo usuário que possibilita um elevado nível de interação com a comunidade que, provavelmente, não seria possível se tivesse decisões a serem feitas sobre códigos, convenções e regras que regem qualquer taxonomia controlada.

Considerando a indexação como ponto central para a recuperação da informação, vale ressaltar que os tipos de linguagem usadas nesse processo são: linguagem de indexação e

linguagem natural. Linguagem natural é usada como meio de comunicação no cotidiano, sendo a maneira como expressamos nosso raciocínio e trocamos informações. Linguagem documentária é uma lista de termos padronizados construída com o objetivo de eliminar possíveis problemas como a ambiguidade que ocorre na linguagem natural e que causa maiores níveis de inconsistência na indexação (LOPES, 2002).

Quando se usa uma linguagem que passa por um processo de indexação, seja ele de duas ou mais etapas, extraem-se conceitos em uma linguagem natural representando o conteúdo do documento. Sendo assim, a linguagem natural se torna um termo padronizado após a conversão em linguagem de indexação com o uso de instrumentos como tesouros, lista de cabeçalhos de assunto e etc. A linguagem controlada ou os descritores são obtidos para que as informações possam ser armazenadas e recuperadas (GIL LEIVA, 2008).

Podemos dizer que a indexação em linguagem natural é uma indexação feita sem o uso de instrumentos de controle de vocabulário. Para o contexto da presente pesquisa, quando o autor faz a submissão de sua tese ou dissertação em um repositório institucional, ele faz uso de uma linguagem natural, sem o uso de instrumentos de controle vocabular, no caso de o sistema não disponibilizar nenhuma ferramenta de auxílio. O uso de linguagem natural pode trazer problemas no que se refere à recuperação do material armazenado. Para Gil Leiva (2008), o principal inconveniente neste caso é que nem a sinonímia e nem a polissemia são controladas.

A linguagem de indexação contribui para que a pesquisa por assunto em sistemas de recuperação da informação seja mais exata e precisa. Para isso, a linguagem de indexação deve ser usada na etapa de tradução dos conceitos em indexação. Além disso, são instrumentos usadas para minimizar a subjetividade do indexador que pode se sobrepôr a uma interpretação que deveria ser mais técnica (SALES; CAFÉ, 2009).

Uma vez que a comunicação acadêmica e científica ocorre por meio de linguagens especializadas codificadas por especialistas de uma determinada área do conhecimento ou por profissionais bibliotecários; e sendo essas linguagens especializadas constituídas por terminologias próprias que, por sua vez, são responsáveis pela transmissão do conhecimento especializado, tem-se que as terminologias também pertencem à linguagem natural, utilizada por um indivíduo comum. “Melhor dizendo, a terminologia pertence a uma linguagem especializada que, por sua vez, pertence à linguagem natural” (SALES, 2009 p. 100).

Com isso, observa-se que uma língua especializada ou técnica é subordinada a linguagem natural. Tal fato leva à consideração de que a comunicação técnico-científica nada mais é do que uma comunicação natural ocorrida dentro dos campos especializados (SALES; CAFÉ, 2009).

As linguagens usadas por especialistas têm como característica viabilizar que a informação comunicada tenha precisão. Porém, para que essa precisão seja sólida são necessários instrumentos que cuidem da questão da ambiguidade do discurso especializado. Diante do exposto, as linguagens de indexação devem ser usadas com o objetivo de recuperar conteúdos informacionais em ambientes específicos (SALES; CAFÉ, 2009).

2.1.2 Práticas Contemporâneas de Indexação

Os fazeres específicos da representação temática têm perdido espaço a partir da automação dos processos de catalogação nas bibliotecas universitárias. Com a prática do intercâmbio de registros entre bibliotecas, sem muitas das vezes ser considerada as características e especificidades da biblioteca que está incorporando o registro em seu sistema, a consistência dos catálogos *online* tende a ser prejudicada. Neste momento, as políticas de gestão de bibliotecas estavam mais preocupadas com a quantidade de registros e não com a representação dos documentos, e isso implicou no prejuízo da qualidade na recuperação da informação por assunto (FUJITA; AGUSTIN LACRUZ; GOMEZ DIAZ, 2012).

Outro fator importante que faz com que os sistemas de busca e recuperação por assunto dos catálogos *online* não funcionem com especificidade e precisão decorre da “[...] perda de relevância temática das ferramentas documentais” (FUJITA; RUBI; BOCCATO, 2009, p. 103).

As recomendações teóricas defendem que cada biblioteca tenha uma política de indexação própria, para que essa situação seja modificada (GIL LEIVA; FUJITA, 2012), além de reforçarem a importância de se ter uma política de indexação documentada e de se fazer a avaliação periódica da indexação (FUJITA, GIL LEIVA, 2015; DAL’EVEDOVE, FUJITA, 2015). Além de uma política de indexação formalizada e registrada, os estudos também afirmam a necessidade da elaboração de um manual de indexação para garantir avaliações e

modificações periódicas, uma vez que a avaliação é parte fundamental do processo de garantia de consistência da indexação nos sistemas de recuperação da informação.

A política de indexação, por exemplo, garantiria que o metadado assunto fosse revisto, em atenção às especificidades da comunidade usuária, visto que com a incorporação de registros de outros sistemas, o trabalho específico da representação de assunto perdeu espaço, sobretudo, no contexto de bibliotecas universitárias. Com isso, perde-se também a validade no uso dos instrumentos de indexação.

Além dos problemas com a automatização dos catálogos e a importação de registros, surgem novas tecnologias e ambientes informacionais, como a *Web*. A *Web 2.0* foi designada como uma segunda geração de comunidades e serviços, tendo como conceito a "*web* enquanto plataforma", envolvendo *wikis*, aplicativos baseados em folksonomia, redes sociais e tecnologias da informação (BRADY, 2006).

A essência da *Web 2.0* foi tornar os usuários parte do cenário, possibilitando a construção coletiva de conhecimento, não apenas como espectadores, mas atores ativos no processo. A *Web 3.0* ou *Web Semântica*, fruto da evolução tecnológica, trouxe consigo a folksonomia e outros formatos de indexação colaborativa – prática colaborativa que revolucionou a forma de organizar, representar, acessar, usar e compartilhar recursos informacionais digitais.

Nesta perspectiva, o usuário pode participar ativamente do processo de descrição dos objetos digitais, por meio da representação colaborativa da informação. Este processo aparece na literatura sob o designo de etiquetagem social, classificação social, etiquetagem colaborativa, folksonomia e indexação social (SANTOS, 2018).

A folksonomia, termo empregado nesta pesquisa, foi inicialmente usado pelo arquiteto de informação Thomas Vander Wal em 2004, tendo como principal característica a criação de descritores a partir da linguagem usada por pessoas. Esta prática classifica e relaciona as categorias de informações na *Web*, pode estar representada por textos, imagens, áudios, vídeos ou qualquer outro formato, disponibilizando de forma organizada toda essa informação. “Embora a sua característica de liberdade para classificar aponte para a ideia de uma falta de estrutura organizacional, o resultado para quem pesquisa é uma maior facilidade para encontrar termos que as demais linguagens de indexação [...]” (SILVA; BLATTMANN, 2007, p. 207).

A recuperação da informação em catálogos de bibliotecas muitas vezes se torna frustrante para os usuários, não porque o assunto pesquisado não exista, mas porque em alguns casos a indexação realizada não é aprofundada, ficando em um nível superficial, tratando de forma genérica os assuntos dos documentos e não representando de fato o que o usuário busca ou necessita. A folksonomia é um sistema orgânico que se baseia na livre escolha das palavras-chave pelos usuários, na qual eles empregam suas experiências pessoais, sendo esse processo de classificação da informação ou de objetos uma nova forma de indexação para os ambientes digitais (WAL, 2007).

Neste sentido, a folksonomia se apresenta como uma alternativa que pode ser utilizada por profissionais bibliotecários como forma de melhorar a indexação nos sistemas de recuperação de informação e nas plataformas colaborativas. Nesta ótica, os repositórios institucionais, além de se constituírem como ambientes digitais de acesso e recuperação de informações, tornam-se plataformas colaborativas com a adoção do autoarquivamento no processo de descrição de recursos informacionais, uma vez que o autor é quem realiza a representação documental.

Segundo Pereira et al. (2007, p. 2) o ponto crítico da folksonomia, em que o usuário faz uso apenas de suas percepções para atribuir termos representativos aos recursos e objetos digitais, é que “[...] a ausência de controles rígidos é considerada a causa tanto do sucesso, quanto dos pontos fracos apresentados por esta técnica”.

Por não haver o envolvimento do profissional bibliotecário durante o processo de indexação e pelo fato de o usuário não fazer uso de um vocabulário controlado, a qualidade da indexação pode ser questionada.

Sabe-se que quando não existe um controle na indexação algumas inconsistências acabam ocorrendo. Algumas dessas inconsistências decorrem do uso de palavras no singular e no plural. Neste caso, o sistema considera as palavras como sendo distintas. A ocorrência de polissemia é outra questão que também deve ser levada em consideração. Na língua portuguesa uma palavra pode ter vários significados, a depender do contexto em que for usada. Um exemplo é a palavra pena, que segundo Aurélio (2018), pode assumir até doze variações de significado.

Por sua vez, a sinonímia refere-se a palavras distintas que compartilham o significado em algum contexto. Esse problema faz com que no processo de busca, usando um termo

escolhido pelo usuário, o sistema não traga todos os documentos referentes ao assunto. Como exemplo dessa situação temos as palavras casa, moradia e lar para representar o lugar onde se mora. Ao ser executada uma busca utilizando apenas um desses termos, o sistema pode não recuperar todo material existente, somente o termo usado na pesquisa e não o assunto no todo (SILVA, 2010).

Outra ocorrência comum na folksonomia é o uso de palavras egoístas, que são termos criados pelo autor, que só fazem sentido para ele, com intenção de rápido acesso futuro (NASCIMENTO; NEVES 2012). Seu uso pode acarretar uma incorreta recuperação da informação, além de dificultar o acesso ao documento por outros usuários. Como consequência, a recuperação da informação se torna frustrada, podendo trazer prejuízos para atividades que necessitam daquela informação (SILVA, 2010). Neste contexto o profissional bibliotecário indexador tem o desafio de tornar a recuperação por assunto em repositórios digitais um processo satisfatório, sem que ocorram variações de termos.

A folksonomia no contexto de bibliotecas universitárias pode ser uma aliada no sentido de se ter a colaboração dos usuários na representação temática. O que se nota nos estudos é um consenso de que a representação temática advinda da folksonomia dos ambientes colaborativos, repositórios institucionais nos quais se tem participação dos usuários no processo de organização da informação, vem contribuir para o fortalecimento e a melhoria na recuperação da informação. Porém, é consenso também a necessidade de se formalizar uma política de indexação com o estabelecimento de diretrizes de acordo com cada instituição. Nesse sentido, Tartarotti, Dal'Evedove e Fujita (2016 p. 11) evidenciam a necessidade da política de indexação neste ambiente colaborativo “a fim de nortear e permitir a convergência dos processos de indexação tradicional e o processo de indexação social realizado pelos usuários [...]”.

Diante de um cenário em constante evolução, Fujita, Agustin Lacruz e Gomez Diaz (2012) ressaltam a importância de os profissionais bibliotecários desenvolverem competências para analisar, representar e buscar informações, valorizar a indexação enquanto atividade intermediadora do conhecimento que impacta socialmente nos ambientes digitais em evolução. Para as autoras, os bibliotecários ainda não despertaram para a importância de se ter políticas voltadas às tomadas de decisões com relação as melhorias

nos processos de indexação, fazendo com que a recuperação da informação se torne cada vez mais sofisticada e satisfatória.

2.1.3 Avaliação da Indexação

Com o aumento, evolução e uso de novas TIC, principalmente tendo a *web* como fonte de pesquisa, a disponibilização de textos científicos em formatos digitais viabiliza o acesso a uma grande quantidade de informações. O grande desafio para os sistemas de informação agora é se manterem e se ajustarem ao ambiente digital.

Apesar dos benefícios e facilidades do ambiente digital, serviços e produtos informacionais devem corresponder as exigências de indivíduos que transitam na cibercultura. Com efeito, os processos de organização e representação da informação devem responder a essas novas exigências incluindo a indexação, elemento chave para o aumento das possibilidades de visibilidade e encontrabilidade de toda a massa de informação disponibilizada nos diversos meios (PEREIRA, 2015).

Para que isso ocorra é necessário o uso de uma linguagem de indexação que beneficie tanto o indexador na atribuição de termos aos conceitos extraídos do documento, quanto o usuário na orientação das suas estratégias de busca. Esta linguagem de indexação pode ter uma estrutura hierárquica e ter relações entre os termos, favorecendo que indexador e usuário possam definir o grau de especificidade e exaustividade na escolha dos termos, para a representação e para a estratégia de busca (FUJITA, 2008).

Um dos problemas que se observa nos catálogos de assuntos são as diversidades de termos empregados na representação dos documentos que tratam do mesmo assunto. Isso leva a inconsistência da indexação. Se for comparado os termos de indexação estabelecidos por dois ou mais indexadores, para um mesmo documento, certamente serão detectadas diferenças de julgamento dos termos. Essa diferença traz incerteza em qualquer sistema de recuperação (FUJITA; GIL LEIVA, 2014).

A avaliação da indexação em sistemas de informação não é um serviço novo. Lancaster (1996) já havia listado quatro razões para manter o sistema de avaliação em bibliotecas, quais sejam: 1) determinar uma escala de desempenho para comparações futuras; 2) comparar

desempenho entre sistemas ou serviços; 3) evidenciar a importância da avaliação; e 4) detectar os problemas no serviço.

Nos anos 1950, surgiram os primeiros ensaios sobre metodologias de avaliação da indexação e também foi a década considerada mais produtiva em termos de publicações sobre o assunto, atingindo o ápice com os estudos de Zunde e Dexter (1969) (OLSON; WOLFRAM, 2006, p. 605).

Em 1960, os procedimentos para realização da avaliação da indexação começaram a ser praticados nas bibliotecas, porém ainda hoje não é costume realizar esse tipo de atividade de controle ou avaliação como rotina. As razões observadas na literatura especializada decorrem, especialmente, da sobrecarga de trabalho, falta de pessoal ou desconhecimento da existência de tais procedimentos. As avaliações deveriam ser realizadas não somente para controle da indexação, mas também para atualização de mudanças e atualizações que ocorrem nas linguagens de indexação.

Fujita e Gil Leiva (2014) trabalharam a exaustividade e a especificidade como sendo dois elementos da política de indexação que influenciam diretamente na revocação e precisão da recuperação da informação. Por esse motivo, podem ser tomados como indicadores de avaliação da indexação. Entende-se por exaustividade a quantidade de termos usados no ato da indexação. Já a especificidade é a atribuição de um nível mais específico para um termo. Pode-se dizer que quanto mais termos forem atribuídos aos documentos, aumenta-se as possibilidades de revocação na recuperação. Além disso, a quantidade de termos mais específicos melhora a precisão na recuperação.

Entretanto, é preciso que o indexador não se limite a indexar somente o assunto principal do documento e aumente o grau de abrangência do conteúdo na representação. Neste sentido, a melhor opção para garantir a recuperação da informação do conteúdo indexado é um equilíbrio entre exaustividade e especificidade² dos termos. Por isso, a

² Fujita (2012) esclarece que a exaustividade não é somente a quantidade de termos, ela é mais do que isso, é o nível de abrangência de representação dos termos. A representação de conteúdo dos termos pode ser representada por três e ter uma grande abrangência. A exaustividade tem duas faces: a quantidade de termos que representa o conteúdo e a capacidade de representação do conteúdo, podendo ter tanto termos específicos como termos genéricos. A especificidade trabalha também no âmbito de cobertura de representação de conteúdo com grande quantidade de termos bem específicos e termos genéricos, neste caso está sendo usando a especificidade e a exaustividade.

linguagem como instrumento de conversão precisará garantir tanto a exaustividade quanto a especificidade (FUJITA; GIL LEIVA, 2014).

Os referidos autores fazem uma discussão da necessidade de se fazer a avaliação da indexação quando existe o uso linguagem natural usada no sistema ou em sistemas com indexação automática, pois pode acarretar dificuldades na recuperação da informação. Quanto aos catálogos tradicionais das bibliotecas, a necessidade é verificar se as linguagens de indexação estão atualizadas e compatíveis com as necessidades dos usuários no momento da busca. Neste mesmo estudo, tem-se uma fundamentação teórica e metodológica da indexação e dos principais métodos de avaliação, com o objetivo de apresentar recomendações para melhorar a recuperação da informação diante da mudança de comportamento informacional dos usuários no contexto *Web* (FUJITA; GIL LEIVA, 2014).

A qualidade da indexação dos documentos pode ser controlada antes e depois dos documentos serem inseridos nas bases de dados (FUJITA; GIL LEIVA, 2014). Lancaster (2004) cita dois tipos de indexação que acontece antes dos documentos serem inseridos nos catálogos: a indexação revisada por especialistas e a avaliada por simulação da realidade. Enquanto a primeira serve para se evitar alguns tipos de erro do indexador, a segunda propõe uma metodologia mais rigorosa, como uma simulação de avaliação real.

Na literatura, Gil Leiva (2008) é citado como um dos autores que desenvolveu metodologias para encontrar a consistência na indexação estabelecendo comparações entre dois índices, baseado nas fórmulas de Hooper (1965). Segundo este mesmo autor, são três tipos de avaliações que podem ser feitas após a inserção dos documentos nas bases de dados: Avaliação Intrínseca Qualitativa, Avaliação Intrínseca Quantitativa e Avaliação Extrínseca.

A Avaliação Intrínseca Qualitativa analisa os elementos que proporcionam a qualidade da indexação. Ela é realizada por dois especialistas que devem ter conhecimento da política de indexação, da linguagem de indexação e das características dos usuários. Seleciona-se uma quantia significativa de documentos que os especialistas reindexam tomando como referência o texto completo, que permite a revisão da indexação baseada no título ou no resumo do documento. Como resultado, comparam com a indexação original elementos como exaustividade, especificidade, erros e omissão de um termo que pode ou não proceder e não foi extraído (GIL LEIVA, 2008).

A Avaliação Intrínseca Quantitativa é feita com um conjunto de documentos repetindo a primeira indexação com reprodução das mesmas condições para conseguir índices de consistência entre as duas indexações por meio de fórmula matemática. Esse tipo de avaliação deve ser feito periodicamente a cada um ou dois anos, transcorrido esse tempo o documento é reindexação para comprovar se há variações com respeito à primeira indexação (FUJITA; GIL LEIVA, 2014).

A Avaliação Extrínseca se divide em Avaliação Extrínseca Mediante a Interconsistência e Avaliação Extrínseca Mediante a Recuperação.

Avaliação Extrínseca Mediante a Interconsistência: mede o resultado da indexação para compará-lo com a indexação de outras bibliotecas que também tenham indexado o mesmo documento (interconsistência).

Avaliação Extrínseca Mediante a Recuperação: permite a comparação entre duas indexações procedentes de um mesmo sistema de recuperação da informação (intraconsistência) ou dois sistemas diferentes. Neste caso, a comparação pode ser feita tanto entre uma indexação manual e uma automática como entre duas indexações automáticas ou duas indexações manuais (GIL LEIVA, 2008, p. 388).

Para uma indexação de qualidade, o principal elemento a ser considerado é o conceito, uma vez que documentos que tratam de um mesmo conceito (tema) devem estar representados igualmente, para não haver prejuízo na recuperação de informações.

Neste cenário, é importante adotar o uso de políticas de indexação com diretrizes claras e bem estabelecidas acerca dos procedimentos de indexação para que não ocorra incompatibilidades na representação de assunto. Sendo assim, o nível de consistência é medido pela semelhança dos termos, quanto mais semelhantes forem os termos, maior será essa consistência (GIL LEIVA; FUJITA, 2010).

A consistência da indexação foi objeto de estudo em Strehl (1998), Gil Leiva, Rubi e Fujita (2008), Inácio (2012), Piovezan (2015), Tartarotti, Dal'Evedove e Fujita (2016), Bandim e Corrêa (2018), Tartarotti, Dal'Evedove e Fujita (2018) entre outros.

Strehl (1998) conduziu um estudo dedicado a avaliação da consistência da indexação realizada em uma biblioteca universitária com acervo especializado em artes plásticas, música e teatro. Como conclusão, o estudo revelou um baixo nível de consistência. Em pesquisa semelhante, Inácio (2012) investigou a aplicabilidade dos métodos da avaliação da indexação em catálogos *online* de bibliotecas universitárias. Os resultados da pesquisa, além de revelar

a carência de investigações dedicadas à avaliação da indexação em Organização do Conhecimento, confirmam a importância e urgência da aplicação dos métodos de avaliação da indexação em catálogos *online* de bibliotecas universitárias, a fim de contribuir em atividades organizacionais, gerenciais e otimização dos recursos e tempo.

Em Gil Leiva, Rubi e Fujita (2008) buscou-se avaliar os índices de consistência entre 30 bibliotecas universitárias brasileiras das regiões sul e sudeste mediante abordagem da avaliação intrínseca quantitativa. Como resultado, o estudo revelou uma consistência entre 34 e 27 % na indexação de catálogos de bibliotecas universitárias do Brasil. Diante desse cenário, as recomendações são para o cumprimento dos procedimentos de indexação, de modo a ter parâmetros para o processo de indexação em bibliotecas universitárias nacionais. Ainda, o estudo apontou a incompatibilidade entre as linguagens documentais, falta de atualização desses instrumentos, ausência de política de indexação como fatores que contribuíram para o baixo índice de consistência nas indexações.

Fujita e Gil Leiva (2014) mostram uma experiência de avaliação da indexação tendo o propósito de verificar se o sistema de recuperação da informação atende às necessidades do usuário, diante do processo de indexação efetuado pelos profissionais no catálogo da rede de bibliotecas da Universidade Estadual Paulista “Júlio de Mesquita Filho (UNESP).

Tartarotti, Dal’Evedove e Fujita (2017) realizaram estudo diagnóstico da representação temática de livros em catálogos coletivos *online* de bibliotecas universitárias federais da região nordeste, trazendo um panorama mais atual e complementar acerca da qualidade da indexação em sistemas de recuperação da informação no contexto brasileiro. Piovezan (2015), por sua vez, comparou a indexação em catálogos *online* de bibliotecas universitárias com utilização do método de avaliação extrínseca mediante a interconsistência, sendo constatada a pertinência do método aos estudos de política de indexação.

Bandim e Corrêa (2018) investigaram a qualidade da indexação automática por atribuição de artigos científicos brasileiros na área de Ciência da Informação via índice de consistência, em comparação com a indexação intelectual. O estudo concluiu que a indexação automática por atribuição com base no Tesouro Brasileiro de Ciência da Informação (TBCI) e uso do Sistema de Indización Semi-Automático (SISA) resulta em uma consistência média satisfatória em comparação com a indexação de artigos científicos selecionados.

O estudo de Tartarotti, Dal'Evedove e Fujita (2018) dedicado a avaliação da indexação de documentos não-textuais revelou a escassez de estudos neste escopo investigativo. Sendo o ambiente digital uma realidade, este resultado demonstra a necessidade da condução de investigações sobre a consistência da indexação praticada em diferentes ambientes e aplicada em diferentes recursos informacionais.

A partir da contextualização das pesquisas nacionais desenvolvidas no tema, a relevância de discussões teóricas e aplicações práticas acerca da qualidade da indexação fica evidente. Para Fujita e Gil Leiva (2015, p. 65) a indexação e a recuperação de informações são interdependentes, sendo elementos essenciais de um sistema de informação.

Esta interdependência fica evidente quando se demonstram os resultados de estudos de avaliação da indexação que se realizam para verificar, basicamente, se o sistema de informação atende as necessidades de seus usuários na recuperação da informação e, em decorrência, se o processo de indexação representou com especificidade e exaustividade a cobertura temática dos conteúdos documentais utilizando linguagem de indexação compatível.

A falta de discussões e reflexões acerca da avaliação da indexação fica mais evidente quando são considerados os ambientes digitais. Para introduzir a importância de pesquisas dessa natureza em Organização do Conhecimento, a seguir são destacados os repositórios institucionais como ambientes contemporâneos de acesso e recuperação de recursos informacionais digitais e como importantes ambientes de investigação para a área.

2.2 REPOSITÓRIOS INSTITUCIONAIS

Podemos dizer que o princípio dos repertórios estava nas ideias visionárias de Paul Otlet e Henri La Fontaine que por meio do *Repertoire Bibliographique Universel* (Repertório Bibliográfico Universal – RBU), organizado pelo Instituto Internacional de Bibliografia no século XIX, realizou a primeira tentativa de desenvolver uma bibliografia do conhecimento mundial.

Nos estudos de Paul Otlet são reconhecidas as tecnologias usadas atualmente como o início do atual *world wide web*, o desenvolvimento de sistemas complexos de organização para integrar dados bibliográficos. Estes estudos impulsionaram as noções de *hiperlinks*, os sistemas de pesquisa, o acesso remoto e as redes sociais (NHACUONGUE, 2015).

Sabe-se que para uma sociedade ser considerada desenvolvida é primordial que a ciência e a tecnologia tenham uma produção ativa e sua comunicação ocorra por meio de canais de informação com a literatura científica divulgando o conhecimento produzido, de modo a ser consumida por outros pesquisadores e favorecer a condução de novas produções científicas (MUELLER, 1996).

Neste contexto o autor participa ativamente no processo de comunicação, uma vez que ele é o ator principal no consumo, na geração e na disseminação do conhecimento. Para gerar o conhecimento este indivíduo tem a necessidade de buscar outras pesquisas fazendo uso do sistema de comunicação. Posteriormente, ocorre a etapa de divulgação do conhecimento gerado em sua pesquisa. A obrigatoriedade da divulgação de pesquisas científicas, além de outros fatores, em muito é associada ao financiamento com recursos públicos, uma vez que as agências de fomento estão instituindo políticas obrigatórias de acesso aberto.

Assim, o autor participa da comunicação em todo processo de construção de conhecimento, da criação à divulgação. Podemos dizer que “existe um complexo sistema de comunicação científica que permeia as comunidades científicas e instituições acadêmicas” (LEITE e COSTA, 2007 p. 93).

Para esta realidade os repositórios de acesso aberto são os meios mais apropriados para a disseminação do resultado de pesquisas no âmbito das instituições acadêmicas. Nos repositórios institucionais os autores depositam sua produção científica com fins de preservação e divulgação de sua produção, em busca de uma maior visibilidade, citação e impacto, que atualmente é um importante índice de desempenho (FREITAS e LEITE, 2019).

Os periódicos científicos são o principal meio utilizado na disseminação desse conhecimento científico produzido, tornando-se importantes canais de informação. No início dos anos 1990 surgiram entraves que ainda hoje são sentidos, tais como o alto custo das assinaturas dos periódicos, as dificuldades em publicar, o aumento significativo da quantidade de títulos, entre outros. Como alternativa à questão, os repositórios foram planejados e inseridos no ambiente acadêmico. Por pressão de alguns membros da academia que apoiavam a disponibilização de acesso aberto da produção científica proveniente de recursos públicos, esses repositórios deveriam ter como características o acesso aberto, disseminação e visibilidade dos resultados das pesquisas científicas.

Mediante os problemas acima citados surge a chamada “crise dos periódicos”, na qual as universidades e os centros de pesquisas começaram a ter dificuldades em manter as assinaturas das coleções de periódicos, a falta de financiamento e o custo crescente cobrado pelas editoras. Esta crise teve início nos Estados Unidos em meados dos anos 1980, havia alguns indícios de problemas no final dos anos 1970. Com esta situação surgiram novas alternativas para publicação dos periódicos científicos. Diante das novas tecnologias e o uso da Internet surgiu um novo ambiente para a disseminação da literatura científica, passando a serem publicadas eletronicamente (MEDEIROS, 2010).

Essas ações se somaram as iniciativas do acesso aberto. Sendo desencadeada pela comunidade científica por meio da Declaração de Budapeste em 2002, e as Declarações de Berlim e de Bethesda em 2003. Duas estratégias foram propostas, o autoarquivamento e o acesso aberto. A comunicação científica deveria ser publicada pelo próprio autor em repositórios (autoarquivamento) e a segunda proposta era que os artigos deveriam ser publicados em periódicos de acesso aberto (VEIGA, et al. 2016, p. 98).

Sendo assim, repositórios digitais nasceram no contexto da academia e foram relacionados com as iniciativas do *Open Access*. Apesar da origem dos repositórios digitais serem em geral para arquivar, compartilhar e preservar vários tipos de documentos, estes sistemas têm a ideia de preservação e de promoção da produção científica em acesso aberto.

O movimento de acesso aberto provocou alterações no processo de produção da ciência trazendo novos paradigmas para a comunicação da literatura científica, para a forma de publicação e para a avaliação dos artigos científicos. Todo processo de produção e comunicação científica sofreu alterações. Repositórios institucionais e periódicos de acesso aberto são os arcabouços deste novo paradigma (SANTOS; CARDOSO, 2013).

Barreto (2010) descreve repositório digital como uma forma de armazenamento e gerenciamento da informação, um sistema de informação aberto e interoperável com capacidade de armazenar arquivos de diversos formatos por longo período de tempo. Além disso, o referido autor entende que os repositórios devem ser criados como uma arquitetura da informação que forneça subsídios na construção de ambientes informacionais adequados às necessidades dos usuários, permitindo desta forma o acesso e usabilidade satisfatórios aos usuários.

Um repositório é determinado pela aplicação e pelos objetivos aos quais ele se destina. Pode-se ter repositórios para diversos objetivos, como repositórios institucionais, temáticos, de dados de pesquisa, para preservar documentos internos de instituições, dentre outros. Hoje, os repositórios digitais recebem o nome de repositórios institucionais, temáticos, centrais ou departamentais, de acordo com seus propósitos. Como o próprio nome indica, os repositórios institucionais são repositórios que servem ao depósito da documentação da instituição. Esta determinação é normalmente formada por intermédio de políticas institucionais de informação, dando as diretrizes do que será depositado.

Marcondes e Sayão (2009, p. 9) definem o repositório institucional como “[...] uma biblioteca digital destinada a guardar, preservar e garantir livre acesso, via internet, à produção científica no âmbito de uma dada instituição”. Verifica-se consonância dessa conceitualização com a apresentada por Leite (2009) ao entender repositórios institucionais como um tipo de biblioteca digital, porém nem toda biblioteca digital é considerada um repositório institucional. Na concepção do autor, repositório institucional é

Um serviço de informação científica - em ambiente digital e interoperável - dedicado ao gerenciamento da produção intelectual de uma instituição. Contempla, por conseguinte, a reunião, armazenamento, organização, preservação, recuperação e, sobretudo, a ampla disseminação da informação científica produzida na instituição (LEITE, 2009, p. 21).

Para Kuramoto, “[...] o desenvolvimento de um repositório institucional não depende apenas de fatores tecnológicos, mas principalmente de fatores relacionados à interoperabilidade humana”. Portanto, o sistema de comunicação científica demanda mecanismos eficientes que comprovem a execução dos processos de produção e gestão do conhecimento científico e a atuação dos profissionais da informação neste processo é fundamental para seu o êxito.

De modo geral, repositórios Institucionais são usados para o registro, disseminação e preservação da produção do conhecimento científico e acadêmica de uma instituição e seu acervo é composto por artigos de periódicos científicos, teses, dissertações, relatório de pesquisa de pós-doutorado, trabalhos apresentados em eventos científicos que tenham sido revisados por pares (VEIGA, et al. 2016).

Repositório institucional é o encontro de repositórios temáticos, sob a responsabilidade técnica e administrativa de uma instituição ou órgão. Portanto, esse tipo de

repositório é multidisciplinar e possui uma variedade de tipos de documentos superior a um repositório temático. Além de agregar o conjunto de informações relativas e à instituição, possuem serviços relacionados à organização, divulgação e acesso ao conteúdo digital (CAFÉ, 2003).

Os repositórios institucionais evoluíram no ambiente universitário para que seus pesquisadores disponibilizassem resultados de pesquisa realizadas a partir de coleções digitais dos departamentos e faculdades. Eles dão visibilidade à produção acadêmica e a agrupam em um sistema de informação que fornece acesso de longo prazo a um preço inferior à publicação em periódicos tradicionais, além de demonstrar seu valor científico, cultural, social e econômico (CROW, 2002 apud SANTAREM, 2010).

Quanto às tecnologias disponíveis para se implantar um repositório institucional, os *softwares* mais usados são o DSpace e o E-Prints desenvolvidos em parceria como o *Massachusetts Institute of Technology – MIT* e a *Hewlett-Packard – HP*.

Fiche, Moura e Maia (2017, p. 61) identificam o DSpace como sendo o gestor de repositórios mais utilizado no Brasil, Portugal e Espanha, sendo obtido os seguintes dados: Portugal – adesão total das universidades; Espanha – adesão de mais de 80% pelas universidades espanholas; e Brasil – o DSpace é implementado pela Universidade de São Paulo, Universidade Federal do Paraná e pelo Instituto Brasileiro de Informação em Ciência e Tecnologia, com adesão de mais de 7% em território nacional, além de ser aderido por outras organizações e entidades fora do ambiente acadêmico e científico.

O DSpace permite ser organizado conforme a estrutura organizacional da instituição, com adoção do padrão de metadados *Dublin Core* e uma interface *Open Archives Initiative – OAI* para colher os metadados armazenados. Permite também disponibilizar o modo de autoarquivamento.

No Brasil, este crescimento foi acelerado por um projeto lançado pelo IBICT-FINEP/PCAL/XBDB, que possibilitou a implantação de vários repositórios institucionais em diversas universidades e instituições de pesquisa. Com a disseminação e consequente conscientização do Movimento de Acesso Aberto à informação científica, diversas instituições brasileiras de ensino superior passaram a desenvolver e a manter repositórios institucionais de acesso aberto.

Em 2009, o IBICT implementou uma política visando à ampla adoção das tecnologias de repositórios nas instituições de ensino superior. Foram distribuídos *kits* tecnológicos compostos por um servidor com os *softwares* DSpace e SEER instalados, assim como a oferta de treinamentos para as instituições atendidas.

A UFSCar participou do edital e foi contemplada com o *kit*, dando início ao processo de implantação do seu repositório. Contudo, esse processo não dependia tão somente dos profissionais bibliotecários para que fosse implantado. Foi um longo processo de conscientização de todo corpo administrativo, docente e técnico da Universidade, pois é preciso que se estabeleçam políticas para a implantação de repositórios institucionais (REPOSITÓRIO, 2015).

No entanto, vários projetos-piloto de repositórios surgiram na UFSCar, em escala modesta e usualmente para atender um dado curso ou projeto específico, o que demonstrou a necessidade de criar condições para que o RI/UFSCar passasse a operar.

O RI/UFSCar nasceu com a coleção de teses e dissertações migradas do sistema da Biblioteca Digital de Teses e Dissertações (BDTD), como a maioria dos repositórios das instituições brasileira de ensino superior. Atualmente o RI/UFSCar atua com a modalidade de autoarquivamento ainda para esta mesma coleção, e já se encontra em estudo o povoamento com outras coleções como artigos, capítulos de livros, arquivos administrativos e outras coleções.

2.2.1 Autoarquivamento em Repositórios Institucionais

O autoarquivamento a princípio foi proposto como uma prática universal por Stevan Harnad em sua publicação *online* de 1994 "*Subversive Proposal*" (mais tarde publicada na *Association of Research Libraries*). Embora cientistas da computação estivessem

praticando autoarquivamento em arquivos FTP anônimos desde a década de 1980 no *CiteSeer*,³ os físicos faziam isso desde o início dos anos 90 na *Web* com *arXiv*⁴.

Conforme Swan (2005), o autoarquivamento não é uma alternativa para publicação em periódicos científicos, mas uma atividade complementar em que o autor publica o seu artigo científico em qualquer periódico e depois autoarquiva uma cópia em um repositório institucional.

A implantação de um repositório institucional em uma determinada comunidade se inicia a partir de uma política institucional de informação na qual deve conter a política de publicação de autoarquivamento, que assinala a conscientização da necessidade de se criar uma cultura de postagem, passando pela implementação técnica do repositório e finalizando com a efetiva utilização do mesmo dentro da comunidade científica (SANTAREM, 2010).

Os repositórios institucionais só conseguem ser criados e manterem-se ativos com o estabelecimento de uma política institucional de informação. É ela, em princípio, que vai estimular os pesquisadores de uma instituição a fazer o depósito de sua produção científica. Esta prática é oriunda da década de 1990, como a Iniciativa dos Arquivos Abertos OAI e o Movimento de Acesso Aberto OAM (*Open Access Movement*), visando promover modelos eficientes de armazenamento, disseminação, visibilidade e acesso aos conteúdos científicos.

Assim como no caso dos repositórios institucionais, o movimento do autoarquivamento expandiu-se fundamentado por manifestos como a Declaração de Budapeste, proposta na reunião da BOAI (*Budapest Open Access Initiative*), em 2002; e as Declarações de Bethesda e de Berlim, em 2003 (WEITZEL, 2006).

³ O CiteSeer foi criado pelos pesquisadores Lee Giles, Kurt Bollacker e Steve Lawrence em 1997, enquanto estavam no NEC Research Institute (agora NEC Labs), em Princeton, Nova Jersey, EUA. O objetivo da CiteSeer era rastrear e colher ativamente documentos acadêmicos e científicos na web e usar indexação de citações autônoma para permitir a consulta por citação ou por documento, classificando-os pelo impacto da citação. GILES, C. L.; BOLLACKER, K. D.; LAWRENCE, S. CiteSeer: an automatic citation indexing system. *Proceedings of the Third ACM Conference on Digital Libraries*. Pittsburgh, Pennsylvania, USA 1998, p. 89–98.

⁴ O arXiv tornou-se possível graças ao compacto formato de arquivo TeX, que permitia que os artigos científicos fossem facilmente transmitidos pela Internet e processados pelo cliente. Por volta de 1990, Joanne Cohn começou a enviar por e-mail preprints de física para os colegas como arquivos TeX, mas o número de artigos sendo enviados logo encheu as caixas de correio. Paul Ginsparg reconheceu a necessidade de armazenamento central e, em agosto de 1991, criou uma caixa de correio central armazenada no Laboratório Nacional de Los Alamos (LANL), que podia ser acessada de qualquer computador. Modos adicionais de acesso foram logo adicionados: FTP em 1991, Gopher em 1992, e a World Wide Web em 1993. O termo e-print foi rapidamente adotado para descrever os artigos. GINSPARG, P. It was twenty years ago today ... arXiv. New York, 2011. https://www.researchgate.net/publication/51936449_It_was_twenty_years_ago_today.

As ferramentas e procedimentos fundamentais que apoiaram a OAI são o autoarquivamento pelos autores, os metadados padronizados para descrição dos documentos, o acesso livre à produção científica e o Protocolo OAI, que viabilizou a interoperabilidade e o compartilhamento das informações. Essas ferramentas garantem maior flexibilidade, interação, auto-gestão e auto sustentabilidade das publicações *online*, dos repositórios digitais e dos provedores de serviços (WEITZEL, 2006).

Foram três princípios básicos que regeram a filosofia do acesso aberto em repositórios digitais, pelo OAI manifestado na Convenção de Santa Fé em 1999, sendo elas o autoarquivamento, a revisão por pares e a interoperabilidade. O autoarquivamento está baseado na postagem espontânea, pelo autor de seu texto em um repositório digital. A revisão por pares no caso específico dos repositórios de *e-prints* difere em relação aos sistemas que não utilizam o *software*. Nesta estrutura, o processo inclui os comentários espontâneos de pesquisadores cadastrados no repositório e a interoperabilidade possibilita a interconexão de vários sistemas e redes, a definição de sua arquitetura e usabilidade, os mecanismos necessários para o estabelecimento da comunicação científica integrar citantes e citados, serviços de alerta e de recuperação da informação (WEITZEL, 2006).

Na literatura nacional e internacional existem várias discussões sobre repositórios institucionais. Porém, sobre o autoarquivamento as pesquisas são escassas, podendo-se citar dois trabalhos que estudam a análise das políticas de autoarquivamento e um estudo exploratório nos repositórios brasileiros com autoarquivamento, sendo eles o apresentado por Melero (2007), Assis (2013) e Veiga (2015).

O trabalho de Assis (2013) verificou a existência de políticas de autoarquivamento dos repositórios institucionais brasileiros e portugueses, registrados no *Directory of Open Access*. Dos 70 repositórios encontrados apenas 26 repositórios eram considerados institucionais e 6 apresentaram alguma política na página inicial do sistema. A pesquisa identificou que dos 28 repositórios institucionais portugueses encontrados, 10 apresentaram alguma política na página inicial do sistema. A autora constata que no contexto português, a grande maioria dos repositórios apresentam políticas de autoarquivamento. A Universidade do Minho possui o maior índice de autoarquivamento, e esse fato está ligado à forma como a política foi implementada na instituição.

No Brasil, apenas duas instituições identificadas adotam o autoarquivamento, porém apenas uma apresenta esse processo em sua política. Nos repositórios brasileiros, as políticas encontradas estão denominadas como institucionais, de informação ou de acesso aberto. Apenas no Repositório Institucional da Universidade Federal da Bahia a política faz referência e adota o autoarquivamento. Ainda, no repositório citado disponibiliza para os seus pesquisadores manuais para a realização do depósito no sistema (ASSIS, 2013).

A autora supracitada considera que para se alcançar o autoarquivamento e a maximização de documentos nos repositórios institucionais, a instituição deve definir uma política de obrigatoriedade de autoarquivamento. Uma das formas de garantir a participação da comunidade científica da instituição é fazer com que os pesquisadores entendam a importância do processo de autoarquivamento para o desenvolvimento de novas pesquisas (ASSIS, 2013).

Veiga (2015) realizou um levantamento dos repositórios institucionais brasileiros no qual detectou que a maioria dos sistemas (64%) tem implementado a função de autoarquivamento. Porém 36% dos repositórios institucionais não habilitam esta função para o autor depositar sua produção científica.

A baixa adesão do autoarquivamento é um fenômeno que tem se repetido em vários países. Os motivos para esta baixa adesão podem ser diversos. O Brasil, apesar de estar em 7º lugar em número de repositórios, somente a metade possui coleção de artigos e em sua maioria foram migrados de outros sistemas. Os fatores que estimulam o autoarquivamento devem ser estudados no âmbito de cada instituição e área do conhecimento. A redução de barreiras, criação de incentivos e principalmente uma política de repositório deve ser estabelecida nas instituições para que este instrumento tenha a possibilidade de sucesso no contexto nacional (VEIGA, 2015).

A literatura sobre o acesso aberto e repositórios institucionais é unânime ao indicar que o principal entrave em relação ao seu sucesso tem sido a falta do depósito pelos autores, apesar da constatação de seus inúmeros benefícios. Desse modo, parece ser senso comum que nos estudos feitos sobre autoarquivamento as diretrizes devem vir da ação conjunta das instâncias superiores para que a mesma ganhe força.

Apesar do autoarquivamento ser recomendado na Via Verde, ainda não é uma prática comum entre os autores. Estudos apontam que cerca de 15-20% dos 2,5 milhões de artigos

publicados anualmente em todo o mundo estão sendo autoarquivados pelos seus autores (GARGOURI et al., 2010; HAJJEM, HARNAD; GINGRAS, 2005 apud GOMES, 2017). Algumas questões são apontadas para a não realização do autoarquivamento, sendo uma delas a preocupação com a qualidade dos artigos (SWAN, 2005).

A Universidade de Southampton foi pioneira na adoção de política de mandato de autoarquivamento, iniciando no Departamento de Eletrônica & Ciência da Computação, em janeiro de 2003 (MELERO, 2007; XIA, et al. 2012; GARGOURI, et al. 2010 apud ASSIS, 2013). Esta mesma conduta foi observada em outras universidades como na Austrália, na Universidade da Tasmânia e na Universidade Tecnológica de Queensland. Universidades europeias como a do Minho, de Liege e Zurique seguiram o modelo de obrigatoriedade (MALERO, 2007).

Melero (2007) cita algumas instituições que também recomendam ou determinam o autoarquivamento e indica que outras universidades optaram por estimular ou recomendar o autoarquivamento, como a Universidade de Lund, Universidade Humboldt, Estocolmo e Oslo.

No início do ano de 2019 a Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP) instituiu a “Política para Acesso Aberto às Publicações Resultantes de Auxílios e Bolsas FAPESP”. De acordo com esta política toda produção científica com apoio financeiro da FAPESP, total ou parcial em projetos e bolsas de pesquisas deverão obrigatoriamente depositar os resultados em repositórios de acesso aberto, no qual a busca possa ser feita por qualquer pessoa acessando a *Web*. O objetivo desta política é de ampliar a visibilidade das pesquisas nacionais e aumentar o impacto (PORTARIA CTA nº 01/2019).

Constata-se que os resultados dos estudos sobre repositórios institucionais se referem ao baixo número de instituições que contemplam políticas institucionais formalizadas, principalmente no âmbito da questão do autoarquivamento, com iniciativas para estimular seus pesquisadores a depositarem seus trabalhos e capacitação dos profissionais diante das especificidades dos repositórios institucionais.

Algumas iniciativas políticas foram consideradas importantes para impulsionar a trajetória do movimento de acesso aberto em território nacional, tais como o Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica no Brasil, a Declaração de Salvador Sobre Acesso Aberto, a Carta de São Paulo, datadas de 2005, além da Declaração de Florianópolis, publicada em 2006 (KURAMOTO, 2006).

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) assumiu o papel como principal entidade, em escala nacional, para articulações políticas e da promoção de ações voltadas para a implantação de repositório nas instituições de ensino e pesquisa. Um dos objetivos é promover a competência e o desenvolvimento de recursos e infraestrutura de informação científica e tecnológica no Brasil para a criação de repositórios (VEIGA, 2015).

A política de informação ainda hoje é um problema persistente, este problema é relatado em vários estudos. A política institucional para os repositórios é uma das principais garantias para a sua criação e manutenção. Hunter e Day (2005), Santarém (2010), Assis, (2013), Veiga, (2015) e outros estudiosos concordam que os repositórios institucionais devem elaborar formalmente suas políticas de desenvolvimento. A política pode incluir informações sobre os fluxos para submissão de trabalhos, incluindo os direitos de propriedade intelectual, acesso e sustentabilidade do repositório (CARIBE, 2008).

Baggio, (2017) fez um panorama das Universidades Federais Brasileiras, no qual foram levantadas 63 instituições. Neste estudo, o autor constatou que 37 das instituições possuíam repositórios implantados e disponibilizados com acesso *online*, mas somente 20 possuíam políticas implantadas e disponibilizadas, o que confirma que a maior dificuldade ainda é com a construção de políticas de informação para os repositórios institucionais.

Conforme constatado por Caribé (2008), vários são os problemas pertinentes a implantação e manutenção de um repositório institucional, dentre os quais tem-se o desenvolvimento e gestão de coleções; garantia de direitos autorais; sustentabilidade das coleções a longo prazo; duplicação de conteúdo, com documentos submetidos em vários repositórios, submissão de versões diferentes do mesmo conteúdo. De modo complementar, as estratégias de desenvolvimento e gerenciamento de coleções devem levar em consideração que alguns repositórios institucionais podem se reestruturar ou fechar, como também podem permitir que indivíduos que submeteram documentos possam retirá-los.

Nesse sentido, acentua-se a importância das políticas nacionais e institucionais para a criação e povoamento de repositórios, uma vez que são essas políticas que garantem o depósito da produção científica das instituições, consolidando o acesso aberto no âmbito nacional e proporcionando maior visibilidade à produção científica das instituições brasileiras.

No Brasil como em outros países, os gestores de repositórios institucionais se deparam com os mesmos problemas, porém, alguns repositórios nacionais se destacaram. Um bom

exemplo é o Repositório Digital da Universidade Federal do Rio Grande do Sul (LUME/UFRGS), que apresenta grande volume de documentos e políticas bem estabelecidas.

Nos repositórios institucionais brasileiros que possibilitam o autoarquivamento observa-se uma baixa adesão do autor. Dos 110 repositórios institucionais brasileiros listados no site do IBICT, apenas 22 estão cadastrados no *Registry of Open Access Repository Mandates and policies* (ROARMAP), dos quais apenas 8 apresentam política de autoarquivamento. Destes, dois apresentam uma política em que o autoarquivamento é opcional podendo ou não ser depositada pelo autor, com status de política mista em que tanto a biblioteca como o autor podem fazer o arquivamento.

A baixa adesão do autor por esta forma de compartilhamento de informação é um fenômeno que tem se repetido em vários países. Os motivos podem ser diversos. O número baixo de políticas institucionais mandatórias e a inexistência de uma política mandatória nacional cooperam para a baixa adesão, porém, como já mencionado, estudos já demonstram que apesar de políticas mandatórias ainda assim há resistência (BAGGIO, 2017).

No quadro a baixo vemos as Instituições que adotam o autoarquivamento e disponibilizam suas políticas.

Quadro 2 - Instituições brasileiras com política de autoarquivamento registradas no ROARMAP

Repositório Institucional	Política de autoarquivamento	Resolução / Portaria	Ano de implantação da política
Universidade Federal do Ceará - UFC	Mista	Art. 3º O depósito de informações referentes à produção técnico-científica será registrado no RI pela comunidade universitária da instituição e seu acesso será livre nos contextos nacional e internacional.	2011
Universidade Federal de Goiás - UFG	Autoarquivamento	Art. 5º O RI-UFG será povoado de forma descentralizada, cabendo sua realização pelas coordenações de curso, coordenações de trabalho de conclusão de curso, direção de departamento (onde existir) ou Unidade Acadêmica. § 1º O RI-UFG permitirá aos autores a submissão de seus próprios documentos, dentro das normas estabelecidas pela comissão gestora e divulgadas na página eletrônica institucional. § 4º As bibliotecas que compõem o Sibi-UFG, de modo a facilitar o povoamento do RI-UFG, poderão promover o registro da produção intelectual da UFG, mediante a autorização do	2014

		autor da referida produção, efetuando a entrada de cada documento no RI ou importando dos dados já registrados em outros repositórios.	
Universidade Federal de São Paulo - UNIFESP	Autoarquivamento	Art. 9º A submissão realizada pelo autor, coautor ou por um mediador compreenderá o registro dos metadados, aceitação da Licença <i>Creative Commons</i> e da Licença de Distribuição Não Exclusiva, sendo de sua total responsabilidade as questões legais de depósito e de permissão, juntamente aos editores e instituições envolvidas na produção da obra.	2017
Universidade Federal do Rio Grande do Sul - UFRG	Autoarquivamento	Art. 12º Os metadados são coletados do Sistema de Automação de Bibliotecas da UFRGS (SABi) e dos sistemas da Universidade. No caso de autoarquivamento, o registro dos metadados é realizado pelo autor ou responsável pela comunidade, sob a orientação da equipe técnica do Lume,	2010
Universidade Federal Rural do Amazônia - UFRA	Autoarquivamento	Art. 23º. É de responsabilidade das comunidades: – Decidir quem poderá autoarquivar documentos na comunidade, essa informação deve ser explícita nas políticas da comunidade.	2018
Universidade Fernando Pessoa - UFP	Mista	Depósito de documentos. a) O depósito de documentos no Repositório Institucional da UFP deve ser efetuado através das opções: arquivo e auto arquivo. Sendo esta última a preferencial. No arquivo o depósito de documentos é realizado pela Biblioteca. No auto arquivo o depósito de documentos é realizado pelos próprios autores. b) No caso da produção acadêmica (tese, dissertação, etc.) O arquivo é sempre realizado pela Biblioteca, após autorização formal dos autores.	2012
Fundação Getúlio Vargas - FGV	Autoarquivamento	Os alunos deverão utilizar o procedimento denominado de “Autoarquivamento”, visando o depósito no website da FGV, na página “Repositório Digital”, de suas teses e dissertações. O aluno realizará o referido procedimento em conformidade com a Norma de Procedimentos, expedida pela FGV, de nº NP-40.001.00, de 29.06.2011;	2018
Fundação Oswaldo Cruz - FIOCRUZ	Autoarquivamento	Parágrafo Único – Será adotado preferencialmente o autoarquivamento, quando os autores preencherem os metadados e farão o depósito do objeto digital. As bibliotecas e os gestores das comunidades atuarão como facilitadores no autoarquivamento.	2014

Fonte: ROARMAP (2019)

Em relação à qualidade dos metadados, uma análise pormenorizada das oito políticas listadas acima revela apenas a da Universidade Federal Rural do Amazônia (UFRA) traz indicações mais explícitas sobre a qualidade dos metadados. Conforme o documento, a

publicação do documento somente é realizada após a verificação dos metadados como etapa fundamental da padronização dos processos de busca e recuperação da informação. As outras políticas mencionam a preocupação em seguir os padrões de metadados nacionais e internacionais, ou com algum tipo de validação antes do depósito, porém nada sobre uma atenção maior com o metadados assunto ou com orientações aos usuários de como proceder a indexação nos repositórios.

A Universidade Federal de São Carlos sinalizava a necessidade de implantar o seu Repositório Institucional, conforme consta em documentos institucionais, como no Plano de Desenvolvimento Institucional de 2013. Entretanto, a concretização se deu pela Resolução ConsUni nº 819, de 26 de agosto de 2015. Algumas diretrizes previstas no Plano de Desenvolvimento Institucional eram:

Estimular o envolvimento dos servidores mais experientes na oferta de atividades de capacitação e de transferência de conhecimentos aos servidores recém-contratados, por meio de mapeamento de atividades de gestão e disponibilização de repositórios de fluxogramas e tutoriais em prol de melhorias;

Implementar e manter o Repositório Institucional da UFSCar, em acesso aberto, definindo critérios e mecanismos de certificação/validação dos conteúdos a serem publicados e de preservação da propriedade intelectual, de maneira a disponibilizar de forma integrada os produtos de pesquisa, ensino e extensão da Instituição e garantir a preservação dessa produção e da memória institucional;

Integrar o repositório institucional digital da UFSCar a outras ferramentas;

Institucionais de acompanhamento, gestão e difusão da produção científica e cultural da Universidade;

Criar mecanismos de incentivo ao registro institucional da produção científica e cultural da UFSCar por parte de toda a comunidade acadêmica, inclusive por meio de critérios que considerem esse registro em processos de progressão na carreira e

Criar condições para o funcionamento de um sistema de arquivos/acervos da UFSCar, com política para gestão de arquivos temporários e permanentes e aprimoramento dos mecanismos de depósito, preservação e acesso a esses arquivos administrativos, científicos, tecnológicos e artístico-culturais. (REPOSITÓRIO, s/p web 2015).

Foi designado pela reitoria um Grupo de Trabalho para elaboração da proposta de implantação do RI/UFSCar. O trabalho se deu de forma colaborativa e interdisciplinar, tendo como resultado deste trabalho a elaboração de um Projeto que serviu de proposta para a criação do repositório institucional (REPOSITÓRIO, 2015).

Em 2017, com o aumento de volume da produção de Teses e Dissertações, falta de profissionais e, sobretudo, pensando nas próximas etapas de implantação da coleção de artigos, capítulo de livros, trabalhos provenientes de eventos científicos e outros materiais semelhantes, houve a necessidade de se habilitar o autoarquivamento no RI/UFSCar. Em atenção aos princípios do manifesto de acesso aberto, o RI/UFSCar passa a cumprir com as normas estabelecidas para um legítimo repositório de acesso aberto. Com a disponibilização do autoarquivamento foi necessário a realização de uma política mínima, além de tutorial, manual e o fluxo do autoarquivamento para a orientação dos usuários e a disponibilização na página do RI/UFSCar.

Porém, o desenvolvimento de um repositório institucional perpassa por várias questões para a sua sustentabilidade, não só questões de políticas de informação, mas questões de preservação digital, gestão da informação, usabilidade, suporte tecnológico, dentre outras. Com efeito, a gestão de repositórios institucionais demanda esforços e o estabelecimento de políticas próprias. A partir da criação do repositório institucional em uma instituição, criam-se várias iniciativas para manter o seu funcionamento.

Por se tratar de um projeto institucional, Veiga (2015, p. 44) salienta que a criação, implantação e manutenção de um repositório envolve diversas áreas, sendo um trabalho coletivo, em que

[...] a visão do gestor do repositório e da própria instituição é fundamental para o avanço do acesso aberto pela via verde no Brasil. Os dirigentes das instituições e os profissionais que ocupam a função de gestores dos repositórios precisam estar alinhados a filosofia do acesso aberto e serem os principais evangelistas do autoarquivamento.

Neste contexto, a biblioteca universitária deve fazer parte no processo do desenvolvimento do repositório institucional, devido ao conhecimento especializado no tratamento e na gestão da informação. Além disso, os bibliotecários são os profissionais adequados para manter a qualidade do conteúdo digital depositado pelos autores e na elaboração de metadados de acordo com padrões internacionais (VEIGA, 2015).

Sendo contextualizado o tema e objeto de estudo a partir da literatura nacional e internacional de Ciência da Informação, a seguir são descritos o método de pesquisa e o conjunto de ações que compõem o desenvolvimento da parte empírica desta dissertação.

3 PROCEDIMENTOS METODOLÓGICOS

A presente pesquisa, segundo sua finalidade é de caráter teórico-aplicada, com objetivos de natureza exploratória e descritiva, possuindo procedimentos para a coleta de dados de ordem bibliográfica, visando a realização de uma análise qualitativa.

Esta escolha metodológica amplia a probabilidade de serem obtidos resultados mais completos e confiáveis, com vistas à validação científica. Diante dos diferentes aspectos presentes no objeto de estudo aqui contemplado, considera-se que “[...] as diferentes perspectivas metodológicas complementam-se no estudo do assunto, um processo que é entendido como a compensação complementar das deficiências e dos pontos obscuros de cada método isoladamente” (FLICK, 2004, p. 274).

Para a execução da pesquisa, a primeira etapa foi composta por um amplo levantamento bibliográfico na literatura nacional e internacional de Biblioteconomia e Ciência da Informação que contemplassem questões acerca do autoarquivamento, repositórios institucionais, prática de indexação em repositórios institucionais, representação de assunto em repositórios institucionais e outros temas correlatos. Para tanto, foram elegíveis fontes primárias, secundárias e terciárias, tendo como resultados manuais técnicos, artigos, teses e dissertações, capítulos de livros e livros. O levantamento bibliográfico foi realizado na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI), P@rthenon, Portal de Periódicos da Capes, Google Acadêmico e na *Web of Science* a partir da aplicação de diferentes termos e estratégias de pesquisa em português, inglês e espanhol, todos no escopo da pesquisa.

Após a coleta dos materiais, realizou-se a análise e separação dos documentos mediante a pertinência ao tema, idioma e atualidade dos documentos. A leitura atenta e cuidadosa dos documentos viabilizou a construção dos referenciais teóricos para elucidação do problema de pesquisa e criou condições para a condução da pesquisa empírica. Esta, por sua vez, decorre da avaliação da qualidade da indexação de Teses e Dissertações depositadas por autoarquivamento no RI/UFSCar.

A metodologia da avaliação da indexação é o método utilizado para analisar o processo de indexação praticado e a qualidade do trabalho desenvolvido. A consistência da indexação

é um elemento característico do processo de indexação, bem como dos resultados da representação e recuperação por assuntos em sistemas de informação.

Tartarotti, Dal'Evedove e Fujita (2016, p. 2) salientam que:

No âmbito do tratamento temático da informação documental, destaca-se a relevância da avaliação do processo de indexação como mecanismo de análise da qualidade dos produtos e serviços gerados no contexto de bibliotecas universitárias. Por meio da metodologia quantitativa da AI, é possível obter um *feedback* da representação da informação com o intuito de aprimorá-la de acordo com a demanda do sistema, contribuindo para a aplicabilidade e sistematização da atividade de indexação em bibliotecas universitárias.

As investigações dedicadas à avaliação da indexação são de natureza teórica, atentando-se para a pertinência e atualização de instrumentos e diretrizes destinados ao processo de indexação, assim como aos aspectos relativos à atuação profissional e especificidades dos sistemas de recuperação de informações, ou de cunho prático, com adoção de equações para a quantificação da consistência. Isto é, à obtenção de índices de consistência mediante a aplicação de fórmulas matemáticas diversas de semelhança entre indexações (GIL LEIVA; RUBI; FUJITA, 2008).

Na literatura especializada de Indexação, o nível de concordância ou consistência consiste no “grau de semelhança na representação da informação documental de um documento por meio de termos de indexação selecionados por um ou vários indexadores, resultando em um índice de consistência” (GIL LEIVA; RUBI; FUJITA, 2008, p. 234).

Em atenção aos propósitos da pesquisa, considera-se que a avaliação intrínseca qualitativa seja o método de avaliação da indexação mais adequado. Na prática, Fujita e Gil Leiva (2014, p. 57) explicam que “trata-se da indexação de um conjunto de documentos repetindo, na medida do possível, o entorno no que se produziu a primeira indexação [...] para conseguir índices de consistência entre as duas indexações por meio de fórmulas matemáticas”. O entorno indicado pelos referidos autores contempla, dentre outros elementos, os próprios indexadores, a política de indexação, a linguagem de indexação, as condições de trabalho e os usuários do sistema.

A avaliação intrínseca quantitativa se subdivide em dois tipos (GIL LEIVA, 2008), quais sejam:

- Avaliação Intrínseca Quantitativa Mediante a Interconsistência ou consistência interindexador: serve para comparar a indexação de um mesmo documento realizada por dois ou mais indexadores.
- Avaliação Intrínseca Quantitativa Mediante a Itraconsistência ou consistência intraindexador: compara a indexação de um mesmo documento realizada pelo mesmo indexador em momentos diferentes.

Para a obtenção do índice de consistência, as pesquisas na área utilizam-se a equação matemática apresentada por Hooper (1965) e adaptada por Gil Leiva (2008), a qual permite a comparação dos índices de consistência, conforme segue:

Figura 1 – Fórmula para obter o índice de consistência entre duas indexações

$$C_i = \frac{T_{co}}{(A + B) - T_{co}}$$

Onde,
C_i = Índice de Consistência
T_{co} = Número de termos comuns nas duas indexações
A = Número de termos usados na Indexação A
B = Número de termos usados na Indexação B

Fonte: Gil Leiva (2008).

Por se tratar da avaliação da qualidade da indexação de documentos provenientes da prática do autoarquivamento em repositório institucional, o uso nesta pesquisa da avaliação intrínseca quantitativa mediante a interconsistência será adaptado, em que a comparação da indexação de um mesmo documento não se dará a partir de dois indexadores, mas entre o autor e o profissional bibliotecário. Assim, a fórmula para obter o índice de consistência entre duas indexações será a mesma, apenas sendo considerado como *Indexação A* os termos

provenientes do autoarquivamento⁵ e a *Indexação B*, termos provenientes da indexação realizada pelo bibliotecário indexador.

Esta modalidade é pioneira nos estudos em Ciência da Informação, o que demonstra o ineditismo da pesquisa e reforça o caráter inovador aplicado aos estudos de avaliação da indexação, em atenção aos novos ambientes de acesso e recuperação de informações e as concepções da folksonomia.

Para tanto, elegeram-se 46 teses e dissertações, sendo 22 provenientes do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade e 24 do Programa de Engenharia de Produção. Para a escolha dos Programas de Pós-Graduação levou-se em consideração o caráter multidisciplinar, de modo que a questão do indexador ser especialista no assunto não fosse um aspecto negativo na avaliação da consistência entre as indexações analisadas na pesquisa. A seleção dos dois programas para a composição do corpus da pesquisa se deve ao baixo número de materiais disponíveis para análise, caso a coleta de dados ocorresse em apenas um programa, visto que o processo de autoarquivamento é recente.

Os documentos de análise da consistência da indexação tiveram o recorte no ano de 2017, ano de início da modalidade de autoarquivamento no RI/UFSCar. Os metadados dos documentos selecionados foram encaminhados para um bibliotecário do Sistema Integrado de Bibliotecas da Universidade Federal de São Carlos – SIBi/UFSCar, especialista no processo de indexação de assuntos e conhecedor dos procedimentos de indexação da instituição. O bibliotecário indexador reindexou cada uma das 46 teses e dissertações selecionadas, tomando como referência o texto completo.

A partir dos resultados obtidos, elaborou-se quadros por programa de pós-graduação contendo os dados de autoria, título, palavras-chave do autor e palavras-chave do bibliotecário especialista, a saber:

⁵ Para efeitos desta pesquisa, será usado como referência o tipo de Indexação orientada pelo autor, na qual o próprio autor do documento utiliza termos que são habitualmente compreendidos e geralmente aceitos em sua comunidade de pesquisa.

Quadro 4: Resultados da indexação por extração de teses e dissertações do PPGCTS.

N.	Título	Autor	Palavras-chave do Autor	Palavras-chave do Bibliotecário Especialista
1	Da linguagem à memória: uma análise dos gêneros discursivos desenvolvidos na história da TV Cultura	Silva, Thiago Rodrigues da	1 Gêneros discursivos 2 Linguagem e cultura 3 TV Cultura de São Paulo 4 Memória	1 TV Cultura de São Paulo 2 Televisão – estações educativas 3 Televisão pública - Análise do discurso
2	A construção dos discursos que circulam entre a Ciência, a Tecnologia e a Sociedade no gênero “revista de atualidades”	Pugliese, Allan Tadeu	1 Discursos 2 Ciência, tecnologia e sociedade 3 Tradução e interpretação 4 Design 5 Bakhtin, Mikhail Mikhailovitch, 1895-1975	1 Análise do discurso 2 Serviços de tradução 3 Periódicos
3	¿De qué hablan? Discursos editoriais em Centroamérica	López, Edi Efrain Bámaca	1 América Central 2 Bakhtin 3 Editoriais 4 Análise do discurso 5 Periodismo 6 Cambio climático	1 Análise crítica do discurso 2 Editoriais – América Central. 3 Mudanças climáticas
4	Contratação de cooperativas de catadores de materiais recicláveis pelo poder público municipal: estudo em três municípios do estado de São Paulo	Santos, Carolina Valente	1 Catadores de materiais recicláveis 2 Cooperativas 3 Políticas públicas 4 Contratação 5 Cadeia produtiva da reciclagem	1 Catadores de lixo – Política governamental 2 Cooperativas de reciclagem – São Paulo (Estado)
5	A construção de um modelo multirreferencial para a coordenação de projetos coletivos em um Núcleo de Economia Solidária	Rufo, Carla Renata	1 Economia solidária 2 Ciência, Tecnologia e Sociedade 3 Tecnologia social 4 Complexidade 5 Transdisciplinaridade	1 Economia social
6	Competência em informação para pesquisa científica de estudantes de cursos tecnológicos do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – IFSP	Dupin, Aline Aparecida da Silva Quintã	1 Competência em informação 2 Educação Tecnológica 3 Ensino Profissionalizante 4 IFSP	1 Letramento informacional 2 Instituto Federal de Educação, Ciência e Tecnologia de São Paulo 3 Ensino profissional
7	Biblioteca como infraestrutura de apoio para a Educação a Distância	Ferreira, Ligia Leite Castelli	1 Biblioteca para EaD 2 Ensino Superior a Distância 3 Biblioteca a distância	1 Ensino à distância 2 Bibliotecas universitárias
8	Indicadores de gênero no periódico Scientometrics (1981-2017)	Godoi, Fábio dos Santos	1 Estudo de gênero 2 Gênero na ciência 3 Sociologia do conhecimento científico	1 Mulheres na ciência – Periódicos 2 Indicadores de ciência 3 Bibliometria

				4 Sociologia do conhecimento
9	Análise relacional intrainstitucional: redes de coautoria e acoplamento de autores	Franco, Nathalia Mendes Gerotti	1 Colaboração científica 2 Acoplamento de autores 3 Análise de redes	1 Pesquisadores – Coautoria 2 Trabalho de grupo na pesquisa
10	Processo de construção do conhecimento científico na educação básica a partir de experiências com robótica pedagógica	Oliveira, Ortenio de	1 Robótica pedagógica 2 Tecnologia educacional 3 Ciência e tecnologia 4 Construcionismo 5 Ensino de ciências	1 Educação básica 2 Ciência – Estudo e ensino 3 Robótica
11	Contribuição à identificação da competência científica de um Instituto Público de Pesquisa por meio de indicadores bibliométricos	Torres, Marcela de Fátima Nascimento de Macedo	1 Competência científica 2 Abordagem das competências 3 Indicadores de ciência e tecnologia 4 Institutos Públicos de Pesquisa	1 Bibliometria 2 Indicadores de ciência 3 Pesquisadores 4 Institutos de pesquisa
12	Gestão do conhecimento: um modelo para unidades de pesquisa do Ministério da Ciência, Tecnologia, Inovações e Comunicações	Rocha, Micherlangela Barroso	1 Gestão do conhecimento 2 Administração pública 3 Modelos de gestão do conhecimento 4 Unidades de pesquisa	1 Administração pública 2 Gestão do conhecimento 3 Institutos de pesquisa
13	Avaliação da Plataforma Lattes como fonte de coleta de metadados para povoamento automatizado em repositórios institucionais	Sarvo, Denilson de Oliveira	1 Acesso Aberto 2 Repositórios Intitucionais 3 Recuperação da informação 4 Plataforma Lattes	1 Repositórios institucionais 2 Metadados 3 Sistemas de recuperação da informação
14	Diagnóstico da Rede Sociotécnica de Inovação em uma Instituição Federal de Ensino Superior	Silva, Daniela Salgado Gonçalves da	1 Acesso aberto 2 Rede sociotécnica 3 Gestão da inovação 4 Ciência, Tecnologia e Sociedade 5 Universidade Federal de São Carlos 6 Gestão pública	1 Administração pública 2 Inovações tecnológicas 3 Institutos de pesquisa
15	Comunicação no processo de inovação tecnológica: relações entre ICT e o setor empresarial através dos NITs	Liberato, Tatiane Furukawa	1 Acesso aberto 2 Divulgação científica 3 Inovação tecnológica 4 Propriedade intelectual 5 Transferência de tecnologia	1 Propriedade intelectual 2 Inovações tecnológicas 3 Comunicação – Pesquisa
16	Acesso aberto na UFSCar: um estudo sobre a produção científica dos docentes dos Programas de Pós-graduação do CCET e CECH	Guirra, Daniel Andre Rigo	1 Produção acadêmica 2 Programas de Pós-graduação 3 Pesquisadores-docentes 4 Periódicos científicos	1 Periódicos acadêmicos 2 Pesquisadores 3 Professores de pós-graduação
17	Representações sociais de autonomia e autogestão de catadores de materiais recicláveis e apoiadores da Coopervida, São Carlos/SP	Oliveira, Letícia Dal Picolo Dal Secco de	1 Representações sociais 2 Catadores de materiais recicláveis 3 Economia solidária 4 Autonomia 5 Autogestão	1 Economia social 2 Cooperativas de reciclagem

18	Diagnóstico da oferta de indicadores bibliométricos nas bibliotecas das universidades públicas do estado de São Paulo	Ramos, Renan Carvalho	1 Indicadores 2 Bibliometria 3 Bibliotecas Universitárias 4 Serviços bibliométricos 5 Aspectos jurídicos da inteligência competitiva	1 Bibliometria 2 Bibliotecas universitárias 3 Universidade e faculdades públicas
19	Aspectos jurídicos da inteligência competitiva	Coxe, Rodrigo Antonio Garcia	1 Aspectos jurídicos 2 Inteligência competitiva 3 Ciclo de IC	1 Inteligência competitiva (Administração)
20	InfoMinerva: interface do usuário para sistemas de software do domínio de patrimônio histórico	Moura, Milene Rosa de Almeida	1 Interação humano-computador 2 Prototipação 3 Patrimônio histórico 4 Ciência, tecnologia e sociedade	1 Patrimônio cultural 2 Interfaces de usuário (Sistemas de computação)
21	Para uma epistemologia da errância: erro, hiância e ciência em Michel Pêcheux	Almeida, João Flávio de	1 Epistemologia 2 Ciência 3 Erro 4 Errância 5 Michel Pêcheux	1 Pêcheux, Michel 1938-1983 2 Teoria do conhecimento 3 Erro
22	Convergência de saberes: a integração dos conhecimentos científicos e populares no campo da Agroecologia	Silva, Anna Carolina Santana da	1 Extensão rural 2 Construção de conhecimentos 3 Conhecimento popular 4 Agroecologia 5 Conhecimento científico	1 Ecologia agrícola

Fonte: Elaboração do autor.

Quadro 5 – Resultados da indexação por extração de teses e dissertações do PPGEP.

N.	Título	Autor	Palavras-chave do Autor	Palavras-chave do Bibliotecário Especialista
1	Risco reputacional em frigoríficos de carne bovina da amazônia brasileira	Galuchi, Tharic Pires Dias	1 Gestão de cadeia de suprimento 2 Risco reputacional 3 Sustentabilidade 4 Frigorífico 5 Pecuária 6 Amazônia	1 Pecuária 2 Frigoríficos 3 Sustentabilidade 4 Avaliação de riscos ambientais
2	Efeito da servitização no desempenho de pequenas e médias empresas fabricantes de bens de capital	Queiroz, Stefânia Aparecida Belute	1 Servitização 2 Pequena e média empresa 3 PMEs 4 Desempenho financeiro 5 Desempenho não-financeiro 6 Empresas de manufatura	1 Processos de fabricação 2 Pequenas e médias empresas
3	Redes, relações e subgrupos: uma análise da estrutura relacional das empresas do setor automotivo	Ferreira, Augusto Squarsado	1 Indústria automotiva 2 Rede interorganizacional 3 Análise de redes sociais 4 Medidas de performance 5 Análise de similaridade 6 Detecção de comunidades	1 Indústria automobilística 2 Governança corporativa 3 Redes sociais

4	Influência da maturidade de Business Intelligence & Analytics sobre usos dos Sistemas de Medição de Desempenho: Estudos de caso	Xavier, José Eduardo Mendonça	1 Medição de Desempenho 2 Capacidades 3 Maturidade	1 Engenharia de Produção 2 Desempenho – Avaliação
5	Disseminando o uso do controle de carga no planejamento e controle da produção: modelagem utilizando um software de simulação comercial	Barreto, Bruno da Silva	1 Controle de carga 2 Simulação 3 Estrutura de conversão	1 Simulação (computadores) 2 Controle de produção
6	Resilience elements to combat counterfeit medicines in supply chain	Lima, Flávia Renata Pinho de	1 Cadeia de suprimentos 2 Elementos de resiliência 3 Medidas anti-contrafação 4 Cadeia de medicamentos	1 Sistemas de distribuição de medicamentos 2 Cadeia de logística integrada
7	Análise do desempenho de gráficos de controle X-Bar considerando diferentes estimadores do desvio padrão	Sobue, Cássio Eduardo Faria	1 Gráfico de controle 2 Simulação 3 Estimadores de desvio-padrão 4 Desempenho	1 Controle de processos – métodos estatísticos 2 Desvios-padrão Simulação por computador
8	Modelos de otimização e métodos de solução para o planejamento da produção e distribuição na indústria de móveis	Lugo, Pedro Luis Miranda	1 Problema de produção e roteamento 2 Planejamento e programação da produção e distribuição 3 Heurísticas de programação matemática 4 Indústria moveleira	1 Canais de distribuição 2 Indústria de móveis 3 Programação heurística 4 Planejamento da produção
9	Sustaining lean thinking in healthcare: case studies	Henrique, Daniel Barberato	1 Hospitais 2 Sustentabilidade 3 Estudo de caso 4 Melhoria contínua	1 Sustentabilidade 2 Hospitais – Administração
10	Compreensão do trabalho dos aeroviários no atendimento aos passageiros com necessidade de assistência especial: uma perspectiva da atividade	Figueiredo, Júlia Pierre	1 Passageiros com deficiência 2 Trabalho de aeroviários 3 Transporte aéreo 4 Ergonomia 5 Atividade	1 Ergonomia 2 Linhas aéreas 3 Pessoas com deficiência – Orientação e mobilidade
11	O espaço social dos “Negócios de Impacto” no Brasil: O caso Artemisia	Lourenço, Bruna Scanavachi	1 Negócio social 2 Negócio de impacto 3 Sociologia econômica 4 Artemisia	1 Economia – aspectos sociológicos 2 Negócios
12	Mecanismos para resiliência na cadeia de suprimentos: uma análise sob a ótica da indústria processadora de carne	Santos, Márcio Gonçalves dos	1 Resiliência 2 Cadeia de suprimentos 3 Indústrias frigoríficas 4 Rupturas	1 Carne – Indústria 2 Cadeia de logística integrada
13	Contribuições do DFMEA na integração entre desenvolvimento de produtos e engenharia da	Bueno, Fernanda Campos	1 Desenvolvimento de produtos 2 Engenharia da qualidade 3 Integração interfuncional 4 DFMEA	1 Indústrias 2 Engenharia de Produção 3 Administração de produtos

	qualidade: casos em empresas de grande porte			4 Qualidade dos produtos
14	Fatores que influenciam o comportamento do consumidor de edificações sustentáveis: um estudo exploratório no contexto brasileiro	Kirby, Regiane Cristina Laveli de Souza	1 Imóvel sustentável 2 Edificações sustentáveis 3 Fatores 4 Comportamento do consumidor	1 Comportamento do consumidor 2 Edifícios sustentáveis
15	Proposta de método para mensuração da qualidade de políticas públicas em sistemas de inovação	Pakes, Paulo Renato	1 Políticas Públicas 2 Sistema de Inovação 3 Qualidade no setor público	1 Política pública – Avaliação 2 Administração pública – Qualidade
16	Análise da estrutura relacional das patentes: o caso domínio tecnológico no seguimento dos veículos elétricos e híbridos	Assis, Filippo Savoi de	1 Inovação 2 Redes 3 Setor automotivo 4 Patentes	1 Veículos elétricos – Patentes
17	A servitização sob a perspectiva das capacidades dinâmicas	Mello, Luciana Torres Correia de	1 Servitização 2 Sistema produto-serviço 3 PSS 4 Capacidades dinâmicas 5 Capacidades operacionais	1 Competências essenciais 2 Serviços ao cliente
18	Modelo de decisão de licitação utilizando sistema fuzzy baseado em regras	Marçola, Jadilson Astorino	1 Licitação 2 Modelo de decisão 3 Sistema fuzzy baseado em regras	1 Sistemas difusos 2 Licitação pública
19	Competências para internacionalização em empresas agroindustriais: estudos de caso em frigoríficos brasileiros	Santos, Alexandre Borges	1 Agronegócio 2 Brasil 3 Estratégia 4 Investimento estrangeiro direto	1 Frigoríficos 2 Agroindústria 3 Investimentos estrangeiros
20	Otimização do planejamento e programação da produção na indústria de polpa moldada	Pérez Martínez, Karim Yaneth	1 Planejamento e programação da produção 2 Problemas de dimensionamento e sequenciamento de lotes 3 Indústria de polpa moldada 4 Programação inteira mista	1 Embalagens – indústria 2 Planejamento da produção
21	Eficiência dos municípios brasileiros: Uma aplicação da Análise Envoltória de Dados (DEA)	Zinidarsis, Izabel Cristina	1 Empreendedorismo 2 Eficiência 3 Análise Envoltória de Dados	1 Pesquisa sobre municípios 2 Análise de envoltória de dados
22	Um levantamento das práticas de Manufatura Enxuta em pequenas e médias empresas no Brasil	Hayashi, Ana Paula	1 Práticas de manufatura enxuta Survey 2 Pequenas e médias empresas	1 Pequenas e médias empresas 2 Eficiência industrial 3 Produção enxuta
23	Sigilo - múltiplas dimensões: ação dos agentes, formas de entendimento e dinâmica organizacional	Busnelo, Homero Cremm	1 Sigilo da informação 2 Transparência 3 Visão hierárquica 4 Processo social 5 Governança 6 Estratégia 7 Nexos de contrato	1 Governança pública 2 Desenvolvimento organizacional

24	Proposta de melhoria do processo tradicional de inteligência e do subprocesso de coleta de documentos de patente: estudo de caso no Núcleo de Informação Tecnológica em Materiais	Bessi, Nayara Cristini	1 Inteligência tecnológica 2 Processo de Inteligência 3 Processo de coleta de documentos de patente 4 Qualidade de dados e informações 5 Melhoria de processo 6 Precisão	1 Patentes 2 Recuperação em materiais
----	---	------------------------	---	--

Fonte: Elaboração do autor.

Os resultados apresentados nos quadros 4 e 5 revelam a diferença na quantidade de termos utilizados na indexação feita por autores em relação à indexação realizada pelo bibliotecário especialista em indexação. Enquanto os autores das teses e dissertações selecionadas usam, em média, 4,5 termos de indexação, o bibliotecário especialista adota 2,5 termos.

De modo geral, verifica-se o uso de termos generalistas ou com multiplicidade de sentidos na indexação realizada pelos autores como, por exemplo, contratação; complexidade; ciência; fatores; redes, entre outros. Enquanto que na indexação do especialista é notória a especificidade dos termos escolhidos, a indexação dos autores tende a ser mais exaustiva.

3.1 Universo da pesquisa

O Repositório Institucional da Universidade Federal de São Carlos (RI/UFSCar) é um sistema de informação que visa armazenar, preservar, organizar e disseminar amplamente a produção intelectual dos diversos setores e segmentos da comunidade universitária, provendo o acesso aberto à informação produzida na instituição.

No ano de 2009, a UFSCar participou de um Edital do IBICT no qual foram distribuídos *kits* tecnológicos compostos por um servidor com o *software* DSpace instalado e treinamento para a instituição. Contudo, a criação do Repositório Institucional não dependia somente dos bibliotecários, sendo realizado um amplo trabalho de conscientização de todo corpo administrativo, docente e técnico da instituição.

Do período da aquisição do *kit* e de sua efetiva implantação surgiram vários projetos-piloto de repositórios na UFSCar, em proporção modesta e normalmente para atender um curso ou um projeto específico, o que demonstrou sua importância e necessidade, sendo

criadas as condições necessárias para que o repositório institucional tivesse sua implantação efetivada. Com a coleção de Teses e Dissertações migrada do sistema da Biblioteca Digital de Teses e Dissertações (BDTD), o RI/UFSCar torna-se uma realidade.

Em 2014, o RI/UFSCar passa a existir oficialmente. Atualmente, equipe do RI-UFSCar é composta por bibliotecários e analistas sendo 2 bibliotecários do Sistema Integrado de Bibliotecas (SIBI) e 2 analistas da Secretaria de Informática da UFSCar. Atualmente o RI-UFSCar conta com um acervo de 3.310 teses de doutorado e 7.550 dissertações de mestrado, num total de 10.860 trabalhos disponibilizados, com uma média de 1.200 acessos semanais.

Figura 2 – Página inicial do Repositório Institucional da UFSCar.

The screenshot shows the homepage of the UFSCar Institutional Repository. At the top left is the logo 'Repositório Institucional UFSCar'. On the top right, there is a language selector set to 'português (Brasil)' and a 'Entrar' button. Below the header is a red navigation bar with 'Página Inicial' and a home icon. The main content area is divided into three columns. The left column contains a search bar and a 'NAVEGAR' menu with options like 'Todo o repositório', 'Comunidades e Coleções', 'Por data do documento', 'Autores', 'Orientador', 'Títulos', 'Assuntos', 'Áreas do CNPq', 'Programa de pós-graduação', and 'Tipo de documento'. The middle column has a 'Sobre o repositório' section with a description of the system, followed by a 'Sobre o autoarquivamento de Teses e Dissertações' section with instructions and links, and a 'Tipos de produção' section with two buttons: 'Dissertação (7422)' and 'Tese (3252)'. The right column is partially visible and contains a 'MINHA CONTA' section.

Desde a sua implantação, diversas discussões e reuniões de trabalho vem sendo realizadas visando a melhoria do funcionamento deste sistema de recuperação de informações, tanto do ponto de vista do usuário como do gestor. A partir do dia 02 de outubro de 2017, o RI/UFSCar passa a operar com o módulo de autoarquivamento em sua interface, na qual o próprio autor faz o depósito de sua tese ou dissertação.

Estão disponíveis na interface inicial os documentos de apoio para o sistema de autoarquivamento, como manual, infográfico, carta comprovante (documento previsto no fluxo de autoarquivamento destinado à validação da versão final de teses e dissertações pelo

orientador) e vídeo-tutorial o qual oferece o passo-a-passo para a realização da submissão por parte dos pós-graduandos da instituição. No entanto, a atual política do RI/UFSCar não contempla questões relacionadas à indexação de assuntos, além de inexistir uma política de indexação própria. Este cenário causa preocupação no que se refere à qualidade do metadado assunto, uma vez que os termos são empregados sem o auxílio de linguagens de indexação e não são corrigidos pelos bibliotecários do sistema para ajustes e conferência.

Contextualizado o método o conjunto de dados utilizados para o desenvolvimento da pesquisa, seguem a análise e discussão dos resultados.

4 ANÁLISE DA CONSISTÊNCIA DA INDEXAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFSCar

Como fonte de análise, foram recuperados no RI-UFSCar 46 teses e dissertações provenientes de dois programas de pós-graduação multidisciplinares, sendo 22 trabalhos do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade e 24 do Programa de Pós-Graduação em Engenharia de Produção, considerados programas multidisciplinares.

Para a análise dos dados coletados aplicou-se a fórmula proposta por Gil Leiva (2008) com o cruzamento dos termos empregados pelo autor com os termos elencados pelo bibliotecário especialista.

O corpus da pesquisa é composto por um total de 314 palavras-chave em linguagem natural e controlada. Destas, 199 são provenientes da indexação dos autores mediante o autoarquivamento, e 115 termos do bibliotecário especialista em indexação. Foi possível observar que a indexação do autor é mais exaustiva, porém não tão específica como a do especialista em indexação. Entende-se por exaustividade para esta pesquisa a quantidade de termos usados no ato da indexação. Já a especificidade é a atribuição de um nível mais específico para um termo (FUJITA; GIL LEIVA, 2014).

A qualidade das indexações foi medida a partir da análise do nível de coincidência entre a indexação do autor em comparação com a indexação do especialista. Conforme já sinalizado, o nível de concordância ou consistência consiste no “grau de semelhança na representação da informação documental de um documento por meio de termos de indexação selecionados por um ou vários indexadores, resultando em um índice de consistência” (GIL LEIVA; RUBI; FUJITA, 2008, p. 234).

Para cumprir com o propósito desta pesquisa, considerou-se a abordagem da avaliação intrínseca quantitativa mediante a interconsistência, aplicada para medir “[...] a consistência de indexação entre dois indexadores ou um grupo de indexadores” (PINHEIRO, 1978, p. 109). Portanto, a finalidade desta abordagem é comparar a indexação de um mesmo documento realizada por dois ou mais indexadores (GIL LEIVA, 2008). No caso desta pesquisa, a intenção foi medir o grau de consistência da indexação do autor com a indexação do bibliotecário, isto é, avaliar o grau de concordância entre as indexações.

O Quadro 6 contempla os termos padronizados do catálogo de assuntos da Biblioteca Nacional usados na indexação do bibliotecário especialista com a similaridade dos termos usados pelos autores no autoarquivamento. Notadamente, alguns dos termos usados pelos

autores para representar o conteúdo de suas pesquisas científicas são remissivos, relacionados ou gerais, o que demonstra uma correspondência entre os termos atribuídos nas indexações.

Quadro 6 – Correspondência entre os termos atribuídos pelo autor e pelo bibliotecário a partir do catálogo de assuntos da Biblioteca Nacional.

Designação do termo	Termos padronizados	Indexador	PPG
Termo tópico	Análise do discurso	Bibliotecário e Autor	PPGCTS
Termo específico	Análise crítica do discurso		
Termo tópico	Análise crítica do discurso	Bibliotecário	PPGCTS
Termo geral	Análise do discurso	Autor	PPGCTS
Termo tópico	Serviços de tradução	Bibliotecário	PPGCTS
Termo geral	Tradução e interpretação	Autor	PPGCTS
Termo tópico	Tradução e interpretação	Autor	PPGCTS
Termo específico	Serviços de tradução		
Termo tópico	Economia social	Bibliotecário	PPGCTS
Remissiva ver (US-UF)	Economia solidária	Autor	PPGCTS
Termo tópico	Letramento informacional	Bibliotecário	PPGCTS
Remissiva ver (US-UF)	Competência em informação	Autor	PPGCTS
Termo tópico	Teoria do conhecimento	Bibliotecário	PPGCTS
Remissiva ver (US-UF)	Epistemologia	Autor	PPGCTS
Termo tópico	Ecologia agrícola	Bibliotecário	PPGCTS
Remissiva ver (US-UF)	Agroecologia	Autor	PPGCTS
Termo tópico	Indústria automobilística	Bibliotecário	PPGEP
Remissiva ver (US-UF)	Indústria automotiva	Autor	PPGEP
Termo tópico	Linhas aéreas	Bibliotecário	PPGEP
Remissiva ver (US-UF)	Transporte aéreo industrial		
	Transporte aéreo (este termo não tem no catálogo da BN)	Autor	PPGEP

Termo tópico	Economia – Aspectos sociológicos	Bibliotecário	PPGEP
Remissiva ver (US-UF)	Sociologia da economia	Autor	PPGEP
Termo tópico	Agroindústria	Bibliotecário (usou o termo Ecologia agrícola)	PPGCTS
Remissiva ver (US-UF)	Agroecologia	Autor	PPGCTS

Fonte: Catálogo de terminologia de assuntos da Biblioteca Nacional, 2019.

No conjunto ora descrito, verifica-se relações semânticas entre termos atribuídos pelos autores e bibliotecário. Alguns dos exemplos são descritos abaixo.

Relação de equivalência:

- Sinônimo: Ecologia agrícola X Agroecologia
- Abreviatura: IFSP X Instituto Federal de São Paulo
- Grafias diferentes: Investimento estrangeiro X Investimentos estrangeiros

Relação hierárquica:

- Todo/parte: Linhas aéreas X Transporte aéreo industrial

Relação associativa:

- Concorrente: Teoria do conhecimento X Epistemologia
- Similaridade de processo: Análise crítica do discurso X Análise do discurso

A referida contextualização visa indicar que o catálogo de assuntos da Biblioteca Nacional serviu de base para a verificação da similaridade entre os termos utilizados pelos autores em relação aos do bibliotecário especialista, a fim de comprovar as relações semânticas e oportunizar a avaliação da consistência da indexação pela abordagem quantitativa.

Os quadros 7 e 8 apresentam os resultados após aplicada a fórmula matemática para a geração dos índices ou ensaios de interconsistência, sendo utilizada a comparação rígida entre as indexações por ser o tipo que considera termos cujo assunto determinado coincidem

completamente. Do total de 314 termos, tem-se 13 termos iguais nas duas indexações (4,14%) e 48 termos semelhantes (15,28%).

Quadro 7 – Índices de consistência entre as palavras-chave e os termos atribuídos às teses e dissertações do PPGCTS.

N.	Palavras-chave (linguagem natural) do autor	Termos (linguagem de indexação) do Bibliotecário Especialista	Índice de Intraconsistência
1	1 Gêneros discursivos 2 Linguagem e cultura 3 TV Cultura de São Paulo 4 Memória	1 TV Cultura de São Paulo 2 Televisão – estações educativas 3 Televisão pública - Análise do discurso	0,17 ou 17%
2	1 Discursos 2 Ciência, tecnologia e sociedade 3 Tradução e interpretação 4 Design 5 Bakhtin, Mikhail Mikhailovich, 1895-1975	1 Análise do discurso 2 Serviços de tradução 3 Periódicos	0,33 ou 33%
3	1 América Central 2 Bakhtin 3 Editoriais 4 Análise do discurso 5 Periodismo 6 Cambio climático	1 Análise crítica do discurso 2 Editoriais – América Central 3 Mudanças climáticas	0,28 ou 28%
4	1 Catadores de materiais recicláveis 2 Cooperativas 3 Políticas públicas 4 Contratação 5 Cadeia produtiva da reciclagem	1 Catadores de lixo – Política governamental 2 Cooperativas de reciclagem – São Paulo (Estado)	0,40 ou 40%
5	1 Economia solidária 2 Ciência, Tecnologia e Sociedade 3 Tecnologia social 4 Complexidade 5 Transdisciplinaridade	1 Economia social	0,20 ou 20%
6	1 Competência em informação 2 Educação Tecnológica 3 Ensino Profissionalizante 4 IFSP	1 Letramento informacional 2 Instituto Federal de Educação, Ciência e Tecnologia de São Paulo 3 Ensino profissional	0,40 ou 40%
7	1 Biblioteca para EaD 2 Ensino Superior a Distância 3 Biblioteca a distância	1 Ensino à distância 2 Bibliotecas universitárias	0,25 ou 25%
8	1 Estudo de gênero 2 Gênero na ciência 3 Sociologia do conhecimento científico	1 Mulheres na ciência – Periódicos 2 Indicadores de ciência 3 Bibliometria 4 Sociologia do conhecimento	0,17 ou 17%
9	1 Colaboração científica 2 Acoplamento de autores 3 Análise de redes	1 Pesquisadores – Coautoria 2 Trabalho de grupo na pesquisa	0%
10	1 Robótica pedagógica 2 Tecnologia educacional 3 Ciência e tecnologia 4 Construcionismo 5 Ensino de ciências	1 Educação básica 2 Ciência – Estudo e ensino 3 Robótica	0,33 ou 33%
11	1 Competência científica 2 Abordagem das competências	1 Bibliometria 2 Indicadores de ciência	0,14 ou 14%

	3 Indicadores de ciência e tecnologia. 4 Institutos Públicos de Pesquisa	3 Pesquisadores 4 Institutos de pesquisa	
12	1 Gestão do conhecimento 2 Administração pública 3 Modelos de gestão do conhecimento 4 Unidades de pesquisa	1 Administração pública 2 Gestão do conhecimento 3 Institutos de pesquisa	0,40 ou 40%
13	1 Acesso Aberto 2 Repositórios Institucionais 3 Recuperação da informação 4 Plataforma Lattes	1 Repositórios institucionais 2 Metadados 3 Sistemas de recuperação da informação	0,40 ou 40%
14	1 Acesso aberto 2 Rede sociotécnica 3 Gestão da inovação 4 Ciência, Tecnologia e Sociedade 5 Universidade Federal de São Carlos 6 Gestão pública	1 Administração pública 2 Inovações tecnológicas 3 Institutos de pesquisa	0,12 ou 12%
15	1 Acesso aberto 2 Divulgação científica 3 Inovação tecnológica 4 Propriedade intelectual 5 Transferência de tecnologia	1 Propriedade intelectual 2 Inovações tecnológicas 3 Comunicação – Pesquisa	0,60 ou 60%
16	1 Produção acadêmica 2 Programas de Pós-graduação 3 Pesquisadores-docentes 4 Periódicos científicos	1 Periódicos acadêmicos 2 Pesquisadores 3 Professores de pós-graduação	0,40 ou 40%
17	1 Representações sociais 2 Catadores de materiais recicláveis 3 Economia solidária 4 Autonomia 5 Autogestão	1 Economia social 2 Cooperativas de reciclagem	0,17 ou 17%
18	1 Indicadores 2 Bibliometria 3 Bibliotecas Universitárias 4 Serviços bibliométricos 5 Aspectos jurídicos da inteligência competitiva	1 Bibliometria 2 Bibliotecas universitárias 3 Universidade e faculdades públicas	0,33 ou 33%
19	1 Aspectos jurídicos 2 Inteligência competitiva 3 Ciclo de IC	1 Inteligência competitiva (Administração)	0,33 ou 33%
20	1 Interação humano-computador 2 Prototipação 3 Patrimônio histórico 4 Ciência, tecnologia e sociedade	1 Patrimônio cultural 2 Interfaces de usuário (Sistemas de computação)	0,20 ou 20%
21	1 Epistemologia 2 Ciência 3 Erro 4 Errância 5 Michel Pêcheux	1 Pêcheux, Michel 1938-1983 2 Teoria do conhecimento 3 Erro	0,33 ou 33%
22	1 Extensão rural 2 Construção de conhecimentos 3 Conhecimento popular 4 Agroecologia 5 Conhecimento científico	1 Ecologia agrícola	0,20 ou 20%

Fonte: Elaborado pelo autor.

Em relação as indexações de teses e dissertações provenientes do PPGCTS, dos 22 materiais analisados, 97 corresponde ao número total de palavras-chave empregadas pelos autores como termos representativos dos conteúdos das pesquisas científicas avaliadas, enquanto que o bibliotecário especialista elencou um total de 57 termos para representar o mesmo conjunto de documentos. Desse modo, tem-se como resultado uma média de 77 termos para o conjunto de 22 teses e dissertações.

Quanto ao grau de concordância entre as indexações provenientes do autoarquivamento e as realizadas pelo bibliotecário especialista, percebe-se que o índice de consistência variou de 0 a 60%. Houve apenas um caso em que o índice de consistência foi nulo, totalmente inconsistente.

Quadro 8 – Índices de consistência entre as palavras-chave e os termos atribuídos às teses e dissertações do PPGEF.

N.	Palavras-chave (linguagem natural) do autor	Termos (linguagem de indexação) do Bibliotecário Especialista	Índice de Intraconsistência
1	1 Gestão de cadeia de suprimento 2 Risco reputacional 3 Sustentabilidade 4 Frigorífico 5 Pecuária 6 Amazônia	1 Pecuária 2 Frigoríficos 3 Sustentabilidade 4 Avaliação de riscos ambientais	0,43 ou 43%
2	1 Servitização 2 Pequena e média empresa 3 PMEs 4 Desempenho financeiro 5 Desempenho não-financeiro 6 Empresas de manufatura	1 Processos de fabricação 2 Pequenas e médias empresas	0,14 ou 14%
3	1 Indústria automotiva 2 Rede interorganizacional 3 Análise de redes sociais 4 Medidas de performance 5 Análise de similaridade 6 Detecção de comunidades	1 Indústria automobilística 2 Governança corporativa 3 Redes sociais	0,12 ou 12%
4	1 Medição de Desempenho 2 Capacidades 3 Maturidade	1 Engenharia de Produção 2 Desempenho – Avaliação	0,25 ou 25%
5	1 Controle de carga 2 Simulação 3 Estrutura de conversão	1 Simulação (computadores) 2 Controle de produção	0,66 ou 66%
6	1 Cadeia de suprimentos 2 Elementos de resiliência 3 Medidas anti-contrafação 4 Cadeia de medicamentos	1 Sistemas de distribuição de medicamentos 2 Cadeia de logística integrada	0%

7	1 Gráfico de controle 2 Simulação 3 Estimadores de desvio-padrão 4 Desempenho	1 Controle de processos – métodos estatísticos 2 Desvios-padrão 3 Simulação por computador	0,40 ou 40%
8	1 Problema de produção e roteamento 2 Planejamento e programação da produção e distribuição 3 Heurísticas de programação matemática 4 Indústria moveleira	1 Canais de distribuição 2 Indústria de móveis 3 Programação heurística 4 Planejamento da produção	0,60 ou 60%
9	1 Hospitais 2 Sustentabilidade 3 Estudo de caso 4 Melhoria contínua	1 Sustentabilidade 2 Hospitais – Administração	0,50 ou 50%
10	1 Passageiros com deficiência 2 Trabalho de aeroviários 3 Transporte aéreo 4 Ergonomia 5 Atividade	1 Ergonomia 2 Linhas aéreas 3 Pessoas com deficiência – Orientação e mobilidade	0,33 ou 33%
11	1 Negócio social 2 Negócio de impacto 3 Sociologia econômica 4 Artemisia	1 Economia – aspectos sociológicos 2 Negócios	0,20 ou 20%
12	1 Resiliência 2 Cadeia de suprimentos 3 Indústrias frigoríficas 4 Rupturas	1 Carne – Indústria 2 Cadeia de logística integrada	0%
13	1 Desenvolvimento de produtos 2 Engenharia da qualidade 3 Integração interfuncional 4 DFMEA	1 Indústrias 2 Engenharia de Produção 3 Administração de produtos 4 Qualidade dos produtos	0%
14	1 Imóvel sustentável 2 Edificações sustentáveis 3 Fatores 4 Comportamento do consumidor	1 Comportamento do consumidor 2 Edifícios sustentáveis	0,50 ou 50%
15	1 Políticas Públicas 2 Sistema de Inovação 3 Qualidade no setor público	1 Política pública – Avaliação 2 Administração pública – Qualidade	0,66 ou 66%
16	1 Inovação 2 Redes 3 Setor automotivo 4 Patentes	1 Veículos elétricos – Patentes	0,25 ou 25%
17	1 Servitização 2 Sistema produto-serviço 3 PSS 4 Capacidades dinâmicas 5 Capacidades operacionais	1 Competências essenciais 2 Serviços ao cliente	0%
18	1 Licitação 2 Modelo de decisão 3 Sistema fuzzy baseado em regras	1 Sistemas difusos 2 Licitação pública	0,25 ou 25%
19	1 Agronegócio 2 Brasil 3 Estratégia 4 Investimento estrangeiro direto	1 Frigoríficos 2 Agroindústria 3 Investimentos estrangeiros	0,40 ou 40%

20	1 Planejamento e programação da produção 2 Problemas de dimensionamento e sequenciamento de lotes 3 Indústria de polpa moldada 4 Programação inteira mista	1 Embalagens – indústria 2 Planejamento da produção	0,20 ou 20%
21	1 Empreendedorismo 2 Eficiência 3 Análise Envoltória de Dados	1 Pesquisa sobre municípios 2 Análise de envoltória de dados	0,25 ou 25%
22	1 Práticas de manufatura enxuta Survey 2 Pequenas e médias empresas	1 Pequenas e médias empresas 2 Eficiência industrial 3 Produção enxuta	0,66 ou 66%
23	1 Sigilo da informação 2 Transparência 3 Visão hierárquica 4 Processo social 5 Governança 6 Estratégia 7 Nexos de contrato	1 Governança pública 2 Desenvolvimento organizacional	0,12 ou 12%
24	1 Inteligência tecnológica 2 Processo de Inteligência 3 Processo de coleta de documentos de patente 4 Qualidade de dados e informações 5 Melhoria de processo 6 Precisão	1 Patentes 2 Recuperação em materiais	0%

Fonte: Elaborado pelo autor.

No âmbito das teses e dissertações provenientes do PPGEP, dos 24 materiais analisados, 102 corresponde ao número total de palavras-chave (linguagem natural) empregadas pelos autores como termos representativos dos conteúdos das pesquisas científicas, enquanto que o bibliotecário especialista elencou um total de 58 termos (linguagem de indexação). Neste cenário, tem-se uma média de 80 termos empregados ao conjunto de 24 teses e dissertações, sendo que o índice de consistência variou de 0 a 66%.

Dentre os índices de consistência obtidos, 5 foram nulos, sendo gerados 0% de coincidência entre as palavras-chaves utilizados pelo autor no processo de autoarquivamento e a indexação conduzida pelo bibliotecário. A ausência de consistência em 6 do total de 46 indexações avaliadas desperta duas importantes considerações. Na primeira, este resultado demonstra que a correção e validação dos termos atribuídos pelos autores a partir do autoarquivamento deve ser considerada em repositórios institucionais, valendo-se que a eficácia da indexação está relacionada com uma recuperação precisa.

Sobre correção na indexação, Gil Leiva (2008) reforça a sua relevância ao considerar que na indexação podem ocorrer falhas, tanto por omissão de termo e/ou por inclusão de algum termo sem necessidade. O referido autor indica que

Para determinar esse tipo de erro, é necessário conhecer os termos ou assuntos mais apropriados àquele documento, para isso o profissional pode recorrer a seus pares e usuários. A correção será observada pela satisfação da recuperação da informação ou se há uma omissão ou atribuição indevida dos termos que representem o documento (GIL LEIVA, 2008, p. 72).

Diante disso, o bibliotecário indexador passa a desempenhar um papel de grande relevância em repositórios institucionais que aderem ao autoarquivamento, de modo a primar por termos que efetivamente satisfaçam a recuperação da informação no sistema. Assim, a garantia literária (atribuição de termos pelo autor de acordo com o conteúdo do documento) e a garantia de uso (recuperação da informação) tendem a ser contempladas.

Para satisfazer a garantia de uso, a especificidade da indexação deve ser favorecida, a fim de se obter profundidade e exatidão entre o termo indexado e o conteúdo do documento. Para tanto, o uso de linguagens de indexação como instrumento de apoio para o momento da atribuição de termos pelo autor é uma opção a ser avaliada pelos repositórios institucionais, conforme o grau de necessidade de especificidade da comunidade usuária. Caso a instituição não disponha de um vocabulário controlado próprio, linguagens de indexação mais gerais ou mesmo tesouros especializados podem ser recomendados aos autores para a atribuição de conceitos às suas pesquisas científicas.

A média da consistência entre as indexações de teses e dissertações depositadas no RI/UFSCar e provenientes do PPGCTS é de 28%, enquanto que no PPGEP esse índice cai para 22%. Gil Leiva (2008) afirma que os índices resultantes da avaliação intrínseca quantitativa oscilam entre 25% e 60%, sendo pouco provável obter 100% de consistência (HUGHES; RAFFERTY, 2011). Contudo, a avaliação das indexações provenientes do autoarquivamento e do bibliotecário especialista revelou um baixo índice de consistência, sendo 25% de similaridade entre as indexações.

Apesar de o índice de consistência obtido entre a avaliação da indexação resultante do autoarquivamento e a indexação realizada pelo bibliotecário especialista ter ficado dentro da margem sinalizada pela literatura (GIL LEIVA, 2008; PINHEIRO, 1978), é importante registrar que o índice de consistência poderia ser melhor caso o RI/UFSCar tivesse uma política de

indexação elaborada a partir das necessidades dos usuários atuantes na busca e recuperação das publicações mas, sobretudo, como colaboradores na representação da informação.

Em pesquisa recente, as autoras Dal'Evedove, Tartarotti, e Fujita (2015) ressaltam que, na literatura nacional, os estudos sobre política de indexação em ambientes colaborativos ainda estão no início e que a maioria das bibliotecas digitais e repositórios institucionais de teses e dissertações não contemplam uma política de indexação, ou seja, não existem diretrizes que norteiem o tratamento da informação nestes ambientes.

Diante da complexidade da indexação de assuntos, orientações previstas na política de indexação acerca da análise do assunto das produções científicas e sua posterior representação em termos de indexação são necessárias durante o processo de autoarquivamento.

Em pesquisa apresentada por Santos e Neves (2018) as autoras observaram as implicações da indexação por extração no processo de representação do conteúdo das produções científicas no Repositório Institucional da Universidade Federal do Rio Grande do Norte (RI/UFRN). No RI/UFRN, os termos empregados no metadado assunto correspondem aos mesmos termos empregados pelo autor nas palavras-chave do resumo do seu trabalho. Como resultados, as autoras constataram que os termos atribuídos em linguagem natural trazem implicações no processo de representação e recuperação por assuntos em ambiente digital, com destaque para o uso de frases, erros ortográficos, termos com polissemia, emprego de termos abrangentes e com multiplicidade de sentidos, uso de siglas e abreviações, dentre outras implicações (SANTOS, NEVES, 2018).

Em atenção ao contexto aqui apresentado, apresentam-se as considerações finais da pesquisa.

5 CONSIDERAÇÕES FINAIS

Diante dos novos ambientes informacionais que se apresentam ao campo da informação, buscou-se avaliar a consistência da indexação de teses e dissertações inseridas no RI/UFSCar mediante a prática do autoarquivamento em comparação com a indexação de um bibliotecário especialista no processo de indexação, integrante do Sistema de Bibliotecas da SIBi/UFSCar.

A revisão de literatura nacional e internacional de Biblioteconomia e Ciência da Informação revelou que são poucas as pesquisas dedicadas à representação e recuperação por assuntos em repositórios institucionais na condição de ambientes colaborativos.

De modo geral, os resultados apontam para a necessidade de adoção por parte do RI/UFSCar de um instrumento de padronização para a representação temática da coleção de teses e dissertações. Atualmente, o ambiente adota o catálogo de assuntos da Biblioteca Nacional, o qual não está disponível para o autor no momento da indicação de termos representativos da sua produção científica, ou alguma menção do uso por parte do RI/UFSCar.

Ao disponibilizar o acesso deste catálogo para os autores no momento do processo de autoarquivamento no sistema do RI/UFSCar, juntamente com orientações de uso do mesmo, poderiam ser minimizados os casos de polissemia, sinonímia, uso de singular e plural, uso de abreviaturas, entre outras implicações identificadas na análise realizada nesta pesquisa. Confirma-se também a importância da atuação de bibliotecários e recomenda-se que este profissional tenha como atribuição validar o metadado assunto. Além disso, o estabelecimento de uma política de indexação formalizada para o RI/UFSCar não pode ser negligenciado, considerando-se a necessidade de avaliações periódicas da indexação para a melhoria da recuperação por assuntos e atualização do índice de assunto do referido ambiente. Com efeito, este cenário tornaria o processo de recuperação da informação mais eficaz e favoreceria a padronização dos termos, no sentido de que assuntos do mesmo tema estariam agrupados.

Como proposta para trabalhos futuros, mostra-se oportuno que outros repositórios institucionais que aderem ao autoarquivamento sejam investigados no que se refere a qualidade da indexação com fins de recuperação por assunto. A elaboração de políticas de indexação em repositórios institucionais também é um nicho de investigação a ser explorado,

com importantes benefícios para a representação e recuperação por assuntos nestes ambientes.

De modo complementar, mostra-se oportuna a construção de um vocabulário controlado próprio para o RI/UFSCar com a participação do usuário enquanto elemento fundamental no processo de construção colaborativa na indexação. Sendo o autor da publicação científica um especialista em sua área de atuação, a construção de um instrumento desta natureza contribuirá para o fortalecimento das relações semânticas entre os termos nos índices de assuntos dos repositórios institucionais.

Diante do exposto, cabe registrar que o autoarquivamento favorece que os termos atribuídos ao metadado assunto estejam mais próximos da realidade e necessidades dos usuários de repositórios institucionais. Este é um movimento importante para a Organização do Conhecimento, em que a perspectiva do usuário é efetivamente considerada no momento da representação da informação, sem depender da figura do bibliotecário para mediar esse processo. Assim, o autoarquivamento favorece o desenvolvimento de repositórios institucionais mais democráticos, tendo como resultado uma representação da informação mais representativa e inclusiva.

Todavia, compete aos pesquisadores dedicados à representação temática conduzirem pesquisas que tomem como referência a avaliação da indexação em tempos de autoarquivamento, a fim de avançar no tema e apresentar novas possibilidades à Organização do Conhecimento.

REFERÊNCIAS

ALVARENGA, L. D. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 8, n. 15, p. 18-40, 2003. Disponível em: <http://www.brapci.inf.br/v/a/1331>. Acesso em: 17 set. 2018.

ANCIB. Associação nacional de pesquisa e pós-graduação em ciência da informação. **Grupos de trabalho da ANCIB**. 2017. Disponível em: <https://www.ancib.org.br/>. Acesso em: 3 jun. 2019.

ASSIS, T. B. **Análise das políticas de autoarquivamento nos repositórios institucionais brasileiros e portugueses**. InCID: Revista Ciência da Informação e Documentação, Ribeirão Preto, v. 4, n. 2, ed. esp. P. 212-227, jul./dez. 2013. Disponível em: DOI: <https://doi.org/10.11606/issn.2178-2075.v4i2p212-227> Acesso em: 5 nov. 2018.

AURÉLIO. **Novo Dicionário Aurélio**: 2018. Disponível em: <https://www.dicio.com.br/pena/> Acesso em: 05 nov. 2018.

BACHA, M. N; ALMEIDA, M. S. G. Vocabulário controlado e palavra-chave em repositórios digitais: relato de experiência do repositório institucional da FGV. XXV **Congresso Brasileiro de Biblioteconomia, Documento e Ciência da Informação** – Florianópolis, SC, Brasil, 07 a 10 de julho de 2013. Disponível em <https://portal.febab.org.br/anais/article/view/1361/1362> Acesso em: 03 dez. 2017.

BAGGIO, C. C.; BLATTMANN, U. **Os Repositórios das Universidades Federais do Brasil e suas Políticas de Informação**. RCI: Revista Ibero-americana Ciência da Informação, Brasília, v. 10, n. 2, p.350-363, jul./dez. 2017. Disponível em: DOI: <https://doi.org/10.22478/ufpb.1981-0695.2017v12n2.37134>. Acesso em: 5 nov. 2018.

BANDIM, M. A. S.; CORRÊA, R. F. A consistência na indexação automática por atribuição de artigos científicos na área de Ciência da Informação. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 23, n. 53, p. 64-77, set. 2018. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p64/37285>. Acesso em: 01 set. 2018

BARITÉ, M. G. Organización del conocimiento: um nuevo marco teórico-conceptual en bibliotecología y documentación. In: CARRARA, Kester (org.). **Educação, universidade e pesquisa**. III Simpósio em Filosofia e Ciências Marília. São Paulo: Unesp-Marília Publicações, Fapesp, 2001.

BARRETO, A. **Blog Aldo Barreto**. Disponível em: <https://aldobarreto.wordpress.com/2010/04/21/bases-de-dados-e-repositorios-de-informacao/> Acesso em: 18 jan. 2018.

- BARTON, J.; CURRIER, S.; HEY, J. M. N. Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice. **International Conference on Dublin Core and Metadata Applications**, [s.l.], p. 39-48, sep. 2003.
- BLATTMANN, U.; SILVA, F. C. C. Colaboração e interação na Web 2.0 e biblioteca 2.0. **Revista ACB**, Florianópolis, v. 12, n. 2, p. 191-215, nov. 2007. Disponível em: <https://revista.acbsc.org.br/racb/article/view/530/664>. Acesso em: 4 nov. 2018.
- BRÄSCHER, M.; GUIMARÃES, J. A. C. Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. **Liinc em Revista**, Rio de Janeiro, v.14, n.2, p. 241-258, nov. 2018.
- BRADY, F. **Controversy about our "Web 2.0" service mark**. 2006. Disponível em: https://web.archive.org/web/20080129105950/http://radar.oreilly.com/archives/2006/05/controverisy_about_our_web_20_s.html. Acesso em: 26 out. 2018.
- BRUNO, A. R.; PESCE, L. Mediação partilhada, dialogia e letramento: contribuições para a docência na contemporaneidade. **Atos de pesquisa em educação PPGE FURB**. v. 7, n. 3, p. 683-706, set./dez. 2012. Disponível em: <http://proxy.furb.br/ojs/index.php/atosdepesquisa/article/viewFile/3461/2175>. Acesso em: 21 jun. 2019
- BUDAPEST open access initiative. Budapest, Hungary. February 14, 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/read> Acesso em: 06 nov. 2018.
- CAFÉ, L. *et al.* Repositórios institucionais: nova estratégia para publicação científica na Rede. 26 **Congresso Brasileiro de Ciências da Comunicação**. Belo Horizonte, 2003. Disponível em: https://www.researchgate.net/publication/237689354_Repositorios_institucionais_nova_es_trategia_para_publicacao_cientifica_na_Rede. Acesso em: 12 set. 2018.
- CARLAN, E.; MEDEIROS, BRASCHER B. M. Sistemas de organização do conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, [S.l.], v. 4, n. 2, fev. 2012. Disponível em: <http://periodicos.unb.br/index.php/RICI/article/view/6209/5102>. Acesso em: 17 set. 2018.
- CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 3, p. 221-241, set. 1985. Disponível em: www.brapci.inf.br. Acesso em: 10 jul. 2019.
- CHAUÍ, M. **Introdução à História da Filosofia**: dos pré-socráticos a Aristóteles. v. 1 2. ed. São Paulo: Companhia das Letras, 2002.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v.33, n.1, p. 11-19, 2006.
- DAL'EVEDOVE, P. R. A política de tratamento da informação na percepção de catalogadores de assunto. *In*: XXIV Congresso Brasileiro de Biblioteconomia e Documentação – CBBDD, 2011,

Maceió. **Anais**. Sistemas de informação, multiculturalidade e inclusão social. Brasília: FEBAB, 2011, p.1-12.

DAL'EVEDOVE, P. R.; FUJITA, M. S. L. O conhecimento profissional do catalogador de assunto sobre política de indexação em bibliotecas universitárias. **Revista digital em Ciência da Informação**. Campinas, v.11, n. 2, p. 21-39, maio/ago. 2013. Disponível em: www.sbu.unicamp.br/seer/osj/index.php. Acesso em: 17 jun. 2019

DAL'EVEDOVE, P. R.; TARTAROTTI, R. C. D.; FUJITA, M. S. L. Concepções de política de indexação na visão de bibliotecas brasileiras. *In: VII ENCUENTRO IBÉRICO EDICIC*, 2015. Perspectivas de investigacion. Disponível em: http://edicic2015.org.es/ucmdocs/actas/tema_autor.html#perspectivas. Acesso em: 04 jul. 2019.

EISENSTEIN, E. L. *The Printing Press as na Agent of Change*, Cambridge University Press, 1980. Disponível em: https://en.wikipedia.org/wiki/Bibliotheca_universalis#CITEREFEisenstein1980. Acesso em: 11 set. 2018.

FICHE, R. H.; MOURA, L. R. C.; MAIA, L. C. G. A utilização dos repositórios digitais encontrados nas organizações. **Revista Analisando em Ciência da Informação**, João Pessoa, v. 5, n. 1, p. 61-80, jan.-jun. 2017. Disponível em: http://www.luizmaia.com.br/docs/2017-racin_v5_n1_artigo04.pdf Acesso em 17 jun. 2019.

FLICK, U. **Uma introdução à pesquisa qualitativa**. 2. ed. Porto Alegre: Bookman, 2004. 405p.

FREITAS, M. A.; LEITE, F. C. L. Atores do sistema de comunicação científica: apontamentos para discussão de suas funções. **Informação & Informação**, Londrina, v. 24, n. 1, p. 273-299, mar. 2019. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/32030>. Acesso em: 04 jul. 2019.

FREIRE, P. **Ciência da informação: temática, histórias e fundamentos. Perspectiva em Ciência da Informação**, Belo Horizonte. v. 11 n. 1, p. 6-19, jan./abr. 2006. Disponível em: <http://www.scielo.br/pdf/pci/v11n1/v11n1a02.pdf>. Acesso em: 2 out. 2017.

FUJITA, M. S. L. A importância teórica e prática da indexação na fundamentação científica da organização e representação do conhecimento. *In: DOBEDEI, V. GUIMARÃES, J. A. C. (org.). Complexidade e organização do conhecimento: desafios de nosso século*. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013. P. 147-159. Disponível em: http://isko-brasil.org.br/?page_id=42. Acesso em: 10/01/2019.

FUJITA, M. S. L. Organização e representação do conhecimento no brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**. v. 1, n. 1, 2008, p. 1-32. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000007781/e96dbf62e63261202e239c4e64970672/> Acesso em: 15 set. 2018.

FUJITA, M. S. L.; AGUSTIN LACRUZ, M. D. C.; GOMEZ DIAZ, R. A situação atual da indexação nas tarefas bibliotecárias. **Perspectivas em Ciência da Informação**. Escola de Ciência da Informação da UFMG, v. 17, n. 1, p. 94-109, 2012. Disponível em: <http://hdl.handle.net/11449/10558> Acesso em: 26 out. 2018.

GIL LEIVA, I. **Manual de indización**: teoría y práctica. Gijón: Ediciones Trea, 2008. 429 p. (Biblioteconomía y Administración Cultural, 193).

FUJITA, M. S. L.; GIL LEIVA, I. Avaliação da indexação por meio da recuperação da informação. **Ciência da Informação**, Brasília, v. 41, n. 1, jun. 2014, p. 50-66. Disponível em: <<http://revista.ibict.br/ciinf/article/view/1418>>. Acesso em: 12 set. 2018.

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, dez. 2008. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862008000300003&lng=en&nrm=iso. Acesso em: 21 jul 2019.

GOMES, H. E. Marcos históricos e teóricos da organização do conhecimento. **Informação & Informação**, [s.l.], v. 22, n. 2, p. 33-66, out. 2017. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/31442>. Acesso em: 11 set. 2018.

GUEDES, R. M.; DIAS, E. J. W. Indexação social: abordagem conceitual. Revista ABC: Biblioteconomia em Santa Catarina, Florianópolis, v. 15, n. 1, p. 39-53 jan./jun. 2010. Disponível em: http://www.brapci.inf.br/_repositorio/2010/06/pdf_fcb17df2cd_0010808.pdf. Acesso em: 11 set. 2018.

GUIMARÃES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, v.41, n. 1, p. 13-21, jan/abr., 2014. Disponível em: Acesso em: 02 jul. 2019.

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, n. 1, p. 77-99, 2008. Disponível em: <http://periodicos.unb.br/index.php/RICI/article/view/2761/2331> Acesso em: 11 set. 2018.

GUIMARÃES, J. A. C.; MORAES, J. B. E.; GUARIDO, M. D. M. Análisis documental de contenido de textos narrativos: bases epistemológicas y perspectivas metodológicas. **Avances y perspectivas en sistemas de información y documentación em entorno digital**. Zaragoza: Prensas Universitarias de Zaragoza, 2007. p. 93-100. Disponível em: <https://www.ibersid.eu/ojs/index.php/ibersid/article/viewFile/3267/3028> Acesso em: 11 set. 2018

HJORLAND, B. Toward a theory of aboutness, subject, topicality, theme, domain, field, content... and relevance. **Journal of the American Society for Information Science and Technology**, 52, p. 249-298, 2001.

HODGE, G. Systems of Knowledge Organization for digital libraries: beyond traditional authority files. 2000.

HUGHES, A. V.; RAFFERTY, P. Inter-indexer consistency in graphic materials indexing at the National Library of Wales. **Journal of Documentation**, v. 67, n. 1, p. 9-32. 2011.

INÁCIO, M. O. **Avaliação da indexação em bibliotecas universitárias**: uma aplicação em catálogos online. 2012. Dissertação (Mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2012. 157p.

INSTITUTO Brasileiro de Informação em Ciência e Tecnologia (IBICT). Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais/historico>. Acesso em 18 jan. 2018.

ISKO – International Society for Knowledge Organization. **ISKO internacional**. 2019. Disponível em: http://isko-brasil.org.br/?page_id=5. Acesso em 02 jul. 2019.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, [S.l.], v. 35, n. 2, ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1144> Acesso em: 2 set. 2018.

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n. 1, p. 41-52, Jan. 2002. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010019652002000100005&lng=en&nrm=iso. Acesso em: 11 set. 2018.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 1993.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: Ibict, 2009. 120p.

LEITE, F. C. L.; COSTA, S. Gestão do conhecimento científico: proposta de um modelo conceitual com base em processos de comunicação científica. **Ciência da Informação**, Brasília, v. 36, n. 1, jan./abr. 2007. Disponível em: http://www.SciELO.br/SciELO.php?pid=S0100-19652007000100007&script=sci_arttext. Acesso em: 05 jul. 2019.

LUCAS, C. R. **Leitura e interpretação em biblioteconomia**. Campinas: Unicamp, 2000. 91 p. (Coleção Pesquisas).

MACHADO, A. M. N. **Informação e controle bibliográfico**: um olhar sobre a cibernética. -São Paulo: Editora UNESP, 2003. Disponível em:

<http://www.dominiopublico.gov.br/download/texto/up000023.pdf>. Acesso em: 11 set. 2018.

MAI, J.E. Analysis in indexing: document and domain centered approaches. **Information Processing and Management**, v. 41, n. 3, p. 599-611. 2005.

MARCONDES, C. H.; SAYÃO, L. F. Introdução: repositórios institucionais e livre acesso. In: SAYÃO, L. F. et al. (org.). **Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação**. Salvador: EDUFBA, 2009. p. 9-22.

MARQUES, F. Comunicação. **Revista Pesquisa FAPESP**, 14 mar. 2019. Disponível em: <https://revistapesquisa.fapesp.br/2019/03/14/fapesp-lanca-politica-para-acesso-aberto/>. Acesso em 25 jun. 2019. (Portaria CTA nº 01/2019).
http://www.fapesp.br/12632?fbclid=IwAR1QVe9sxxDWIWMXTxtCLTChzvhsTaEwSxU2Ce65lQM2-EF_WaQyR0AlcN0.

MARTINHO, N. O.; GUEDES, E. F. **Charles Ammi Cutter**: sua contribuição para organização da informação. [s.d.]. Disponível em: <http://enancib.ibict.br/index.php/enancib/xenancib/paper/viewFile/3152/2278>. Acesso em: 11 set. 2018.

MARTINS, G. K. **Institucionalização cognitiva e social da Organização e Representação do Conhecimento na Ciência da Informação no Brasil**. 2014. 184 f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2014. Disponível em: <http://hdl.handle.net/11449/114036> Acesso em: 11 set. 2018.

MARTINS, G. K.; MORAES, J. O. B. E. Organização e representação do conhecimento: institucionalização como disciplina científica no âmbito da ciência da informação. **Encontro Nacional de Pesquisa em Ciência da Informação**, v. 16, 2015. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000017692/3d2abf9e88efd50417ec9696cb0c49ea>. Acesso em: 11 set. 2018.

MEDEIROS, G. M. **Organização da informação em repositórios digitais: implicações do auto-arquivamento na representação da informação**. 2010, 273 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Informação. Universidade Federal de Santa Catarina, 2010. Disponível em: Acesso em: 11 set. 2018.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, [S.l.], v. 35, n. 2, ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1138/1293>. Acesso em: 08 set 2018.

MURAKAMI, T.; FAUSTO, S. Panorama atual dos Repositórios Institucionais das Instituições de Ensino Superior no Brasil. InCID: **Revista de Ciência da Informação e Documentação**, v. 4, n. 2, p. 185-201, 20 dez. 2013. Disponível em: <http://www.revistas.usp.br/incid/article/view/69327>. Acesso em: 11 set 2018.

NASCIMENTO, G. F.; NEVES, D. A. Folksonomia: um estudo das tags dos bibliotecários

brasileiros no del.icio.us. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 11, 2010. Rio de Janeiro. Anais... Rio de Janeiro: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 2010.

NAVES, M. M. L. **Princípios e técnicas de indexação, com vistas a recuperação da informação**. Belo Horizonte, 2004, 24 p. (Curso de indexação). Disponível em: <https://pt.slideshare.net/telmasobrinho/principios-tecnicas-deindexacao1>. Acesso em 13 jul. 2019.

NAVES, M. M. L. **Princípios e técnicas de indexação, com vistas à recuperação da informação**. UFMG, Belo Horizonte, 2004. Curso de indexação. Disponível em: <https://pt.slideshare.net/telmasobrinho/principios-tecnicas-deindexacao1>. Acesso em 13 jul. 2019.

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: ABNT, 1992. 4p.

NHACUONGUE, J. A. **O campo da Ciência da Informação: contribuições, desafios e perspectivas da mineração de dados para o preenchimento pós-moderno**. 2015. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. 2015.

NORMAN'S, J. Henry E. **Bliss Develops the Bliss Bibliographic Classification (1929-1953)**. HistoryofInformation.com. 2004 Jeremy Norman & Co., Inc.A Disponível em: <http://www.historyofinformation.com/expanded.php?id=4850>. Acesso em: 11 set. 2018.

PALAVITSINIS, N.; MANOUSELIS, N.; SANCHEZ-ALONSO, S. Metadata quality in digital repositories: empirical results from the cross-domain transfer of a quality assurance process. **Journal of the Association of Information Science and Technology**, v. 65, n. 6, p. 1202–1216, 2014.

PANDO, D. B. **Epistemologia da Organização da Informação: uma análise de sua cientificidade no contexto brasileiro**. 2018, 463p. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista – UNESP. Marília, 2018.

PINHEIRO, L. V. R. Medidas de consistência da indexação: interconsistência. **Ciência da Informação**, Rio de Janeiro, 1978, v. 7, n.2, p. 109-114.

PINHO, F. A. **Aspectos éticos em representação do conhecimento: em busca do diálogo entre Antônio García Gutiérrez, Michele Hudon e Clare Beghtol**. 2006. 123 f. Dissertação (Mestrado), Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 2006. Disponível em: <http://hdl.handle.net/11449/93693>. Acesso em 18 jan. 2018.

POLÍTICA de indexação para biblioteca: elaboração, avaliação e implantação. Mariângela Spotti Lopes Fujita. (org.). Marília: Oficina Universitária, 2016. 142p.

RAFFERTY, P.; HEDDERLEY, R. Flickr and democratic indexing: dialogic approaches to indexing. **Aslib Proceedings**, v. 59, Issue 4/5, 2007. p. 397-410. Disponível em: <https://www.emeraldinsight.com/doi/pdfplus/10.1108/00012530710817591>. Acesso em: 20 jun. 2019.

REPOSITÓRIO Institucional da UFSCar: proposta de implantação, disseminação e uso. Projeto de Repositório institucional UFSCar. São Carlos: UFSCar, 2015. (Relatório)

RIBEIRO, F. N. Edgar Morin, o pensamento complexo e a educação. **Pró-Discente: Caderno de Produções Acadêmicas -Científicas Programa de Pós-Graduação Educação**, Vitória - ES, v. 17, n. 2, jul./dez. 2011. Disponível em: <http://www.portaldepublicacoes.ufes.br/PRODISCENTE/article/viewFile/5804/4245>. Acesso em: 28/02/2019.

ROARMAP. http://roarmap.eprints.org/view/policymaker_type/research=5Forg.html. 2019

ROUSIDIS, D. *et al.* Metadata for Big Data: a preliminary investigation of metadata quality issues in research data repositories. **Information Services & Use**, n. 34, p. 279-286, 2014.

RUBI, M. P. Política de indexação para construção de catálogos coletivos em bibliotecas universitárias. 2008. 166. **Tese** (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/103388/rubi_mp_dr_mar.pdf?se Acesso em: 10 set. 2018.

SALES, R.; CAFE, L. Diferenças entre tesouros e ontologias. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 14, n. 1, p. 99-116, abr. 2009. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362009000100008&lng=en&nrm=iso. Acesso em: 10 set. 2018.

SALES, R. A. Classificação de Livros de William Torrey Harris: influências de Bacon e Hegel nas classificações de biblioteca. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 22, n. 50, p. 188-204, set. 2017. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p188>. Acesso em: 15 set. 2018.

SANTOS, B. S. Um discurso sobre as ciências na transição para uma ciência pós moderna. **Estudos avançados**, v. 2, n. 2, p. 46-71, 1988. Disponível em: http://www.scielo.br/scielo.php?pid=S0103-40141988000200007&script=sci_arttext. Acesso em: 10 set. 2018.

SANTOS, R. A.; CARDOSO, R. C. Avaliação de repositórios institucionais: o Brasil no ranking webométrico. XXV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação Florianópolis, SC, Brasil, 07 a 10 de julho de 2013. Disponível em: [file:///C:/Users/bco16/Downloads/1387-1400-1-PB%20\(3\).pdf](file:///C:/Users/bco16/Downloads/1387-1400-1-PB%20(3).pdf). Acesso em 24 jun 2019.

SANTOS, R. F.; NEVES, D. A. B. Práticas de indexação em repositórios digitais de acesso aberto: análise do metadado assunto do repositório institucional da UFRN. **Encontro**

Nacional de Pesquisa em Ciência da Informação, n. XIX ENANCIB, 2018. Disponível em: <http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/view/1117>. Acesso em: 18 jan. 2019.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1 p. 41-62, jan./jun. 1996. ISSN 19815344. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>. Acesso em: 17 jun. 2019.

SARACEVIC, T. A natureza interdisciplinar da ciência da informação. **Ciência da Informação**, [S.l.], v. 24, n. 1, apr. 1995. ISSN 1518-8353. Disponível em: <http://revista.ibict.br/ciinf/article/view/608>. Acesso em: 15 sep. 2018.

SILVA, M. B. A aplicação da folksonomia em sistemas de informação. **Encontro Brasileiro de Arquitetura da Informação**, 4, 2010, São Paulo, 2010. Disponível em: https://www.researchgate.net/publication/267723472_a_aplicacao_da_folksonomia_em_sistemas_de_informacao_a_aplicacao_da_folksonomia_em_sistemas_de_informacao. Acesso em: 03 jan. 2018.

SILVA, M. B.; FUJITA, M.S. L. A prática de indexação: análise da avaliação de tendências teóricas e metodológicas. **Transinformação**, Campinas, v.16, n. 2, p. 133-161, ago. 2004 .. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862004000200003&lng=en&nrm=iso. Acesso em: 3 jan. 2019.

SILVA, J. L. C.; FREIRE, G. H. A. Um olhar sobre a origem da ciência da informação: indícios embrionários para sua caracterização identitária. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 17, n. 33, p. 1-29, abr. 2012. ISSN 1518-2924. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n33p1/21708>. Acesso em: 25 jun. 2019.

SOUZA, A. L. *et al.* Importância dos repositórios institucionais na preservação intelectual: em foco a gestão do conhecimento. Encontro Nacional de Estudantes de Biblioteconomia, Documentação, Ciência da Informação e Gestão da Informação, 35., 2012. Belo Horizonte. **Anais [...]**. Belo Horizonte: UFMG, 2012. P. 1-17. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/moci/article/view/1696/1145>. Acesso em: 6 nov. 2018.

SOUZA, R. T. B. Os princípios da teoria da classificação e o processo de organização de documentos de arquivo. **Arquitetura & Administração**, Rio de Janeiro, v. 6, n. 1, jan./jun. 2007. Disponível em: <http://www.brapci.inf.br/index.php/article/download/62321> Acesso em: 10 set. 2018.

STREHL, L. Avaliação da consistência da indexação realizada em uma biblioteca universitária de artes. **Ciência da Informação**, Brasília, v. 27, n. 3, p.329-335, Set./dez. 1998. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000300011&lng=en&nrm=iso. Acesso em: 01 jul. 2019.

SUBER, P. The primacy of authors in achieving open access. *Nature*, 2004. Disponível em: <http://www.nature.com/nature/focus/accessdebate/24.html>>. Acesso em: 15 abr. 2018.

TARTAROTTI, R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Avaliação da indexação de documentos não-textuais: uma análise da literatura. XIX **Encontro Nacional de Pesquisa em Ciência da Informação** – ENANCIB, 2018, Londrina. P. 547-563. Disponível em: <http://www.brapci.inf.br/index.php/res/v/103723>. Acesso em:

TARTAROTTI, R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. **A metodologia quantitativa da avaliação da indexação: um estudo comparativo da representação temática em catálogos coletivos online e em bases de dados de bibliotecas universitárias**. 17 Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB, 2016. P. 1-19.

VEIGA, V.; Macena, L. G. O. O Autoarquivamento nos Repositórios Institucionais Brasileiros: Um Estudo Exploratório. **Ponto de Acesso**: Revista do Instituto de Ciência da Informação da UFBA, Salvador, v.9, n.3, p. 35-47, dez. 2015. Disponível em: <http://dx.doi.org/10.9771/1981-6766rpa.v9i3.15107>. Acesso em: 1 set 2018.

VEIGA, V. S. O. *et al.* Panorama do autoarquivamento nos repositórios institucionais portugueses. **Cadernos BAD**. 7º Conferencia luso-brasileira de acesso aberto, Instituto Politécnico de Viseu, Portugal, 2016. Disponível em: <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1586/pdf>. Acesso em: 10 set. 2018.

WAL, T. V. Folksonomia definition and Wikipedia. 2005. Disponível em: <http://www.vanderwal.net/random/entrysel.php?blog=175>. Acesso em: 10 set. 2018.

WEITZEL, S. R. **Os repositórios de e-prints como nova forma de organização da produção científica: o caso da área das Ciências da Comunicação no Brasil**. Tese (Doutorado em Ciência da Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2006. doi:10.11606/T.27.2006.tde-14052009-133509. Acesso em: 2 set. 2018.

WIKIPEDIA. Anthony Panizzi. Disponível em: https://ipfs.io/ipfs/QmXoyvizjW3WknFiJnKLwHCnL72vedxjQkDDP1mXWo6uco/wiki/Anthony_Panizzi.html. Acesso em: 10 set. 2018.