



UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE CIÊNCIAS HUMANAS, COMUNICAÇÃO E ARTE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Valter dos Santos Andrade

ANÁLISE DOS METADADOS DE ASSUNTO NO REPOSITÓRIO
INSTITUCIONAL DA UNIVERSIDADE FEDERAL DE ALAGOAS – RIUFAL

Maceió

2023

VALTER DOS SANTOS ANDRADE

**ANÁLISE DOS METADADOS DE ASSUNTO NO REPOSITÓRIO INSTITUCIONAL
DA UNIVERSIDADE FEDERAL DE ALAGOAS – RIUFAL**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Alagoas, área de concentração em Informação, Tecnologia e Inovação, linha de pesquisa em Informação, Comunicação e Processos Tecnológicos, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Orientadora: Profa. Dra. Robéria de Lourdes Vasconcelos Andrade.

Maceió

2023

Catálogo na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 – 1767

A553a Andrade, Valter dos Santos.
 Análise dos metadados de assunto no Repositório Institucional da
 Universidade Federal de Alagoas / Valter dos Santos Andrade. – Maceió,
 2023.
 80 f. : il.

Orientadora: Robéria de Lourdes de Vasconcelos Andrade.
Dissertação (Mestrado em Ciência da Informação) – Universidade
Federal de Alagoas. Instituto de Ciências Humanas, Comunicação e Artes.
Maceió, 2023.

Bibliografia: f. 74-80.

1. Universidade Federal de Alagoas. Repositório Institucional. 2.
Representação da informação. 3. Indexação. 4. Autoarquivamento. I. Título.
I. Título.

CDU: 025.4


VALTER DOS SANTOS ANDRADE

**ANÁLISE DOS METADADOS DE ASSUNTOS NO REPOSITÓRIO
INSTITUCIONAL DA UNIVERSIDADE FEDERAL DE ALAGOAS**


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

Aprovado em: 17 / 08 / 2023.


BANCA EXAMINADORA

Documento assinado digitalmente
 ROBERIA DE LOURDES DE VASCONCELOS,
Data: 17/08/2023 16:30:29-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Roberia de Lourdes de Vasconcelos Andrade – PPGCI/UFAL
(Orientadora/Presidente)

Documento assinado digitalmente
 EDIVANIO DUARTE DE SOUZA
Data: 20/08/2023 23:28:19-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Edivanio Duarte de Souza – PPGCI/UFAL (Membro Titular Interno)

Documento assinado digitalmente
 LEYDE KLEBIA RODRIGUES DA SILVA
Data: 17/08/2023 17:04:21-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Leyde Klebia Rodrigues da Silva –PPGB/UFCA (Membro Titular Externo)

A Ana Rosa,
espelho de toda a nossa família.

AGRADECIMENTOS

A Deus, por ter traçado esse caminho na minha vida;

Aos meus pais, José Bento de Andrade (*in memoriam*) e Olívia dos Santos Andrade, pessoas humildes, de pouco letramento, que nos deixou como herança o saber, a educação e o mapa para buscar o mundo do conhecimento;

À toda minha família e a família Holanda pelo amor, carinho, apoio, incentivo, torcida e por acreditar no meu potencial; a Valmir e Shirley pela ajuda no protocolo da inscrição em plena festividade de final de ano. “Ufa, haja correria!”;

A Tarlane Tenório por ter dado a mão e me puxado para fazer o mestrado; a Girlaine Santos por ser a outra incentivadora para concretização dessa ideia, além de dividir os aprendizados na sala de aula, as dúvidas, angústias, incertezas, mas, principalmente, as alegrias de cada etapa conquistada. “Bora!”

À Profa. Dra. Robéria de Lourdes de Vasconcelos Andrade, minha orientadora, pela disponibilidade, paciência, compreensão e aprendizado mútuo;

Aos professores doutores Edivanio Duarte de Souza (PPGCI/UFAL) e Leyde Klebia Rodrigues da Silva (PPGB/UFCA), pelas valorosas contribuições na qualificação e na defesa;

Aos docentes do PPGCI/UFAL pela contribuição na minha formação e aprendizado na Ciência da Informação, em especial à professora Dra. Nelma Camelo de Araújo, pelos primeiros encaminhamentos na orientação;

Aos companheiros do repositório da UFAL e, principalmente, à bibliotecária Helena Lessa, pela dedicação e paixão pelo RIUFAL, que me contagiou;

Aos queridos amigos da Biblioteca Central da UFAL, pelo carinho e incentivo;

Aos meus demais amigos que sempre me apoiaram e estiveram ao meu lado, compartilhando alegrias e tristezas;

Aos colegas do PPGCI/UFAL, pelas experiências, aprendizados compartilhados e companheirismo, mesmo distanciados um bom tempo pela COVID-19;

A todos que, direta ou indiretamente, colaboraram comigo nessa jornada, por meio de palavras, incentivos ou ações.

Vim, vi e venci.
(Júlio César, Imperador)

Pois,

Tudo posso naquele que me fortalece.
(Filipenses 4:13)

RESUMO

A representação temática é fator fundamental para o processo de organização, acesso e recuperação da informação, principalmente nos ambientes digitais como os repositórios institucionais. Nesse sentido, a presente pesquisa tem como objetivo geral, analisar a representação da informação no Repositório Institucional da Universidade Federal de Alagoas (RIUFAL), no que concerne aos metadados “assunto”, e como objetivos específicos: a) coletar os TCCs de Enfermagem e Serviço Social cadastrados através de autoarquivamento; b) avaliar os metadados “assunto” nos TCCs coletados; c) Comparar os metadados “assunto” utilizados com os cabeçalhos autorizados do catálogo de assunto do Sistema de Biblioteca da UFAL (SIBI/UFAL); e, d) identificar se, dentre os metadados, existe algum que apresente ruído que possa interferir no processo de busca e recuperação da informação no sistema. Trata-se de uma pesquisa exploratória, descritiva e com abordagem qualitativa. A análise se deu sobre a representação temática dos metadados “assunto” presente nas monografias depositadas através do autoarquivamento pelos graduandos dos cursos de Enfermagem e Serviço Social, no período de 2020 a 2022. Foram coletadas 160 monografias, 84 do curso de Enfermagem e 76 do curso de Serviço Social. A coleta dos dados foi realizada diretamente na base de dados do repositório, sendo os dados agrupados, organizados em quadros, analisados, tratados e apresentados em gráficos. Foram coletados 252 metadados de assunto nas monografias do curso de Enfermagem, dos quais 201 são metadados de assunto autorizados, correspondendo a 79,8% de conformidades com o catálogo de assunto do SIBI/UFAL, identificando apenas 29 metadados de assunto em linguagem natural (11,5%); 15 com erros de padronização (6%); 3 com erros de grafia (1,2%) e 4 com erros de digitação (1,5%). Nas monografias de Serviço Social foram coletados 210 metadados de assunto, apresentando 84 metadados autorizados, ou seja, 40% em conformidade com o catálogo; 86 em linguagem natural, (41%); 23 metadados de assunto sem padronização, (11%), 2 metadados “assunto” com erros de grafia, (1%); e 15 metadados com erros de digitação, (7%). Ressalta que os alunos do curso de Enfermagem são familiarizados com o uso da Base de Dados de Descritores de Saúde para escolhas de termos para seus trabalhos acadêmicos, além da obrigatoriedade do uso da ficha catalográfica elaborada pelo setor Catalogação na Fonte da Biblioteca Central da UFAL, não ocorrendo o mesmo com o curso de Serviço Social, o qual não adotava a obrigatoriedade da ficha catalográfica nem a utilização de ferramentas autorizadas, como tesouros ou índices especializados, fato que provoca a recorrente utilização da linguagem natural, e, com isso, em alguns ruídos que interferem na recuperação da informação no sistema do Repositório. Percebe-se que o uso autorizado de ferramentas como tesouros bem como a orientação do profissional bibliotecário, contribuem para a padronização e qualidade dos metadados de assunto, proporcionando uma maior eficácia e eficiência na busca documental. Conclui, enfatizando a importância para o repositório adotar uma política de indexação, disponibilizando meios para auxiliar os usuários na utilização de ferramentas como tesouros e índices que possam guiá-los nas escolhas dos termos, possibilitando assim, elevar o nível da qualidade de descrição dos conteúdos documentais, visando à precisão no processo de busca e recuperação da informação.

Palavras-chave: repositório institucional; representação da informação; indexação; autoarquivamento.

ABSTRACT

Thematic representation is a fundamental factor in the process of organizing, accessing and retrieving information, especially in digital environments such as institutional repositories. In this sense, the present research has as its general objective, to analyze the representation of information in the Institutional Repository of the Federal University of Alagoas (RIUFAL), with regard to “subject” metadata, and as specific objectives: a) collect the Nursing TCCs and Social Service registered through self-archiving; b) evaluate the “subject” metadata in the collected TCCs; c) Compare the “subject” metadata used with the authorized headers of the subject catalog of the UFAL Library System (SIBI/UFAL); and, d) identify whether, among the metadata, there is any that presents noise that could interfere with the process of searching and retrieving information in the system. This is an exploratory, descriptive research with a qualitative approach. The analysis took place on the thematic representation of the “subject” metadata present in the monographs deposited through self-archiving by undergraduates of Nursing and Social Work courses, in the period from 2020 to 2022. 160 monographs were collected, 84 from the Nursing course and 76 from the Social Service course. Data collection was carried out directly in the repository's database, with the data grouped, organized into tables, analyzed, processed and presented in graphs. 252 subject metadata were collected in Nursing course monographs, of which 201 are authorized subject metadata, corresponding to 79.8% of compliance with the SIBI/UFAL subject catalog, identifying only 29 subject metadata in natural language (11.5%); 15 with standardization errors (6%); 3 with spelling errors (1.2%) and 4 with typing errors (1.5%). In Social Service monographs, 210 subject metadata were collected, presenting 84 authorized metadata, that is, 40% in accordance with the catalogue; 86 in natural language, (41%); 23 subject metadata without standardization, (11%), 2 “subject” metadata with spelling errors, (1%); and 15 metadata with typing errors (7%). It emphasizes that Nursing course students are familiar with the use of the Health Descriptors Database to choose terms for their academic work, in addition to the mandatory use of the catalog form prepared by the Source Cataloging sector of the UFAL Central Library, the same did not occur with the Social Service course, which did not adopt the mandatory catalog form nor the use of authorized tools, such as thesauri or specialized indexes, a fact that causes the recurrent use of natural language, and, therefore, in some noises that interfere with the retrieval of information in the Repository system. It is clear that the authorized use of tools such as thesauruses, as well as the guidance of the librarian professional, contribute to the standardization and quality of subject metadata, providing greater effectiveness and efficiency in document search. It concludes, emphasizing the importance for the repository to adopt an indexing policy, providing means to assist users in the use of tools such as thesauruses and indexes that can guide them in choosing terms, thus making it possible to raise the level of quality of content description documents, aiming at precision in the process of searching and retrieving information.

Keywords: institutional repository; information representation; indexing; self-archiving.

LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura do RIUFAL.....	55
Figura 2 – Página inicial do RIUFAL	56
Figura 3 – Campos de busca do RIUFAL	56
Figura 4 – Acesso às comunidades e coleções	57
Figura 5 – Quantitativo de documentos com os metadado própolis vermelho e própolis vermelha	68

LISTA DE GRÁFICOS

Gráfico 1 – Metadados cadastrados (discentes de Enfermagem)	61
Gráfico 2 – Metadado de assunto cadastrados (discentes de Serviço Social)	62
Gráfico 3 – Comparativo dos metadados de assunto cadastrados (discentes de Enfermagem e Serviço Social).....	63

LISTA DE TABELAS

Tabela 1 – Índice de concordância de assunto cadastrados no RIUFAL	59
Tabela 2 – Metadados de assunto cadastrados no RIUFAL pelos cursos de Enfermagem e Serviço Social	60
Tabela 3 – Metadados de assunto cadastrados no RIUFAL (Enfermagem).....	61
Tabela 4 – Metadados de assunto cadastrados no RIUFAL (Serviço Social)	62
Tabela 5 – Metadados de assunto em linguagem natural cadastrados no RIUFAL	64

LISTA DE QUADROS

Quadro 1 – Variáveis da pesquisa	59
Quadro 2 – Metadados de assunto cadastrados com variações de plural, singular ou siglas..	66
Quadro 3 – Variações dos metadados de assunto cadastrados.....	66
Quadro 3 – Variações dos metadados de assunto cadastrados.....	67
Quadro 4 – Metadados de assunto com erros de grafia ou digitação (discentes de Enfermagem)	68
Quadro 5 – Metadado de assunto com erros de grafia ou digitação (discentes de Serviço Social).....	69
Quadro 6 – Metadados de assuntos com variações de entradas no DeCS, Catálogo da BN e no SIBI/UFAL	70

LISTA DE ABREVIATURAS E SIGLAS

BDTD	Biblioteca Digital Brasileira de Teses e Dissertações
BN	Fundação Biblioteca Nacional
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CEPAL	Comissão Econômica para América Latina
CI	Ciência da Informação
CONSUNI	Conselho Universitário
CTIC	Coordenação de Tecnologia da Informação e Comunicação
DeCS	Descritores em Ciências da Saúde
DRIVER	<i>Digital Repositories Infrastructure Vision for European Research</i>
DSpace	<i>DSpace Institutional Digital Repository Systems</i>
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IES	Instituição de Ensino Superior
IFLA	<i>International Federation of Library Associations and Institutions</i>
LA REFERENCIA	Rede Federada de Repositórios Institucionais de Publicações Científicas na América Latina
LTI	Laboratório de Tecnologia da Informação
MEC	Ministério de Educação
OA	<i>Open Access</i>
OAI	<i>Open Archives Initiative</i>
OAI-PHM	<i>Open Archives Initiative – Protocol for Metadata Harvesting</i>
OASISBR	Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto
OEA	Organização dos Estados Americanos
OPENDOAR	Diretório de Repositórios Acadêmicos Mundiais de Acesso Aberto
PPGCI	Programa de Pós-Graduação em Ciência da Informação
PROPEP	Pró-Reitoria de Pós-Graduação e Pesquisa
RENERE	Rede Nordeste de Repositórios Digitais das Instituições de Ensino Superior, Profissional e Tecnológico
RI	Repositório Institucional
RIIs	Repositórios Institucionais
RIAA	Rede Brasileira de Repositórios Institucionais de Publicações Científicas em Acesso Aberto
RIUFAL	Repositório Institucional da Universidade Federal de Alagoas

SEER	Sistema Eletrônico de Editoração de Revista
SIBI/UFAL	Sistema de Bibliotecas da Universidade Federal de Alagoas
Sinaes	Sistema Nacional de Avaliação da Educação Superior
SOCs	Sistemas de Organização do Conhecimento
SRI	Sistema de Recuperação de Informação
TCCs	Trabalhos de Conclusão de Curso
TICs	Tecnologias de Comunicação e Informação
UFAL	Universidade Federal de Alagoas
Unesco	<i>United Nations Educational, Scientific and Cultural Organization,</i>
WWW	<i>World Wide Web</i>

SUMÁRIO

1	INTRODUÇÃO.....	16
2	A UNIVERSIDADE E A BIBLIOTECA UNIVERSITÁRIA NO CONTEXTO DA SOCIEDADE DA INFORMAÇÃO.....	22
2.1	Universidade e Biblioteca Universitária	22
2.2	A biblioteca universitária.....	25
2.3	A biblioteca universitária brasileira	29
3	O MOVIMENTO DE ACESSO ABERTO E O PAPEL DOS REPOSITÓRIOS INSTITUCIONAIS.....	32
3.1	Repositórios Institucionais	32
3.2	Repositórios Institucionais no Brasil	38
3.3	Autoarquivamento nos repositórios	40
4	REPRESENTAÇÃO DA INFORMAÇÃO, FATOR DETERMINANTE PARA O ACESSO	44
4.1	Representação da informação no ambiente digital.....	44
4.2	A indexação dos cabeçalhos de assunto nos repositórios	46
5	PERCURSO METODOLÓGICO.....	52
5.1	Delineamento da pesquisa	52
5.2	Universo e amostra da pesquisa	53
5.3	Coleta, sistematização e análise dos dados	58
6	ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	60
6.1	Metadado de assunto autorizado.....	61
6.2	Metadado de assunto linguagem natural.....	64
6.3	Metadado de assunto sem padronização	65
6.4	Metadado de assunto com erros de grafia e digitação.....	68
7	CONSIDERAÇÕES FINAIS	71
	REFERÊNCIAS.....	74

1 INTRODUÇÃO

O desenvolvimento das Tecnologias de Comunicação e Informação (TICs) possibilitaram a produção e disponibilização de conteúdos informacionais em uma escala de proporções gigantescas no ambiente informacional – a Internet –, sendo o acesso a esses conteúdos de forma instantânea e ilimitada. A grande quantidade de informação disponível no ambiente da *World Wide Web* (WWW), criou o caos informacional, em que informações de toda natureza, relevantes ou irrelevantes, ocupem o mesmo espaço.

Nesse ambiente, o acesso à informação tem se apresentado como um grande desafio da Ciência da Informação (CI), principalmente quando se trata da informação científica, a qual é, em sua maioria, oriunda das universidades. Para esses espaços acadêmicos, os desafios consistem em como organizar, preservar e, principalmente, disseminar essas informações.

As universidades carregam consigo três princípios fundamentais: a pesquisa, o ensino e a extensão. São elas que detêm uma grande concentração de pesquisadores de alto nível, responsáveis pela realização de pesquisa científica, desenvolvendo projetos em diversas áreas do conhecimento, os quais são materializados através de Trabalhos de Conclusão de Curso (TCCs), Dissertações, Teses, Artigos de periódicos e demais produções acadêmicas. Sendo assim, é fundamental que as informações produzidas no seu âmbito sejam divulgadas, pois possibilita o seu reconhecimento no contributo à produção do conhecimento para o desenvolvimento da sociedade, além de garantir o acesso facilitado aos resultados de pesquisas.

Dentre as inovações desse ambiente informacional surgem os Repositórios Institucionais (RIs), oriundos do movimento da Ciência Aberta, os RIs foram concebidos com o intuito de disponibilizarem as produções científicas em acesso aberto, sem qualquer barreira ou entrave. O pesquisador tem acesso aos conteúdos informacionais contidos nos documentos, sendo o acesso de maneira direta, em formato digital.

Leite (2009, p. 21) define RI como “um serviço de informação científica – em ambiente digital e interoperável – dedicado ao gerenciamento da produção intelectual de uma instituição”. Nesse cenário, os RIs se apresentam como uma forma relevante para as instituições de pesquisa e ensino reunirem a sua produção científica, potencializando a disseminação e o acesso à informação científica.

O RI é um órgão de mediação da informação produzida pela instituição e seus usuários, é através dele que os usuários terão acesso à produção intelectual da instituição, suas descobertas científicas e inovações tecnológicas. Para cumprir verdadeiramente a sua função o RI precisa prover a interoperabilidade e o autoarquivamento. Segundo Villalobos e Gomes

(2015, p. 4), a interoperabilidade consiste na “comunicação entre diferentes sistemas de informação”, disponibilizando os metadados que identificam, descrevem e localizam os recursos informacionais armazenados no RI. Sendo assim, ela possibilita a interação do sistema com o usuário, permite que tenha acesso aos dados, podendo compartilhá-los com outros sistemas de diferentes *hardwares* e *softwares*.

O autoarquivamento vem se tornando uma das etapas operacionais fundamentais, ele possibilita ao autor realizar o depósito do seu documento diretamente na base de dados. Além disso, esse procedimento agiliza o processo de visibilidade e disponibilidade da publicação. No entanto, esse procedimento pode interferir na qualidade e, principalmente, na confiabilidade do sistema. Para Freitas (2019), sobre a questão do autoarquivamento, esse procedimento pode comprometer a qualidade dos descritores de assunto, dificultando no processo de busca da informação.

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 (Freitas, 2019, p. 15).

Essa preocupação é pertinente, visto que o autor do trabalho depositado pode não ter utilizados padrões metodológicos para a escolha dos metadados “assunto”, além de uso inadequado de alguns termos.

É preciso que os RIs garantam a descrição de documentos com metadados padronizados. Os padrões de metadados podem auxiliar na representação e organização de qualquer ambiente informacional digital. Segundo Sanchez, Silva e Vechiato (2018, p. 48): “[...] a descrição dos objetos digitais e dos dados que compõem a pesquisa são determinantes para a recuperação e encontrabilidade de cada um deles”.

Nesse aspecto, faz-se necessário uma análise criteriosa da forma de representação da informação nos ambientes digitais, principalmente nos RIs. A representação pode ser subdividida em representação descritiva (descrição física) e representação temática (descrição de conteúdo). A descritiva se caracteriza pelo processo que descreve as características específicas do documento, visando a sua individualização, podendo ser através da autoria, do título, formato e período, por exemplo, enquanto a temática consiste na análise do conteúdo de um documento, ou seja, o tema ou assunto abordado por ele, objetivando a sua disseminação. Para nosso objeto de estudo, focaremos apenas na representação temática da informação.

Segundo Albuquerque, Gaudêncio e Santos (2019, p. 19), a representação da informação “consiste em uma tarefa indispensável para a organização, recuperação, o acesso e uso dos

registros do conhecimento. Para tanto, são criados processos, produtos e instrumentos que corroborem para (sic) essa finalidade”. Dentre os instrumentos estão os tesouros (ou *thesaurus*), listas de cabeçalhos de assuntos e glossários, os quais podem ser aperfeiçoados, atualizados e implantados nas estruturas dos sistemas de recuperação da informação na rede mundial de computadores (web). Esses instrumentos auxiliam na padronização dos metadados “assunto”. Portanto, a representação da informação é etapa fundamental para o acesso e a recuperação dos conteúdos informacionais, e é a partir dela que se dá condições para identificar e reforçar um recurso informacional.

Entretanto, é preciso que os RIs garantam a descrição de documentos com metadados padronizados. Metadado é comumente definido na literatura como “dados sobre dados”, ou seja, são descrições de dados armazenados em banco de dados, descrevem os atributos e o conteúdo de um documento original. Os metadados podem ser de autoria, título, assunto, formato, conteúdo, língua, local, data, coautoria, dentre outros. Conforme Souza, Catarino e Santos (1997, p. 94), a “finalidade principal dos metadados é documentar e organizar de forma estruturada os dados das organizações com o objetivo de minimizar duplicação de esforços e facilitar a manutenção dos dados”.

Por isso, é fundamental que sejam estabelecidos padrões de configuração desses metadados, possibilitando a identificação individual de cada item com toda a sua especificidade e pertinência, descrevendo de forma fidedigna o seu conteúdo, principalmente no tocante aos metadados “assunto”, fato muito relevante no ambiente dos RIs.

Para Freitas (2019, p. 13), “se uma publicação não puder ser encontrada eletronicamente, ela é pouco útil e, em termos práticos, também não existe”. Nesse sentido, é fundamental estabelecer critérios para padronizar os descritores dos metadados que compõem o sistema de busca.

Dentre o padrão do metadado “assunto”, destaca-se a escolha do uso dos termos de busca para representá-los, uso de linguagem controlada e natural, palavras no singular ou plural, uso ou não de siglas, variações semânticas e sintaxe. Todos esses fatores são responsáveis pela recuperação da informação. Logo, é fundamental que eles apresentem confiabilidade, coerência e obedeça a critérios metodológicos.

O Repositório Institucional da Universidade Federal de Alagoas (RIUFAL) teve sua política de informação regulamentada a partir de 07 de novembro de 2016, por meio da Resolução nº 45/2016-CONSUNI/UFAL. Apesar de sua regulamentação, alguma das competências do RIUFAL ainda não pode ser colocada em prática na sua totalidade, boa parte devido ao pequeno número de funcionários no setor, problemas operacionais e tecnológicos,

além dos cortes orçamentários que atingem as universidades federais. Buscando intensificar a interoperabilidade entre os usuários do repositório, o RIUFAL começou a testar de modo piloto, a partir do ano de 2020, a utilização da prática do autoarquivamento nos cursos de Enfermagem e Serviço Social. Os alunos concluintes dos referidos cursos foram autorizados a fazerem o autoarquivamento dos TCCs diretamente na base de dados do RIUFAL. O processo foi efetuado através do cadastramento dos *e-mails* dos alunos no sistema, possibilitando, a criação de senhas de acesso individual a qual permitia a inserção dos metadados na plataforma. São inseridos os metadados descritivos como: autor(es), colaborador(es), título, paginação, curso, unidade acadêmica, ano e instituição, forma de acesso (aberto ou embargado), e o metadado temático: assunto. Para tanto, foi elaborado um tutorial orientando os alunos em como preencher os campos na base de dados do repositório.

Nesse contexto, essa pesquisa busca analisar a representação temática no metadado “assunto” contido nos TCCs dos referidos cursos cadastrados no RIUFAL, através do autoarquivamento, efetuados no período de 2020 a 2022.

O interesse pelo tema se deve à prática como bibliotecário no RIUFAL, onde me deparei com a necessidade de padronização e normatização dos metadados no que concerne aos descritores de assunto, nesta pesquisa abordada como “metadado ‘assunto’”, visto que, em alguns casos, foram detectados trabalhos com termos apresentando desconformidades com os padrões metodológicos dos tesauros especializados em determinadas áreas do conhecimento, esse fato ocorre devido ao uso excessivo da linguagem natural, que possibilita aos usuários utilizarem termos livres, sem a preocupação de se buscar o vocabulário utilizado em sua área específica de conhecimento.

Nessa perspectiva, a presente pesquisa levanta o seguinte questionamento: os metadados de assunto cadastrados pelos usuários no RIUFAL, no momento do autoarquivamento de seus trabalhos, foram qualificados e validados tendo como referencial os cabeçalho autorizado utilizados pelo Catálogo de Assunto do Sistema Integrado de Biblioteca da Universidade Federal de Alagoas (SIBI/UFAL)?

A hipótese deste estudo é de que o cadastramento do metadado “assunto” por parte dos autores pode não seguir uma padronização e apresentar ruídos, vindo a comprometer a recuperação dos dados relevantes no momento da pesquisa na Base de Dados do RIUFAL.

O RI é uma unidade de informação responsável pelo armazenamento, preservação, disseminação dos conteúdos científicos produzidos pela instituição ao qual faz parte, portanto, é fundamental que garanta o acesso a esse conteúdo de uma forma precisa, possibilitando

agilidade, consistência e qualidade das informações sobre sua guarda, principalmente por se tratar de um ambiente prioritariamente digital.

É através de sua produção que a universidade proporciona os avanços científicos, tecnológicos, culturais e sociais, contribuindo para o desenvolvimento da sociedade, desse modo, permitir o acesso a seus conteúdos científicos se torna imprescindível para o reconhecimento da instituição no âmbito científico, sobretudo pela sociedade.

Esta pesquisa se justifica pela importância do RI para a universidade, pois ele se constitui em uma via que interliga a produção científica às comunidades acadêmica, científica e, principalmente, à sociedade. Nesse aspecto, promover estudos que viabilizem a melhoria do acesso às informações presentes no RI é de suma importância para contribuir para a visibilidade e o reconhecimento local, nacional e internacional da instituição. Justifica-se para o RIUFAL, por ser a pesquisa mais um instrumento para a utilização da unidade e para melhorar e aperfeiçoar os mecanismos de busca na sua base de dados, no tocante ao uso dos cabeçalhos de assunto. No que concerne ao Programa de Pós-Graduação em Ciência da Informação (PPGCI) da UFAL, justifica-se pela pertinência do tema abordado, enquadrando-se na área de concentração em Informação, Comunicação e Processos Tecnológicos, sendo parte dos estudos desenvolvidos pelo referido Programa, além da contribuição para a CI como mais um estudo relacionado aos RIs e às questões de indexação e acesso da informação nesses ambientes.

Diante do exposto, o objetivo geral da presente pesquisa foi analisar a representação da informação no RIUFAL, no que concerne ao metadado “assunto”. Tendo como objetivos específicos:

- a) Coletar os TCC dos cursos de Enfermagem e Serviço Social cadastrados através do autoarquivamento;
- b) Avaliar os metadados de assunto cadastrados nos TCC coletados;
- c) Comparar os metadados utilizados com cabeçalhos autorizados do catálogo de assunto do SIBI/UFAL;
- d) Identificar se, dentre os metadados cadastrados, existe algum que apresente ruídos que possam interferir no processo da busca e recuperação da informação no sistema.

Para proporcionar melhor compreensão, a pesquisa foi estruturada da seguinte forma: Seção 1, com a contextualização sobre a temática, questão-problema, justificativas e objetivos; Seção 2, que aborda a universidade e a biblioteca universitária no contexto da sociedade da informação; Seção 3, que trata sobre os RIs, sua criação e importância, bem como as questões

referentes ao autoarquivamento; Seção 4, em que aborda a representação da informação e a questão da indexação nos RIs e sua importância, contextualiza o RIUFAL no âmbito da UFAL; Seção 5, que descreve os procedimentos metodológicos, como a caracterização da pesquisa, universo e amostra, coleta e sistematização dos dados; Seção 6, a qual trata da análise e discussão dos resultados; seguida das Considerações finais; e, por fim, as referências que foram adotadas para o desenvolvimento teórico desta pesquisa.

2 A UNIVERSIDADE E A BIBLIOTECA UNIVERSITÁRIA NO CONTEXTO DA SOCIEDADE DA INFORMAÇÃO

Esta seção discorre acerca da universidade e sua transição para a sociedade da informação, algo que começa a acontecer a partir da segunda metade do século XX e transcorre ao longo do século XXI. Aborda a transformação do processo de ensino com o uso de novas tecnologias, a questão da internacionalização através dos intercâmbios institucionais, a necessidade da formação continuada dos professores e a oferta de produtos e serviços oferecidos no âmbito da instituição.

2.1 Universidade e Biblioteca Universitária

O século XX proporcionou o desenvolvimento das universidades, os avanços tecnológicos introduzidos contribuíram para o desenvolvimento de produtos e serviços ofertados, marcando profundas transformações em relação ao ensino superior, atingindo vários países no mundo, conseqüentemente, intensifica-se a automação, com o uso do computador, o qual vai impactar na produção de produtos e serviços produzidos e oferecidos por essas universidades. Segundo Rossato (2005, p. 171), o “século XX registrou grandes avanços nos diversos campos sociais. Nesse período, especialmente na segunda metade, também a universidade conheceu notável crescimento, seja em áreas tradicionais seja nos novos países”.

Surgem novos cursos, dentre eles o de Computação, instituindo uma nova hegemonia no meio acadêmico, assim como as Ciências Sociais começam a ter mais visibilidade, as Engenharias, a Química e a Medicina, bem como as demais disciplinas, foram reestruturadas incorporando recursos tecnológicos. O computador passa a ser presença constante no dia a dia das universidades, desde os processos administrativos, aos laboratórios, sala de aulas e sistema de bibliotecas, interligando em rede toda a estrutura da instituição.

Nos Estados Unidos, começa a consolidar o estilo “pragmático” americano, introduzindo a visão do mercado capitalista e instituindo os ensinamentos de universidade-empresa, com foco no empreendedorismo, o qual vai influenciar a forma de ensinamento e direcionamento das universidades. O sistema capitalista passa a interferir no fazer da universidade, direcionando os cursos de seu interesse. Desse modo, a ciência estava a serviço do mercado, conforme Diógenes e Cunha (2017, p. 108) destacam: “[...] no século XX, encontrava-se em pleno desenvolvimento da indústria científica. Ciência e produção expressam, cada vez mais claramente, a inter-relação e as influências mútuas que as unem”.

Ficando evidente assim, a correlação entre a universidade e o mercado, o qual determina o foco dos seus interesses, implementando recursos e insumos em áreas específicas do conhecimento científico.

As universidades estabelecem uma nova modalidade de ensino, o “ensino a distância”, e passa a explorar os recursos oferecidos pela internet, amplia o espaço físico de atuação, levando a educação superior a diferentes localidades. Nesse sentido, essa ampliação da área de atuação das universidades extrapola o seu espaço físico de localização, possibilitando o acesso a um maior público e, conseqüentemente, universalizando-se, no sentido literal.

Dessa maneira, é a partir do final do século XX que se começa delinear a sociedade da informação, a qual lançará a universidade a um outro paradigma no que concerne à produção e transferência do conhecimento.

O século XXI traz consigo o desenvolvimento das TICs, ampliando ainda mais a atuação das universidades. Surgem novas ferramentas informacionais (as mídias sociais como o *Facebook*, *Instagram*, *Twitter*, *Blogs* e *Youtube*, as plataformas científicas e os repositórios), que dinamizam as formas de interação, produção, consumo e acesso à informação. Nessa nova sociedade, o advento da internet 2.0 proporciona o acesso a conteúdos informacionais em um volume nunca imaginado, estando esses conteúdos disponíveis de maneira acessível, independente do suporte e da localização. As instituições não estão agora limitadas aos seus espaços físicos, caem-se as barreiras de tempo e espaço, as informações integram uma rede mundial.

Sites, portais, blogs, produção audiovisual, 3D, celulares multifuncionais, ao mesmo tempo em que geram novas formas de reunião de saberes e pessoas, geram também novas formas de dispersão, novas polifonias, novas formas em que os quadros normativos ficam obscurecidos pelas novas expressões polinômicas de redes sem fronteiras institucionais (González de Gómez, 2011, p. 237).

González de Gómez (2011) ainda enfatiza que, a universidade dessa sociedade precisa ser plural, comunicativa e interativa, atendendo às condições e possibilidades que vão além das infraestruturas de acesso.

Amplia-se a modalidade do ensino a distância. São incorporados a mobilidade e o intercâmbio entre docentes e discentes em instituições de ensino por todo o mundo, consolida a internacionalização do conhecimento acadêmico e se intensifica o incentivo à educação continuada.

Para Stallivieri (2015, p. 10), sobre a internacionalização e a mobilidade acadêmica, as universidades:

Passaram a protagonizar um novo modelo de internacionalização, através do desenho de agendas positivas em relação às ações de cooperação internacional, colocando recursos, investindo na formação dos quadros, dialogando sobre novas práticas pedagógicas que dessem conta das questões da internacionalização curricular, com um melhor desenho de programas de mobilidade de alunos, de professores e de pesquisadores, com base na reciprocidade (Stallivieri, 2015, p. 10).

Dessa forma, pretende atender às necessidades globais com possibilidades de inserção em um mercado multicultural e de um exigente e competitivo mercado de trabalho, visando a troca de experiências e de conhecimento.

No bojo dessas mudanças, as universidades, mais uma vez, reformulam-se para acompanhar essa nova sociedade que se apresenta, a sociedade da informação. Nessa sociedade, a universidade assume um diferenciado papel, ampliando sua atuação, incorporando além do ensino, pesquisa e extensão, a questão da inovação.

A sociedade da informação é alicerçada na ciência, tecnologia, pesquisa e na educação, tomando-as como fatores insubstituíveis de transformações. Para González de Gómez (2011, p. 227), esses fatores “junto as ofertas vigorantes das tecnologias digitais – interativas, nômades e hipermidiáticas –, que se mostrariam habilitadas para mediar novos experimentos econômicos, sociais e culturais de inovação”.

Diante desse cenário, as universidades além de produzirem conhecimentos, precisam disseminá-los de forma que sejam reconhecidas como atuantes inovadoras e provedoras de desenvolvimento científico e tecnológico, contribuindo para a evolução da humanidade. Conforme enfatiza Assis (2013), sobre a importância do acesso das informações produzidas nas universidades:

O acesso à informação científica é considerado fundamental para o desenvolvimento tecnológico e científico de uma comunidade. Uma sociedade desenvolvida é aquela na qual a ciência é mais desenvolvida e que possui uma produção crescente de pesquisas científicas (Assis, 2013, p. 212).

Prover o acesso às suas produções passa a ser o ponto focal das universidades diante dessa nova sociedade interligada, contribuindo assim para o desenvolvimento da sociedade, disponibilizando o que foi produzido em prol do desenvolvimento tecnológico, econômico, cultural e, principalmente, social.

Nesse contexto, os RIs passam a ser uma ferramenta primordial para a organização, preservação, disseminação e, principalmente, para o acesso à produção institucional, constituindo a ponte que possibilita a integração da universidade com a sociedade da informação.

Dentre os desafios para a universidade no contexto da sociedade da informação está a questão da produção do conhecimento, a qual, com o desenvolvimento da inteligência artificial (IA) e das ferramentas como Chat-GPT, passa a contar também com máquinas autônomas, o conhecimento deixa de ser uma produção intelectual exclusivamente humana, proporcionando com isso questionamentos sobre os direitos autorais e a produção do conhecimento.

2.2 A biblioteca universitária

Com a chegada do século XX, especificamente após o período da Segunda Guerra Mundial, é que a biblioteca universitária passa a ser vista como parte integrante dos três princípios norteadores da universidade: ensino, pesquisa e extensão. Para Diógenes e Cunha (2017), é nesse momento que a biblioteca universitária começa a ser valorizada.

Foi no decorrer do século XX, a partir da Segunda Guerra até, aproximadamente, 1970, que no capitalismo organizado, o ensino começa sua caminhada para o ensino de massa, que a biblioteca universitária se torna uma parte importante na universidade para contribuir com seus objetivos de ensino, pesquisa e extensão (Diógenes; Cunha, 2017, p. 125).

A biblioteca universitária é um órgão de apoio à universidade, é uma unidade fundamental na concretização desses princípios. Suas coleções são formadas para atingir esses objetivos, tendo como foco atender às necessidades informacionais dos professores, alunos, técnicos administrativos e pesquisadores da instituição, além de assistir também à comunidade externa.

Diante disso, a biblioteca universitária começa a refletir como espelho a universidade na qual está inserida, necessitando de investimentos para melhorar seus produtos e serviços, e oferecer o apoio necessário para o desenvolvimento científico, cultural, tecnológico, econômico e social da sociedade.

Assim, as bibliotecas, principalmente as universitárias, apresentam uma forma diversificada, incorporam produtos e serviços, intensifica o uso de bases de dados digitais e de recursos eletrônicos e digitais, seus acervos não são mais exclusivamente constituídos de livros e periódicos impressos.

Nunes e Carvalho (2016) destacam a busca da modernização das bibliotecas universitárias, enquadrando-se na evolução da sociedade a qual está inserida, e ressaltam que:

Influenciadas pelo advento da internet na década de 90 do Século XX e a inserção das tecnologias de informação e comunicação nas universidades e na sociedade de uma maneira geral, as bibliotecas universitárias buscam também

sua modernização, e as experiências passam a priorizar o uso dessas tecnologias no gerenciamento das bibliotecas e na elaboração e disponibilização de catálogos on-line, motivados pelas diretrizes lançadas pela *The International Federation of Library Associations and Institutions* (IFLA) no fim do Século XX (Nunes; Carvalho, 2016, p. 188).

Com o aproveitamento dessas novas tecnologias, as bibliotecas universitárias passaram a constituir redes cooperativas de trocas de informações, compartilhando catálogos online e acessos a bases de dados, a fim de interligar cada vez mais recursos e usuários para possibilitar uma maior visibilidade da instituição. Para Santa Anna (2016), a biblioteca é uma unidade adaptativa que incorpora as tendências e exigências dos usuários e as inovações tecnológicas vigentes.

Sousa e Fujita (2012, p. 66) abordam o impacto e a importância das novas tecnologias no processo de armazenar, tratar e recuperar a informação, e ressaltam que “uma mudança perceptível está no fato dos catálogos, que por um grande período foram construídos para uma comunidade local, passaram a ser disponibilizados na web em formato on-line, tornando-se disposto para acesso do público em geral”.

A biblioteca universitária tem a possibilidade de extrapolar o seu espaço físico, ela poderá oferecer seus produtos e serviços em uma esfera global. Conforme Nunes e Carvalho (2016, p. 190), “[...] cuja utilização projeta a biblioteca ao nível de cooperação e visibilidade mundial, facilitando o acesso à informação pelos usuários e a formação de redes de informação”.

Com o advento do século XXI, intensifica-se o uso das TICs, por conseguinte, o acesso direto às redes sociais se torna algo fundamental. A internet concentra um volume inimaginável de informações de acesso aberto disponíveis para uma sociedade cada vez mais conectada às novas tecnologias. Sousa e Fujita (2012, p. 67) ressaltam que, “atualmente, com o uso da internet também é possível ter acesso a uma imensa massa de documentos e outros recursos de informação e esses documentos podem ser tratados e disponibilizados em bases de dados ou nos catálogos de bibliotecas”.

Como se percebe, os avanços tecnológicos aliados ao uso da internet acarretam uma nova possibilidade para as bibliotecas universitárias, necessitando cada vez mais que os bibliotecários dessas instituições se familiarizem com o uso dessa tecnologia, explorando-a e extraíndo todos os recursos que venham a dinamizar o acesso e a produção de conteúdos informacionais. Além disso, esses avanços transformaram o modo como a biblioteca universitária interage com a produção, disseminação da informação e a forma de acesso e

disponibilização aos usuários. Sobre a evolução das bibliotecas universitárias na sociedade da informação, Nunes e Carvalho (2016) destacam que:

As bibliotecas universitárias, assim como todas as demais unidades de informação, têm evoluído com o passar do tempo a fim de atender não apenas às necessidades de informação do público, como também no sentido de acompanhar as mudanças no campo das tecnologias da informação e comunicação, assim como as mudanças de nível comportamental dos usuários, cada vez mais conectados (Nunes; Carvalho, 2016, p. 174).

As bibliotecas, principalmente as universitárias, acompanham o desenvolvimento da sociedade da qual faz parte, desenvolvendo produtos e serviços adequados aos avanços tecnológicos vigentes. Desde a utilização de livros impressos, passando pelos recursos eletrônicos em CD-ROM e bases de dados remotos, chegando à biblioteca 2.0 de acesso à internet e consolidando, no século XXI, a parceria com os RIs de acesso aberto.

Os repositórios institucionais (RI) têm-se tornado uma alternativa relevante para as instituições de pesquisa e ensino reunirem a produção científica em um único local, potencializando a disseminação e acesso à informação científica. Um dos pré-requisitos para que os RI cumpram esse papel é que eles garantam a descrição de documentos com metadados padronizados (Gomes, 2016, p. 93).

Para Alves (2017), a facilidade com o uso da internet no sentido de disponibilizar diretamente os conteúdos informacionais em rede não garante por si só a qualidade ou relevância da informação ali contida.

As possibilidades abertas pela internet com seus mecanismos de publicação direta na rede tornam o acesso a um documento digital uma mera questão de conhecer sua URL. No entanto, esta facilidade de acesso tem como contrapartida a grande dificuldade de encontrar informação relevante, as atividades de *Information Discovery*. Encontrar a informação relevante é fundamental para que ela possa ser utilizada (Alves, 2017, p. 35).

Entende-se que um dos grandes desafios da sociedade da informação é acessar a informação pertinente em meio a grande diversidade de informação, nem sempre confiável ou verídica como é o caso das *fake news*.

Na sociedade da informação, foi transformado radicalmente o papel das unidades de informação (biblioteca e repositório) no que concerne ao ciclo de intermediação e acesso ao documento. Apesar da similitude de alguns serviços e produtos entre a biblioteca universitária e o repositório institucional, o que os diferenciam principalmente é a forma de acesso e a fonte produtora.

Segundo Leite (2009), apesar das diversidades de suportes informacionais contidos nos RIs, alguns similares aos da biblioteca digital, o que caracteriza a sua natureza e o que diferencia delas é que seus conteúdos são referentes à produção intelectual de uma instituição.

Todo repositório institucional de acesso aberto pode ser considerado um tipo de biblioteca digital, mas nem toda biblioteca digital pode ser considerada um repositório institucional. [...] repositórios institucionais de acesso aberto à informação científica lidam exclusivamente com a produção intelectual de uma instituição. Portanto, não se prestam à aquisição e ao armazenamento de conteúdos de outra natureza [...], como pode ser o caso de bibliotecas digitais (Leite, 2009, p. 21).

As bibliotecas universitárias tradicionais ou digitais são responsáveis pela organização, acesso e disseminação das informações científicas e não científicas, oriundas de qualquer parte, de suportes com acessos físico e, às vezes, remoto dos seus conteúdos. Enquanto os RIs são responsáveis pela organização, acesso e disseminação das informações científicas geradas pelas instituições das quais fazem parte, sendo essas informações em formato digital de acesso aberto disponibilizadas em rede.

Por apresentarem essas similitudes, a biblioteca universitária e o repositório institucional, em sua maioria, ocupam o mesmo espaço físico, ou seja, o repositório funciona e desempenha suas atividades dentro da biblioteca universitária, contudo, de forma independente.

Para Santa Anna (2016), as bibliotecas redefinem constantemente seus produtos e serviços e, por isso, são consideradas unidades adaptativas, adquirindo novas características visando às exigências dos usuários e das inovações tecnológicas. Nesse aspecto, a parceria entre a biblioteca universitária e o repositório vem contribuir para a maior visibilidade da instituição tornando-se uma única unidade informacional em prol do desenvolvimento, educacional, cultural, tecnológico e, sobretudo, social.

Sobre o papel das bibliotecas, sejam elas universitárias ou não, Oliveira, Alves e Maia (2013), Silveira (2014) e Santa Anna (2015) ressaltam que as bibliotecas devem oferecer produtos e serviços diversificados, não se limitando apenas aos acervos, às técnicas de tratamento da informação, mas tendo como foco principal a satisfação da comunidade, dos grupos sociais e, enfim, de todos os cidadãos, tendo os usuários sempre como seu objetivo final.

Santa Anna (2016) destaca que a biblioteca na sociedade da informação se constituirá em um diferente ambiente informacional.

Na contemporaneidade, a biblioteca extrapolou sua função meramente informacional, exercendo também funções sociais. Com essa constatação, presume-se que no futuro, a biblioteca ampliará suas funções deixando de ser

um ambiente meramente informacional para transformar-se em espaços de convivência (Santa Anna, 2016, p. 243).

Portanto, todas as bibliotecas, incluindo também a biblioteca universitária, nessa sociedade da informação, devem ser necessariamente tecnológicas – explorando todos os recursos informacionais existentes –, inclusivas – adaptadas às necessidades informacionais dos usuários –, diversificadas em produtos e serviços, bem como, um lugar de convivência e interação social.

Diante disso, é necessário que o profissional bibliotecário tenha sempre atitude proativa, buscando uma formação continuada, “atenado” com as TICs, tendo como foco o usuário final, reafirmando o caráter social do bibliotecário e da biblioteca universitária.

2.3 A biblioteca universitária brasileira

A partir da segunda metade do século XX, a biblioteca universitária brasileira começa a implementar uma política de apoio ao desenvolvimento científico e tecnológico, apoiado por incentivos de organismos internacionais com *United Nations Educational, Scientific and Cultural Organization* (UNESCO), Organização dos Estados Americanos (OEA), Comissão Econômica para América Latina (CEPAL) e *International Federation of Library Associations and Institutions* (IFLA), que foram predominante em toda a América Latina. Conforme Santos e Araújo (2022), essas organizações têm como meta fomentar os planos nacionais de serviços bibliotecários visando promover o apoio ao desenvolvimento do sistema de informação, ciência e tecnologia. Para esses autores, a biblioteca universitária brasileira:

Começa por se reinventar, alterando sua forma de organização para atender a uma universidade cada vez mais exigente por informação através de seus cursos de Pós-Graduação, que buscavam o desenvolvimento de pesquisas e a geração de informação científica nacional sobre os mais variados aspectos do desenvolvimento que era exigido naquele momento (Santos; Araújo, 2022, p. 10).

Buscando atender às demandas informacionais e para complementar os acervos das bibliotecas universitárias, foi nesse contexto que surge, em 2000, a criação do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Ele se constitui em um dos maiores acervos científicos digitais do Brasil, reunindo e disponibilizando conteúdos nacionais e internacionais das instituições de ensino e pesquisa brasileiras. Seu acervo reúne quase 50 mil periódicos com texto completo e, aproximadamente, 455 bases de dados de conteúdos diversos. Além de periódicos, disponibiliza também patentes, referências,

dados estatísticos, material audiovisual, normas técnicas, teses e dissertações. Segundo a CAPES (2022), esse Portal “foi criado para reunir material científico de alta qualidade e disponibilizá-lo à comunidade acadêmica brasileira”, tendo como objetivo, reduzir as desigualdades regionais no que concerne ao acesso à informação científica, cobrindo todo território nacional.

Para os autores Almeida, Guimarães e Alves (2010), o Portal de Periódicos da CAPES possibilitou as bibliotecas suprirem as lacunas dos periódicos incompletos de seus acervos, além de permitir o acesso a novos títulos.

O Portal de Periódicos nasceu como forma de otimizar a política de acesso atualizado ao conhecimento científico. Sua implantação permitiu uma redução do custo médio do acesso por usuário e um direcionamento, em médio prazo, dos investimentos feitos pelo governo brasileiro por meio de uma análise precisa da utilização dos periódicos pela comunidade atendida. Ao mesmo tempo, buscou promover crescentemente o acesso universal a um acervo amplo e atualizado de textos completos publicados em periódicos internacionais e a bases de referência, sem limitações geográficas e de horário. Permitiu, finalmente, o preenchimento das então enormes lacunas nas coleções das bibliotecas, devido às eternas irregularidades no aporte de recursos destinados ao setor (Almeida; Guimarães; Alves, 2010, p. 228).

A criação do Portal de Periódicos Capes foi fundamental para o enriquecimento das bibliotecas universitárias, no sentido de que foi possível oferecer um acervo completo e atualizado, algo que seria impossível devido aos altos preços dos periódicos, além dos constantes cortes orçamentários infligidos às universidades, os quais tinham reflexos nas bibliotecas, pois ficavam sem recursos para aquisição e manutenção de assinaturas de periódicos, comprometendo com isso a qualidade do acervo.

No ano de 2002, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) lançou a Biblioteca Digital de Teses e Dissertações (BDTD), posteriormente, renomeada para Biblioteca Digital Brasileira de Teses e Dissertações. A BDTD tem como objetivo ser uma base de dados nacional de teses e dissertações em formato digital de textos completos, disponível em acesso aberto, contribuindo dessa forma, para maior visibilidade da produção científica nacional. Para Cunha e Diógenes (2016), essas ações colocaram o Brasil em consonância com a filosofia do acesso aberto, disponibilizando, gratuitamente, a literatura científica brasileira no ambiente web.

O IBICT também é responsável pela criação do Sistema Eletrônico de Editoração e Revista (SEER), em 2003, um *software* para construção e gestão de um periódico eletrônico. Conforme o SEER (2022), essa criação “é resultado da prospecção tecnológica realizada pelo IBICT para identificar aplicativos que possibilitassem o tratamento e a disseminação da

produção científica na web”. O seu processo editorial permite uma melhora na avaliação da qualidade dos periódicos e uma maior rapidez nos fluxos das informações, possibilitando a disponibilização imediata do periódico na web.

Outra importante contribuição para as bibliotecas universitárias foi incentivo à criação pelo IBICT dos repositórios institucionais, em 2009, os quais passam a atuar juntos com as bibliotecas no processo de armazenamento, guarda e, principalmente, de disseminação das produções científicas oriundas das instituições das quais fazem parte.

Diferentemente da BDTD, os repositórios armazenam uma gama variada de produtos informacionais, desde teses e dissertações, como também dados estatísticos, patentes, relatórios, vídeos dentre outros.

Um dos pontos preponderantes para a melhora da qualidade dos acervos e equipamentos nas bibliotecas universitárias foi a sua inclusão na avaliação do Ministério da Educação (MEC), nos critérios de qualidade dos cursos superiores, o que provoca a necessidade de também investir nesses ambientes, ampliar os acervos, adquirir equipamentos e mobiliários, melhorando toda a sua infraestrutura.

As bibliotecas universitárias têm, ao longo dos anos, desenvolvido projetos de expansão e qualificação de seus recursos informacionais, seja para auxiliar os cursos de graduação ou pós-graduação, aos quais estão vinculadas, seja para acompanhar a realidade da sociedade, possibilitando, assim, a sua contribuição junto aos RIs, no âmbito da ciência aberta para promover o acesso livre e gratuito ao conhecimento.

3 O MOVIMENTO DE ACESSO ABERTO E O PAPEL DOS REPOSITÓRIOS INSTITUCIONAIS

Nesta seção, será contextualizado o Movimento de Acesso Aberto à Informação Científica e suas estratégias (vias verde e dourada) para propiciar o Acesso Aberto à informação, bem como o surgimento dos repositórios, sua importância e contribuição para a visibilidade institucional, através da organização, preservação e disseminação da produção científica em acesso aberto, além de abordar a questão da necessidade do autoarquivamento nos repositórios.

3.1 Repositórios Institucionais

Tradicionalmente, as universidades têm sido reconhecidas como espaços de produção e transferência de conhecimento científico. O acesso à informação tem se apresentado como um grande desafio, principalmente quando se trata da informação científica. O alto preço imposto às publicações de artigos científicos em periódicos renomados passou a se constituir como uma grande barreira no processo de divulgação e acesso à informação científica.

Segundo Mueller (2006), Rosa e Gomes (2010), os altos preços praticados impactaram diretamente os orçamentos das bibliotecas, dificultando a assinatura de novos periódicos, assim como a continuidade das assinaturas existentes e, com isso, ocasionando prejuízos para os pesquisadores, professores e estudantes, enfim, toda a comunidade acadêmica, dificultando o acesso às informações científicas.

As implicações dessa prática forçada pelos editores geraram a denominada crise dos periódicos científicos. Conforme Mueller (2006), a crise foi imposta devido à impossibilidade das bibliotecas universitárias e de pesquisas americanas manterem suas coleções de periódicos e atenderem à crescente demanda de seus usuários, dispendo de pouco recursos financeiros.

A aparente estabilidade de que gozava o sistema de comunicação científica mundial foi abalada quando estourou a chamada crise dos periódicos, em meados da década de 1980, que já vinha se anunciando desde a década de [19]70. O gatilho da crise foi a impossibilidade de as bibliotecas universitárias e de pesquisa americanas continuarem a manter suas coleções de periódicos e a corresponder a uma crescente demanda de seus usuários, impossibilidade decorrente da falta de financiamento para a conta apresentada pelas editoras, cada ano mais alta [...] (Mueller, 2006, p. 31).

No Brasil, ela fica evidente na década de 1990, seus reflexos perduraram ainda por várias décadas. Não obstante a crise dos periódicos, os pesquisadores para terem respaldo e

reconhecimento de seus trabalhos necessitam publicá-los em periódicos científicos renomados e aceitos no meio acadêmico. Dessa forma, as universidades além de produzirem a informação científica, ainda ficam obrigadas a pagar para que esse conhecimento seja disponibilizado por editora ou grupos de editores, os quais acabavam sendo os detentores dos direitos de publicações.

Alves (2008, p. 128) ressalta que

[...] a produção de tais informações na sua maioria financiada pelo governo ou instituições não-governamentais sem fins lucrativos, fomenta o desenvolvimento da pesquisa e a geração do conhecimento, e é publicada em periódicos especializados, geralmente com alto custo para instituições e pesquisadores.

De uma forma contraditória, as universidades se viam ainda obrigadas a comprar o próprio conhecimento produzido por ela para poder disponibilizá-los aos seus usuários, através de assinaturas de periódicos ou bases de dados.

Conforme Alves (2008, p. 128) destaca que, buscando superar essa crise, pesquisadores de diversas partes do mundo se reuniram e deram início a um grandioso movimento em direção ao acesso aberto à informação científica em nível global. O movimento foi denominado de Movimento de Acesso Aberto à Informação Científica, *Open Access* (OA), Acesso Aberto. Esse movimento preconiza que a informação científica seja de livre acesso, disponibilizada na internet sem nenhuma restrição e, principalmente, se essa informação for oriunda de financiamento público. Esses princípios norteadores foram reafirmados e reforçados nos manifestos e declarações de Bethesda (2002), Budapeste (2003) e Berlim (2003).

Descrevem Andrade e Muriel-Torrado (2017) sobre as Declarações que discutiam as questões do acesso livre das publicações científicas:

À Declaração de Budapeste como a disponibilização gratuita das obras criadas na internet, possibilitando que as pessoas leiam, copiem, imprimam, ou que utilizem a obra com qualquer propósito. A Declaração de Budapeste cita uma única barreira ao uso irrestrito da informação: o *copyright*. [...] A Declaração de Bethesda, é desenvolvida por uma diversa gama de *expertises* como pesquisadores, cientistas ou editores entre outros, com o intuito de assegurar o acesso livre a literatura científica, especificamente a informação na área biomédica. [...] Já a Declaração de Berlim, escrita por pesquisadores de universidades europeias, trata sobre o acesso livre ao conhecimento com foco nas ciências e humanidades. Tal conhecimento é caracterizado como universal e parte do alicerce científico (Andrade; Muriel-Torrado, 2017, p. 2-3).

Outro fator determinante foi o advento da internet nos anos de 1990 com o formato *World Wide Web* (WWW), o qual se apresenta como um campo promissor para operacionalização do acesso aberto das publicações científicas.

Conforme Santo Junior (2010, p. 50), Stevan Harnad, um dos principais pesquisadores responsáveis pelo movimento de livre acesso, junto com outros estudiosos, desenvolveram duas estratégias de ação para tornar um documento livre de barreiras de acesso, a qual ele denominou de vias: a via dourada (*goldenroad*) e a via verde (*greenroad*).

Para Santo Junior (2010, p. 50), “na via verde, a iniciativa de tornar-se um documento de livre acesso parte do próprio pesquisador, o autor do documento. A via verde consiste no depósito realizado pelo próprio autor, por sua livre iniciativa, num repositório de acesso livre, seja temático ou institucional”.

Alves (2008, p. 133) destacou que a via verde se refere ao arquivamento que poderá ser realizado pelos próprios autores de artigos científicos publicados ou aceitos para publicação, obtendo autorização dos editores que os aceitaram para que possam disponibilizar em um servidor de arquivo aberto.

Ainda conforme Santos Junior (2010, p. 50), “na via dourada, a iniciativa de tornar os documentos de livre acesso parte do próprio editor, através da construção de periódicos científicos de acesso livre baseados na iniciativa dos arquivos abertos”, isto é, a via dourada se refere ao arquivamento dos periódicos científicos eletrônicos, cujo acesso aberto a seus conteúdos é garantido pelos próprios editores, sendo esse direito assegurado no próprio periódico. Nesse sentido, a via dourada se refere à produção e ampla disseminação de periódicos eletrônicos de acesso aberto em rede, enquanto a via verde se refere ao autoarquivamento.

Kuramoto (2009, p. 7) enfatiza a importância de se publicar em periódicos de acesso aberto:

Ao publicarem em periódicos de acesso aberto, os pesquisadores potencializam a comunicação científica, já que esta via possibilita a ampliação do diálogo entre os pares. As barreiras econômicas enfrentadas pelos centros de pesquisas e unidades de informação desaparecem, e é possível estabelecer um fluxo direto de comunicação de novidades que podem vir a representar importantes avanços científicos (Kuramoto, 2009, p. 7).

A instituição da política de acesso aberto contribui, sobremaneira, para o acesso e a disseminação da comunicação científica, colabora para a visibilidade das instituições provedoras e mostra a sua retribuição para a ciência e toda a humanidade.

Para Costa (2006), a filosofia do acesso aberto se apoia em três vertentes:

- Uso de *software* livre, para desenvolvimento de aplicações em computador;
- Arquivos abertos, para interoperabilidade em nível global;
- Acesso aberto, para a disseminação ampla e irrestrita de resultados da pesquisa científica.

É nesse panorama que surgem os RIs, com a finalidade de agilizar tanto o processo de comunicação científica quanto a gestão do conhecimento científico.

O acesso livre à informação científica, nos moldes da filosofia dos arquivos abertos, poderá ser realizado de três formas: através dos periódicos eletrônicos, das BDTDs e dos RIs.

Os RIs utilizam o *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH), um protocolo que efetua a coleta de dados da iniciativa de arquivos abertos, a qual transfere metadados de um sistema para outro, independentemente de *hardwares*, *softwares*, oferecendo, assim, um ambiente de interoperabilidade. De acordo com Freitas (2019, p. 14), a 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”.

Quanto à sua tipologia, os RIs se diferem dependendo de sua aplicação. Segundo Leite (2009, p. 19), no contexto de acesso aberto, repositórios digitais são denominados conforme os vários tipos de aplicações, podendo ser de três tipos:

- **Repositórios temáticos ou disciplinares:** voltados a uma comunidade científica específica;
- **Repositórios de teses e dissertações:** voltados exclusivamente para esses materiais;
- **Repositórios institucionais:** voltados à produção intelectual de uma instituição, especificamente, universidades e institutos de pesquisa.

No caso dos repositórios temáticos, estão presentes nos institutos de pesquisa, os quais abordam temas específicos de assunto; os repositórios de teses e dissertações, tratam somente com esses tipos de recursos informacionais, como é o caso da BDTD; e os RIs, por sua natureza mais complexa, englobam além das teses e dissertações, todas as publicações acadêmicas, bem como todos os tipos de informações produzidas no âmbito da instituição à qual estão vinculados.

Marcondes e Sayão (2009), em suas concepções de RIs, definem-nos em:

[...] Base de dados na Web na qual uma instituição de pesquisa deposita sistematicamente sua produção acadêmica e a disponibiliza de forma ampla para as comunidades interessadas. Sobre essa base de dados é oferecido um conjunto de serviços voltados para a gestão e para a disseminação de informações em formato digital. Esses serviços incluem captura, armazenamento, tratamento técnico, organização, preservação e entrega de conteúdos digitais de toda a natureza – texto, imagens, vídeo, áudio, apresentações, programas de computador, datasets etc. (Marcondes; Sayão, 2009, p. 23).

Dentre as suas finalidades, os RIs são utilizados para melhorar a comunicação científica interna e externa à instituição, maximizar a acessibilidade, o uso, a visibilidade, contribuir no impacto da produção científica, possibilitar a retroalimentação das atividades de pesquisa, contribuir para preservação dos conteúdos digitais científicos ou acadêmicos e, principalmente, contribuir para o aumento do prestígio da instituição e do pesquisador.

Xavier e Vargas (2019, p. 2) destacam que, nos RIs, o “surgimento destas coleções está diretamente atrelado ao desenvolvimento de tecnologias de informação e comunicação e seu impacto no processo de comunicação científica”. Os RIs são ferramentas adequadas para a gestão do conhecimento científico, potencializam a condução de processos que maximizam a criação, o comportamento, a disseminação e o uso do conhecimento científico, bem como, agilizam os processos de comunicação científica.

Conforme Dodebei (2009, p. 91), o RI “é uma base de dados digital e virtual (*web-based database*), de caráter coletivo e cumulativo (memória da instituição), de acesso aberto e interoperável que coleta, armazena, dissemina e preserva digitalmente a produção intelectual da instituição”. Marques (2020, p. 143) também ressalta que a visibilidade potencializada pelo acesso à informação, possibilita “a criação de indicadores da qualidade da produção científica nas instituições”. Nesse sentido, os RIs se constituem como uma verdadeira vitrine, expondo toda a produção acadêmica e científica da instituição, dispondo em um único espaço e disponibilizando de forma direta em acesso aberto na web.

Miranda e Delfino (2016), definem os RIs em:

Ferramentas de acesso aberto que têm o objetivo de reunir, preservar, divulgar e garantir o acesso confiável e permanente aos documentos acadêmicos, científicos, artísticos e administrativos gerados na instituição, bem como às suas coleções históricas, e outros documentos relevantes para a organização (Miranda; Delfino, 2016, p. 53).

Portanto, um valioso instrumento para projetar a instituição em âmbito local e global, revelando todo o seu potencial através das suas produções científicas e acadêmicas, capaz de aproximar a sociedade através das inovações oriundas dessa instituição.

Costa e Leite (2017, p. 103) definem como características essenciais para os RIs: “[...] ser institucionalmente definido, tratar da produção acadêmica, ser coleções digitais de textos completos, ser interoperáveis e ser uma iniciativa de acesso aberto”.

Conforme enfatiza Kuramoto (2009, p. 107), “[...] os RIs para funcionar e atender às expectativas do movimento do acesso livre dependem de algum instrumento que garanta o depósito”. Dentre esses instrumentos, também destacado por Kuramoto (2009), está a política de informação institucional. A política informacional é extremamente importante, pois é ela que vai direcionar o fluxo informacional da instituição, definindo quais conteúdos informacionais serão armazenados, preservados e disseminados no repositório, além de ser um instrumento legal.

Apesar de toda potencialidade oferecida pelos RIs, no que concerne à disseminação e ao acesso aos conteúdos informacionais acadêmicos e científicos, ainda é pouco explorado pela comunidade acadêmica. Em sua maioria, seus conteúdos informacionais são constituídos de TCC, dissertações e teses. Isto ocorre devido à resistência por partes dos pesquisadores em depositarem seus artigos, pelo desconhecimento sobre o movimento de acesso aberto e, principalmente, pelo desconhecimento da existência do repositório e sua finalidade.

Para Veiga e Macena (2015, p. 36), “o entendimento dos objetivos do Acesso Aberto e da estratégia da via verde em mobilizar os pesquisadores retomando o comando no processo de comunicação científica é requisito fundamental para dirigentes de instituições que possuem repositórios e gestores de RI”. Assim, ressalta-se que os dirigentes e administradores de RIs devem promover um amplo debate sobre os RIs e sua importância para a visibilidade da instituição, seus pesquisadores e suas produções.

Nesse sentido, deve-se fazer ampla divulgação do repositório, sua finalidade, contribuições e importância para toda a comunidade acadêmica, explorando todos os canais de comunicação, tais como: *sites* da universidade, biblioteca, unidades, institutos, faculdades, centros ou departamentos dos cursos, usar as redes sociais, elaborar tutoriais para treinamento no uso do repositório. É fundamental também o engajamento dos gestores das universidades para dar o aporte necessário para o funcionamento do repositório, reconhecendo a importância para a divulgação da produção científica da instituição e o reconhecimento pela sociedade, elevando cada vez mais o nível das universidades.

3.2 Repositórios Institucionais no Brasil

No Brasil, quem encabeçou o movimento de acesso aberto foi o IBICT. O Instituto foi responsável na elaboração de encontros e congressos para debater o tema do acesso livre, bem como elaboração de documentos, além de participar na construção da Carta de São Paulo (2005) e da Declaração de Florianópolis (2006).

O IBICT, em 2005, lançou o Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica, o qual constituiu como um dos principais documentos institucionais em apoio ao acesso livre. O instituto também foi responsável pelo apoio às Instituições de Ensino Superior (IES) na implantação e infraestrutura, através do seu Laboratório de Tecnologia da Informação (LTI), disponibilizando o acesso a um *software* livre para elaboração da BDTD, a qual tinha como função principal, centralizar a recuperação da informação das teses e dissertações das IES depositadas, contribuindo com o controle bibliográfico da produção científica nacional.

A partir de 2009, o IBICT, implementando a política de desenvolvimento de apoio à criação e manutenção de repositório, intermediou a aquisição de um *software* livre com interoperabilidade e compatibilidade de transferência de dados entre os repositórios, o *DSpace Institutional Digital Repository Systems* (DSpace).

Conforme Weitzel (2019):

Editais da Finep em convênio com o IBICT foram lançados em 2009 e tiveram grande relevância para o financiamento e fortalecimento tanto do Acesso Aberto Dourado quanto Verde no país. As instituições contempladas pelo edital receberam “kits tecnológicos” e apoio técnico-operacional com o objetivo de superar as dificuldades em relação à capacitação de recursos humanos e de infraestrutura que afetavam a implementação do sistema de publicação digital e de repositórios (Weitzel, 2019, p. 109).

Surge, a partir daí, a implantação dos RIs brasileiros, focando na inclusão da produção científica nacional, nos parâmetros da ciência aberta, tornando-a acessível a todos em qualquer lugar, livres de barreiras e imposição do mercado editorial.

Weitzel (2019, p. 120) argumenta que “[...] a etapa inicial de implantação de repositórios liderada pelo governo brasileiro – incluindo as instituições de ensino superior e de pesquisa – já foi concluído”. Necessita agora, de novas ações mobilizando também os envolvidos para que seja alcançado um outro patamar de desenvolvimento dos repositórios.

Apesar do IBICT ter lançado a BDTD com a finalidade de disponibilizar as teses e dissertações produzidas pelas instituições de ensino superior em formato digital, o repositório

surgiu como uma alternativa de ampliar os produtos e serviços oferecidos, incorporando além de teses e dissertações, vídeos, dados de pesquisa, livros, artigos de periódicos, entre outros. Diante disso, algumas instituições englobaram as suas BDTDs ao repositório. É o caso da Universidade Federal de Alagoas (UFAL), enquanto uma pequena parcela optou em ter as duas bases de dados separadamente, por estarem com as suas BDTDs mais consolidadas.

Conforme o IBCT (2023), o Brasil conta com cerca de 118 repositórios institucionais, 67 bibliotecas digitais de teses e dissertações e 4 bibliotecas digitais de monografias, sendo, em grande parte, oriundos de órgãos ou instituições públicas. Desses, apenas 49 fazem parte da rede de provedores de dados, serviços e agregadores como o Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr), o qual permite acesso gratuito à produção científica de autores vinculados às universidades e aos institutos de pesquisa brasileiros e de fontes de informação portuguesas.

Na região sudeste é onde se concentra a maior quantidade de repositórios, principalmente no estado de São Paulo. A região nordeste vem, ao longo do tempo, apresentando uma boa adesão à implantação de repositórios, conforme Cassé (2022, p. 176) destacou, que das 20 universidades federais da região, 14 possuem repositórios.

Conforme relata Xavier (2019), os repositórios são predominantemente institucionais com foco multidisciplinar, apresentando forte papel como biblioteca de teses e dissertações, porém têm estabelecido guarda de artigos científicos e são, majoritariamente, em língua portuguesa.

Os RIs no Brasil, em grande parte, têm nos trabalhos acadêmicos (teses, dissertações e TCCs) a sua maior massa documental, enquanto as outras tipologias documentais (artigos, dados de pesquisas, relatórios técnicos) aparecem em baixa quantidade. Ao constituir, assim, um acervo substancial de teses e dissertações, Xia *et al.* (2012) e Leite (2009) ressaltam que os primeiros documentos a serem depositados nos repositórios são as teses e dissertações defendidas na instituição. Isso decorre do fato de que as teses e dissertações, em sua maioria, não requerem restrição para publicação em acesso livre, salvo os casos em que são passíveis de gerar uma patente, impossibilitando assim, a não publicização imediata. Além disso, é normal a obrigatoriedade do depósito desses materiais nas bibliotecas universitárias ou repositórios, como parte do requisito para obtenção do certificado de conclusão do curso.

A partir de 2017, os TCCs da graduação começaram também a povoar os repositórios, visto que essa presença passou a pontuar no sistema nacional de avaliação do ensino superior (Sinaes), do MEC. A inclusão dos trabalhos na base do repositório em meio digital acessível em rede via internet aumenta o escore do curso a ser avaliado.

O Trabalho de Conclusão de Curso está institucionalizado e considera carga horária, formas de apresentação, orientação e coordenação, a divulgação de manuais atualizados de apoio à produção de trabalhos e a disponibilização dos TCC em repositórios institucionais próprios, acessíveis pela internet. Sinaes, 2017, p. 14).

Diante desses argumentos, fica evidente que a maior parte de documentos disponíveis nos RIs brasileiros é composta por teses, dissertações e TCCs. A razão das outras tipologias documentais aparece em pequena escala, decorre ainda do desconhecimento da importância do repositório para os pesquisadores no processo de dar visibilidade às suas produções científicas e, principalmente, pela falta de política informacional na instituição para intensificar a contribuição dos pesquisadores.

Outro ponto a se destacar na atuação dos RIs, refere-se às atividades técnicas e operacionais. Sousa (2019, p. 514) identificou que “[...] a maioria dos gestores tem dificuldades em atender às demandas de suas instituições relacionadas ao repositório institucional [...]”, em parte se deve principalmente à quantidade reduzida de profissionais na equipe e, por outra, na questão tecnológica dos equipamentos e sua manutenção.

Buscando dirimir problemas comuns nos RIs e compartilhar boas práticas, foram estabelecidas as redes de colaboração de repositórios. Sob a coordenação do IBICT, a Rede Brasileira de Repositórios Institucionais de Publicações Científicas em Acesso Aberto (RIAA), a qual é integrada por suas associadas, as redes de cooperação regionais de RIs: Rede Sul de Repositórios Institucionais, Rede Centro-Oeste de Repositórios Institucionais, Rede Sudeste de Repositórios Institucionais, Rede Nordeste de Repositórios Institucionais e Rede Norte de Repositórios Institucionais.

Segundo Marques (2020, p. 157), “[...] atuar como membro de uma das redes de repositórios regionais para troca de experiência e colaborar com os estudos e avanços dos repositórios ajuda bastante na manutenção do sistema”. Essas redes são fundamentais para suprir as lacunas dos RIs, que são muitas, servindo de ponto de apoio em que os gestores compartilham ideias e tiram dúvidas relativas às atividades e atribuições.

3.3 Autoarquivamento nos repositórios

Com a “via verde”, o Movimento de Acesso Aberto estabelece a estratégia do autoarquivamento, o qual tem como objetivo incentivar os autores a disponibilizem diretamente as suas produções científicas em repositórios de acesso aberto.

Para Assis (2013, p. 212),

O autoarquivamento é o mecanismo de depósito em repositórios institucionais recomendado pelo movimento de acesso aberto à informação científica, seguindo as estratégias da Via Verde. O autoarquivamento é aquele realizado pelo próprio autor. A adoção do autoarquivamento por parte das instituições tem como objetivo garantir os depósitos nos repositórios institucionais, maximizando assim, a sua quantidade de documentos (Assis, 2013, p. 212).

O movimento de acesso aberto à informação científica preconiza que os resultados das pesquisas financiadas com dinheiro público devem ser de acesso aberto, justificando o uso desses recursos, e que essas pesquisas estejam disponíveis gratuitamente na web. Conforme enfatizam Freitas, Dal'Evedove e Tartarotti (2021, p. 169):

O avanço das tecnologias digitais culminou em uma rápida e crescente disponibilização de conteúdos na internet. Esta evolução tecnológica alterou o processo de comunicação científica, com destaque para o movimento mundial em favor do acesso livre e gratuito ao conhecimento gerado no ambiente das universidades e institutos de pesquisa, promovendo benefícios significativos ao desenvolvimento da ciência a partir do acesso livre de barreiras aos resultados de pesquisas conduzidas mediante a concessão de recursos públicos (Freitas; Dal'Evedove; Tartarotti, 2021, p. 169).

Esse entendimento parte do princípio de que tudo que é financiado com recurso público deverá voltar para a sociedade de forma gratuita, contribuindo para o desenvolvimento de todos. Apesar dessa recomendação do acesso aberto, através da via verde, a prática do autodepósito ou autoarquivamento, ainda não foi consolidada. Segundo Assis (2013, p. 214), “apenas cerca de 15-20% dos 2,5 milhões de artigos publicados anualmente em todo mundo estão sendo autoarquivados pelos seus autores”, o que implica no monopólio das editoras, conforme argumenta Assis (2013, p. 213):

As editoras científicas formam um grande monopólio dos resultados das pesquisas científicas, detendo os direitos de acesso a esses documentos. Os pesquisadores, em sua grande maioria, para terem suas pesquisas validadas, ao finalizá-la repassam os direitos da obra aos editores científicos para que estes publiquem o trabalho realizado (Assis, 2013, p. 213).

Isso demonstra ainda o poder de interferência do mercado editorial científico, das grandes editoras, com suas bases de dados e plataformas científicas, as quais, detém os direitos autorais de boa parte das publicações científicas, não as disponibilizando em acesso livre nos repositórios. Entretanto, o autoarquivamento nos RIs se apresenta como a alternativa viável para os pesquisadores confrontarem a exclusividade imposta pelas grandes editoras e plataformas científicas para dar visibilidade às suas produções.

Segundo Santos e Oliveira (2019), o autoarquivamento permite ao pesquisador a autonomia sobre o seu conteúdo científico. Daí a grande importância de incentivar os

pesquisadores à prática do autoarquivamento de seus trabalhos nos repositórios, pois possibilitará o acesso direto aos trabalhos científicos, sem as exigências impostas pelo mercado editorial.

Outras questões que interferem no autoarquivamento, segundo autores como Assis, (2013), Gargouri *et al.* (2010), Keefer (2007), Hajjem, Harnad e Gingras (2005) e Swan (2005), são: preocupações com a qualidade dos artigos, tendo em vista que não seriam avaliados por pares, possibilitando baixar a qualidade do sistema de informação de acesso aberto; o desconhecimento sobre as questões de direitos autorais por parte dos pesquisadores; desconhecimento do próprio movimento de acesso aberto; além da falta de tempo para alimentar o sistema de informação; há outros fatores citados também, como a resistência à obrigação de depositar seus trabalhos; medo de perder o controle de suas obras; e, além de possível plágio.

Conforme argumenta Assis (2013, p. 217), “apesar de não ser ainda uma prática comum nos RIs de todo o mundo, o arquivamento não demanda esforço de trabalho excessivo para o autor”. O processo é bastante prático, cabendo ao autor preencher os metadados solicitados na base de dados do repositório, haja vista que os RIs, incorporam em suas políticas informacionais, a obrigatoriedade do depósito em suas bases de dados, as publicações científicas acadêmicas produzidas nos âmbitos dessas instituições, aparecendo as teses e dissertações como os tipos mais comuns de documentos depositados.

Conforme enfatiza Veiga e Macena (2015, p. 45):

A diversidade de tipologia no RI, como dados de pesquisa, Recursos Educacionais, teses e dissertações, e outros, complementa e maximiza o acesso aberto ao Conhecimento, mas sem perder de vista o coração do movimento de acesso aberto que se iniciou na luta contra a restrição do acesso imposto pelo alto custo das assinaturas de periódicos (Veiga; Macena, 2015, p. 45).

Apesar dos RIs terem sido idealizados com foco principalmente na publicação de artigos científicos, esse tipo de documento ainda aparece em quantidade reduzida nos seus acervos. É evidente que a disputa para o reconhecimento no campo científico cobra ainda aos pesquisadores priorizarem a publicação de seus artigos científicos nas grandes editoras e bases de dados fechadas, fazendo com que seus trabalhos apareçam em periódicos renomados reconhecidos internacionalmente, propiciando, com isso, o reconhecimento entre os seus pares.

Deve-se ressaltar, também, que os repositórios são ainda muito recentes, fato que justifica os pesquisadores ainda não estarem familiarizados com suas práticas e políticas informacionais de acesso aberto. Nesse sentido, é muito importante que seja difundido entre os membros da

instituição a importância dos RIs no contexto do acesso aberto, além de seu papel para o acesso e a disseminação do conhecimento gerado no âmbito dessa instituição.

O autoarquivamento se torna relevante também por agilizar o processo de visibilidade, devido à inclusão imediata no repositório, não sujeita à morosidade que, às vezes, decorre do acúmulo de trabalhos executados pela equipe do repositório, geralmente, bastante reduzida. A prática do autoarquivamento também aproxima a interação entre o autor, a instituição e um universo de usuários que utiliza o ambiente web para produção e compartilhamento de pesquisas científicas.

Freitas, Dal'Evedove e Tartarotti (2021) enfatizam que:

Para além do movimento de acesso à informação científica, o autoarquivamento é fruto da nova configuração do ambiente Web, apresentando-se como 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 (Freitas; Dal'Evedove; Tartarotti, 2021, p. 170).

Entretanto, para as autoras, “apesar das diversas possibilidades oferecidas atualmente para o acesso à informação científica, a agilidade no processo nem sempre é sinônimo de qualidade e relevância das informações recuperadas” (Freitas; Dal'Evedove; Tartarotti, 2021, p. 173). De fato, a possibilidade do próprio autor autoarquivar a sua produção, dá a liberdade de escolher os descritores que nem sempre estão abalizados por uma fonte de referência autorizada, como catálogos de assuntos ou tesouros, além das inconsistências de uso de plural ou singular, termos muito genéricos e até mesmo erro de grafia. Fatos esses que interferem na busca, possibilitando a uma não exatidão no momento da pesquisa e interferindo diretamente na qualidade e confiabilidade da base. Diante disto, cabe aos repositórios buscarem meios para dirimir esses acontecimentos, garantindo a qualidade e confiabilidade do sistema.

4 REPRESENTAÇÃO DA INFORMAÇÃO, FATOR DETERMINANTE PARA O ACESSO

Nesta seção, será abordada a representação da informação como elemento primordial a fim de recuperar a informação, principalmente nos ambientes digitais e a necessidade de padronização do metadado “assunto”, através de índices e tesouros de assunto para uma boa indexação.

4.1 Representação da informação no ambiente digital

A representação da informação é fator fundamental para o processo de organização, acesso e recuperação dessa informação, principalmente nos ambientes digitais, como os RIs. Através da *web*, a ampliação do acesso e uso às tecnologias vem tornando o ambiente digital o principal meio de busca, acesso, disseminação e produção de conteúdos informacionais.

Arakaki e Castro (2020) argumentam que a intensa produção e a disseminação de dados através do meio digital proporcionaram inúmeros benefícios para a sociedade no que diz respeito ao acesso e à democratização da informação, na medida em que se ampliam as formas dessa disponibilização e disseminação dos recursos informacionais.

Para Brascher e Café (2008, p. 5), a representação da informação pode ser “[...] entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico”. É através dela que se pode individualizar um item, caracterizá-lo, descrevê-lo, tornando-o acessível e útil aos usuários.

Conforme Silva e Santarém Segundo (2019), os metadados, os vocabulários controlados e os tesouros são componentes de representação da informação, os quais possibilitam a interação nos sistemas da arquitetura da informação, proporcionando ao usuário uma boa compreensão da informação disponibilizada no ambiente digital. Para os autores, esses elementos da arquitetura da informação digital propiciam alternativa para tornar a informação mais compreensível aos usuários.

Para Tartarotti (2019, p. 70), “um repositório institucional como um sistema de recuperação da informação documental precisa representar o conteúdo intelectual dos recursos informacionais para prover acesso aos mesmos baseando-se no seu assunto”.

A boa representação da informação em ambiente digital poderá proporcionar ao usuário uma experiência mais satisfatória, pois será capaz de suprir as necessidades informacionais em tempo hábil, disponibilizando o(s) item(ns) documental(is) pertinente(s), entregando ao usuário

a informação relevante com maior clareza, rapidez e com satisfação ao interagir com o sistema de informação no ambiente digital.

No contexto da CI, Albuquerque, Gaudêncio e Santos (2019) enfatizam que a representação da informação se constitui como uma atividade pragmática, ampliada como campo de estudos teóricos e aplicados, objetivando responder os problemas e efeitos da informação, a fim de tornar os registros informacionais acessíveis à sociedade.

Para que tudo isso aconteça é necessário integração e interação entre o usuário e o sistema de informação. Para Martins e Córtez (2019), o usuário será sempre o foco principal.

É fundamental considerar o usuário como protagonista nos sistemas de informação, levando em consideração os marcadores sociais nos quais está inserido, com respeito à alteridade para que suas necessidades informacionais sejam satisfeitas e os conteúdos de seu interesse estejam representados adequadamente (Martins; Córtez, 2019, p. 178).

Nesse sentido, a representação deve levar em conta todos os contextos que fazem parte do cotidiano do usuário, os aspectos sociais, culturais, econômicos, incluindo também os aspectos cognitivos.

Para Arakaki e Castro (2020, p. 9), a “representação da informação é uma atividade fundamental para o tratamento informacional, pois é a partir dela que se dá condições para identificar e recuperar um recurso informacional”. De fato, num sistema informacional como RI, em que os usuários têm características múltiplas e interesses informacionais variados, definir a forma de representação da informação é de suma importância para o sucesso do encontro da informação pertinente.

Nos ambientes tradicionais de informação, como biblioteca, arquivo, centro de documentação e museus, a representação da informação é um processo que pode utilizar linguagem específica oriundas de tesouros, índices e catálogos específicos. Maimone, Silveira e Tálamo (2011), utilizam-se de “linguagens” que lhes são próprias, através desses instrumentos, padronizando e criando assim bases para recuperação e acesso à informação. Para as autoras, funcionam como princípio organizativo do conhecimento, controle terminológico, que visa ao tratamento e à recuperação de informações, remetendo o usuário ao material de interesse e, ainda, oferecendo conhecimento de novos termos.

Entretanto, nos ambientes digitais, o uso desses instrumentos ainda é muito incipiente, a liberdade de interação dos usuários com o sistema de informação leva à utilização na maioria das vezes do uso da linguagem natural, sem preocupação com a padronização dos termos registrados através do autoarquivamento.

Conforme argumenta Lancaster (2004), os problemas principais da representação da informação no Sistema de Recuperação de Informação (SRI) são definir o conteúdo informacional dos documentos e determinar sua relevância nas buscas.

No tocante ao sistema padronizado em RIs, Villalobos e Gomes (2015) consideram fundamental o processamento técnico do recurso informacional para o processo de comunicação, permitindo buscar e recuperar a informação de modo eficiente e eficaz. Entretanto, enfatizam que apesar da tendência para disponibilização de conteúdos digitais, existe ainda carência de ações, a fim de padronizar os metadados na representação dos documentos armazenados em ambientes digitais.

Nesse contexto, fica evidente que os ambientes digitais que adotam os componentes de representação padronizados, possibilitam aos usuários uma busca e navegação mais confiável e oferecem ao público a informação de forma compreensiva.

4.2 A indexação dos cabeçalhos de assunto nos repositórios

O acesso aos conteúdos informacionais é uma das preocupações constantes da CI, dispor de mecanismos eficientes para a localização e recuperação dos registros documentais é fator preponderante para o sucesso de uma pesquisa. Dentre esses mecanismos, encontra-se a indexação, ponto determinante para o sucesso ou insucesso em uma busca.

Freitas (2019) define a indexação como um processo fundamental para a identificação dos conteúdos de um documento. Para ela, a indexação caracteriza-se em:

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 às necessidades dos usuários na busca por informação (Freitas, 2019, p. 26).

Para Santos (2017, p. 2), “a indexação consiste na descrição do conteúdo de um documento, de forma concisa e condensada, por meios de emprego de termos – também denominados como palavras-chave ou descritores”. A indexação tem como finalidade caracterizar de forma concisa os conteúdos informacionais de um item, sendo ainda uma ponte de acesso entre o usuário e o item.

Daí a importância na escolha dos termos ou descritores que identificam um determinado item, pois eles possibilitam a sua mensuração quanto à pertinência e relevância para uma

determinada pesquisa, principalmente nos ambientes digitais, em que se encontra um grande volume de informação.

Lancaster (2004, p. 9) considera a indexação em duas etapas: análise conceitual e tradução. A análise conceitual se refere à extração de conceitos que representem o conteúdo de um documento, “implica decidir do que trata um documento – isto é, qual o seu assunto”; enquanto a tradução consiste na representação dos documentos em termos oriundos de instrumentos de indexação, “envolve a conversão da análise conceitual de um documento num determinado conjunto de termos de indexação”.

Os instrumentos de indexação auxiliam os indexadores a escolherem termos padronizados pertencentes a um campo específico de conhecimento, numa linguagem controlada comum entre os pares. Segundo Dias e Naves (2007, p. 6), esses instrumentos de indexação são chamados de linguagens de indexação ou linguagens documentárias e “depois de traduzidos, para uma linguagem de indexação, passam a ser chamados de descritores de assunto, rótulos, cabeçalhos de assunto, termos de indexação ou enunciados”.

No que concerne à tradução, Lancaster (2004) propõe a tipificação em indexação por extração, ou indexação derivada, e indexação por atribuição. No primeiro caso, as palavras ou expressões empregadas pelo autor no item documental são selecionadas pelo indexador para representar seu conteúdo. Ao utilizar-se desse recurso, o indexador faz uso do que é chamado de uma linguagem natural, ele seleciona palavras encontradas no título, nas palavras-chave ou no próprio conteúdo informacional do documento. No segundo caso, indexação por atribuição, como o próprio nome diz, são atribuídos descritores ao documento se utilizando de uma linguagem oriunda de fontes externas, não necessariamente o próprio documento, sendo assim, o indexador se utiliza de vocabulários controlados, tesouros, entre outros, fazendo uso de uma linguagem controlada e/ou artificial, ou seja, buscando uma padronização.

Para Santos (2017), em relação à importância do controle dos cabeçalhos:

São determinados os cabeçalhos a serem empregados objetivando-se controlar sinônimos, diferenciar homógrafos e ligar termos por meios de relações hierárquica, associativas e de equivalência (ou seja, utiliza-se uma linguagem artificial/linguagem controlada (Santos, 2017, p. 7).

Nesse sentido, a linguagem controlada se constitui em linguagem concebida de acordo com as necessidades dos usuários, enquanto a linguagem natural se configura como discurso comum ou coloquial, uma linguagem empregada naturalmente na escrita e na fala.

No espaço da universidade, apesar de predominar uma linguagem padronizada por vocabulários específicos de cada área do conhecimento que constituem os cursos, o seu uso cotidiano acaba por torná-la uma linguagem natural entre seus pares.

De acordo com Freitas (2019), em relação à comunicação acadêmica e científica:

[...] 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 (Freitas, 2019, p. 31).

Isso quer dizer que, tanto a linguagem natural como a especializada são passíveis de utilização por qualquer indivíduo atuante num ambiente informacional.

No âmbito dos RIs, em que se busca uma maior interoperabilidade entre o produtor informacional e o sistema, o uso da indexação por extração, ou seja, a linguagem natural, é predominante. Nesse ambiente, os usuários participam da produção, organização, representação e disseminação da informação, não necessariamente seguindo padrão de escolha dos termos indexados.

Conforme Santos (2017, p. 3), “é evidenciada uma nova forma de organização e representação das informações no âmbito dos Repositórios Digitais em que não se adotam regras de indexação envolvendo controle de vocabulário na descrição dos recursos”.

O uso da indexação por extração possibilita maior aproximação do pensamento do autor e o conteúdo por ele produzido, além de agilizar e facilitar o processo de indexação e, conseqüentemente, a rapidez para tornar o documento disponível no sistema. Entretanto, essa prática por si só não garante a qualidade e confiabilidade do RI.

Lopes (2002), Santos (2017), Freitas (2019) e Tartarotti (2019) abordam em suas pesquisas estudos relacionados à importância da indexação e às questões que envolvem mais escolhas dos cabeçalhos de assunto e suas representações temáticas nos conteúdos documentais.

Conforme Lopes (2002, p. 48), há vantagens e desvantagens na utilização da indexação por extração. Para a autora, as vantagens se caracterizam por “permitir o imediato registro da informação em um repositório digital”, bem como a “não necessidade de treinamento específico no uso da linguagem de controle”; ressaltando ainda que “os termos são extraídos diretamente dos documentos”, além disso, constitui-se em uma única forma de comunicação entre os indexadores e os usuários acessando os mesmos termos. Dentre as desvantagens, a autora elenca: maior esforço intelectual dos usuários no processo de busca na escolha dos termos;

possibilidade de alta incidência de respostas negativas ou de relações incorretas; aumento dos custos de acesso com a inserção de termos aleatórios; necessidade de uma estratégia de busca que arrole todos os principais conceitos e, principalmente, a perda da confiança dos usuários em uma possível resposta negativa.

Santos (2017, p. 8) ressalta que “as informações disponíveis nos RI são descritas com palavras provenientes da linguagem natural sem o auxílio de um instrumento de representação da informação”, ou seja, não seguem uma padronização determinada.

Os usuários representam conteúdos com termos que usam e/ou acreditam que entendem, sendo assim alguns termos podem fazer sentido apenas a quem os elencou e nem sempre para a comunidade ou grupo de pessoas que utilizam o sistema. Com a liberdade da atribuição dos descritores um mesmo termo pode ter diferentes significados para aqueles que atribuíram as etiquetas/descriptores para fins de representação do conteúdo (Santos, 2017, p. 8).

Fica evidente que o uso da linguagem natural, geralmente, apresenta a visão “subjetiva” dos usuários produtores dos conteúdos informacionais, muitas vezes não se atentando para a linguagem do sistema e, com isso, descredibilizando o sistema visto que a utilização de termos não padronizados, ou não controlados em uma determinada área do conhecimento, poderá dificultar a recuperação entre os pares.

A falta de padronização nos cabeçalhos de assunto, tanto nos ambientes digitais como nos repositórios, pode ocasionar implicações negativas, pois a representação da informação, às vezes, apresenta distorções com variações no que concerne ao uso de palavras no singular ou plural, erros gramaticais, termos com mais de um significado, termos distintos com o mesmo significado, além da questão da profundidade ou especificidade da descrição, fato em que o documento não é tratado em sua totalidade.

Sobre essas questões, Noruzi (2007), Amstel (2007), Catarino e Baptista (2009), Santos (2013) e Freitas (2019) abordaram em seus estudos o uso da linguagem comum (linguagem natural) em ambientes de representação colaborativa da informação, enfatizando a falta de padronização.

Freitas (2019, p. 31) ressalta que o “uso de linguagem natural pode trazer problemas no que se refere à recuperação do material armazenado”, isto pode ocorrer devido à falta de padronização e à liberdade de escolha de termos de indexação, o que torna a pesquisa muito abrangente, cuja busca no sistema de informação necessitará de um esforço maior do usuário na escolha dos termos, prejudicando a busca e o acesso ao item pesquisado.

Conforme Tartarotti (2019):

A inadequada representação documental em repositórios institucionais refletirá na qualidade da recuperação de seus recursos informacionais, visto que a qualidade da indexação de assuntos está relacionada à capacidade de reconstruir o assunto tratado em um determinado recurso informacional em conceitos para posterior recuperação por assuntos pelos usuários (Tartarotti, 2019, p. 26).

Diante dessa afirmação, é fundamental que o usuário, ao escolher as palavras-chave ou os termos para indexação de sua produção científica, atente-se para a linguagem do sistema, aos vocabulários específicos de sua área de conhecimento, não só fazendo o uso indiscriminado da linguagem natural.

Tartarotti (2019, p. 26) ainda enfatiza que, “o uso adequado da linguagem documental é fundamental, pois possibilita a representação dos conteúdos dos recursos informacionais compatível com as solicitações de pesquisas dos usuários em uma determinada área científica especializada”. De fato, o uso de termos usuais em uma determinada área científica garante a maior facilidade na escolha dos termos da busca, bem como proporciona a agilidade no processo de recuperação da informação desejada, tornando o sistema de informação confiável e reconhecido entre os usuários que dele utilizam, ou seja, garantindo sua eficácia e eficiência.

Apesar de, no ambiente web, a linguagem natural ocorrer com maior intensidade, devido à facilidade e liberdade das escolhas dos termos, faz-se necessário a busca da padronização para melhor aproveitamento entre a linguagem dos usuários e a linguagem dos sistemas, e isso é possível através do uso dos vocabulários controlados.

Para Santos (2017, p. 10), os vocabulários controlados “são instrumentos que visam facilitar a comunicação mediante a padronização de termos para descrição de conteúdos”, eles possibilitam controlar as dispersões lexicais da linguagem natural, auxiliar o indexador na escolha dos termos, indicar os mais pertinentes para a representação do documento, além de atuar como instrumento de comunicação entre a linguagem do usuário e a linguagem do sistema.

Segundo Fujita e Tolare (2019, p. 93), “o objetivo do controle do vocabulário é a consistência na representação e na recuperação de modo a evitar a dispersão de recursos informacionais relacionados aos termos do vocabulário controlado”.

A prioridade em buscar a padronização de cabeçalhos de assunto em RIs se torna cada vez mais necessária, posto que, com a utilização do autoarquivamento, os usuários do sistema podem determinar seu próprio cabeçalho ou palavras-chave, a inclusão não tem limites nem obedece a critérios de escolha.

Diante desse dilema, Fujita e Tolare (2019, p. 98) argumentam sobre a necessidade e importância dos Sistemas de Organização do Conhecimento (SOCs) e destacam que “os SOCs são importantes porque realizam o controle de vocabulário em sistemas de informação que necessitam prover a recuperação da informação armazenada, seja de forma retrospectiva ou atualizada”.

Dentre os SOCs estão tesouros, ontologias, listas de cabeçalhos, diretórios, dicionários, glossários, arquivos de autoridades, anéis de sinônimos, entre outros.

De fato, prover meios para a padronização dos cabeçalhos de assunto nos RIs é de suma importância, no ambiente em que esses RIs estão inseridos, geralmente em universidades, a multiplicidade de informações, a interdisciplinaridade e o linguajar técnico-científico denotam o uso de termos técnicos e especializados para cada área do conhecimento.

5 PERCURSO METODOLÓGICO

Esta seção tem a função de apresentar as características de como foi delineado o percurso metodológico desta pesquisa, sua natureza, abordagem, objetivos e procedimentos, bem como o universo, a amostra, as técnicas utilizadas para coleta de dados e os métodos empregados na análise e na interpretação dos dados coletados.

Elemento essencial no processo investigativo, o processo metodológico é uma tarefa preponderante para que se possa atingir os objetivos da pesquisa, é através dele que serão apresentados, de forma descritiva, os métodos utilizados pelo pesquisador durante o estudo. Conforme Praça (2015, p. 73), “o caminho pelo qual se propõem a obter o conhecimento científico deve sempre ser direcionado por procedimentos técnicos e metodológicos bem definidos visando fornecer subsídios necessários na busca de um resultado [...]”.

Marconi e Lakatos (2018) consideram que o caminho na obtenção de resultados de uma pesquisa se dá pela aplicação dos métodos científicos e de como será realizada. Segundo Chauí (1994, p. 354), *methodos* significa uma investigação que segue um modo ou uma maneira planejada e determinada para conhecer alguma coisa; procedimento racional para o conhecimento seguindo um percurso fixado. Para Oliveira (1998, p. 17), o método assinala, portanto, um percurso escolhido entre outros possíveis, e enfatiza ainda que o método não representa tão somente um caminho qualquer entre outros, mas um caminho seguro, uma via de acesso que permita interpretar com maior coerência e correção possíveis as questões sociais propostas num determinado estudo.

Severino (2013, p. 88) argumenta ainda que: no entanto, não basta seguir um método e aplicar técnicas para se completar o entendimento do procedimento geral da ciência. Esse procedimento precisa ainda se referir a um fundamento epistemológico que sustenta e justifica a própria metodologia praticada.

5.1 Delineamento da pesquisa

Quanto ao alcance dos objetivos propostos, deu-se por meio de um estudo de caráter exploratório e descritivo na Base de Dados do RIUFAL, com a análise do metadado “assunto”, especificamente os assuntos cadastrados através do autoarquivamento pelos alunos concluintes dos cursos de Enfermagem e Serviço Social da UFAL.

Segundo Creswell (2010), a pesquisa exploratória traz uma aproximação com o seu objeto de estudo, baseando-se nas ideias dele com a utilização de técnicas mais adequadas para

desenvolver um maior entendimento do assunto. Severino (2013, p. 107) afirma que a pesquisa exploratória busca apenas levantar informações sobre um determinado objeto, delimitando assim um campo de trabalho, mapeando as condições de manifestação desse objeto.

No que concerne à pesquisa descritiva, Lakatos e Marconi (2017) descrevem como um tipo de estudo no qual o pesquisador tem o interesse em descrever a natureza e a forma de relação que possui seu objeto de pesquisa. Para Prodanov e Freitas (2013, p. 52), o pesquisador apenas registra e descreve os fatos observados sem interferir neles. Visa descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis.

No que se refere à modalidade, a pesquisa fez uma abordagem qualitativa e quantitativa a partir de informações coletadas na Base de Dados do RIUFAL, no tocante ao metadado “assunto”.

Para Laville e Dionne (1999, p. 224), na abordagem quantitativa, após ter reunido os elementos retirados dos conteúdos em categorias, o pesquisador constrói distribuições de frequência e outros índices numéricos. Nesse sentido, Fernandes e Gomes (2003) ressaltam que o método qualitativo descreve uma relação entre o objetivo e os resultados que não podem ser interpretados através de números, nomeando-se como uma pesquisa descritiva. Todas as interpretações dos fenômenos são analisadas indutivamente.

5.2 Universo e amostra da pesquisa

No que tange ao universo da pesquisa, Lakatos e Marconi (2002) compreendem que se refere a indivíduos, fatos, eventos ou conjunto de objetos que apresentam algumas características em comum. Nesse sentido, o universo desta pesquisa foi composto pelos TCCs depositados, via autoarquivamento, no RIUFAL.

O RIUFAL teve sua implantação ocorrida no ano de 2010. No período que vai de 2010 ao início de 2015, o RIUFAL atuava paralelamente com a BDTD no armazenamento, preservação e disseminação das dissertações e teses produzidas na instituição, nesse ínterim, com problemas operacionais e estruturais, a BDTD foi praticamente desativada e o RIUFAL foi afetado diretamente, o que impossibilitou o seu pleno desenvolvimento, chegando ao ponto de quase parar suas atividades.

Após ter sua minuta de criação atualizada, analisada e deliberada pela Câmara Acadêmica foi aprovada e homologada pelo Conselho Universitário (CONSUNI) e, em 2016, foi oficialmente instituída a criação do RIUFAL.

Conforme preconiza a Resolução nº 45/2016-CONSUNI/UFAL, de 07 de novembro de 2016, o RIUFAL tem como missão: reunir a produção científica e acadêmica gerada no âmbito da UFAL com o fim de preservá-la e disseminá-la; disponibilizar o acesso livre às produções acadêmico-científicas da Universidade em acesso aberto; promover a visibilidade e o impacto dos resultados de pesquisa, dos pesquisadores e da própria instituição. O RIUFAL tem com objetivos principais: reunir em um só local virtual toda a produção científica da instituição; promover o acesso livre às produções acadêmicas e científicas, bem como preservar a memória intelectual da UFAL.

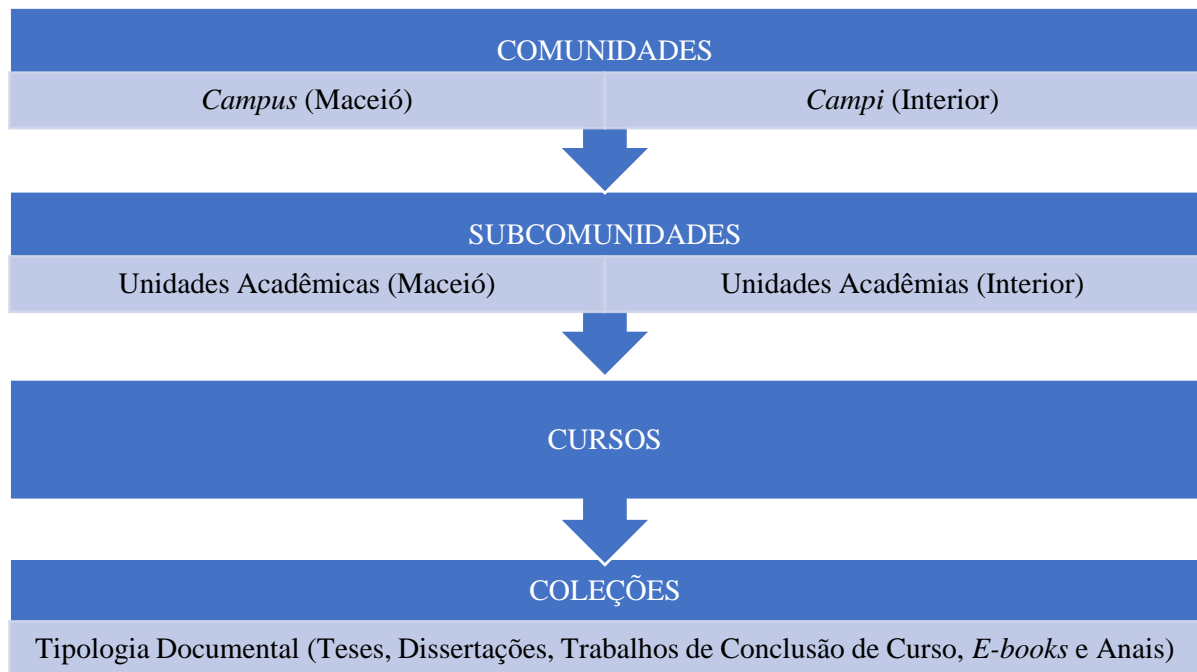
Nesse sentido, o repositório é fundamental para dar visibilidade à instituição, é através dele que a UFAL expõe sua produção científica, possibilitando o acesso às suas pesquisas, sendo a ponte que liga a universidade à comunidade científica e à sociedade de modo geral, revelando o que ela produziu em prol do desenvolvimento humano através da ciência, tecnologia, arte e cultura.

Inserido no movimento da ciência aberta, o RIUFAL faz parte de redes de provedores de dados, serviços e agregadores, como o Oasisbr, o OpenDOAR que é um diretório de repositórios acadêmicos mundiais que monitora a qualidade e consistência da informação nos repositórios acadêmicos de acesso aberto e a LA Referencia, Rede Federada de Repositórios Institucionais de Publicações Científicas na América Latina, a qual apoia estratégias nacionais de acesso aberto na América Latina através de plataforma que prioriza a interoperabilidade, o compartilhamento e a visibilidade da produção científica produzida em instituições de ensino superior e pesquisa científica.

De fato, a participação em provedores renomados credencia cada vez mais a qualidade do repositório, abalizando os pesquisadores e os seus trabalhos. O RIUFAL também faz parte da Rede Nordeste de Repositórios Digitais das instituições de Ensino Superior, Profissional e Tecnológico (RENERE), o qual tem como objetivo reunir as instituições de ensino superior da região Nordeste para consolidar parcerias com ações em prol de desenvolvimento e da sustentabilidade de seus repositórios.

O RIUFAL conta com uma equipe de 2 (dois) bibliotecários, 2 (dois) técnicos administrativos e tem o suporte técnico na parte de informática da Coordenação de Tecnologia da Informação e Comunicação (CTIC). O RIUFAL é responsável pela guarda, preservação e disseminação dos conteúdos informacionais da instituição.

Sua estrutura organizacional é formada por: Comunidades (os *Campi*), SubComunidades (as Unidades Acadêmicas e seus respectivos cursos) e as Coleções (as tipologias documentais), conforme é representado na Figura 1.

Figura 1 - Estrutura do RIUFAL

Fonte: Dados da pesquisa, 2023.

Os *campi* são assim distribuídos: *Campus* Aristóteles Calazans Simões (cursos da UFAL em Maceió), *Campus* Deuza Getaí (cursos da UFAL em Rio Largo), *Campus* Arapiraca (cursos da UFAL em Arapiraca, Palmeira dos Índios, Penedo e Viçosa), *Campus* Sertão (cursos da UFAL em Delmiro Gouveia, Santana do Ipanema).

Seu acervo até o ano de 2022 era composto por, aproximadamente, 8.453 títulos, entre teses, dissertações, TCCs, memoriais acadêmicos, *ebooks* e anais. No tocante aos TCCs, no RIUFAL consta cerca de 2.550 títulos.

Salientando, porém, que os títulos do *Campus* Arapiraca ainda não foram integrados ao RIUFAL, estando em um repositório à parte na Biblioteca Central de Arapiraca, o qual será migrado brevemente.

O acesso ao RIUFAL se dá através do endereço eletrônico <www.repositorio.ufal.br>. Na Figura 2, abaixo, visualiza-se a página inicial do RIUFAL.

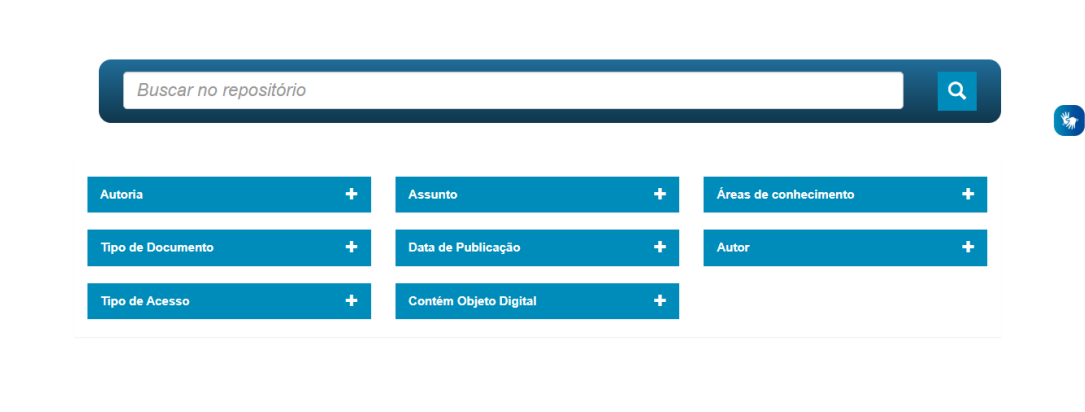
Figura 2 – Página inicial do RIUFAL



Fonte: RIUFAL, 2023.

O RIUFAL tem como mecanismo de busca através do título, autor, assunto, cursos, programas, ano de publicação e tipologias documentais, conforme Figura 3.

Figura 3 – Campos de busca do RIUFAL



Fonte: RIUFAL, 2023.

Clicando nas opções “navegar”, “comunidades e coleções”, o usuário terá acesso às comunidades, subcomunidades e coleções que compõe o RIUFAL. Conforme apresentado na Figura 4.

Figura 4 – Acesso às comunidades e coleções



Fonte: RIUFAL, 2023.

Quanto à inserção dos trabalhos no RIUFAL, os alunos concluintes enviam os seus trabalhos em formato digital PDF, contendo ficha catalográfica e ata assinada para o RIUFAL, os quais são inseridos pela equipe do repositório, com exceção dos cursos de graduação em Enfermagem e Serviço Social que estão fazendo, de forma experimental, o autoarquivamento dos seus trabalhos. Para isso, foi realizado um treinamento através de um tutorial, ensinando aos alunos a inserirem os metadados diretamente na plataforma do RIUFAL.

Salientamos que a obrigatoriedade da ficha catalográfica era até 2021, exclusivamente dos cursos de pós-graduação, passando a ser obrigatória para todos os cursos de graduação da UFAL a partir de 2022, entretanto, desde muito tempo, os cursos da área de saúde a adotam como obrigatória em seus trabalhos, tanto da graduação como da pós-graduação.

O RIUFAL recomenda aos alunos, ao inserirem seus trabalhos na base de dados, no que se refere ao metadado assunto, além das palavras-chave, do resumo e *abstract*, inserir também os cabeçalhos de assunto que constam na ficha catalográfica. Essa medida se justifica pelo fato de que os cabeçalhos de assunto da ficha catalográfica são elaborados por bibliotecários, obedecendo critérios de indexação e padronização.

Buscando aplicar definitivamente o processo do autoarquivamento para todos os cursos, o RIUFAL está elaborando um manual de autoarquivamento, visando guiar os estudantes e pesquisadores da instituição para o preenchimento dos metadados no ato do depósito de seus trabalhos científicos e acadêmicos diretamente no sistema, promovendo, assim, a interoperabilidade, e proporcionar a rapidez da inclusão dos trabalhos, contribuir para a disponibilidade mais rápida desses trabalhos e, conseqüentemente, a sua visibilidade.

Conforme Lakatos e Marconi (2002), a amostra se refere a uma parcela do universo selecionado com diversas finalidades e intenções do pesquisador. Nesse sentido, a amostra foi constituída por 160 monografias, depositadas via autoarquivamento, sendo 84 de Enfermagem, e 76 de Serviço Social.

5.3 Coleta, sistematização e análise dos dados

Os procedimentos metodológicos adotados para coleta e sistematização dos dados foram divididos nas seguintes fases:

Primeira fase: busca direta na base de dados do RIUFAL, na qual foram mapeadas as monografias depositadas através do autoarquivamento, no período de 2020 a 2022. Através da pesquisa nas comunidades “curso de Enfermagem e Serviço Social”, dentro da subcomunidades “TCC”, no qual resultou na amostra de 160 monografias. Foram recuperados 475 metadados de assunto, sendo 260 de Enfermagem e 215 de Serviço Social.

Segunda fase: os dados foram transcritos em planilhas do *Excel* e *Word*, ambos programas da *Microsoft*, possibilitando assim a organização, tabulação e análise dos dados.

Terceira fase: os dados foram tratados e depurados, eliminando o metadado “assunto” repetido, perfazendo um total de 462 cabeçalhos de assunto, 252 nas monografias de Enfermagem e 210 nas monografias de Serviço Social.

Esse processo de reunião de dados é fundamental para execução da pesquisa, pois é através dele que o pesquisador, após tratar e tabular, poderá retratar a realidade do universo pesquisado. Segundo Teixeira (2003):

A coleta e a análise de dados são realmente um processo contínuo que pode estender-se indefinidamente. Como saber se os dados coletados são suficientes? A decisão para finalizar este processo pode estar fundada nos seguintes critérios: esgotamento de fontes; saturação de categorias; aparecimento de regularidades e hiperextensão (Teixeira, 2003, p. 198).

O sucesso da pesquisa dependerá da confiabilidade dos dados coletados, pois todo o trabalho pode ser arruinado, caso os dados não sejam confiáveis ou não puderem ser obtidos para análise.

Na Quarta fase: foram feitas as análises dos metadados “assunto” (Cabeçalhos de assunto), com os cabeçalhos de assunto autorizados pelo catálogo de assunto do SIBI/UFAL.

Foram definidas as seguintes variáveis: metadado de assunto autorizado, metadado de assunto em linguagem natural, metadado de assunto sem padronização, metadado de assunto com erro de grafia e metadado de assunto com erro de digitação.

Quadro 1 – Variáveis da pesquisa

VARIÁVEL	DESCRIÇÃO
Metadado de assunto autorizado	Utilizado pelo SIBI/UFAL
Metadado de assunto linguagem natural	Usado pelo discente (extração do resumo)
Metadado de assunto sem padronização	Registrado com várias variantes
Metadado de assunto com erro de grafia	Ortografia
Metadado de assunto com erro de digitação	Digitação

Fonte: Dados da pesquisa, 2023.

Essas variáveis foram escolhidas pelo fato de serem os ruídos mais comuns apresentados nos metadados “assunto” presentes nos RIs.

Também foi estipulado o índice de concordância dos metadados cadastrados, o qual possibilitará a mensuração de pertinência dos metadados com os utilizados pelo RIUFAL. Conforme a Tabela 1:

Tabela 1 – Índice de concordância de assunto cadastrado no RIUFAL

REFERENCIAL	INDICADOR DE COMPATIBILIDADE
80% a 100%	EXCELENTE
70% a 79%	ALTO
50% a 69	MÉDIO
Abaixo de 50%	BAIXO

Fonte: Dados da pesquisa, 2023.

É a partir desses valores que podemos definir se a linguagem empregada pelo usuário está compatível com a linguagem do sistema do RIUFAL, além de determinar o grau de compatibilidade.

6 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Cumpridas todas as fases estipuladas para a pesquisa, procedeu-se à análise e interpretação dos resultados, os quais serão apresentados em quadros, tabelas e gráficos para maior compreensão do estudo, de modo a facilitar a sua interpretação.

Dos 462 metadados assunto (cabeçalho de assunto), foram identificados 285 metadados assunto autorizado, 115 em linguagem natural, 38 sem padronização, 5 com erro de grafia e 19 com erro de digitação. Conforme Tabela 2, a seguir.

Tabela 2 – Metadados de assunto cadastrados no RIUFAL pelos cursos de Enfermagem e Serviço Social

METADADOS “ASSUNTO” DOS TCCs DE ENFERMAGEM E SERVIÇO SOCIAL	TOTAL DE TCCs	SIBI/UFAL %
Metadado de assunto autorizado	285	61,7%
Metadado de assunto em linguagem natural	115	25%
Metadado de assunto sem padronização	38	8,3%
Metadado de assunto com erro de grafia	5	1%
Metadado de assunto com erro de digitação	19	4%
Total	462	100%

Fonte: Dados da pesquisa, 2023.

O resultado demonstra que os metadados “assunto” cadastrados pelos alunos concluintes dos cursos de Enfermagem e Serviço Social, apresentam um índice mediano quanto à compatibilidade dos metadados “assunto” pertencentes ao SIBI/UFAL (61,7%), ficando aquém de um índice considerado excelente. Os metadados de assunto têm por objetivo descrever e identificar um documento, a fim de facilitar o processo de busca para a recuperação da informação.

Segundo Fujita e Tolare (2019, p. 94), “a falta de controle de vocabulário afeta também a visibilidade das publicações arquivadas no repositório bem como a interação do sistema com o usuário”. Nesse sentido, descrevê-los de forma padronizada utilizando com maior frequência a linguagem do sistema de informação possibilitará a recuperação de informação relevante dentre esses vastos universos de informação que se apresentam nos ambientes digitais.

6.1 Metadado de assunto autorizado

Os instrumentos de indexação têm como função auxiliar nas escolhas dos termos padronizados pertencentes a um campo específico de conhecimento numa linguagem controlada comum entre os pares, facilita a comunicação entre o usuário e o sistema de informação.

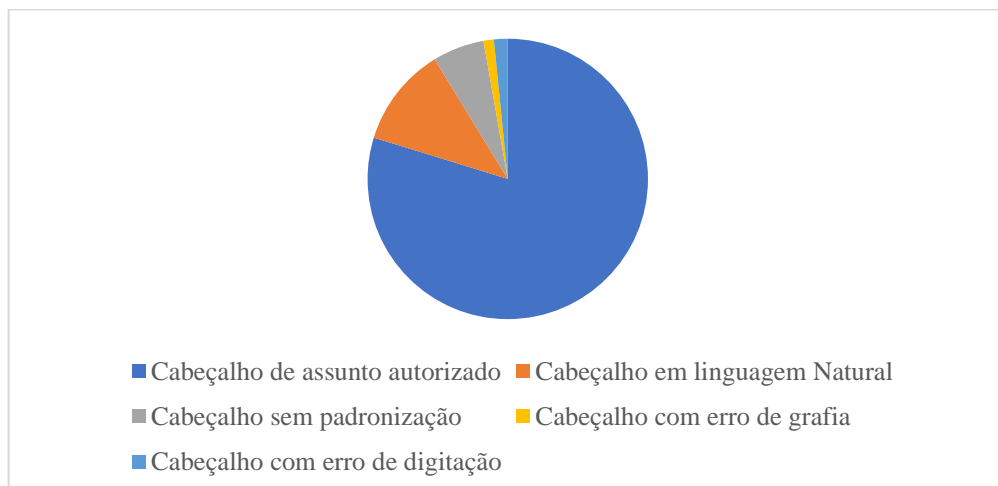
Quanto ao uso de metadados de assunto autorizados, as monografias de Enfermagem apresentaram um excelente resultado, dos 252 metadados de assunto analisados, 201 apresentaram concordância com o catálogo de assunto do SIBI/UFAL, correspondendo a 79,8% e 20,2% sem concordância. Conforme apresentados na Tabela 3 e no Gráfico 1:

Tabela 3 – Metadados de assunto cadastrados no RIUFAL (Enfermagem)

METADADO “ASSUNTO” DOS TCCs DE ENFERMAGEM	TOTAL DE TCCs	SIBI/UFAL %
Metadado de assunto autorizado	201	79,8%
Metadado de assunto em linguagem natural	29	11,5%
Metadado de assunto sem padronização	15	6%
Metadado de assunto com erro de grafia	3	1,20%
Metadado de assunto com erro de digitação	4	1,5%
Total	252	100%

Fonte: Dados da pesquisa, 2023.

Gráfico 1 – Metadados cadastrados (discentes Enfermagem)



Fonte: Dados da pesquisa, 2023.

Tais valores demonstram, assim, o alto grau de compatibilidade dos metadados assunto extraídos dos TCCs de Enfermagem com os utilizados no catálogo de assunto do SIBI/UFAL.

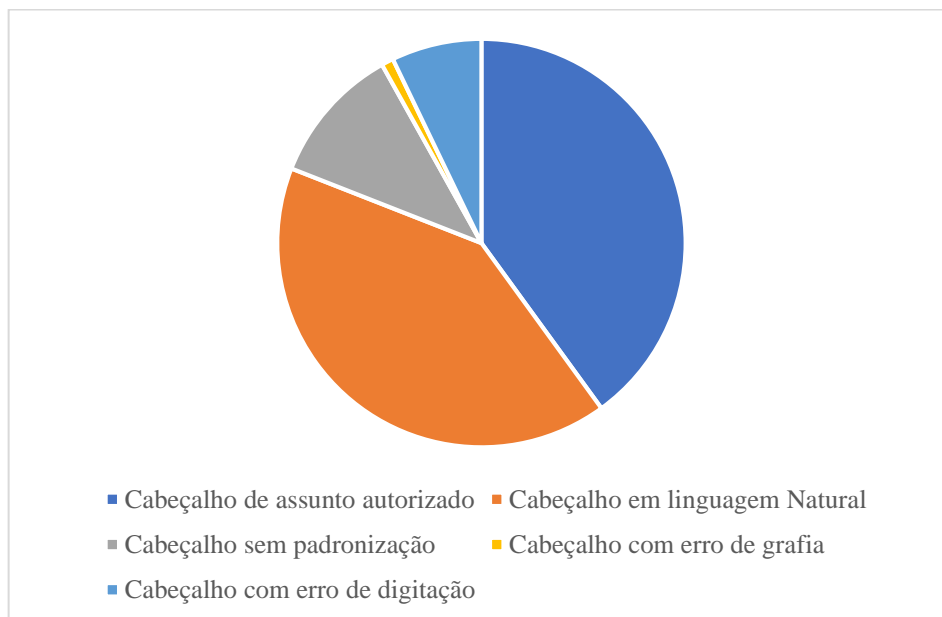
Nos metadados de assunto autorizados apresentados nos TCCs de Serviço Social, dos 210 metadados de assunto analisados, 84 foram compatíveis com os cabeçalhos do catálogo de assunto do SIBI/UFAL, correspondendo a 40% de concordância, e 86 nas demais situações, correspondendo a 60% de discordância. Conforme apresentado na Tabela 4 e no Gráfico 2:

Tabela 4 – Metadados de assunto cadastrados no RIUFAL (Serviço Social)

METADADO “ASSUNTO“ DOS TCCs DE SERVIÇO SOCIAL	TOTAL DE TCCs	SIBI/UFAL %
Metadado de assunto autorizado	84	40%
Metadado de assunto em linguagem natural	86	41%
Metadado de assunto sem padronização	23	11%
Metadado de assunto com erro de grafia	2	1%
Metadado de assunto com erro de digitação	15	7%
Total	210	100%

Fonte: Dados da pesquisa, 2023.

Gráfico 2 – Metadados de assunto cadastrados (discentes de Serviço Social)

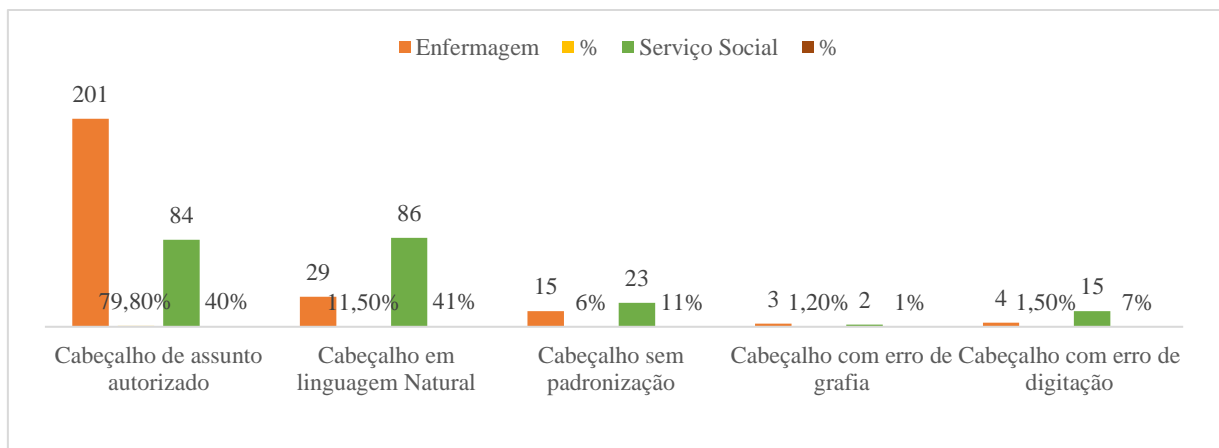


Fonte: Dados da pesquisa, 2023.

Quanto aos metadados de assunto extraídos dos TCCs de Serviço Social, percebe-se que esses trabalhos apresentaram um baixo índice de concordância com os metadados assunto do catálogo de assunto do SIBI/UFAL.

Comparando as duas situações, percebe-se a discrepância entre os dois cursos. Enquanto os TCCs de Enfermagem apresentam um alto índice de compatibilidade entre os metadados assunto extraídos deles com metadados assunto pertencentes ao catálogo de assunto do SIBI/UFAL, o mesmo não ocorre com os TCCs de Serviço Social, conforme apresentado no Gráfico 3, abaixo:

Gráfico 3 – Comparativo dos metadados de assunto cadastrados (discentes de Enfermagem e Serviço Social)



Fonte: Dados da pesquisa, 2023.

Os alunos de Serviço Social utilizam mais de metadados “assunto” em linguagem natural, desse modo, apresenta um baixo grau de concordância com o catálogo de assunto do SIBI/UFAL. Algo relevante para destacar o bom grau de concordância entre os metadados cadastrados pelos alunos de Enfermagem é o fato de que os alunos desse curso utilizam com frequência a base de dados DeCS, para coletar termos da área da saúde, bem como o uso dos cabeçalhos de assunto que consta na ficha catalográfica, os quais são elaborados pelos bibliotecários do SIBI/UFAL, sendo item obrigatório em seus TCCs.

Ao fazerem o autoarquivamento, os alunos do curso de Enfermagem foram orientados a incluírem também no momento do preenchimento do metadado assunto, além das palavras-chave do seu resumo, os cabeçalhos de assunto da ficha catalográfica. O curso de Serviço social não adotava a inclusão da ficha catalográfica nos seus TCCs, algo que passou a ser obrigatório a partir de 2023.

Santos (2017) e Fujita e Tolare (2019) enfatizam que o uso do vocabulário controlado, evita a dispersão de recursos informacionais, facilita a comunicação com a padronização, controla as dispersões lexicais da linguagem e indica os mais pertinentes para representação da informação. Desse modo, ficou evidente que o auxílio de ferramentas de indexação contribuiu para uma maior padronização dos cabeçalhos de assunto nos TCCs de Enfermagem.

O índice de discordância em relação aos termos autorizados fica muito baixo quando o usuário do sistema tem algum auxílio que o oriente nas escolhas dos termos, apesar do autoarquivamento possibilitar a livre escolha do usuário para inserção dos cabeçalhos, a opção por termos ou cabeçalhos não padronizados desse grupo é muito pequena.

6.2 Metadado de assunto linguagem natural

Apesar de ser comum o uso da linguagem natural nos ambientes digitais com os RIs, há preocupação por parte dos pesquisadores pela falta de padronização que isso acarreta. Os autores Santos (2017), Freitas (2019) e Tartarotti (2019) ressaltam que o uso indiscriminado de termos pode gerar problemas na representação da informação, além de não garantir a qualidade do sistema.

No que se refere ao uso da linguagem natural (por extração) nos TCCs de Enfermagem, apenas 29 cabeçalhos de assunto apareceram nessa situação, o que correspondeu a 11,5%, isso fica evidente. E o baixo aparecimento se deve à utilização de ferramentas de indexação e o uso da linguagem natural é ocasionada por termos que não constam nessas ferramentas. Enquanto os TCCs de Serviço Social apresentaram muitos cabeçalhos em linguagem natural (por extração), 86 cabeçalhos, o que corresponde a 41%. Os dados completos são apresentados na Tabela 5:

Tabela 5 – Metadados de assunto em linguagem natural cadastrados no RIUFAL

METADADOS “ASSUNTO”	ENFERMAGEM		SERVIÇO SOCIAL	
	QUANTIDADE	%	QUANTIDADE	%
Linguagem natural	29	11,5%	86	41%
Demais situações	223	88,5%	124	59%
Total	252	100%	210	100%

Fonte: Dados da pesquisa, 2023.

A utilização de linguagem natural nesses ambientes é muito comum, no entanto, vários estudos comprovam que é necessário a busca de uma padronização, não deixando o uso indiscriminado dessa linguagem, pois devido à sua natureza heterogênea e muito subjetiva, poderá criar ruídos que interferem na qualidade dos sistemas informacionais, comprometendo a recuperação da informação e, conseqüentemente, provocando dúvidas quanto à credibilidade do sistema. O fato de os discentes de Serviço Social não utilizarem nenhuma ferramenta de suporte para auxílio na indexação e escolha dos termos para o autoarquivamento proporciona o uso exclusivo da linguagem natural (por extração) e, essa, na maioria das vezes, não é padronizada e boa parte, muito subjetiva.

A possibilidade de interação entre o usuário, o sistema de informação e o uso da indexação por extração ocasiona uma maior aproximação do pensamento do usuário e o conteúdo por ele produzido, entretanto, é fundamental que ele fale uma linguagem mais próxima da linguagem do sistema para maior aproveitamento na busca dos conteúdos informacionais.

6.3 Metadado de assunto sem padronização

Nesta categoria, enquadram-se os metadados de assunto com variações de apresentação, como o uso no plural e no singular, de letras todas em maiúsculas, de letras todas em minúsculas e só com a primeira letra em maiúscula. Esses casos não interferem diretamente na busca, visto que o sistema recupera as variações, entretanto, para elaboração de relatório causa um grande problema devido ao sistema DSpace cadastrar cada metadado assunto como uma entrada, agrupando palavras semelhantes.

Quanto ao uso de metadado assunto no singular/plural foram detectados quatro (4) casos e, em relação a siglas, foram encontrados dois (2). No tocante às siglas, devido à possibilidade de instituições e organizações terem as mesmas siglas, é recomendado a utilização por extenso para evitar erros de associações de assuntos diferentes, o uso da sigla só é recomendado nos casos em que a sigla é reconhecida exclusivamente por ela, como é o caso da Unesco. Abaixo o quadro referente aos metadados de assuntos cadastrados com variações de plural, singular ou siglas.

Quadro 2 – Metadados de assunto cadastrados com variações de plural, singular ou siglas.

CADASTRADOS NO AUTOARQUIVAMENTO	UTILIZADO NO CATÁLOGO SIBI/UFAL
Enfermeiro	Enfermeiros
Estudante	Estudantes
Unidades de Terapias Intensivas	Unidade de terapia intensiva
Adolescências	Adolescência
SUS	Sistema Único de Saúde
PENAE	Programa Nacional de Alimentação Escolar (Brasil)

Fonte: Dados da pesquisa, 2023.

Os TCCs de Enfermagem apresentaram 15 metadados assuntos sem padronização, ou seja 6%. Os cadastrados pelos alunos concluintes do Curso de Serviço Social, são 23 metadados assuntos sem padronização, correspondendo a 11%. No Quadro 2 estão descritos os metadados assuntos com variações.

Quadro 3 – Variações dos metadados de assunto cadastrados

(Continua)

VARIAÇÕES DOS CABEÇALHOS		
	Cabeçalho sem padronização	Cabeçalho autorizado SIBI/UFAL
ENFERMAGEM	Cobertura Vacinal	Cobertura vacinal
	Covid-19	COVID - 19 (Pandemia)
	Covid -19 (Doença)	COVID-19
	Enfermagem no Consultório	Enfermagem no consultório
	Enfermeiro	Enfermeiros
	Estratégia Saúde da Família	Estratégia Saúde da Família (ESF)
	Estudante	Estudantes
	Infecção dos Ferimentos	Infecção dos ferimentos
	Infecções Relacionadas à Assistência em Saúde	Infecções relacionadas à assistência em saúde
	Própolis vermelho	Própolis vermelho Própolis vermelha
	Saúde da Mulher	Saúde da mulher
	Saúde do trabalhador	Saúde do Trabalhador
	Transtorno do Espectro Autista	Transtorno do espectro autista
	Unidades de Terapias Intensivas	Unidade de terapia intensiva
	Unidades de Terapia Intensiva Neonatal	Unidade de terapia intensiva neonatal

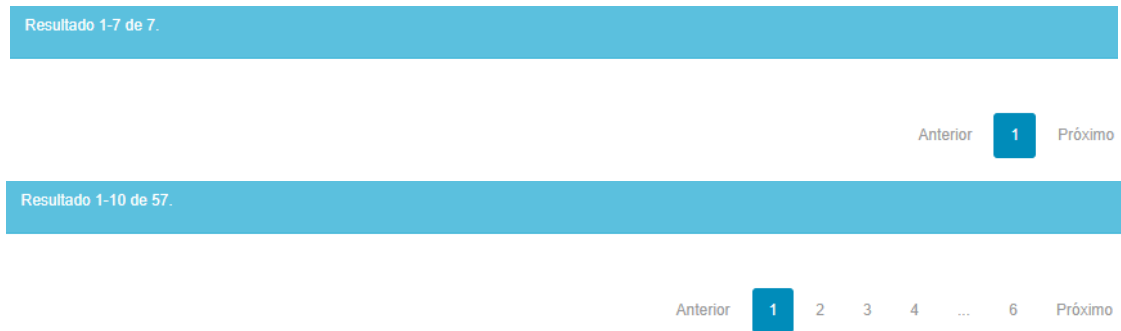
Quadro 3 – Variações dos metadados de assunto cadastrados

(Conclusão)

VARIAÇÕES DOS CABEÇALHOS		
SERVIÇO SOCIAL	Cabeçalho sem padronização	Cabeçalho autorizado SIBI/UFAL
	Adolescências	Adolescência
	Assistência Social	Assistência social
	Assistente Social	Assistente social
	capitalismo	Capitalismo
	Covid-19	COVID-19
	Capitalismo Contemporâneo	Capitalismo contemporâneo
	Crise do capital	Crise do capital
	dependentes químicos	Dependentes químicos
	Direitos Humanos	Direitos humanos
	Educação Superior	Educação superior
	Estudante	Estudantes
	Pandemia da Covid-19	COVID - 19 (Pandemia)
	PNAE	Programa Nacional de Alimentação Escolar (Brasil)
	Proteção Social	Proteção social
	Restrição de Direitos	Restrição de direitos
	Relações de Trabalho	Relações de trabalho
	Reestruturação Produtiva	Reestruturação produtiva
	Questão Social	Questão social
	Saúde Mental	Saúde mental
Serviço Social	Serviço social	
Segurança Alimentar e Nutricional	Segurança alimentar e nutricional	
SUS	Sistema Único de Saúde	
Transporte Público	Transporte público	

Fonte: Dados da pesquisa, 2023.

Outro ponto que leva a grande preocupação e o cadastramento de metadado de assunto com variações linguísticas, como: “própolis vermelha e própolis vermelho”, na busca na base de dados RIUFAL por esses termos geram no caso de própolis vermelho a recuperação de sete documentos, enquanto própolis vermelha recupera 57 documentos. Conforme figura 5.

Figura 5 – Quantitativo de documentos com os metadado própolis vermelho e própolis vermelha

Fonte: RIUFAL, 2023.

Essa falta de padronização proporciona um grande problema para a base de dados do RIUFAL, pois poderá levar o usuário pesquisador ao acesso de uma quantidade limitada de documentos dependendo da escolha do termo, além de passar a visão deque o sistema só possui essa quantidade de documentos sobre o assunto.

Conforme (Santos, 2017; Freitas, 2019; Tartarotti, 2019), o uso indiscriminado da linguagem natural pode gerar problemas na recuperação da informação, propicia a não padronização, força uma busca mais abrangente a além de não garantir a qualidade do sistema.

Nestes aspectos, fica evidente que os ruídos gerados com esses metadados assuntos no RIUFAL provocam sérios problemas para a integridade do sistema.

6.4 Metadado de assunto com erros de grafia e digitação

Quanto ao metadado de assunto com erros de grafia, foram detectados apenas cinco (5), sendo três (3) em Enfermagem e dois (2) em Serviço Social, enquanto com erro de digitação foram detectados 19 cabeçalhos, quatro (4) de Enfermagem e 15 de Serviço Social.

Quadro 4 – Metadados de assunto com erros de grafia ou digitação (discentes de Enfermagem)

METADADO DE ASSUNTO DE ENFERMAGEM	
ERROS DE GRAFIA	CORRETO
Enfermagem prioperatoria	Enfermagem perioperatória
Prenatal	Pré-natal
Enfermera	Enfermeira
ERRO DE DIGITAÇÃO	
Epidemiologia descritiva.,	
Esgotamento profissional*	
Estratégia Saúde da Família.	
Infecção.	

Fonte: Dados da pesquisa, 2023.

Percebe-se que a maior parte dos erros de digitação consistem no uso do ponto no final no metadado de assunto, esse fato é gerado devido aos alunos concluintes ao “copiar e colar” as palavras-chave dos resumos, copiam como está no TCC e colocam no campo indicado para cadastramento do metadado “assunto”, no sistema do RIUFAL.

Quadro 5 – Metadado de assunto com erros de grafia ou digitação (discentes de Serviço Social)

CABEÇALHOS DE SERVIÇO SOCIAL	
ERROS DE GRAFIA	CORRETO
Política Nacional de Humanização	Política Nacional de Humanização
Sociojuridico	Sociojurídico
ERRO DE DIGITAÇÃO	
Crise estrutural do capital.	
Direitos humanos.	
Direitos sociais.	
Neoliberalismo.	
Programa Nacional de Assistência Estudantil.	
Reforma da Previdência.	
Sistema Prisional.	
Responsabilidade civil Cidadania	
Encarceramento Feminino.	
Serviço Social.	
Mulheres.	
Políticas Sociais.	
Reação Burguesa.	
Serviço Social.	
Surdez.	

Fonte: Dados da pesquisa, 2023.

Outro aspecto relevante, referente às escolhas dos metadados de assunto cadastrados dos TCCs de Enfermagem e Serviço Social, é que os alunos concluintes, na maioria das vezes, utilizam-se de uma quantidade resumida de metadados, não explorando totalmente o conteúdo informacional em seus trabalhos.

O RIUFAL, apesar de usar o catálogo de assunto do SIBI/UFAL e ter como referência também os catálogos do DeCS, para assuntos na área de saúde, e o da Fundação Biblioteca Nacional (BN) para os demais assuntos nas escolhas dos metadados “assunto” para compor sua base, não garante a padronização desses metadados, visto que, às vezes, o mesmo metadado “assunto” tem forma de entrada diferente no DeCS e na BN, conforme apresenta o Quadro 6 abaixo:

Quadro 6 – Metadados de assunto com variações de entrada no DeCS, Catálogo da BN e no SIBI/UFAL

VARIAÇÕES DE ENTRADA DE METADADOS “ASSUNTO” NOS CATÁLOGOS		
Decs	Catálogo da BN	Catálogo do SIBI/UFAL
COVID-19	Covid-19 (Doença) Covid-19, Pandemia de, 2020	COVID-19 COVID-19 (Pandemia) Covid-19 (Doença)
Cuidados de Enfermagem	Cuidados de enfermagem	Cuidados de enfermagem
Diagnóstico de Enfermagem	Diagnóstico em enfermagem	Diagnóstico de enfermagem
Estratégia de Saúde Nacional	Estratégia de Saúde da Família (Brasil)	Estratégia de Saúde da Família Estratégia de Saúde da Família (ESF)
Síndrome de Imunodeficiência Adquirida	AIDS (Doença)	AIDS
Transtorno do Espectro Autista	Transtornos do espectro autista	Transtorno do Espectro Autista Transtorno do espectro autista
Sistema Único de Saúde	Sistema Único de Saúde (Brasil)	Sistema Único de Saúde Sistema Único de Saúde - SUS

Fonte: Dados da pesquisa, 2023.

A falta de uma política de indexação no RIUFAL faz com que não haja padronização dos metadados “assunto”, pois, se os metadados utilizados são qualificados e validados por catálogos e tesouros específicos, as variações de entrada desses metadados compromete a qualidade do sistema de informação, devido ao indexador não definir qual será a entrada autorizada do metadado “assunto” no RIUFAL, utilizando ora metadados “assunto” de um ou de outro catálogo, causando com isso um ruído no sistema informacional.

7 CONSIDERAÇÕES FINAIS

O RIUFAL se constitui como um instrumento fundamental para preservação, disseminação e divulgação das produções acadêmicas e científicas, ele cumpre o seu papel enquadrado na filosofia da Ciência Aberta. Entretanto, para que seus conteúdos sejam disseminados é preciso que sejam descritos e representados de forma que os usuários do sistema possam localizá-los em suas buscas, nesse sentido, a representação da informação é fator fundamental para o processo de organização, acesso e recuperação da informação. Outro ponto preponderante, é a utilização do autoarquivamento como um processo contínuo, possibilitando a interação do autor, seu conteúdo informacional e o sistema de informação.

Diante disso, a pesquisa buscou analisar a representação da informação nos metadados “assunto” cadastrados no RIUFAL através do autoarquivamento dos TCCs, realizados pelos alunos dos cursos de Enfermagem e Serviço Social. No que concerne aos objetivos propostos, quanto ao objetivo A, a coleta se deu de forma direta, visto que, apenas os dois cursos foram os únicos a fazerem o autoarquivamento; no objetivo B, que tratou da análise dos metadados “assunto”, percebeu-se que o uso da linguagem natural foi mais predominante no curso de Serviço Social, o qual apresentou 41%, enquanto o curso de Enfermagem 11,5%, demonstrando que, quem tem como base o uso de ferramentas auxiliares como catálogos autorizados, tesouros especializados e orientação do profissional bibliotecário para a elaboração dos metadados “assunto” não utilizam tanto da linguagem natural. No que se refere à falta de padronização, 6% foi detectado no curso de Enfermagem, aparecendo novamente o curso de Serviço Social com o maior percentual, 11%; no tocante ao erro de grafia e digitação, foi apresentado um baixo índice, apenas 1% para erro de grafia e 4% para erro de digitação.

No tocante ao objetivo C: comparar os metadados utilizados com cabeçalhos autorizados do catálogo de assunto do Sistema de Bibliotecas da Universidade Federal de Alagoas (SIBI/UFAL), dos 462 metadados de assunto analisados, 285 apresentaram concordância com os metadados de assunto do catálogo do SIBI/UFAL, perfazendo o percentual de 61,7%, o que equivale a um índice mediano. Quando analisados os dados isolados por cada curso, o índice de concordância apresenta uma grande variação, enquanto o curso de Enfermagem dos 252 metadados de assunto, 201 apresentaram concordância com o catálogo SIBI/UFAL, ou seja, 79,8%, um alto índice; enquanto o curso de Serviço Social, dos 210 metadados “assunto”, apenas 84 apresentaram concordância, 40%, ou seja, um índice baixo.

E, finalmente, no objetivo D: identificar se, dentre os metadados cadastrados, existe algum que apresente ruídos que possam interferir no processo da busca e recuperação da

informação no sistema, foram constatados vários ruídos, desde variações de uso de palavras no plural e singular, erros de digitação e grafia, bem como variações de forma de apresentação do mesmo assunto, caso dos metadados de assunto “própolis vermelha e própolis vermelho”, no qual o primeiro apresenta a recuperação de 57 documentos, enquanto o segundo a recuperação de sete (7) documentos, ocasionando, assim, uma grande falha que pode prejudicar a busca de conteúdos informacionais, visto que, dependendo da escolha do metadado “assunto”, poder-se-á deixar de fora determinados documentos na pesquisa recuperada.

Ficou muito evidente, nesta pesquisa, que a utilização da linguagem natural, em grande proporção, apresenta uma maior ocorrência, com variações de cabeçalhos com falhas de padronização, erro de digitação e, principalmente, não concordância com o catálogo do SIBI/UFAL, na outra vertente, monografias que utilizaram ferramentas como tesouros e os cabeçalhos de assunto, presentes na ficha catalográfica, apresentaram um excelente grau de concordância e confiabilidade, contribuindo para a qualidade do sistema. Ainda assim, não há dúvidas de que o uso da linguagem natural é importante também para a aproximação do autor com o conteúdo produzido, contudo, a padronização se torna algo essencial para a qualidade e credibilidade de qualquer sistema de informação.

Das dificuldades encontradas ao longo da pesquisa, destacam-se, além da pandemia da Covid-19 que afetou toda a sociedade, desde o aspecto clínico como também psicológico, causando um período de incerteza, o RIUFAL apresentou problemas operacionais, os quais o deixou fora do ar por vários meses, prejudicando, com isso, a coleta de dados.

Com base nos dados analisados e nos resultados obtidos, pode-se afirmar que a presente pesquisa alcançou os objetivos propostos e respondeu a hipótese do trabalho, pois foi possível constatar os argumentos através das análises dos dados coletados.

Ressalta-se que a falta de uma política de indexação por parte do RIUFAL também contribui para a falta de padronização dos metadados de assunto, pois foi constatado que o catálogo de assunto do SIBI/UFAL, ora usava termos retirados do Catálogo da BN, ora da base DeCS, apresentando o mesmo metadado “assunto” registrado com diferentes formas de entrada.

Acredita-se, que esta pesquisa contribua para o desenvolvimento da área da Ciência da Informação, sendo mais um elemento no estudo de Repositório Institucional, autoarquivamento, representação e indexação da informação; constitua-se, também, de um subsídio teórico para a melhoria do RIUFAL no tocante aos metadados “assunto” utilizados no sistema, pois apresenta um estudo com base nos metadados “assunto” cadastrados no RIUFAL, através do autoarquivamento.

Recomenda-se a implantação de uma política de indexação no RIUFAL, criação de mecanismos que contribua para a padronização, qualidade e pertinência dos metadados “assunto”, tais como índice ou catálogos de assunto, elaboração de manuais e guias para inserção de metadados, tutoriais para treinamentos dos usuários e a implantação definitiva do autoarquivamento para a toda comunidade.

Por fim, que este estudo estimule pesquisas futuras que contribuam para o desenvolvimento do RIUFAL e, principalmente, dos temas “representação da informação e indexação nos repositórios”, de suma importância para a recuperação da informação em ambientes digitais.

REFERÊNCIAS

- ALBUQUERQUE, M. E. C.; GAUDÊNCIO, S. M.; SANTOS, R. F. Reflexões teóricas em representação da informação. *In*: ALBUQUERQUE, M. E. C.; MARTINS, G. K.; MOTA, D. A. R. (orgs.). **Organização e representação da informação e do conhecimento**: intersecções teórico-sociais. João Pessoa: Editora UFPB, 2019.
- ALMEIDA, E. C. E.; GUIMARÃES, J. E.; ALVES, I. T. G. Dez anos do Portal de Periódicos Capes: histórico, evolução e utilização. **RBPG**, Brasília, v. 7, n. 13, p. 218-246, nov. 2010.
- ALVES, V. **Informação científica em Biblioteconomia**: o livre acesso nas universidades federais. Curitiba: Appris, 2017.
- ALVES, V. B. A. Open archives: via verde ou via dourada? **Ponto de Acesso**, Salvador, v. 2, n. 2, p. 127-137, ago./set. 2008. Disponível em: <http://www.pontodeacesso.ici.ufba.br>. Acesso em: 23 set. 2022.
- AMSTEL, F. V. **Folksonomia**: vocabulário descontrolado, anarquitectura da informação ou samba do crioulo doido? 2007. Disponível em: https://www.usabilidoido.com.br/arquivo/folksonomia_anarquitectura.pdf. Acesso em: 23 set. 2022.
- ANDRADE, R. M.; MURIEL-TORRADO, E. C. Declarações de acesso aberto e a lei de direitos autorais brasileira. **Reciis**: Rev. Eletron. Comun. Inf. Saúde, Rio de Janeiro, n. 11 (supl.), p. 1-5, nov. 2017. Disponível em: <http://www.reciis.icict.fiocruz.br>. Acesso em: 12 jan. 2022.
- ARAKAKI, A. C. S.; CASTRO, F. F. Aspectos da representação da informação na curadoria digital. **Encontros Bibli.**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 25, p. 01-21, 2020.
- ASSIS, T. B. Análise das políticas de autoarquivamento nos repositórios institucionais brasileiros e portugueses. **InCID**: Rev. Ci. Inf. Doc, Ribeirão Preto, v. 4, n. 2, Ed. Esp., p. 212-227, jul./dez. 2013.
- BRASCHER, M.; CAFÉ, L. M. A. Organização da informação ou organização do conhecimento? *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: USP, 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/176535>. Acesso em: 15 maio 2023.
- CAPES. Portal de Periódicos da Capes. Disponível em: <http://www.periodicos-capes-gov-br.ezlp.periodicos.capes.gov.br/>. Acesso em: 15 jun. 2022.
- CASSÉ, A. P. O. M. **Gestão da informação e da comunicação em repositórios institucionais**: estudo comparativo entre a Universidade Federal de Alagoas e a Universidade Federal do Ceará. 2023. 293 f. Dissertação em (Ciência da Informação) – Instituto de Ciências Humanas, Comunicação e Artes, Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Alagoas, Maceió, 2022.

CASTRO, F. F.; SANTOS, P. L. V. A. C. Elementos de interoperabilidade na perspectiva da catalogação descritiva. **Informação & Sociedade**, João Pessoa, v. 24, n. 3, p. 13-25, set./dez. 2017. Disponível em: <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/16660>. Acesso em: 15 dez. 2022.

CASTRO, M. F. Biblioteca universitária: desafios diante das tecnologias da informação e da comunicação no Brasil. **Bibl. Univ.**, Belo Horizonte, v. 4, n. 2, p. 4-17, jul./dez. 2017.

CATARINO, M.; BAPTISTA, A. Folksonomias: características das etiquetas na descrição de recursos da web. **Inf. Inf.**, Londrina, v.14, n. esp, p. 46-67, 2009.

COSTA, M. P.; LEITE, F. C. L. **Repositórios institucionais da América Latina e o acesso aberto à informação científica**. Brasília: IBICT, 2017.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Tradução de Magda França Lopes, 3. ed. Porto Alegre: Artmed, 2010.

CUNHA, M. B.; DIÓGENES, F.C. B.A trajetória da biblioteca universitária no Brasil: no período de 1901 a 2010. **Encontros Bibli.**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 21, n. 47, p. 100-123, set./dez. 2016.

DIAS, E. W.; NAVES, M. M. **Análise de assunto: teoria e prática**. Brasília: Thesaurus, 2007.

DIÓGENES, F. C. B.; CUNHA, M. B. Desenvolvimento das universidades e bibliotecas universitárias na idade média até à modernidade. **RDBCI: Rev. Digit.Bibliotecon.Cienc.Inf.**, Campinas, v. 15, n. 1, p. 99-129, jan./abr. 2017.

DODEBEI, V. Repositórios institucionais: por uma memória criativa no ciberespaço. *In: SAYÃO, L. et al. Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. Salvador: EDUFBA, 2009.

FREITAS, M. P. **Autoarquivamento e representação de assunto: estudo analítico de teses e dissertações do Repositório Institucional da UFSCar**. 2019. 89 f. Dissertação (Mestrado em Ciência da informação) – Centro de Educação e Ciências Humanas, Programa de Pós-Graduação em Ciência da informação, Universidade Federal de São Carlos, São Carlos, 2019.

FUJITA, M. S. L.; RUBI, M. P.; BOCCATO, V. R. C. O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. **DataGramaZero**, v. 10, n. 2, p. 1-22, abr. 2009.

FUJITA, M. S. L.; TOLARE, J. B. Vocabulários controlados na representação e recuperação da informação em repositórios brasileiros. **Inf. Inf**, Londrina, v. 24, n. 2, p. 93-125, maio/ago. 2019.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GOMES, F. A. Padronização de metadados na representação da informação em repositórios institucionais de universidades federais brasileiras. **Ponto de Acesso**, Salvador, v. 10, n. 2, p. 93-94, ago. 2016. Disponível em: <http://www.pontodeacesso.ici.ufba.br>. Acesso em: 15 dez. 2020.

GOMES, M. J.; ROSA, F. Comunicação científica: das restrições ao acesso livre. *In*: GOMES, M. J.; ROSA, F (org.). **Repositórios institucionais**: democratizando o acesso ao conhecimento. Salvador: EDUFBA, 2010.

GONZÁLEZ DE GÓMEZ, M. N. A universidade e a “sociedade de informação”. **Rev. Dig. Bibl. Ci. Inf.**, Campinas, v. 9, n. 1, p. 225-242, jul./dez. 2011.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Classificação Decimal universal (CDU). **Publicações IBICT**, jun. 2019. Disponível em: <https://antigo.ibict.br/publicacoes-e-institucionais/classificacao-decimal-universal-cdu>. Acesso em: 15 jan. 2022.

KURAMOTO, H. Prefácio. *In*: LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009.

KURAMOTO, H. Repositórios institucionais. *In*: SAYÃO, L *et al.* (org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Briquet de Lemos, 2004.

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

LOPES, I L. Uso das linguagens controlada e natural em base de dados: revisão de literatura. **Ci. Inf**, Brasília, v.31, n. 1, p. 41-52, jan. /abr., 2002. Disponível em: <https://www.scielo.br/j/ci/a/WhYch5gHnWYPCPY3NPwbkcR/?format=pdf&lang=pt>. Acesso em: 25 jul. 2022.

MAGALHÃES, A. M. **A identidade do ensino superior**: política, conhecimento e educação, numa época de transição. Porto: Fundação Calouste Gulbenkian, 2004. (Série Textos Universitários de Ciências Sociais e Humanas).

MAIMONE, G. D.; SILVEIRA, N. C.; TÁLAMO, M. F. G. M. Reflexões acerca das relações entre representação temática e descritiva. **Informação & Sociedade**, João Pessoa, v. 21, n. 1, p. 27-35, jan./abr. 2011. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/7367/5596>. Acesso em: 15 maio. 2023.

MARCONDES, C. H.; SAYÃO, L. F. Softwares livres para repositórios institucionais: alguns subsídios para a seleção. *In*: SAYÃO L. F. *et al.* (org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

MARCONI, M.; LAKATOS, E. M. **Técnicas de pesquisa**. 8. ed. São Paulo: Atlas, 2018.

MARQUES, C. A. G. Gerenciamento de repositórios digitais: apontamentos práticos para o desenvolvimento dos repositórios institucionais. **ConCI**: Conv. Ciênc. Inform, Aracaju, v. 3, n. 2, p. 135-162, maio/ago. 2020.

MARTINS, G. K.; CÔRTEZ, G. R. A representação da informação do conhecimento e as representações sociais: intersecções e limites. *In*: ALBUQUERQUE, M. E. C.; MARTINS, G. K.; MOTA, D. A. R. (orgs.). **Organização e representação da informação e do conhecimento**: intersecções teórico-sociais. João Pessoa: Editora UFPB, 2019.

MIRANDA, I A. A.; DELFINO, J, G. M. Repositórios institucionais: novos desafios para as bibliotecas e para os bibliotecários. **Bibl. Univ**, Belo Horizonte, v. 3, n. 1, p. 52-64, jan./jun. 2016.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ci. Inf.**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <https://www.scielo.br/j/ci/a/nGD3MkKfNxtjnnWshf3YVjP/?format=pdf&lang=pt>. Acesso em: 20 set. 2022.

NORUZI, A. Folksonomies: why do need controlled vocabulary? **Webology**, v. 4, n. 2, 2007. Disponível em: <http://www.webology.org/2007/v4n2/editorial12.html>. Acesso em: 23 set. 2022.

NUNES, M. S. C.; CARVALHO, K. As bibliotecas universitárias em perspectiva histórica: a caminho do desenvolvimento durável. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 21, n. 1, p. 173-193, jan./mar. 2016.

OASISBR. Portal Brasileiro de Publicações e Dados Científicos de Acesso Aberto. Disponível em: <https://oasisbr.ibict.br/vufind/datasources/home>. Acesso: 15 jun. 2023.

OLIVEIRA, M. A. D.; ALVES, M. V.; MAIA, M. A. Q. A função social do profissional da informação numa biblioteca inclusiva. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 25., 2013, Florianópolis. **Anais [...]**. Florianópolis: FEBAB, 2013. Disponível em: <https://portal.febab.org.br/cbbd2019/article/view/1600>. Acesso em: 17 maio 2023.

RAMPAZZO, L.; JOSLIN, É. B. Identidade, história e desafios da universidade. **Revista de Ciência Contemporânea**, Guaratinguetá, v. 1, n. 1, p. 75-87, jan./jun. 2017. Disponível em: http://uniesp.edu.br/sites/guaratingueta/revista.php?id_revista=31. Acesso em: 09 jan. 2022.

RIETJEANS, M. H.; KROEFF, M. S.; PINHEIRO, L. V. Ciclo de vida dos recursos eletrônicos: enfoque na gestão de coleções em bibliotecas universitárias. **Ponto de Acesso**, Salvador, v. 13, n. 1, p. 3-29, abr. 2019. Disponível em: <http://www.pontodeacesso.ici.ufba.br>. Acesso em: 27 dez. 2022.

ROSSATO, R. **Universidade**: nove séculos de história. Passo Fundo: UPF, 2005.

SANCHEZ, F. A.; SILVA, N. B. P.; VECHIATO, F. L. Padrão de metadados para representação e organização da informação em repositórios de dados de pesquisas. **Informação & Tecnologia (ITEC)**, Marília/João Pessoa, v. 5, n. 1, p. 37-51, jan./jun. 2018.

SANTA ANNA, Jorge. O bibliotecário em face das transformações sociais: de guardião a um profissional desinstitucionalizado. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 20, n. 1, p. 138-157, jan./abr. 2015. Disponível em: http://revista.acb.org.br/racb/article/view/985/pdf_118. Acesso em: 17 maio 2023.

SANTA ANNA, J. A redefinição da biblioteca no século XXI: de ambientes informacionais a espaços de convivência. **RDBCI: Rev. Digit. Bibliotecon. Cienc. Inf**, Campinas, v. 14, n. 2, p. 232-246, maio/ago. 2016.

SANTOS, E. T. G.; ARAÚJO, E. A. A Biblioteca Universitária no Brasil: o desafio de participação no desenvolvimento científico e tecnológico na segunda metade do século XX. **RICI: R. Ibero-amer. Ci. Inf. Brasília**, v. 15, n. 1, p. 6-24, jan./abr. 2022.

SANTOS, H. P. Etiquetagem e folksonomia: usuário e sua motivação para organizar e compartilhar informação na web 2.0. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 1, n. 2, p. 91-104, abr./jun. 2013.

SANTOS, R. F. Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. **Rev. Inf. na Soc. Contemp.**, Natal, n. Especial, p. 11-22. 2017.

SANTOS, S. R. O.; OLIVEIRA, D. A. Autoarquivamento na Ciência da Informação: uma análise dos documentos depositados no repositório digital e-LIS. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, p. 1-12. 2019.

SANTO JUNIOR, E. R. **Repositórios institucionais de acesso livre no Brasil**: estudos Delfos. 2022. 89 f. Dissertação (Mestrado em Ciência da informação) – Faculdade de Ciência da Informação, Programa de Pós-Graduação em Ciência da informação, Universidade de Brasília, Brasília, 2010.

SAYÃO, L. F. *et al.* (org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

SCHWARTZMAN, S. **Pesquisa acadêmica, pesquisa básica e pesquisa aplicada em duas comunidades científica.** 1979. Disponível em: <http://www.schwartzman.org/simon/1979academ>. Acesso em: 22 set. 2022.

SEER. Serviço Eletrônico de Editoração de Revista. Disponível em: <http://www.sitistorico.ibict.br/sistema-eletronico-de-editor>. Acesso em: 15 jun. 2022.

SILVA, F. F.; SILVA, R. T. P. História da educação superior: das primeiras universidades à criação da Universidade Federal de Rondônia. **Humanidades & Inovação**, Palmas, v. 5, n. 8, p. 39-47, 2018.

SILVA, L. C.; SANTAREM SEGUNDO, J. E. Componentes de representação da informação em ambientes de informação digital: estudo do sistema de organização do software Tainacan. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, 20., 2019, Florianópolis. **A CIÊNCIA DA INFORMAÇÃO E A ERA DA CIÊNCIA DE DADOS**, 2019.

SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO SUPERIOR (SINAES). Instrumento de Avaliação de Cursos de Graduação: presencial e a distância. Brasília: Inep, 2017.

SILVEIRA, Fabrício José Nascimento da. Biblioteca pública e identidade: percepções intersubjetivas enraizadas em torno da Luiz de Bessa. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, número especial, p.128-150, out./dez. 2014. Disponível

em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/2283/1474>. Acesso em: 17 ago. 2015.

SOARES, M. S. A. O sistema de apoio às atividades das instituições de ensino superior. *In*: SOARES, M. S. A, (org.). **Educação superior no Brasil**. Brasília: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, 2002.

SOUSA, B. P.; FUJITA, M. S. L. Do catálogo impresso ao on-line: algumas considerações e desafios para os bibliotecários. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 17, n. 1, p. 59-75, jan./jun. 2012.

SOUSA, J. A. G. Repositórios institucionais brasileiros e a transição para a próxima geração. **Ci. Inf**, Brasília, v. 48, n. 3 (Supl.), p. 514-515, set./dez. 2019. Disponível em: <https://scholar.archive.org/work/tlm5z323brbxbjzoo43mkcohm/access/wayback//revista.ibict.br/ciinf/article/download/4990/4475>. Acesso em: 25 jun. 2023.

SOUZA, E. D. **A ciência da informação: fundamentos epistêmico-discursivos do campo científico e do objeto de estudo**. Maceió: EDUFAL, 2015.

STALLIVIERI, L. Prefácio. *In*: PEREIRA, E. M. A.; HEINZLE, M. R. S. (orgs.). **Internacionalização na educação superior: políticas, integração e mobilidade acadêmica**. Blumenau: Edifurb, 2015.

TARTATOTTI, R. C. D. **Avaliação do processo de indexação de assuntos em repositórios institucionais pela abordagem da recuperação da informação**. 2019. 369 f. Tese (Doutorado em Ciência da informação) – Faculdade de Filosofia e Ciências. Programa de Pós-Graduação em Ciência da informação, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2019.

TEIXEIRA, E. B. A análise de dados na pesquisa científica: importância e desafios em estudos organizacionais. **Desenvolvimento em questão**, v. 1, n. 2, jul./dez. 2003.

UNIVERSIDADE FEDERAL DE ALAGOAS. Conselho Universitário. **Resolução nº 45/2016 – CONSUNI/UFAL, de 07 de novembro de 2016**. Regulamenta a Política de Informação do Repositório Institucional da UFAL (RI/UFAL). Maceió: CONSUNI, 2016. Disponível em: <https://ufal.br/resolucoes/2016/resolucao-no-45-2016-de-07-11-2016>. Acesso em: 20 dez. 2020.

VILLALOBOS, A. P. O.; GOMES, F. A. Padronização de metadados na representação da informação em repositórios institucionais de universidades federais brasileiras. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/187093>. Acesso em: 17 abr. 2023.

WEITZEL, S. R. O mapeamento dos repositórios institucionais brasileiros: perfil e desafios. **Encontros Bibli**, Belo Horizonte, v. 24, n. 54, p. 105-123, jan./abr. 2019. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2019v24n54p105/38092>. Acesso em: 14 jan. 2022.

XAVIER, R. F. Repositórios de acesso aberto brasileiros: características, crescimento e possibilidades futuras. **BIREDIAL-ISTEC**, p. 1-13, São Paulo, 2019. Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/27782>. Acesso em: 15 maio 2023.

XIA, J. *et al.* A review of open access self-archiving mandate policies. **Libraries and the Academy**, v. 12, n. 1, p. 85-102, jan. 2012.