

UFRRJ

INSTITUTO MULTIDISCIPLINAR

**PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR EM
HUMANIDADES DIGITAIS**

DISSERTAÇÃO

**O Aporte das Humanidades Digitais na Gestão da BDTD: Rumo à
Sistematização no Processamento das Teses e Dissertações Eletrônicas da
Universidade Federal Rural do Rio de Janeiro**

ELIANE MARICÁ SOARES

2023



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO MULTIDISCIPLINAR
PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR EM
HUMANIDADES DIGITAIS**

**O APORTE DAS HUMANIDADES DIGITAIS NA GESTÃO DA BDTD:
RUMO À SISTEMATIZAÇÃO NO PROCESSAMENTO DAS TESES E
DISSERTAÇÕES ELETRÔNICAS DA UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO**

ELIANE MARICÁ SOARES

Sob a Orientação do Professor
Alexandre Fortes

Dissertação submetida como requisito parcial para a obtenção do grau de **Mestre em Humanidades Digitais**, no Programa de Pós-Graduação em Humanidades Digitais. Área de Concentração em Análise Qualitativa e Quantitativa de Dinâmicas Sociais.

Seropédica, RJ
Novembro de 2023

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

SS676a Soares, Eliane Maricá, 1976-
O aporte das Humanidades Digitais na gestão da
BDTD: rumo à sistematização no processamento das teses
e dissertações eletrônicas da Universidade Federal
Rural do Rio de Janeiro / Eliane Maricá Soares. -
Seropédica, 2023.
120 f.: il.

Orientador: Alexandre Fortes.
Dissertação (Mestrado). -- Universidade Federal Rural
do Rio de Janeiro, Programa de Pós-graduação
Interdisciplinar em Humanidades Digitais, 2023.

1. Bibliotecas digitais. 2. Teses. 3. Organização
da informação. 4. Humanidades Digitais. 5. Ciência da
Informação. I. Fortes, Alexandre, 1966-, orient. II
Universidade Federal Rural do Rio de Janeiro.
Programa de Pós-graduação Interdisciplinar em
Humanidades Digitais III. Título.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO MULTIDISCIPLINAR
PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR EM
HUMANIDADES DIGITAIS**

ELIANE MARICÁ SOARES

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre em Humanidades Digitais**, no Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais.

DISSERTAÇÃO APROVADA EM: 13/11/2023

Emitido em 13/11/2023

ATA Nº 5766/2023 - PPGIHD (11.39.00.16)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 06/12/2023 10:23)

ALEXANDRE FORTES
PROFESSOR DO MAGISTERIO SUPERIOR
DeptH/IM (12.28.01.00.00.88)
Matrícula: ###084#6

(Assinado digitalmente em 06/12/2023 13:52)

SERGIO MANUEL SERRA DA CRUZ
PROFESSOR DO MAGISTERIO SUPERIOR
DCOMP (11.39.97)
Matrícula: ###24#6

(Assinado digitalmente em 06/12/2023 14:11)

ELIANE MARICA SOARES
DISCENTE
Matrícula: 2021#####7

(Assinado digitalmente em 06/12/2023 11:44)

ROSANA PORTUGAL TAVARES DE MORAES
ASSINANTE EXTERNO
CPF: ###.###.797-##

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: **5766**, ano: **2023**, tipo: **ATA**, data de emissão: **06/12/2023** e o código de verificação: **4efd690207**

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me dar forças para não desistir, mesmo diante dos diversos desafios que enfrentei ao longo dessa jornada.

À Universidade Federal Rural do Rio de Janeiro, meu sincero agradecimento pelo incentivo em abrir vagas destinadas aos servidores da instituição em seus cursos de pós-graduação. Essa oportunidade foi fundamental para o meu crescimento acadêmico e profissional.

Aos membros da banca, os professores Sérgio Manuel Serra da Cruz e Rosana Portugal Tavares de Moraes, sou grata pela atenção dedicada à análise da minha proposta de qualificação. Suas valiosas contribuições foram essenciais para a conclusão desta pesquisa.

Ao professor e orientador Alexandre Fortes, agradeço pela confiança em me permitir escolher o tema da dissertação.

Aos meus amigos Ana Paula Lima dos Santos e André Luiz Martins Costa, expresso minha gratidão por estarem ao meu lado e por me proporcionarem confiança para seguir adiante nesta jornada acadêmica.

À minha família, especialmente a minha mãe Cândida Maricá Soares, minha irmã Teresa Cristina da Silva e minhas sobrinhas Maria Paula da Silva, Mariana Cristina da Silva, Anna Beatriz da Silva, Anna Júlia da Silva e Brenda Costa da Silva, meu profundo agradecimento pela compreensão e incentivo durante todo o processo. Vocês são minha inspiração.

Meu agradecimento a toda a equipe da Biblioteca Central da UFRRJ, em particular à Leticia Schettini, Gracielle Rosa Faria de Oliveira Penedo, Josy Soares da Silva Mendes de Moraes de Souza, Alessandra Castro Fiorini Bessa, Sandra Regina Pereira, Marisa Gonçalves Bahia e Heloisa Assis de Almeida (*in memoriam*), pelo apoio e incentivo.

Aos colegas do curso do PPGIHD, agradeço pelas trocas de experiências ao longo do curso. Vocês tornaram essa trajetória acadêmica ainda mais enriquecedora.

Aos meus amiguinhos de quatro patas, Pingo e Magali (*in memoriam*), por fazerem companhia e trazerem alegria aos momentos de estudo no meu quarto.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Este tem sido um de meus mantras – foco e simplicidade. O simples pode ser mais difícil do que o complexo: é preciso trabalhar duro para limpar seus pensamentos de forma a torná-los simples. Mas no final, vale a pena, porque quando chegamos lá, podemos mover montanhas (Steve Jobs).

RESUMO

SOARES, Eliane Maricá. **O aporte das Humanidades Digitais na gestão da BDTD: rumo à sistematização no processamento das teses e dissertações eletrônicas da Universidade Federal Rural do Rio de Janeiro.** 2023. 120 f. Dissertação (Mestrado em Humanidades Digitais) – Instituto Multidisciplinar de Nova Iguaçu, Universidade Federal Rural do Rio de Janeiro, Nova Iguaçu, RJ, 2023.

O avanço tecnológico das últimas décadas tem permitido que a informação digital, disponível no ambiente virtual, se torne um meio eficiente de transmitir conhecimento registrado. Nesse contexto, os repositórios digitais (RD) despontam como uma base de dados virtual que armazena, preserva, organiza e proporciona acesso ao acervo eletrônico, desempenhando também um papel importante ao fornecer visibilidade à produção documental das instituições. Em 2005, a Universidade Federal Rural do Rio de Janeiro (UFRRJ) aderiu à Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), um repositório digital exclusivo desses documentos. Com o crescimento do volume desses trabalhos acadêmicos da UFRRJ, surge uma questão preocupante: como gerenciar, de forma eficiente, esse valioso acervo eletrônico da instituição no seu repositório BDTD, com o intuito de aumentar a visibilidade de seu conteúdo e, conseqüentemente, promover um maior uso pela comunidade acadêmica e científica? O objetivo geral desse estudo foi apresentar propostas para melhorar o tratamento das teses e dissertações digitais da UFRRJ, buscando soluções para diversas questões relacionadas ao processamento dos itens na BDTD. A dissertação mostra que a Ciência da Informação não enfrenta o presente desafio de maneira isolada, ressaltando a importância do aporte das Humanidades Digitais, que está presente na organização da informação tanto de maneira explícita quanto implícita. A primeira etapa da pesquisa consistiu em uma revisão da literatura relevante para o problema de pesquisa, com destaque para os tópicos: gestão de repositórios digitais, processamento das teses e dissertações eletrônicas, organização da informação, Leis de Ranganathan, Princípios FAIR e o estudo da inter-relação entre as Humanidades Digitais e a Ciência da Informação. Na segunda fase, baseada nos conhecimentos adquiridos das fontes selecionadas, empreendeu-se uma análise minuciosa na BDTD, com o intuito de aplicar teorias, procedimentos e técnicas que permitissem a sistematização do tratamento informacional das teses e dissertações nesse sistema de informação. Assim, empregou-se uma abordagem de pesquisa-ação, englobando observação e intervenção ativa no objeto de estudo. Concluiu-se que o repositório digital necessita de uma abordagem mais ampla e interdisciplinar para garantir uma melhor experiência para os usuários em sua busca por conhecimento.

Palavras-chave: repositórios digitais; teses e dissertações eletrônicas; Organização do Conhecimento; Humanidades Digitais; Leis de Ranganathan.

ABSTRACT

SOARES, Eliane Maricá. **The contribution of Digital Humanities to the management of the BDTD: towards systematization in the processing of electronic theses and dissertations at the Federal Rural University of Rio de Janeiro.** 2023. 120 p. Dissertation (Master in Digital Humanities) – Instituto Multidisciplinar de Nova Iguaçu, Universidade Federal Rural do Rio de Janeiro, Nova Iguaçu, RJ, 2023.

Technological advances in recent decades have allowed digital information, available in the virtual environment, to become an efficient means of transmitting recorded knowledge. In this context, digital repositories (RD) emerge as a virtual database that stores, preserves, organizes and provides access to the electronic collection, also playing an important role in providing visibility to the institutions' documentary production. In 2005, the Federal Rural University of Rio de Janeiro (UFRRJ) joined the Brazilian Digital Library of Theses and Dissertations (BDTD), an exclusive digital repository of these documents. With the growth in the volume of these academic works at UFRRJ, a worrying question arises: how to efficiently manage this valuable electronic collection of the institution in its BDTD repository, with the aim of increasing the visibility of its content and, consequently, promoting a greater use by the academic and scientific community? The general objective of this study was to present proposals to improve the treatment of digital theses and dissertations at UFRRJ, seeking solutions to various issues related to the processing of items in the BDTD. The dissertation shows that Information Science does not face this challenge in isolation, highlighting the importance of the contribution of Digital Humanities, which is present in the organization of information both explicitly and implicitly. The first stage of the research consisted of a review of literature relevant to the research problem, with emphasis on the topics: management of digital repositories, processing of electronic theses and dissertations, organization of information, Ranganathan's Laws, FAIR Principles and the study of interrelationship between Digital Humanities and Information Science. In the second phase, based on the knowledge acquired from the selected sources, a thorough analysis was undertaken in the BDTD, with the aim of applying theories, procedures and techniques that would allow the systematization of the informational treatment of theses and dissertations in this information system. Thus, an action research approach was used, encompassing observation and active intervention in the object of study. It was concluded that the digital repository needs a broader and interdisciplinary approach to guarantee a better experience for users in their search for knowledge.

Keywords: digital repositories; electronic theses and dissertations; Organization of knowledge; Digital Humanities; Ranganathan's Laws.

LISTA DE FIGURAS

| | |
|--|-----------|
| Figura 1 - Exemplo simples de taxonomia | 35 |
| Figura 2 – BDTD/UFRRJ | 65 |
| Figura 3 – Relatório do Pergamum..... | 68 |
| Figura 4 – Página da PPGEA da UFRRJ..... | 69 |
| Figura 5 – Estatística da Dissertação Recuperada (BDTD/UFRRJ) | 70 |
| Figura 6 – Inclusão de novos dados (BDTD/UFRRJ) | 71 |
| Figura 7 – Busca pela referência bibliográfica no Google..... | 72 |
| Figura 8 – Nuvem de palavras da BDTD/UFRRJ, antes da correção (reprodução)..... | 72 |
| Figura 9 – Navegar por assunto (BDTD/UFRRJ) | 73 |
| Figura 10 – Navegar por orientador (BDTD/UFRRJ) | 74 |
| Figura 11 – Currículo Lattes | 75 |
| Figura 12 – Pesquisa na BDTD (dados com acento) | 76 |
| Figura 13 – Pesquisa na BDTD (dados sem acento)..... | 77 |
| Figura 14 – Busca pelo ORCID no Google..... | 78 |
| Figura 15 – <i>Links</i> do CNPq na BDTD/UFRRJ..... | 79 |

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1 - Características da CI | 26 |
| Quadro 2 - Principais pontos em comum entre HD e CI | 29 |
| Quadro 3 - Resumo das principais características dos SOC | 32 |
| Quadro 4 - Tipos de repositórios digitais | 42 |
| Quadro 5 - Acesso ao RD | 44 |
| Quadro 6 - Esquema de metadados do DC | 48 |
| Quadro 7 - Princípios FAIR | 58 |
| Quadro 8 - BDTD/UFRRJ – Ações x Resultados | 66 |
| Quadro 9 – Questões não resolvidas X Soluções propostas (BDTD/UFRRJ) | 76 |

LISTA DE GRÁFICOS

| | |
|--|-----------|
| Gráfico 1 – Mineração de texto no Sobek | 39 |
|--|-----------|

LISTA DE SIGLAS E ABREVIATURAS

| | |
|---------|---|
| AACR2 | Código de Catalogação Anglo-Americano – 2ª edição |
| ABNT | Associação Brasileira de Normas Técnicas |
| AIFF | Audio Interchange File Format |
| ANSI | American National Standards Institute |
| BC | Biblioteca Central |
| Bepress | Berkeley Electronic Press |
| BIREME | Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde |
| BDTD | Biblioteca Digital Brasileira de Teses e Dissertações |
| BRAPCI | Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação |
| CAPES | Coordenação de Aperfeiçoamento de Pessoal de Nível Superior |
| CC | Ciência da Computação |
| CD-ROM | Compact Disc Read Only Memory |
| CDD | Classificação Decimal de Dewey |
| CI | Ciência da Informação |
| CPDA | Programa de Pós-Graduação de Ciências Sociais em Desenvolvimento, Agricultura e Sociedade |
| CPF | Cadastro de Pessoas Físicas |
| CNPq | Conselho Nacional de Desenvolvimento Científico e Tecnológico |
| DC | Dublin Core |
| DOI | Identificador de Objeto Digital |
| DSI | Disseminação Seletiva da Informação |
| EAD | Encoded Archival Description |
| E-LIS | EPrints in Library and Information Science |
| ETD | Electronic Theses and Dissertations |
| FAIR | Findable, Accessible, Interoperable e Reusable |
| Fedora | Flexible Extensible Digital Object and Repository Architecture |
| Fiocruz | Fundação Oswaldo Cruz |
| Funcate | Fundação de Ciência, Aplicações e Tecnologia Espaciais |
| IBICT | Instituto Brasileiro de Informação em Ciência e Tecnologia |
| IBM | International Business Machines Corporation |
| IMNI | Instituto Multidisciplinar de Nova Iguaçu |
| ITR | Instituto Três Rios |
| HD | Humanidades Digitais |
| HP | Hewlett-Packard |
| IFES | Instituições Federais de Ensino Superior |
| ISBN | International Standard Book Number |
| ISO | International Organization for Standardization |
| ISSN | International Standard Serial Number |
| JPEG | Joint Photographic Experts Group |
| KOS | Knowledge Organization System |
| LOM | Learning Object Metadata |
| MARC | Machine Readable Cataloging |
| MIT | Massachusetts Institute of Technology |
| MODS | Metadata Object Description Schema |
| MPEG | Moving Picture Experts Group |

| | |
|----------|---|
| MP4 | MPEG Layer 4 |
| MTD3-BR | Padrão Brasileiro de Metadados da Biblioteca Digital Brasileira de Teses e Dissertações |
| NDLTD | Networked Digital Library of Theses and Dissertations |
| NISO | National Information Standard Organization |
| OAI-PMH | Open Archives Initiative Protocol for Metadata Harvesting |
| ORCID | Open Researcher and Contributor ID |
| PDF | Formato Portátil de Documento |
| PID | Persistent Identifier |
| PLN | Processamento de Linguagem Natural |
| PPGEA | Programa de Pós-Graduação em Educação Agrícola |
| PPGIHD | Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais |
| PROPPG | Pró-Reitoria de Pesquisa e Pós-Graduação |
| RD | Repositório Digital |
| RI | Repositório Institucional |
| RIMA | Repositório Institucional do Instituto Multidisciplinar |
| SOC | Sistemas de Organização do Conhecimento |
| SPT | Seção de Processamentos Técnicos |
| TEDE | Sistema de Publicação Eletrônica de Teses e Dissertações |
| THATCamp | The Humanities and Technology Camp |
| TIC | Tecnologias de Informação e Comunicação |
| TIFF | Tagged Image File Format |
| UFRGS | Universidade Federal do Rio Grande do Sul |
| UFRRJ | Universidade Federal Rural do Rio de Janeiro |
| UFU | Universidade Federal de Uberlândia |
| UnB | Universidade de Brasília |
| UTF-8 | 8-bit Unicode Transformation Format |
| VRA | Visual Resources Association |

SUMÁRIO

| | |
|--|------------|
| 1 INTRODUÇÃO | 15 |
| 2 METODOLOGIA..... | 20 |
| 3 HUMANIDADES DIGITAIS E CIÊNCIA DA INFORMAÇÃO: UMA DISCUSSÃO DE SEUS PONTOS DE CONTATO | 23 |
| 4 ABORDAGENS DE TRATAMENTO INFORMACIONAL PARA COLEÇÕES DE TESES E DISSERTAÇÕES ELETRÔNICAS: UM ESTUDO DAS PRINCIPAIS ESTRATÉGIAS..... | 30 |
| 4.1 Sistemas de Organização do Conhecimento (SOC) aplicáveis ao tratamento de teses e dissertações: classificação, tesouro, ontologia e taxonomia | 30 |
| 4.2 A mineração de texto como recurso auxiliar na indexação das teses e dissertações .. | 36 |
| 5 REPOSITÓRIOS DIGITAIS DE CONHECIMENTO | 41 |
| 5.1 O software Dspace | 45 |
| 5.2 Metadados e a interoperabilidade..... | 47 |
| 5.3 Princípios para os repositórios digitais..... | 50 |
| 5.3.1 As Cinco Leis de Ranganathan..... | 50 |
| 5.3.2 Princípios FAIR..... | 57 |
| 6 BDTD DA UFRRJ: PESQUISA-AÇÃO..... | 62 |
| 6.1 Resultados e discussão..... | 66 |
| 6.2 O repositório em transição: propostas para o futuro..... | 79 |
| 7 CONSIDERAÇÕES FINAIS..... | 82 |
| REFERÊNCIAS | 86 |
| ANEXO A – Manual da BDTD da UFRRJ..... | 101 |
| ANEXO B – MTD3-BR (Metadados para Teses e Dissertações Digitais Brasileiras).119 | |

1 INTRODUÇÃO

A humanidade, ao longo da sua história, percebeu que era imprescindível guardar os registros de informação para o seu acesso posterior. Em vista disso, construiu bibliotecas, arquivos e museus com a importante missão de preservar acervos de conhecimentos.

Com a invenção da imprensa de Gutenberg, no século XV, foi possível a reprodução em série do conhecimento registrado, o que desencadeou o crescimento do volume de publicações editadas no mundo. No século passado, durante a Segunda Guerra Mundial, o investimento governamental em pesquisa e desenvolvimento aumentou, contribuindo para a promoção de avanços científicos e tecnológicos, que impulsionaram o processo de comunicação científica formal, resultando na multiplicação acelerada das publicações (Weitzel, 2002). O fenômeno da explosão bibliográfica proporcionou um grande desafio para a organização e a recuperação dos documentos impressos. No entanto, com o progresso tecnológico das últimas décadas, o formato físico não é mais o principal meio de disseminação da informação. O conteúdo digital disponível no ambiente virtual tornou-se um expressivo meio de transmissão do conhecimento.

No caso de um livro físico, o exemplar pode ser preservado em uma estante da biblioteca. E quanto à integridade do documento eletrônico para consulta futura? Como lidar com essa grande quantidade de documentos eletrônicos? Os repositórios digitais (RD) despontam, portanto, como uma importante ferramenta que armazena, preserva, organiza e proporciona acesso aberto ao acervo eletrônico, desempenhando também um papel fundamental ao fornecer visibilidade à produção documental das instituições.

No Brasil, as ações do Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (BIREME) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) se destacam como as principais iniciativas para promover o acesso aberto ao conhecimento no país (Weitzel, 2019). Um marco importante do IBICT foi a criação da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), que tem como objetivo registrar e disseminar as teses e dissertações do ensino superior. Cada instituição fica com a sua base local para depositar a sua produção, a plataforma do Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE), e a BDTD centraliza as informações na sua biblioteca (Kuramoto, 2010). Funciona como um provedor de serviço que “colhe” os metadados dos repositórios das instituições, com o objetivo de oferecer um ponto único de busca por informação (Rodrigues *et al.*, 2013).

Em 2005, a Universidade Federal Rural do Rio de Janeiro (UFRRJ) aderiu à BDTD e passou a depositar na sua plataforma os arquivos em Formato Portátil de Documento (PDF) das

teses e dissertações. No presente momento, o sistema da Rural disponibiliza mais de seis mil teses e dissertações em sua base, dos cursos de pós-graduação de sua instituição. A Seção de Processamentos Técnicos (SPT), da Biblioteca Central (BC) da UFRRJ, é a responsável pela gestão desse sistema de informação.

A produção acadêmica em formato eletrônico da UFRRJ, em especial as teses e dissertações, aumenta cada vez mais ao longo dos anos, conforme ampliação do número de pós-graduandos da instituição. Diante desse crescimento, uma questão provoca preocupação: como gerenciar de forma eficiente esse valioso acervo eletrônico da instituição na BDTD, com o intuito de aumentar a visibilidade de seu conteúdo e, conseqüentemente, promover um maior uso pela comunidade acadêmica e científica? Seguem os principais desafios observados para alcançar a gestão ideal do repositório:

- Incorrecções e falta de padronização nos registros de objetos digitais: falhas no preenchimento dos metadados e não estabelecimento de normas apropriadas;
- Ocorrência de múltiplos registros para um mesmo orientador, palavra-chave e outros dados: mesmo pequenas variações nos cadastros, como uma grafia idêntica, criam novas entradas;
- A representação temática das teses e dissertações eletrônicas apresenta inadequações: a única forma de representar esses documentos no repositório é por meio de palavras-chave, porém essa abordagem nem sempre funciona como um meio eficaz de indexação de conteúdo;
- Documentos pendentes, que não estão disponíveis para consulta na plataforma: considerando que o repositório visa canalizar as teses e dissertações da instituição, é necessário encontrar uma forma de recuperar esses objetos digitais indisponíveis na BDTD.

O objetivo geral desta pesquisa é, portanto, aprimorar o tratamento da coleção de teses e dissertações digitais da UFRRJ, por meio da adoção de uma abordagem sistemática para o seu processamento na BDTD, que depende de várias ações para garantir uma melhor organização, acessibilidade e eficiência na disponibilização desses documentos eletrônicos nesse ambiente virtual. O estudo reconhece que a Ciência da Informação (CI) não enfrenta sozinha a complexa tarefa de organizar o conhecimento registrado em ambientes digitais, sendo necessário incorporar soluções fundamentadas nos domínios das Humanidades Digitais (HD), a fim de lidar efetivamente com os desafios inerentes à gestão desses documentos no RD. Dessa forma, como objetivos específicos, almeja:

- Identificar os pontos de convergência entre os estudos da CI e das HD, ressaltando, assim, a estreita relação entre as duas disciplinas. Isso representa o enfoque teórico central da dissertação;
- Examinar as melhores práticas destinadas ao tratamento de documentos técnico-científicos em repositórios digitais, visando à identificação das abordagens mais apropriadas e viáveis para o processamento eficiente desses objetos digitais em ambiente virtual;
- Sistematizar os procedimentos em forma de uma proposta a ser aplicada na BDTD/UFRRJ. Após explorar a base teórica, a dissertação detalhará a implementação de alguns processos no repositório, além de apresentar sugestões para o futuro;
- Estudar os vários aspectos do RD, com o propósito de compreender como essa plataforma tecnológica contribui para a disseminação do conhecimento.

Na sequência, são apresentados três pontos fundamentais que justificam o estudo proposto para esta pesquisa:

- A principal justificativa para o tema está associada aos benefícios práticos para o atendimento às demandas da comunidade acadêmica da UFRRJ, como também aos usuários externos;
- A preservação digital das teses e dissertações da UFRRJ;
- O presente estudo desempenha um papel importante, atuando tanto como referência teórica para pesquisas futuras quanto como uma valiosa fonte de consulta para profissionais da informação. Especificamente, esta pesquisa é de grande relevância para bibliotecários responsáveis por repositórios digitais em suas instituições, assim como para especialistas no campo da tecnologia da informação. A escassez de materiais de apoio abordando a gestão de teses e dissertações digitais nesse ambiente virtual torna esse estudo significativo.

No que diz respeito à estrutura da dissertação, o presente estudo está organizado em sete seções. A primeira seção é a introdução, que tem como objetivo apresentar, de forma sucinta, os principais pontos abordados na pesquisa. Já a segunda, aborda a metodologia utilizada, descrevendo de forma detalhada as etapas necessárias para alcançar os objetivos propostos.

A terceira seção trata da inter-relação entre as HD e a CI. São abordados tanto o histórico quanto a definição desses dois campos, com enfoque na exploração dos pontos convergentes entre ambos. A escolha desse tema encontra justificativa no fato de que o RD é considerado um elemento situado na fronteira desses dois campos de estudo.

A quarta seção se dedica à análise das diversas abordagens para o tratamento do conhecimento no RD, com enfoque na coleção de teses e dissertações. O processamento informacional em um Sistema de Recuperação da Informação (SRI), como no caso dos repositórios, compreende duas ações fundamentais: a representação descritiva e a representação temática. Enquanto a representação descritiva foca nos aspectos físicos do documento, como autor, título, editora e dados de publicação, a representação temática aborda o conteúdo intelectual do material, incluindo seus assuntos e conceitos. Essa seção concentra-se na análise temática dos documentos, examinando tanto os Sistemas de Organização do Conhecimento (SOC) quanto a viabilidade da aplicação da mineração de texto para aprimorar a indexação dos objetos digitais. Isso destaca a importância de não se limitar apenas à tecnologia digital no ambiente virtual, mas também considerar a estrutura semântica necessária para viabilizar a recuperação da informação na BDTD.

A quinta seção aborda o RD em profundidade, analisando aspectos como o seu funcionamento, o *software* utilizado, a estrutura dos metadados e a interoperabilidade. Além disso, explora princípios orientadores essenciais para garantir uma gestão eficaz desse recurso. As Leis de Ranganathan oferecem orientações valiosas para sistemas de informação, inclusive os digitais. Apesar de terem sido concebidas há décadas, suas teorias continuam a ser aplicáveis na organização do conhecimento em ambientes digitais, como repositórios. Portanto, as suas "Cinco Leis da Biblioteconomia" direcionam decisões sobre a gestão eficiente dessa ferramenta. Mas, é importante adaptar suas ideias para aprimorar a eficiência do sistema de informação, garantindo que as necessidades dos usuários sejam plenamente atendidas. Assim, a pesquisa mostra a interseção entre conhecimentos das Ciências Humanas e Sociais e a tecnologia digital. De forma paralela, o estudo também discorre sobre os Princípios FAIR, que desempenham uma função de grande relevância na gestão de RD. Embora, originalmente, tenham sido desenvolvidas apenas para o gerenciamento de dados de pesquisa, a dissertação demonstra que tais diretrizes podem ser aplicadas com sucesso a qualquer repositório de informações.

Na penúltima seção, são destacados os resultados da pesquisa, como a retificação da nuvem de palavras e a normalização das autorias no cadastro, incluindo os nomes dos autores, orientadores e demais membros da banca. O estudo apresenta os progressos em andamento, que se tornaram possíveis após a elaboração do Manual da Biblioteca Digital de Teses e Dissertações pela equipe técnica da SPT. Também é abordada a inclusão de novos dados, como o Identificador de Pesquisador e Colaborador Aberto (Open Researcher and Contributor ID –

ORCID), que funciona como a identidade do pesquisador. Por fim, a pesquisa aponta sugestões de melhorias para a BDTD, as quais podem ser incorporadas no Repositório Institucional de Múltiplos Acervos (RIMA) da UFRRJ no futuro, considerando a iminente fusão do BDTD para o Repositório Institucional (RI). Consta-se que algumas ações se respaldam no saber da CI, enquanto outras emanam das HD, demonstrando, assim, o alinhamento das teorias, procedimentos e técnicas que convergem para a resolução dos problemas no tratamento informacional.

Para concluir, a última seção sintetiza os principais pontos abordados ao longo do estudo, com o posicionamento da autora. Adicionalmente, como anexos relevantes, o Manual da BDTD/UFRRJ e o Padrão Brasileiro de Metadados da Biblioteca Digital Brasileira de Teses e Dissertações, o MTD3-BR.

A pesquisa apresenta a visão de uma bibliotecária para as HD, que enxerga esse campo como uma ciência em constante construção. Uma profissional da informação formada no início do século XXI, que vivencia as transformações de sua área de atuação na era digital.

Diante do apresentado, o estudo se enquadra no Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais (PPGIHD), da UFRRJ, e se insere na área de concentração Análise Qualitativa e Quantitativa de Dinâmicas Sociais, na linha de pesquisa Mineração de Dados Digitais.

2 METODOLOGIA

De maneira sucinta, a abordagem metodológica representa o percurso necessário para alcançar os objetivos da pesquisa, englobando os passos fundamentais para sua concretização. Segundo Prodanov e Freitas (2013), a metodologia envolve o uso de procedimentos e técnicas empregadas na construção do conhecimento, com o propósito de demonstrar sua utilidade e validade em diferentes contextos sociais. A fim de ser considerado um discurso científico, é importante que ele esteja fundamentado em bases sólidas, apresente uma estrutura lógica e sistemática, bem como mantenha uma coerência argumentativa.

Partindo da concepção de que método é um procedimento ou caminho para alcançar determinado fim e que a finalidade da ciência é a busca do conhecimento, podemos dizer que o método científico é um conjunto de procedimentos adotados com o propósito de atingir o conhecimento (Prodanov; Freitas, 2013, p. 25).

Dado o exposto, a presente dissertação seguirá uma abordagem qualitativa empírica, sendo dividida em duas fases. A primeira fase consistirá em uma revisão da literatura relevante para o problema de pesquisa. O levantamento bibliográfico é, portanto, a etapa preliminar do projeto de investigação, que proporciona o aprofundamento do conhecimento sobre o tema proposto para o estudo. Na segunda etapa, com base nos conhecimentos obtidos a partir da literatura selecionada, foi conduzida uma análise detalhada do objeto de estudo, com o objetivo de identificar os principais desafios relacionados ao processamento de teses e dissertações na BDTD/UFRRJ. Após essa análise, foram propostas soluções para resolver os problemas identificados. Dessa forma, foram adotados procedimentos característicos de uma pesquisa-ação, envolvendo a observação e a intervenção ativa no objeto de estudo.

Para atingir o primeiro objetivo específico desta dissertação, que explora a conexão entre CI e HD, foi conduzida uma ampla revisão da literatura nos dois domínios, selecionando, principalmente, os textos que discutem a inter-relação entre essas disciplinas. Esses elementos de convergência foram, então, sintetizados e incorporados ao texto.

Em relação ao segundo objetivo, que diz respeito à organização de documentos técnicos científicos em repositórios digitais, além da revisão bibliográfica correspondente sobre o tratamento informacional voltado a recuperação da informação de forma eficaz, foram realizados exames das estratégias adotadas para teses e dissertações eletrônicas em alguns repositórios digitais de destaque no Brasil, como o Arca, da Fundação Oswaldo Cruz (Fiocruz). Essa avaliação permitiu identificar as melhores práticas a serem aplicadas na BDTD/UFRRJ.

A realização do penúltimo objetivo específico, concentrado na sistematização dos procedimentos em uma proposta aplicável à BDTD/UFRRJ, tornou-se possível através da combinação dos conhecimentos obtidos tanto na revisão da literatura quanto na experiência dos servidores responsáveis pelo tratamento das teses e dissertações eletrônicas na BDTD/UFRRJ. Ou seja, foi necessário incorporar uma valiosa colaboração por parte desses indivíduos engajados no processo. Uma das propostas já implementadas foi a criação do manual, que desempenha o papel de guia para o registro dos documentos digitais na plataforma.

No último objetivo específico, que aborda o estudo sobre o RD, foi necessária a exploração da literatura para compreender a sua definição, os tipos existentes e o seu funcionamento. Além disso, durante a revisão bibliográfica, buscou-se também identificar os princípios que podem nortear a gestão dessa ferramenta, incluindo as Leis de Ranganathan e os Princípios FAIR.

No processo de levantamento bibliográfico, as pesquisas foram conduzidas em diversas bases de dados e periódicos pertinentes ao estudo. As fontes consultadas englobaram:

- Arca (Repositório Institucional da Fiocruz);
- Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI);
- BDTD;
- Google Acadêmico;
- Revista Brasileira de Computação Aplicada;
- Revista de Humanidades Digitais;
- Science Direct.

Ao conduzir a pesquisa, utilizaram-se as expressões de busca em português (como também suas traduções em inglês), tais como: “gestão de repositório digital”, “gestão da BDTD”, “gestão de teses e dissertações eletrônicas”, “Humanidades Digitais”, “Humanidades Digitais e Ciência da Informação”, “Humanidades Digitais e a Organização da Informação”, “Leis de Ranganathan”, “Princípios FAIR” e “mineração de texto”. Na seleção de documentos, foi atribuída prioridade às teses, dissertações e artigos científicos. A ferramenta digital Zotero¹ foi importante na organização das referências bibliográficas selecionadas para a pesquisa.

De acordo com Prodanov e Freitas (2013), a pesquisa-ação não se limita a uma simples coleta de dados ou relatórios a serem arquivados. Através desse método, os pesquisadores têm o objetivo de desempenhar um papel ativo na realidade dos fatos observados. Essa abordagem

¹ Disponível em: <https://www.zotero.org/>. Acesso em: 9 set. 2023.

oferece subsídios para organizar a pesquisa de forma convencional, envolvendo observação, processamento de dados, experimentação, entre outros procedimentos relevantes.

A pesquisa-ação é uma abordagem que visa não apenas entender uma situação ou problema, mas também criar mudanças práticas e melhorias por meio da colaboração entre pesquisadores e participantes. Ela envolve um ciclo contínuo de planejamento, ação, observação e reflexão, com o objetivo de resolver problemas específicos em um contexto real. Nesse tipo de pesquisa, os participantes atuam como copesquisadores, contribuindo com suas perspectivas e conhecimentos para a definição dos problemas, o planejamento das ações e a avaliação dos resultados. Isso promove a participação ativa das pessoas afetadas pelo problema e leva a soluções mais contextualmente relevantes e sustentáveis. Esse método é frequentemente utilizado em áreas como educação, saúde, desenvolvimento comunitário e gestão organizacional, onde o foco está em melhorar práticas, processos ou condições específicas por meio da colaboração e da ação direta. Portanto, é uma abordagem dinâmica e interativa que busca não apenas compreender a realidade, mas também influenciar positivamente essa realidade por meio de intervenções planejadas e monitoramento contínuo (Tripp, 2005).

Assim, a abordagem da pesquisa-ação revela-se como uma alternativa apropriada para conduzir as propostas de aprimoramento do tratamento das teses e dissertações eletrônicas da UFRRJ, em seu RD. Isso leva a uma finalidade prática aplicada ao estudo.

Na sequência, a próxima seção abordará a inter-relação entre os campos HD e CI.

3 HUMANIDADES DIGITAIS E CIÊNCIA DA INFORMAÇÃO: UMA DISCUSSÃO DE SEUS PONTOS DE CONTATO

O pioneiro da “Computação nas Humanidades” foi o Padre Roberto Busa, que iniciou, em 1949, o seu projeto Index Thomisticus. Com o suporte da empresa International Business Machines Corporation – IBM, Busa conseguiu indexar palavras dos textos das obras de Tomás de Aquino, com o uso de cartões perfurados para torná-los legíveis por máquina. O trabalho demorou mais de 30 anos para ser concluído (Busa, 1980). O fato ficou registrado na história como uma importante aplicação da computação para resolver um problema da área de humanas. Já Alves (2016) considera que a “Computação nas Humanidades” teve seu começo quando os primeiros computadores surgiram, em meados do século XX, e compreende que o seu objetivo não é substituir as humanidades tradicionais, e sim complementá-las, facilitando a cooperação entre elas.

É relevante lembrar que os primeiros computadores ocupavam salas inteiras e projetos dependentes dessas máquinas podiam levar décadas para serem concluídos, como foi o caso do projeto Index Thomisticus, de Busa. No entanto, ao longo do tempo, principalmente nas últimas décadas, ocorreram transformações tecnológicas significativas que tiveram um impacto amplo em diversas áreas.

Diante desse cenário, observa-se que várias atividades inicialmente executadas por humanos estão sendo redirecionadas para a computação. Um exemplo disso é a substituição dos catálogos de cartões pelos mecanismos de busca. A história da automação reforça a ideia de que, sempre que possível, as tarefas serão transferidas dos seres humanos para os computadores (Sula, 2013).

O campo HD, termo como o conhecemos hoje, se encontra em processo de desenvolvimento, por isso alguns autores relutam em estabelecer uma definição. Há, ainda, dificuldade para descobrir o seu alcance, visto que está em constante transformação. Andrade e Dal’Evedove (2020) descrevem bem essa ideia quando falam que a disciplina está em plena configuração. Os autores acreditam que “há certo consenso da comunidade científica de que surge como um campo transdisciplinar que envolve pesquisas e práticas das Ciências Humanas e Sociais ao se relacionar com o ambiente digital” (2020, p. 441). Para Brayner (2021, p. 234)

Ainda que não haja um consenso estabelecido do que venham a ser as práticas e o escopo das Humanidades Digitais (HDs), o tema vem ganhando cada vez mais espaço nas publicações e debates acadêmicos. Via de regra, o termo HD é empregado em relação às novas metodologias de pesquisa surgidas nas últimas décadas, baseadas no uso das tecnologias de informação (TICs), ferramentas e técnicas computacionais para criação, seleção, extração, descrição, mapeamento e análise de dados apresentados em formato eletrônico. Todavia, e apesar da sua recente popularização no campo das

Ciências Humanas, é incontestável o fato de que as HDs estão ampliando os horizontes das pesquisas realizadas no campo das Humanidades, permitindo-nos elaborar novos questionamentos e chegar a novas descobertas nunca antes imaginadas.

Nessa linha de pensamento, Moura (2019, p. 60-61) tem o seguinte posicionamento sobre o campo HD:

As humanidades digitais constituem um território científico relativamente novo, decorrentes da interseção entre as Humanidades, as Ciências Sociais e as tecnologias digitais. Essa composição parece desafiar o modo contemporâneo de produzir ciência na medida em que a seara de produção, armazenamento e circulação do conhecimento está sobreposta em um mesmo ambiente digital [...] O campo de pesquisa em HDs se tornou possível devido à intensa digitalização de massas de dados (big data), à necessidade de compreensão da natureza dos dados obtidos, a ampliação de pesquisas colaborativas a distância e a alteração da cadeia de produção científica.

Por outro lado, as HD também são consideradas um empreendimento social, que “abriga redes de pessoas que trabalham juntas, compartilhando pesquisas, discutindo, competindo e colaborando há muitos anos [...]” (Kirschenbaum, 2010, p. 56, tradução nossa). Segundo Brayner (2021), a colaboração com pesquisadores de diversas áreas, como CI, Informática, Sociologia, Economia e Geografia, entre outras, possibilita que os humanistas digitais expandam continuamente seu campo de investigação, buscando alcançar metas compartilhadas que atendam aos objetivos de pesquisa específicos de cada uma dessas áreas. Além disso, é importante notar que existe também o entendimento de que as HD não constituem um campo unificado, como revelado na citação a seguir:

As Humanidades Digitais não são um campo unificado, mas um conjunto de práticas convergentes que exploram um universo em que: a) a impressão não é mais o meio exclusivo ou normativo em que o conhecimento é produzido e / ou disseminado; em vez disso, a impressão é absorvida por novas configurações de multimídia; e b) as ferramentas, técnicas e mídias digitais alteraram a produção e a disseminação do conhecimento nas artes, nas Ciências Humanas e Sociais [...] (The Digital, [20--], p. 2, tradução nossa).

As HD também estão proporcionando novos *insights* ao explorar o conhecimento acumulado do passado e ao contextualizar informações e dados emergentes no presente. Adicionalmente, estão conseguindo abrir um amplo campo de pesquisa, que engloba o estudo da influência das Tecnologias de Informação e Comunicação (TIC) na formação do indivíduo contemporâneo, bem como a compreensão do crescente impacto social, cultural, político e econômico causado pelo constante avanço tecnológico que afeta as relações que mantemos com a produção do conhecimento na sociedade informatizada (Brayner, 2021).

Considera-se como marco importante para a consolidação do termo, o documento Manifesto das Humanidades Digitais, de 2011, que surgiu a partir da reunião de pesquisadores

e atores sociais da área no evento The Humanities and Technology Camp (THATCamp) (Andrade; Dal'Evedove, 2020). Segue a definição que consta no documento:

I. Definição

1. A opção da sociedade pelo digital altera e questiona as condições de produção e divulgação dos conhecimentos.
2. Para nós, as *digital humanities* referem-se ao conjunto das Ciências Humanas e Sociais, às Artes e às Letras. As Humanidades Digitais não negam o passado, apoiam-se, pelo contrário, no conjunto dos paradigmas, *savoir-faire* e conhecimentos próprios dessas disciplinas, mobilizando simultaneamente os instrumentos e as perspectivas singulares do mundo digital.
3. As *digital humanities* designam uma transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências Humanas e Sociais (THATCamp, 2011).

É importante destacar o seguinte trecho no documento: “[...] as Humanidades Digitais não negam o passado, apoiam-se, pelo contrário, no conjunto dos paradigmas, *savoir-faire* e conhecimentos próprios dessas disciplinas [...]”. Em outras palavras, esse campo se baseia nos conhecimentos do passado e valoriza os saberes das Ciências Humanas e Sociais.

Para Guerreiro e Borbinha (2014), o campo HD não abrange apenas a simples transferência do analógico para o digital, vai além, “centrando-se no desafio epistemológico e na articulação com os conhecimentos e os métodos utilizados nas ciências humanas com o mundo digital”.

Rollo (2020, p. 21) destaca os seguintes eixos mais importantes do campo HD: “salvaguarda e preservação do patrimônio digital; educação e formação; organização da informação; articulação com outras áreas científicas; acessibilidade, disseminação e partilha do conhecimento; compromisso com a sociedade”.

Apontado como um dos eixos do campo HD, a organização da informação/conhecimento tem um papel fundamental na evolução da humanidade, e a CI também se destaca nessa missão. Segundo Borko (1968), a CI estuda as propriedades e o comportamento da informação, as forças que direcionam os fluxos de informação e os significados do processamento da informação para alcançar a acessibilidade e a usabilidade ideal. A disciplina lida com a “origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação” (1968, p. 1).

Isto inclui a pesquisa sobre a representação da informação em ambos os sistemas, tanto naturais quanto artificiais, o uso de códigos para a transmissão eficiente da mensagem, bem como o estudo do processamento e de técnicas aplicadas aos computadores e seus sistemas de programação. É uma ciência interdisciplinar derivada de campos relacionados, tais como a Matemática, Lógica, Linguística, Psicologia, Ciência da Computação, Engenharia da Produção, Artes Gráficas, Comunicação, Biblioteconomia, Administração, e outros campos científicos

semelhantes. Têm ambos componentes, de ciência pura visto que investiga seu objeto sem considerar sua aplicação, e um componente de ciência aplicada, visto que desenvolve serviços e produtos (Borko, 1968, p.1-2).

Assim, a CI é entendida como o estudo da informação em formato físico ou digital, independentemente do seu suporte (livro impresso, *e-book*, Compact Disc Read Only Memory – CD-ROM, artigo científico eletrônico, PDF e outros documentos), a partir da sua concepção até a sua utilização, tendo como propósito a disseminação do conhecimento.

De acordo com Saracevic (1996), a disciplina possui três características, a saber: tem natureza interdisciplinar, está ligada à tecnologia da informação e faz parte da evolução da sociedade da informação, conforme pode ser observado no quadro a seguir:

Quadro 1 - Características da CI

| | |
|--|---|
| Interdisciplinar | “A CI é, por natureza, interdisciplinar, embora suas relações com outras disciplinas estejam mudando. A evolução interdisciplinar está longe de ser completada.” |
| Ligada à tecnologia da informação | “Está inexoravelmente ligada à tecnologia da informação. O imperativo tecnológico determina a CI, como ocorre também em outros campos. Em sentido amplo, o imperativo tecnológico está impondo a transformação da sociedade moderna em sociedade da informação, era da informação ou sociedade pós-industrial.” |
| Presente na evolução da sociedade da informação | “[...] a CI é, juntamente com muitas outras disciplinas, uma participante ativa e deliberada na evolução da sociedade da informação. A CI teve e tem um importante papel a desempenhar por sua forte dimensão social e humana, que ultrapassa a tecnologia. Essas três características ou razões constituem o modelo para a compreensão do passado, presente e futuro da CI e dos problemas e questões que ela enfrenta.” |

Fonte: Saracevic (1996, p. 42).

Não há consenso entre os autores sobre o surgimento da CI, já que vários eventos contribuíram para o seu nascimento. Mas, de acordo com Oliveira (2005), é possível destacar o período após a Segunda Guerra Mundial, quando se iniciou o esforço de enfrentar os problemas de organização, crescimento e disseminação do conhecimento registrado. Para a autora (2005, p. 13), a CI “nasceu para resolver um grande problema, [...] que é o de reunir, organizar e tornar acessível o conhecimento cultural, científico e tecnológico produzido em todo o mundo”.

Observa-se, portanto, que tanto a CI como as HD são campos de estudo comprometidos com a organização e a disseminação da informação/conhecimento. Assim, a presente seção propõe-se a estudar e a responder os seguintes questionamentos: é possível encontrar uma conexão entre HD e CI? As HD estão na CI? Qual a contribuição dos dois campos na organização da informação/conhecimento?

Segundo Andrade (2021, p. 13), “[...] as Humanidades Digitais são um tema emergente para a Ciência da Informação, e plenamente capaz de colaborar para o domínio e apropriação de seus recursos, tornando ambas inter-relacionadas”. Para Santos (2012, p. 63) “[...] a Ciência da Informação está frequentemente permeada por conceitos, noções e ideias interdisciplinares, estabelecendo pontes e fronteiras com múltiplas disciplinas”. Há, então, uma ponte com as HD? Segundo Moura (2019, p. 65), “o exame das interfaces entre a Ciência da Informação e as Humanidades Digitais revelou paralelismos, pontos de convergência e esforços cruzados”. A autora (2019, p. 66) observou, também, que as duas disciplinas possuem “[...] potencialidades teórico-metodológicas e inúmeras conexões relacionadas aos valores, métodos científicos e a agenda de pesquisas”.

Já Pimenta (2020) acredita que HD e CI são campos relacionados em termos de atuação metodológica e familiaridade com que pesquisadores compartilham discursos interdisciplinares, e considera que suas fronteiras têm se tornado mais permeáveis devido ao desenvolvimento das TIC. Para o autor, os discursos das duas disciplinas

[...] parecem convergir cada vez mais para um campo de interesses em comum onde o uso das tecnologias digitais nos diversos processos de mediação da informação ou de seu respectivo acesso e recuperação parecem indicar uma espécie de ‘encontro’ entre as duas (2020, p. 2).

Já que

[...] se por um lado é possível afirmar que a categoria de informação esteve e está presente nos debates de cunho metodológico, teórico ou mesmo epistemológico existentes entre os *players* das Humanidades Digitais, por outro, tais debates parecem estar cada vez mais em ressonância com questões que são centrais para o campo disciplinar da CI, como: acesso aberto, repositórios, uso e reuso de dados, usuários, divulgação, visualização de dados e informação. Estes são alguns exemplos de como aquilo que move algumas pesquisas no âmbito das Humanidades Digitais são igualmente ‘dínamos’ de ações de pesquisa na CI (2020, p. 8).

De acordo com Rio Riande (2018), HD e CI compartilham objetos de fronteira (como banco de dados, repositórios digitais e etc.), e possuem uma epistemologia compartilhada. Em suma, são ciências interessadas no “tratamento informacional”. Andrade e Dal’Evedove (2020, p. 441) descrevem bem a ideia de parceria entre as disciplinas quando afirmam que as “Humanidades Digitais passam a integrar a Biblioteconomia e a Ciência da Informação na medida em que produtos e serviços provenientes de sistemas de recuperação de informações convergem com o ambiente digital [...]”.

[...] observamos que as Humanidades Digitais possuem duas grandes áreas de atuação dentro da Ciência da Informação. A primeira diretamente relacionada à representação e acesso aos acervos de humanidades, presente em arquivos, bibliotecas, museus e demais instituições que favorecem a elaboração, divulgação e preservação de bens e

artefatos culturais. Neste sentido é que se colocam também as Humanidades Digitais, no contexto da facilitação do acesso à informação em que o usuário é o pesquisador ou investigador da área de Ciências Sociais e Humanas, como foco principal. A segunda atuação está relacionada às tecnologias computacionais no cenário contemporâneo do excesso de informações e dados digitais, e de como outros campos do conhecimento podem se relacionar com a Ciência da Informação para a promoção de soluções semânticas no ambiente *Web* (Andrade; Dal'Evedove, 2020, p. 445).

Dessa forma, as HD estão intrinsecamente ligadas à CI em um amplo e complexo universo. Tanto os debates teóricos quanto as práticas de organização, seleção, disseminação e curadoria digital contribuem para essa conexão (Salcedo; Bezerra, 2018).

Assim, a curadoria digital desempenha um papel fundamental como um ponto de contato entre essas duas disciplinas. Essa prática, de acordo com Tavares (2014), abrange uma série de atividades, como preservação, seleção, manutenção, coleta e arquivamento de objetos digitais, visando garantir, manter e agregar valor aos repositórios digitais para seu uso presente e futuro. Em geral, é realizada por profissionais qualificados, como bibliotecários, arquivistas, cientistas, historiadores ou museólogos.

Todas as atividades envolvidas no gerenciamento de dados, desde o planejamento de sua criação, as melhores práticas de digitalização e documentação e a garantia de sua disponibilidade e adequação para descoberta e reutilização no futuro, fazem parte da curadoria digital. A curadoria digital também pode incluir o gerenciamento de vastos conjuntos de dados para uso diário, por exemplo, garantindo que eles possam ser pesquisados e continuem legíveis. A curadoria digital é, portanto, aplicável a uma ampla gama de situações profissionais desde o início até o fim do ciclo de vida da informação; digitalizadores, criadores de metadados, financiadores, formuladores de políticas e gerentes de repositório, para citar alguns exemplos (Abbott, 2008, tradução nossa).

Batista, Silva e Jorente (2021) defendem que a curadoria digital é uma área emergente no campo da CI. Já Siebra, Borba e Miranda (2016) argumentam que o conceito de curadoria digital começou a ser explorado, tanto na CI quanto na Ciência da Computação (CC), a partir de 2003, impulsionado pelo crescimento exponencial das informações digitais. Esses pesquisadores acreditam que a curadoria é uma prática naturalmente interdisciplinar, envolvendo aspectos tecnológicos, comunicacionais, gerenciais, cognitivos, de geração de conhecimento e informações.

Por isso tudo, considerando os aspectos observados pelos autores citados na presente seção, foi possível elaborar um quadro com os principais pontos em comum entre HD e CI, conforme segue:

Quadro 2 - Principais pontos em comum entre HD e CI



Fonte: elaborado pela autora (2023).

Com base nas informações apresentadas, é possível afirmar que as HD e a CI são campos de pesquisa transdisciplinares, que compartilham objetivos comuns. Essa inter-relação tem sido efetiva na busca por soluções para desafios complexos, relacionados ao gerenciamento de informações em ambientes virtuais. Nesse contexto, os repositórios digitais desempenham um papel importante como objeto de fronteira entre os dois campos.

Dessa forma, as perguntas existentes na seção já receberam respostas. A próxima parte abordará o tema tratamento informacional das teses e dissertações eletrônicas, em relação a diversos aspectos.

4 ABORDAGENS DE TRATAMENTO INFORMACIONAL PARA COLEÇÕES DE TESES E DISSERTAÇÕES ELETRÔNICAS: UM ESTUDO DAS PRINCIPAIS ESTRATÉGIAS

4.1 Sistemas de Organização do Conhecimento (SOC) aplicáveis ao tratamento de teses e dissertações: classificação, tesauro, ontologia e taxonomia

Indiscutivelmente, a "organização" é o requisito principal para localizar, com agilidade, um documento físico, independentemente do local em que esteja (casa, escritório, biblioteca, arquivo etc.). No caso do acervo eletrônico, em particular a coleção de teses e dissertações eletrônicas, surge a seguinte questão: como ordená-lo para permitir sua utilização no RD?

Conforme aponta Dziekaniark (2010), atualmente é possível navegar por enormes acervos na Internet, acessar milhares de sítios sobre diversos temas e interesses, além de poder obter uma obra por *download* sem se preocupar com o prazo da devolução ou com a cobrança de multa. Diante dessa situação, a autora aborda os seguintes questionamentos sobre a informação no meio eletrônico:

[...] como tratar o grande volume informacional disponibilizado na rede diariamente? Serão as técnicas tradicionais da Biblioteconomia que darão conta dessa nova demanda? Será o bibliotecário o único sujeito a realizar tarefas de processamento/tratamento/difusão/recuperação da informação? [...] (Dziekaniark, 2010, p. 51).

As universidades desempenham um papel de destaque nessa história, atuando, conforme mencionado por Fujita (2005), como organismos geradores, transmissores e receptores de conhecimento. As bibliotecas digitais das instituições, repositórios das teses e dissertações em formato eletrônico, enfrentam, segundo Bastos (2005), a falta de uma estrutura de categorias adequada para organizar seu conteúdo informacional. Apesar de passados quase 20 anos desde essa constatação, a situação não mudou, e são várias as dificuldades relacionadas à organização dessa vasta quantidade de objetos digitais em ambientes virtuais.

Como mencionado anteriormente, a CI não está sozinha na busca por soluções para lidar com questões diversas relacionadas à organização da informação digital. As HD emergiram com o potencial para ajudar a enfrentar esses desafios. De acordo com Andrade (2021), seu campo está presente, implícita ou explicitamente, na organização do conhecimento, com foco na recuperação da informação. Segundo Mattos (2019), o seu trabalho na organização do conhecimento trouxe novas perspectivas para as humanidades tradicionais ao possibilitar o uso da tecnologia digital para organizar grandes volumes de texto.

De acordo com Broughton (2006), a organização da informação no ambiente virtual é diferente. A autora defende que a linearidade pouco importa, e que as inquietações na gestão do armazenamento da informação digital não são com a organização do material e sim com a descrição correta dos metadados, a disponibilização das ferramentas de busca para auxiliar na navegação e localização, e em um escopo limitado, a apresentação dos resultados.

Por outro lado, Sayão (2009, p.16) compreende que “o digital não é o antagônico do impresso, como o rolo de papiro não é o antagônico do livro [...]”. É, então, apropriado usar o conhecimento do passado para resolver o problema atual de gerenciamento de coleções digitais, como o caso das teses e dissertações eletrônicas?

A organização do conhecimento tem origem nos primeiros registros de informações na forma de catálogo. Depois disso, produtos e tecnologias decorrentes da necessidade de organizar esses elementos foram surgindo ao longo da história, sendo que os marcos mais influentes datam da primeira metade do século XIX. Com o tempo, essa mudança começou a eclodir cada vez mais, especialmente com o surgimento da profissão de bibliotecário, a profissionalização da organização da informação, o desenvolvimento dos cursos de Biblioteconomia e, subsequentemente, o surgimento da CI como um campo científico. Portanto, a necessidade de ordenar o conhecimento sempre foi uma preocupação, mas no passado parecia ser apenas domínio de bibliotecários e filósofos (Andrade, 2021).

O campo da Organização do Conhecimento denota uma das mais genuínas angústias da humanidade: processar o conhecimento humano de forma que seja bem aplicado à geração de novos conhecimentos e ao desenvolvimento contínuo da sociedade. Além disso, sabemos que a aprendizagem humana se baseia na capacidade de analisar, representar e organizar dados, informações e conhecimentos, daí a necessidade premente de organizá-los a fim de recuperá-los (Andrade, 2021, p. 42).

Segundo Paletta e Moreira-González (2019), o conhecimento organizado é registrado como informação que pode ser recuperada para produzir novos conhecimentos, visto que a informação precisa de suporte para transmissão (digital, impresso, áudio, visual) e usa os sistemas de linguagem para decodificação (linguagem documentária). É produzido por uma pessoa que gera conhecimento para outro indivíduo. Já Vignoli, Souto e Cervantes (2013) defendem que a “informação” é considerada o insumo básico para a construção do “conhecimento” no indivíduo, mas nem sempre se torna conhecimento, porque depende da sua assimilação.

Para Andrade (2021, p. 41), “[...] as informações registradas, quando bem organizadas e representadas, configuram-se como a principal ferramenta para o desenvolvimento do conhecimento humano”. Para a autora, a dificuldade em organizar o conhecimento está

relacionada à volumosa quantidade de dados não estruturados e à crescente dificuldade de acesso e recuperação de grandes quantidades de informações.

Na era digital, a necessidade de organizar sistematicamente as informações tornou-se universal (Golub; Kamal; Vekselius, 2022). Mas como realizar essa tarefa complexa? Chega-se à conclusão que, além de adotar soluções tecnológicas, é necessário buscar formas adequadas para estruturar os dados nos sistemas de informação, como o caso dos repositórios das teses e dissertações eletrônicas.

Os SOC (no original inglês, Knowledge Organization System – KOS), são ferramentas semânticas com um vocabulário estruturado e formalizado usado para o processamento e a recuperação de informações em ambientes virtuais e tradicionais. Compreendem os vários tipos de esquemas de organização e representação do conhecimento (Carlan; Medeiros, 2011).

Os SOC são instrumentos que dependem de padronização para melhor interagirem, integrarem e interoperarem entre eles e entre humanos, sejam usuários, provedores de informações ou máquinas. Esses sistemas contribuem para diminuir o abismo na comunicação entre pesquisadores, bancos de dados e documentos armazenados, criados pelas dificuldades na organização das informações e do conhecimento (Carlan; Medeiros, 2011, p. 71-72).

Seguem alguns exemplos de SOC:

Quadro 3 - Resumo das principais características dos SOC

| | |
|-----------------------|--|
| Classificações | Amplamente utilizadas para organizar os acervos das bibliotecas. |
| Taxonomias | Têm grande relevância na organização de informações em empresas e instituições, especialmente no contexto do desenvolvimento de portais no ambiente <i>web</i> . |
| Tesauros | Usados no campo da organização da informação. Sua existência é confirmada e padronizada por normas internacionais (International Organization for Standardization – ISO, American National Standards Institute – ANSI/ National Information Standard Organization – NISO). |
| Ontologias | Referem-se à capacidade de representação do conhecimento de forma complexa e completa. |

Fonte: Carlan; Medeiros (2011).

Os SOC são necessários porque realizam o controle de vocabulário em sistemas de informação que fornecem recuperação de informações armazenadas (retrospectiva ou atualizada). Os sistemas tratam diretamente da organização e representação do conhecimento, e caso isso não aconteça, a busca da informação é realizada através da linguagem natural presente no conteúdo da informação (Fujita; Tolare, 2019).

No caso da classificação de documentos, o recurso é considerado “[...] um processo utilizado na organização da informação para facilitar sua busca e recuperação”, conforme

Barrera (2016, p. 34, tradução nossa). Gomes (2017) completa essa ideia quando diz que a “ordenação” é considerada a base da Organização do Conhecimento, e a “classificação”, o componente principal. Segundo a autora, a organização

[...] significa classificação em seu sentido mais amplo, incluindo indexação (atribuição de uma classe de assunto ou aspecto a um documento/recurso), esquemas de classificação com ou sem notação, terminologia, tesouros, taxonomias e assemelhados (Gomes, 2017, p. 35).

A classificação muito popular no mundo ocidental é a Classificação Decimal de Dewey (CDD), utilizada em várias bibliotecas do Brasil e do mundo.

Hoje, a classificação em seus aspectos teórico e prático é fundamental para a organização do conhecimento como também para a gestão do conhecimento nas organizações; para produtos e serviços como portais, intranets, arquitetura de informação; para a criação de instrumentos semânticos; e até mesmo criação de metadados, na maioria das vezes associados a facetas, os quais têm igualmente, papel na organização do conhecimento (Gomes, 2017, p. 61).

Já o tesouro é considerado uma ferramenta composta por uma lista alfabética de palavras-chave e referências cruzadas, destinada a mostrar uma hierarquia de conceitos; como também a própria estrutura hierárquica representa aspectos e/ou categorias. Ajuda o usuário na busca e acesso às informações do seu interesse (Santos; Lima; Araújo, 2017).

Por exemplo, diferentes usuários podem expressar as mesmas necessidades de informação usando uma linguagem diferente, com sinônimos, abreviações, acrônimos, entre outros, e o tesouro serve para resolver esses problemas característicos da linguagem natural, relacionando os termos que representem o mesmo conceito, selecionando apenas um termo como padrão, enquanto os outros termos restantes serão considerados sinônimos, além de estabelecer relações entre esses termos (Santos; Lima; Araújo, 2017, p. 44).

Outro recurso aliado à CI, em relação à organização do conhecimento, é a ontologia, considerada modelo de linguagem documentária, segundo Dziekaniak (2010). Sob esse prisma, a autora entende ser necessário incorporar a matéria ontologia aos currículos da Biblioteconomia, a fim de promover e fomentar novas pesquisas que vinculem o tema ao processamento, à representação e à recuperação da informação.

Segundo Sales, Campos e Gomes (2006), as ontologias são construídas por termos, definições e relações. Os autores defendem a ideia de que a ontologia não pode ser definida apenas como linguagem documentária, já que possui também a funcionalidade que possibilita ao computador processar o raciocínio automatizado por meio de regras e inferências.

Um importante elemento das ontologias é a representação do conhecimento, que envolve análise semântica, em especial para a organização de sistemas de conhecimento que, nas ontologias, visam otimizar a recuperação, ou seja, prover o computador de mais inteligência. Trata-se de uma área de pesquisa estreitamente

ligada à organização do conhecimento, em especial no que tange à rede semântica. Esta representa o conhecimento na forma de nós (conceitos, atributos) ligados por arcos (relacionamentos). [...] Nas ontologias, o conjunto de relações é mais rico do que nas tabelas de classificação bibliográfica ou nos tesauros, o que permite maior estrutura representativa do conhecimento registrado em um discurso, para que possa ser manipulado pelo computador. É importante ressaltar que tais representações são igualmente relevantes na formulação de buscas nas ontologias (Sales; Campos; Gomes, 2006, p.64-65).

Para a construção de ontologias, Magalhães e Souza (2019) consideram indispensável um estudo multidisciplinar, um aprofundamento nas questões semânticas, axiomas e sinônimos de determinado domínio do conhecimento. Para os autores, são necessárias também as parcerias entre as áreas da CI e da CC para concretizar recursos tecnológicos que garantam a precisão, a relevância e a eficácia dos Sistemas de Recuperação da Informação e, como consequência, possibilitar a satisfação do usuário.

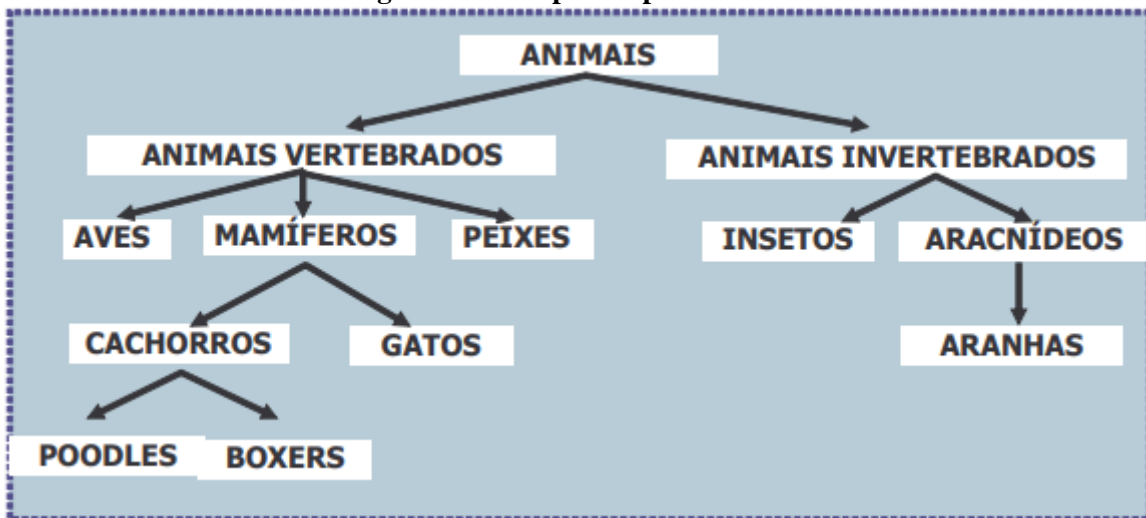
Acrescenta-se também o importante papel das taxonomias na organização da informação/conhecimento no ambiente virtual, conforme descrito por Campos e Gomes (2008): “[...] a organização das informações através do conceito de taxonomia permite alocar, recuperar e comunicar informações dentro de um sistema de maneira lógica através de navegação”.

Estas taxonomias se caracterizam por:

- Conter uma lista estruturada de conceitos/termos de um domínio.
- Incluir termos sem definição, somente com relações hierárquicas.
- Possibilitar a organização e recuperação de informação através de navegação.
- Permitir agregação de dados, diferentemente das taxonomias seminais, além de evidenciar um modelo conceitual do domínio.
- Ser um instrumento de organização intelectual, atuando como um mapa conceitual dos tópicos explorados em um Sistema de Recuperação de Informação.
- Ser um novo mecanismo de consulta em portais institucionais, através de navegação (Campos; Gomes, 2008).

Vignoli, Souto e Cervantes (2013) enfatizam que as ontologias e as taxonomias são ferramentas de organização e representação do conhecimento em SOC destinadas aos usuários finais, para possibilitar a busca e a recuperação de informações.

Figura 1 - Exemplo simples de taxonomia



Fonte: Terra *et al.* (2005, p. 2).

Para Terra *et al.* (2005, p. 1), a taxonomia tem a seguinte definição:

A taxonomia é um sistema para classificar e facilitar o acesso à informação, e que tem como objetivos: representar conceitos através de termos; agilizar a comunicação entre especialistas e entre especialistas e outros públicos; encontrar o consenso; propor formas de controle da diversidade de significação; e oferecer um mapa de área que servirá como guia em processos de conhecimento. É portanto, um vocabulário controlado de uma determinada área do conhecimento, e acima de tudo um instrumento ou elemento de estrutura que permite alocar, recuperar e comunicar informações dentro de um sistema, de maneira lógica.

De acordo com Nascimento e Pinho (2019, p. 111),

[...] assim como outros sistemas que formam os SOCs, a taxonomia se virtualiza e é, geralmente, desenvolvida em computadores, podendo manter seu funcionamento em rede, interagindo com base de dados na Internet, por exemplo.

No ambiente virtual, o objetivo das taxonomias é simplificar as buscas e a navegação e designar a responsabilidade de avaliar, organizar, deletar e arquivar informações. A taxonomia ideal deve facilitar a recuperação de documentos, permitindo que o usuário pesquise por diferentes critérios, por exemplo, autor, data, formato do arquivo, área de dados etc. (Terra *et al.*, 2005).

A taxonomia é considerada uma ferramenta de representação para mapeamento da área de conhecimento que é frequentemente usada em sistemas de informação para entrada de dados, organização e classificação de registros do conhecimento, visando à recuperação da informação. Já o tema “taxonomias facetadas navegacionais” é estudado tanto pela CI como pelas áreas da CC, incluindo Engenharia de Software, Arquitetura da Informação e Informática (Maculan, 2011).

A taxonomia facetada apresenta uma interface que mostra a categoria principal e as facetas dentro delas, numa distribuição hierárquica dos conceitos. Uma vantagem do uso da taxonomia é que sua estrutura em árvore mostra, em seus primeiros níveis, os conhecimentos mais genéricos. Já nos níveis mais avançados (nos 'ramos'), encontram-se informações mais específicas das buscas. Em cada nó da árvore se insere um conjunto de informações, que aborda sobre determinada temática e que é resultado da compilação do conhecimento construído por agentes diversos, dentro da biblioteca digital. Recursos audiovisuais como textos, imagens, mapas, sons e vídeos podem ser inseridos, o que potencializa a eficácia na transmissão de informação aos que almejam acessá-la. Além disso, as taxonomias facetadas são pós-coordenadas e permitem a representação de conceitos inter-relacionados, isto é, com elementos semânticos. Elas têm sido ferramentas muito utilizadas em portais corporativos (Maculan; Lima; Penido, 2010, p. 243).

Conforme apresentado, os instrumentos de representação do conhecimento, como os tesouros, desempenham um papel importante no processo de identificação do conteúdo dos documentos digitais, proporcionando padronização que resulta em maior precisão na busca por conhecimento. Por outro lado, a classificação e a taxonomia são reconhecidas como ferramentas de categorização conceitual, que estabelecem estruturas que facilitam o acesso à informação pelos usuários. Já as ontologias conferem eficiência à organização da informação/conhecimento no ambiente virtual, pois conseguem representar, estruturar e relacionar os termos, o que aprimora a recuperação da informação. Mas, infelizmente, não é uma tarefa fácil construir ontologias para um RD, já que demanda tempo e o envolvimento de pessoal qualificado de diversas áreas (como bibliotecários, cientistas da computação e linguistas) para executar o projeto.

Além das opções apresentadas no quadro 3 da dissertação, há outras formas de tratamento informacional no ambiente virtual. O objetivo desta seção é estudar as opções mais adequadas para aprimorar o tratamento das teses e dissertações eletrônicas no RD. Conclui-se, portanto, que a adição dos SOC no repositório digital terá um impacto positivo na organização dos documentos depositados nesse sistema de informação.

4.2 A mineração de texto como recurso auxiliar na indexação das teses e dissertações

Ao processar os objetos digitais no sistema de informação, o bibliotecário não dispõe de tempo para ler integralmente todo o conteúdo do documento, como uma tese eletrônica, a ser registrada na base de dados. A tarefa de cadastrar informações básicas, como título, autor e número de páginas (ou seja, a descrição física), demanda tempo e atenção especial à padronização, uma vez que qualquer falha, seja um erro de digitação ou uma atribuição equivocada de autoria, pode comprometer a eficácia na recuperação da informação. Por outro lado, o desafio relacionado à representação temática é ainda mais complexo. A compreensão

dos principais tópicos abordados em um documento é uma atividade desafiadora. Assim, surge a pergunta: há um método automático de processamento de documentos digitais que forneça uma descrição temática do texto que está sendo tratado?

Diversos estudos tentam encontrar soluções para os problemas de organização e processamento da informação digital, em busca de métodos que visem facilitar e melhorar atividades como indexação de documentos, identificação de semelhanças entre documentos, criação automatizada de tesauros, geração de resumos automáticos, taxonomia, classificação, pesquisa de informações de texto completo, extração de informações, identificação de relacionamentos e termos, entre outros. As pesquisas propõem técnicas e métodos baseados em: aprendizado de máquina, reconhecimento de padrões, métodos estatísticos, métodos de redes neurais artificiais e técnicas de Processamento de Linguagem Natural (PLN) (Barrera, 2016).

Segundo a percepção de Santos (2017), a indexação consiste em atividades intelectuais consideradas como o limite das máquinas. A autora acredita que, apesar dos avanços tecnológicos, o computador ainda é incapaz de cumprir a tarefa humana de abstrair, distinguir e entender o texto, ou seja, para Santos, indexadores humanos ainda são requeridos para essa difícil tarefa.

A indexação fundamenta-se em descrever o conteúdo dos documentos de forma concisa e compacta por meio do uso de termos (também conhecidos como descritores ou palavras-chave) que desempenham funções que permitem a identificação e a recuperação dos documentos. Dessa forma, para executar tal tarefa, existem as linguagens documentárias/vocabulários controlados de documentos, que são ferramentas tradicionais de representação da informação. Seu objetivo é facilitar a comunicação através da padronização dos termos utilizados para descrever o conteúdo dos documentos (Santos, 2017).

Conforme destacado na citação abaixo, o processo de indexação abrange dois estágios:

[...] o analítico, em que é realizada a compreensão do texto como um todo, a identificação e a seleção de conceitos válidos para a indexação e o estágio de tradução, que consiste na representação de conceitos por termos de uma linguagem de indexação: • Determinação do assunto: estabelecimento dos conceitos tratados num documento; • Representação de conceitos por termos de uma linguagem de indexação: a tradução dos conceitos nos termos da linguagem de indexação (Fujita, 2003, p. 63).

Levando em consideração a explicação sobre a função de cada estágio da indexação, pode-se confirmar que uma das etapas consideradas mais importantes na tarefa do indexador é a “análise de assunto”. Essa fase tem como propósito identificar e escolher os conceitos que representam a essência de um documento. É um procedimento complexo, pois demanda que o profissional da informação adote uma metodologia apropriada para alcançar resultados

satisfatórios. A eficácia desse processo pode ser avaliada pelos resultados obtidos pelo usuário ao buscar informações (Fujita, 2003).

Com o avanço tecnológico, tornou-se viável o desenvolvimento de programas capazes de obter informações essenciais e analisar quantidades crescentes de dados. A mineração de textos ocorre automaticamente através desses *softwares* que identificam padrões e singularidades na estrutura de textos escritos por humanos. Inclusive, a mineração de textos é considerada uma ferramenta digital de pesquisa e promoção das HD (Matos; Jacintho; Alvarez, 2019).

Tanto no sentido da aprendizagem de formas mais reflexivas e aprofundadas de leitura, quanto do ponto de vista da facilidade em manejar montantes muito grandes de textos, com vistas à organização e sistematização, a mineração de textos tem apresentado resultados promissores, a julgar pelo que os estudos vêm apresentando. O comportamento de computadores digitais que analisam textos e favorecem a ‘descoberta de conhecimento’ parece muito mais interessante quando é descrito em termos mentais. A capacidade de um programa, de fornecer um gráfico com os termos de maior ocorrência em um texto, e a estrutura das relações entre os termos, por exemplo, tem toda a aparência de um comportamento organizado e inteligente (Matos; Jacintho; Alvarez, 2019, p. 136).

O processo de mineração de texto é responsável pelo uso e desenvolvimento de métodos automatizados para processar de forma rápida e eficiente documentos digitais não estruturados, com o propósito de analisá-los e organizá-los. O recurso permite o tratamento de documentos digitais usando métodos de extração e classificação de informações, o que possibilita organizar os documentos (Barrera, 2016).

Os autores Araújo Júnior e Tarapanoff, na publicação de 2006, já defendiam a ideia de que a mineração de texto poderia ajudar na construção da representação temática dos documentos, como pode ser notado na seguinte citação:

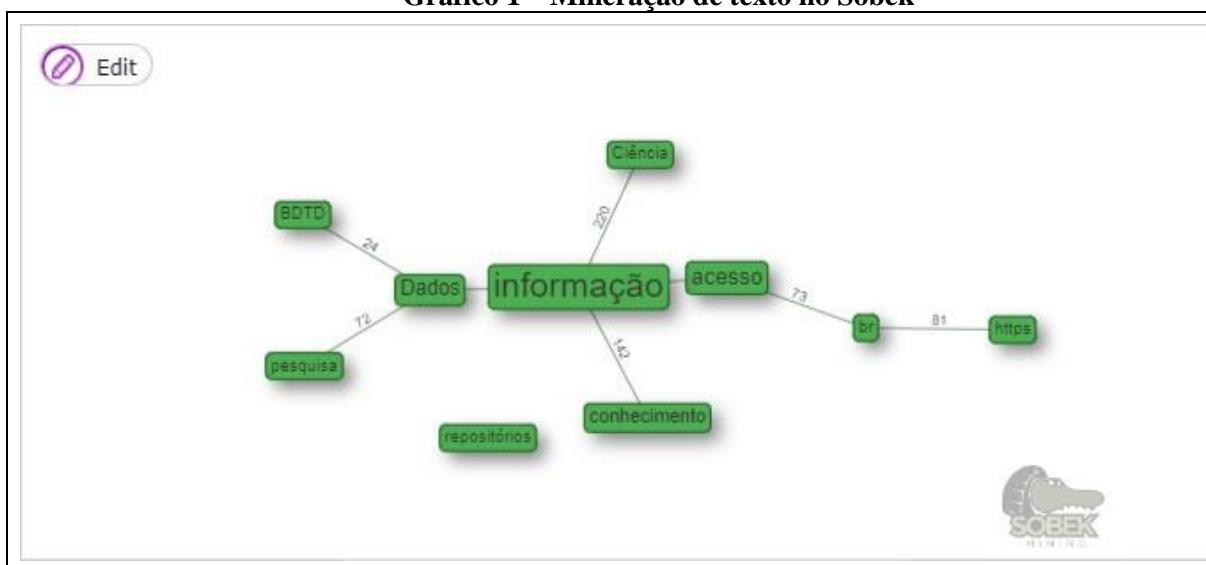
[...] a cautela aponta na direção da utilização da mineração de textos para auxiliar o processo de indexação, já que a grande potencialidade da ferramenta está na sua capacidade de sumarizar grandes conjuntos de documentos em agrupamentos, apresentando-os sob a forma de listas de palavras que mais ocorrem por documento, ou por resultado de pesquisas (conjuntos de documentos), e em alguns casos com gráficos indicativos das relações semânticas entre os termos (Araújo Júnior; Tarapanoff, 2006, p. 239).

Há diversos *softwares* de mineração de texto disponíveis. Um exemplo é o Sobek, desenvolvido com a finalidade de criar representações gráficas dos termos mais relevantes extraídos de documentos de texto. A ferramenta digital foi concebida pelo Programa de Pós-Graduação em Informática da Universidade Federal do Rio Grande do Sul (UFRGS) e está disponível nas versões *online* e para *download* (Medeiros; Pinho; Corrêa, 2019).

A mineração de texto desenvolvida pelo Sobek faz a tarefa de sumarização de texto, de forma não supervisionada, baseada em modelo de grafo e análise estatística do texto, sendo aplicado a textos individuais, visando gerar um sumário ou resumo gráfico do texto. [...] Assim como os mapas conceituais e mentais, as apresentações em grafos geradas pelo Sobek têm sido utilizadas como ferramentas para compreensão de conteúdos diversos, na sua assimilação e estudo, também no âmbito educacional. De maneira muito simples, o software exige apenas a inserção do conteúdo escrito na sua caixa de texto e automaticamente, com base em cálculos estatísticos [...] extrai e relaciona os principais termos (Medeiros; Pinho; Corrêa, 2019, p. 154-155).

A ferramenta Sobek realiza a extração de informações do texto e as apresenta de forma gráfica, utilizando um grafo em que os principais termos extraídos são dispostos como vértices (nodos) e as ligações entre os termos são representadas como arestas. Essa representação gráfica auxilia o usuário a visualizar os principais termos do texto e suas relações (Epstein, 2017). A seguir, encontra-se um experimento realizado no *site* do Sobek, utilizando como *corpus* a presente dissertação (documento completo):

Gráfico 1 – Mineração de texto no Sobek



Fonte: *site* Sobek (<http://sobek.ufrgs.br/sobekonline/index.html>).

O experimento conduzido no portal do Sobek não passou por qualquer modificação manual, isto é, o botão "Edit" não foi acionado, revelando alguns termos e suas inter-relações. Apesar de ser uma ferramenta digital simples, ela consegue destacar alguns tópicos que podem contribuir para a descrição temática das teses e dissertações. No entanto, possui limitações que impedem a realização da indexação completa do documento. Por exemplo, o tópico "Humanidades Digitais" não é representado no gráfico, e a palavra "repositórios" está isolada, sem conexões com outros termos. Adicionalmente, é necessário desconsiderar o "br" e "https". É importante destacar que, na mineração de texto, as análises são quantitativas e baseiam-se na repetição dos termos.

Dessa forma, é evidente que, ao utilizar esse minerador de texto, a supervisão de um profissional bibliotecário continua sendo imprescindível. Apesar da capacidade da ferramenta em agilizar o processo, a análise humana mantém sua relevância para realizar os ajustes necessários.

Após abordar as diversas formas de tratamento informacional para a coleção de teses e dissertações, a próxima seção apresentará o tópico repositórios digitais.

5 REPOSITÓRIOS DIGITAIS DE CONHECIMENTO

Ao longo do tempo, as bibliotecas físicas têm sido reconhecidas como repositórios de conhecimento e cultura, com o objetivo de armazenar, organizar e disseminar informações (Varela; Barbosa; Guimarães, 2009). Nesse contexto, Sayão e Marcondes (2009) sustentam que os RD têm suas bases em diversos serviços tradicionais de biblioteca. Segundo os autores, os RD foram criados com o propósito de garantir a disponibilidade e a acessibilidade de documentos digitais de pesquisa por um longo período, possibilitando a preservação da memória acadêmica das instituições.

O primeiro RD (ArXiv.org)² foi desenvolvido pelo físico Paul Ginsparg, em 1991, no EUA. A ideia nasceu com o movimento de livre acesso da informação, quando a comunidade acadêmica, incluindo associações de pesquisadores e bibliotecas universitárias e especializadas, estava procurando uma solução para o alto custo de assinaturas de publicações comerciais (Marcondes; Sayão, 2009). Acredita-se, portanto, que o acesso aberto maximiza a disponibilidade da pesquisa e, assim, aumenta sua visibilidade ao fornecer informações sem barreiras legais e financeiras (Cerrao; Castro, 2018).

No âmbito universitário, as bibliotecas carregam a responsabilidade de coletar, organizar, armazenar, preservar e compartilhar os conteúdos técnico-científicos gerados nas suas instituições. Dentro desse cenário, a preservação digital ganha uma importância fundamental no ciclo de vida das teses e dissertações, devido aos progressos tecnológicos que ampliaram sua exposição e, por conseguinte, o alcance dessa significativa produção intelectual, tanto em nível nacional quanto internacional (Gonçalves *et al.*, 2022).

A possibilidade de disponibilizar tais recursos digitais através dos repositórios constitui um avanço significativo, conferindo aos RD um papel de destaque na preservação e disseminação dos conhecimentos gerados pelas universidades, sendo reconhecidos como ferramentas fundamentais para a condução de pesquisas em escala global.

Conforme a definição do IBICT (2016), os RD são:

[...] bases de dados online que reúnem de maneira organizada a produção científica de uma instituição ou área temática. Os RDs armazenam arquivos de diversos formatos. Ainda, resultam em uma série de benefícios tanto para os pesquisadores quanto às instituições ou sociedades científicas, proporcionam maior visibilidade aos resultados de pesquisas e possibilitam a preservação da memória científica de sua instituição.

² Disponível em: <http://arxiv.org> Acesso em: 9 set. 2023.

Ainda, de acordo com o IBICT (2016), os RD podem ser institucionais ou temáticos: “[...] institucionais lidam com a produção científica de uma determinada instituição. Os repositórios temáticos com a produção científica de uma determinada área, sem limites institucionais”. Na realidade, de acordo com Barroso, Sales e Melo (2017), os repositórios institucionais são formados pela reunião de micro repositórios temáticos.

Do ponto de vista de Queiroz e Rodrigues (2022), são quatro tipos de RD, a saber: institucionais, temáticos (ou disciplinares), dados de pesquisa (que reúnem resultados das pesquisas) e *preprints* (a prévia publicação antes da aprovação dos pares). Já Ferreira *et al.* (2017) afirmam que, no início do século XXI, os repositórios digitais eram classificados conforme o seu conteúdo, e poderiam ser temáticos, institucionais, governamentais e de dados.

Para Leite (2009, p. 20), há três tipos de repositórios digitais, a saber:

Quadro 4 - Tipos de repositórios digitais

| | |
|---|--|
| Repositórios Institucionais | Voltados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa. Exemplo: e-Prints Soton – repositório de Pesquisa da Universidade de Southampton (http://eprints.soton.ac.uk/); |
| Repositórios Temáticos “ou” Disciplinares | Voltados a comunidades científicas específicas. Tratam, portanto, da produção intelectual de áreas do conhecimento em particular. Exemplo: EPrints in Library and Information Science – E-LIS (http://eprints.rclis.org/) e arXiv. org (http://arxiv.org/): |
| Repositórios de Teses e Dissertações (Electronic Theses and Dissertations – ETD) | Repositórios que lidam exclusivamente com teses e dissertações. Muitas vezes a coleta das muitas ETD é centralizada por um agregador. Exemplo: BDTD/UnB – Biblioteca Digital de Teses e Dissertações da Universidade de Brasília – UnB (http://btdt.bce.unb.br/tesesimplificado/) e BDTD (http://btdt.ibict.br/). |

Fonte: Leite (2009, p. 20).

Quanto ao RI, Marcondes e Sayão (2009, p. 9) sustentam que se trata de “[...] uma biblioteca digital destinada a guardar, preservar e garantir livre acesso, via internet, à produção científica no âmbito de uma dada instituição”. O RI também contribui para a prestação de contas à sociedade sobre os estudos que são realizados no Brasil, já que o financiamento das pesquisas provém em grande parte de recursos públicos (Sobral; Santos, 2017).

Os repositórios institucionais são, sem dúvida, alternativa recomendada para que as instituições sejam protagonistas e tenham relevância no ciclo da comunicação, não se trata de rejeitar o sistema estabelecido, mas operar em consonância com a missão que lhes é atribuída: preservação e divulgação [...] (Sobral; Santos, 2017, p. 182).

Para Ferreira *et al.* (2017), o RI presta diversos serviços para a comunidade, e representa uma via de mão dupla, conforme pode ser observado:

Sendo assim, há uma via de mão dupla, pois em um momento é o sujeito (autor vinculado à instituição) que está prestando serviços e enriquecendo o conteúdo intelectual do repositório e em outro momento, é o repositório, como ambiente informacional digital, que dá respaldo ao conteúdo, organizando-o e preservando-o para futuros acessos (Ferreira *et al.*, 2017, p. 612).

A visão de repositórios digitais como meros depósitos estáticos de informações digitais, com a habilidade de recuperação, foi rapidamente superada. O conceito original evoluiu para um modelo mais complexo de sistema de informação que integra facilidades de comunicação, colaboração e várias formas de interação dinâmica entre os usuários em um amplo contexto (Sayão; Marcondes, 2009).

Além disso, os repositórios possibilitam o depósito de uma ampla variedade de documentos, como dissertações, teses, artigos, livros, vídeos, imagens, anais, entre outros. Quanto ao formato, considera-se a questão da preservação em longo prazo. Entre as opções destacam-se: Tagged Image File Format (TIFF) para imagens, PDF/A para documentos, MPEG Layer 4 (MP4) para vídeos e Audio Interchange File Format (AIFF) para áudio. (Queiroz; Alves, 2018, p. 22).

A ferramenta precisa dos seguintes indicadores para acompanhamento, segundo Blattmann e Weber (2008, p. 481): “acessos, downloads, uploads, análise bibliométrica, entre outros, e desta forma, proporcionar visibilidade de quem acessa, o que acessa, quando e com que ferramentas”.

Quanto ao acesso no RD, são três modalidades:

Quadro 5 - Acesso ao RD

| | |
|------------------------------------|---|
| Acesso aberto | Acesso ao documento integral, direto e imediato. |
| Acesso restrito (embargado) | Não é permitido acesso imediato ao documento integral do documento durante um determinado período de tempo. |
| Acesso fechado | Não é permitido acesso aos metadados e ao documento integral. |

Fonte: Queiroz; Alves (2018, p. 21).

Geralmente, a responsável pela criação, implementação, funcionamento e administração dos repositórios de uma universidade é a biblioteca,

[...] por ser legitimada a trabalhar com o material institucional e por ser o órgão mais capacitado a gerenciar a informação e trabalhar com assuntos de comunicação e pesquisa científica, de forma a atender necessidades informacionais (Cerraio; Castro, 2018, p. 95).

Inclusive, os RI “são uma das maiores justificativas para o envolvimento do bibliotecário e cientista da informação com as tecnologias de informação e comunicação de forma mais intensa [...]” (Dziekaniak, 2010, p. 56).

Para os profissionais da informação o desafio é constante: adaptar-se às novas formas de trabalho, e integrar-se ao contexto dos repositórios digitais e contribuir com esse novo cenário que envolve a comunicação científica, no que se refere a atender os princípios do acesso aberto à informação por meio da gestão e disponibilização da produção científica e acadêmica de uma instituição (Marques, 2020, p. 137).

Com base no exposto, fica evidente que os RD possuem múltiplas utilidades tanto para as instituições que os disponibilizam quanto para os usuários, fornecendo meios para fomentar o intercâmbio de conhecimento entre os pesquisadores. Um importante exemplo é a coleção de teses e dissertações eletrônicas provenientes das universidades, acessíveis através de bibliotecas digitais ou RI, cujo alcance transpassa fronteiras geográficas. Em outras palavras, esses repositórios desempenham um papel central no cenário acadêmico e científico, ao facilitar o acesso aberto ao conhecimento, contribuindo, por conseguinte, para a contínua progressão da pesquisa em nível mundial.

No entanto, para garantir seu funcionamento adequado, é necessário contar com uma equipe composta por profissionais qualificados, com destaque para os papéis desempenhados pelo bibliotecário e pelo especialista em Tecnologia da Informação (TI). A interseção dos

saberes provenientes dessas duas áreas, ou seja, CI e CC, certamente resultará em progressos para a ferramenta digital.

5.1 O software Dspace

Como funciona um RD? No primeiro momento, é inevitável pensar no *software*, ou seja, no programa que atenderá as demandas da instituição que pretende implantar o RD. Há vários tipos de *softwares* de repositório disponíveis no mercado, com diferentes características. Seguem algumas opções que estão em destaque no ranque do Registry of Open Access Repositories: DSpace, E-Prints, Berkeley Electronic Press (Bepress) e Flexible Extensible Digital Object and Repository Architecture (Fedora). Segundo Borges *et al.* (2019), no caso específico das universidades federais do Brasil, o DSpace é o programa mais usado pelas 50 instituições examinadas.

O programa do RD geralmente possui as seguintes operações e funcionalidades, segundo Leite (2009, p. 58):

- Capturar e descrever conteúdos digitais por meio de interface de autoarquivamento;
- Tornar público, por meio da Internet, o acesso a esses conteúdos (ou quando necessário ao menos aos seus metadados);
- Armazenar, organizar e preservar digitalmente conteúdos a longo prazo;
- Compartilhar os metadados com outros sistemas na Internet.

O DSpace,³ o foco da presente pesquisa, é um *software* livre de código aberto, flexível e personalizável, que aceita diversos formatos de conteúdo digital, como PDF, Word, Joint Photographic Experts Group (JPEG), Moving Picture Experts Group (MPEG) e TIFF. O programa já foi implantado em mais de 3.000 organizações, principalmente como RI em bibliotecas de pesquisa. No momento, a versão 7.4 está disponível no *site* para *download*.

O *software* foi desenvolvido por duas organizações, o Massachusetts Institute of Technology (MIT) e a Hewlett-Packard (HP). Não tem fins lucrativos ou comerciais para criar repositórios digitais livres, e pode ser adaptado de acordo com as necessidades da instituição (Borges *et al.*, 2019). Quanto a sua instalação,

[...] os sistemas operacionais recomendados são aqueles baseados em UNIX/Linux.
[...] Os bancos de dados apropriados são o Oracle (www.oracle.com) e o PostgreSQL (www.postgresql.org). O segundo, por ser gratuito, não acarreta ônus adicional para instituição (Blattmann; Weber, 2008, p. 474).

³ Disponível em: <https://dspace.lyrasis.org/features/>. Acesso em: 9 set. 2023.

O DSpace suporta os padrões comuns de interoperabilidade usados no domínio dos repositórios institucionais, como o Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), Open Archives Initiative Object Reuse and Exchange (OAI-ORE), SWORD, WebDAV, OpenSearch, OpenURL, RSS, ATOM. Versões mais recentes do DSpace também suportam a pesquisa facetada e a funcionalidade de navegação usando o Apache Solr (Sobral; Santos, 2017, p. 179).⁴

O programa pode realizar o gerenciamento dos mais diversos documentos digitais (como livros, teses, imagens, arquivos de áudio e vídeo, páginas da web e etc.), dando ênfase à proteção e à disponibilidade dos conteúdos arquivados ao longo do tempo, o que por si só é uma característica muito relevante, pois a rapidez com que os recursos eletrônicos “são implementados está em paridade com a possibilidade real de perdermos capacidade de recuperar a informação contida nos mais diversos suportes em que se encontra depositada” (Sobral; Santos, 2017, p. 168).

A organização dos objetos digitais no RD do Dspace apresenta a seguinte estrutura:

O conteúdo do repositório deve ser organizado em Comunidades, que são as entidades administrativas conhecidas como coleções. Cada Comunidade pode conter um número ilimitado de sub-comunidades. Cada sub-comunidade poderá ser uma coordenação de áreas temáticas, ou mesmo um núcleo ou grupo de ação e manifestação cultural que atue no repositório. Os documentos digitais serão denominados de Itens e estarão agrupados em Coleções. As Comunidades e Sub-Comunidades podem ter também um número ilimitado de Coleções. Os Itens serão indexados seguindo os padrões já pré-estabelecidos de metadados do próprio software pautado no Dublin Core. As Comunidades são mostradas em página própria e terão suas informações, notícias e links próprios, além de mostrar a lista com suas Sub Comunidades e Coleções (Blattmann; Weber, 2008, p. 479-480).

No DSpace, a “Comunidade” é o nível mais alto da hierarquia de conteúdo. Equivale à parte da instituição como departamento, laboratório, centro de pesquisa ou escola. Outrossim, é importante observar mais alguns detalhes técnicos, como: os metadados, incluindo informações de acesso e configuração, são guardados em um banco de dados relacional; o formato de metadados padrão para o DSpace é o Dublin Core (DC), mas qualquer campo pode ser adicionado ou modificado para se adequar a base conforme a necessidade; o *software* suporta todos os esquemas de nomes simples não hierárquicos, apesar de que é possível inserir outros esquemas hierárquicos de metadados (por exemplo, Machine Readable Cataloging – MARC e Metadata Object Description Schema – MODS) (Sobral; Santos, 2017).

A arquitetura modular do DSpace permite a criação de grandes repositórios multidisciplinares que, em última instância, podem ser expandidos através das

⁴ ATOM: Acesso à Memória; OpenURL: Open Uniform Resource Locator; RSS: Really Simple Syndication; SWORD: Simple Web Service Offering Repository Deposit; WebDAV: Web-based Distributed Authoring and Versioning.

fronteiras institucionais. [...] A interface do usuário final oferece suporte à navegação e pesquisa nos arquivos. Uma vez localizado um item, os arquivos formatados na web podem ser exibidos em um navegador, enquanto outros formatos podem ser baixados e abertos com um programa de aplicação adequado (Sobral; Santos, 2017, p. 177).

Observando os aspectos analisados pelos autores, conclui-se que o DSpace se popularizou no Brasil por ser um *software* livre, simples, oferecem diversas vantagens para a organização e para o usuário que utiliza o ambiente virtual para suas pesquisas. Uma característica muito importante a considerar, ao escolher o programa, é a capacidade de personalizá-lo de acordo com as necessidades da instituição.

5.2 Metadados e a interoperabilidade

A pergunta a se fazer é “O que são metadados?” Há várias definições para este termo na literatura, e aqui segue a explicação do National Information Standard Organization – NISO (2004, p.1, tradução nossa):

[...] são informações estruturadas que descrevem, explicam, localizam ou facilitam a recuperação, uso ou gerenciamento de um recurso de informação. Metadados são frequentemente chamados de dados sobre dados ou informações sobre informações.

Em outras palavras, os metadados conseguem representar a informação no meio digital para promover o encontro entre o conhecimento registrado e o usuário (Alves, 2005). Dessa forma, caracteriza-se como o instrumento que possibilita encontrar o documento digital desejado dentro da base de dados.

Inclusive, é possível fazer uma analogia entre os metadados usados no RD e a catalogação tradicional realizada na biblioteca, ou seja, são os dados necessários para efetivar a recuperação da informação no sistema, conforme pode ser observado na seguinte citação:

A catalogação tradicional organiza todo um conjunto de dados: autor, título, data, assunto, etc., armazenados em um sistema que contém as coleções de uma biblioteca. Os metadados correspondentes a estes dados são considerados componentes lógicos centrais dos esquemas utilizados em bibliotecas e repositórios digitais. Os esquemas de metadados garantem a efetiva busca, acesso e recuperação de recursos de informação no ambiente digital. Pode-se optar pela utilização de um esquema único de metadados ou de um perfil de aplicação onde diferentes elementos de um ou mais esquemas de metadados são selecionados e combinados num esquema composto (Pavão *et al.*, 2015, p. 104-105).

A padronização dos metadados vai possibilitar a troca de informações com outras instituições que usam o mesmo padrão, como também com organizações que usam outro modelo. “Isso é importante, pois além de diminuir o trabalho de descrição de recursos, permite que um usuário possa, em uma única pesquisa, buscar informações em diferentes instituições”

(Gracio, 2002, p. 25). Em resumo, o uso adequado de metadados melhora a recuperação de informações em bases digitais, aumentando a probabilidade de que os usuários localizem informações importantes para atender suas necessidades (Alves, 2010).

Há vários padrões de metadados, como DC, Learning Object Metadata (LOM), Visual Resources Association (VRA), Encoded Archival Description (EAD) e MARC (Pavão *et al.*, 2015). Conforme já foi mencionado anteriormente, DC é o modelo padrão do DSpace.

O DC é considerado simples, facilmente compreensível e utilizável por qualquer usuário do RD (Cerrao; Castro, 2018). Acrescentam-se também as considerações de Blattmann e Weber (2008) para esse padrão:

O padrão de metadados Dublin Core (DC) - <http://www.dublincore.org/> é o formato recomendado internacionalmente e o mais adotado pelo protocolo OAI-PMH, criado pela iniciativa dos arquivos abertos. Seus elementos permitem expressar o conteúdo de um documento digital de maneira bibliográfica e particularizá-lo dentre vários outros materiais armazenados. A importância dos metadados é justificada pela necessidade dos documentos disponíveis estarem codificados em formatos e linguagens de marcação padronizadas, de modo a facilitar a interoperabilidade, o acesso e a manutenção, a recuperação e localização eficaz do acervo digital para o usuário (Blattmann; Weber, 2008, p. 473).

Seguem os 15 metadados simples do DC para descrição de recursos na Web:

Quadro 6 - Esquema de metadados do DC

| Dublin Core (15 metadados simples) |
|---|
| Title: Título–título (dc.title) |
| Creator: Autor (dc.creator) |
| Subject: Assunto/ palavras-chave (dc.subject) |
| Description: Descrição (dc.description) |
| Publisher: Editor (dc.publisher) |
| Contributor: Contribuidor/colaborador (dc.contributor) |
| Date: Data (dc.date) |
| Type: Tipo do recurso (dc.type) |
| Format: Formato (dc.format) |
| Identifier: Identificador do recurso (dc.identifier) |
| Source: Fonte (dc.source) |
| Language: Idioma (dc.language) |
| Relation: Relação (dc.relation) |
| Coverage: Abrangência/ Cobertura (dc.coverage) |
| Rights: Gerenciamento de Direitos Autorais (dc.right) |

Fonte: *site* da DublinCore (<http://dublincore.org/>).

O quadro 6 apresenta o modelo padrão dos metadados do DC. Os elementos podem ser alterados, repetidos e excluídos, para adaptar a estrutura de acordo com as necessidades da instituição. De acordo com Pavão *et al.* (2015, p. 115), as

Ações cooperativas e a observância das normas e padrões na definição de esquemas e perfis de metadados para descrição de objetos digitais são imprescindíveis para proporcionar a uniformidade na descrição e a interoperabilidade entre repositórios [...].

Destaca-se como exemplo significativo o padrão de metadados para teses e dissertações desenvolvido pelo IBICT, denominado MTD3-BR. O documento correspondente pode ser encontrado anexado (ANEXO B).

Quanto à interoperabilidade, é vista como “[...] a capacidade de múltiplos sistemas com diferentes plataformas de hardwares e de softwares, estruturas de dados e interfaces para trocar dados com perda mínima de conteúdo e funcionalidade [...]” (NISO, 2004, p. 02, tradução nossa). De maneira resumida, é quando sistemas distintos conseguem se comunicar e compartilhar dados entre si (Rodrigues *et al.*, 2013).

Já Santarem Segundo, Silva e Martins (2019) compreendem a interoperabilidade como:

[...] a capacidade de organizações trocarem informações e dados entre si, independentemente dos sistemas e/ou plataformas que utilizem. Entende-se também a interoperabilidade como aspecto tecnológico fundamental para a evolução e desenvolvimento dos mais variados tipos de acervos digitais, permitindo integração de conteúdo de diversas fontes e promovendo a navegação de conteúdos de vários acervos, antes isolados, em ambientes integrados (Santarem Segundo; Silva; Martins, 2019, p. 68-69).

Para Sayão (2007), o conceito de interoperabilidade não é novo para as bibliotecas tradicionais. Sempre se soube que as unidades informacionais não são ilhas e que precisam da troca de informações, da criação de serviços colaborativos e do intercâmbio de documentos para cumprir sua missão.

Entretanto, com a consolidação e a concretização do conceito de bibliotecas digitais - que se localiza na interseção entre Biblioteconomia, Ciência da Computação e tecnologias de rede -, aliados ao crescente interesse da indústria de conteúdos nas formas de disseminação dos repositórios digitais como meio de distribuição de seus produtos no ambiente de uma nova economia da informação, a interoperabilidade se torna um foco de grande interesse para muitos atores (Sayão, 2007, p. 42).

No momento atual, a informação e o conhecimento são elementos fundamentais para contribuir com o progresso da humanidade. No entanto, a abundância de tecnologias recentes, a constante criação de conteúdo, as redes sociais, o paradigma do *big data*, o uso amplo de serviços em nuvem, as áreas de estudo ligadas às HD e todos os procedimentos de mudança de formatos de materiais do analógico para o digital, têm deixado evidente a necessidade de

resolver o desafio enfrentado pela humanidade de conseguir acessar e localizar informações de maneira mais eficaz. A integração passa a ser essencial nos procedimentos indispensáveis para converter informações em conhecimento de maior abrangência (Santarem Segundo; Silva; Martins, 2019).

Dessa forma, assegurar a interoperabilidade entre os repositórios digitais demanda a implementação de uma padronização dos metadados. Essa medida possibilita a troca de dados entre sistemas diversos, de modo a satisfazer, de maneira eficiente, as necessidades dos usuários no que se refere à busca por documentos.

5.3 Princípios para os repositórios digitais

5.3.1 As Cinco Leis de Ranganathan

Shialy Ramamrita Ranganathan, conhecido no Brasil como o pai da Biblioteconomia, nasceu na Índia, em 1892. Era politizado e preocupava-se com o ensino e a pesquisa em seu país. Graduou-se em Matemática e exercia a função de professor na Universidade de Madras, quando surgiu uma vaga na biblioteca da instituição em que trabalhava. Para ocupar o cargo de bibliotecário, especializou-se em Biblioteconomia no College University, em Londres. Quando iniciou seu novo ofício, o indiano começou a observar o complexo mundo da Biblioteconomia: “[...] mundo que se colocava entre aquele que produzia e registrava o conhecimento e aquele que necessitava de informação/conhecimento contido nos documentos” (Campos, 1999). Ao visitar cerca de 100 bibliotecas, ele encontrou diversas práticas biblioteconômicas, e ficou intrigado com os seguintes questionamentos:

[...] Será que todos esses agregados empíricos de informações e práticas não seriam redutíveis a um punhado de princípios fundamentais? Será que não se poderia adotar, neste caso, o processo indutivo? Não seria possível deduzir, a partir dos princípios fundamentais, todas as práticas conhecidas? [...] (Ranganathan, 2009, p. 2).

Após suas observações e sua experiência como bibliotecário, Ranganathan conseguiu idealizar, em 1928, as Cinco Leis da Biblioteconomia (Ranganathan, 2009). Elas são simples, claras e objetivas, conforme segue:

- 1^a Lei - Livros são para o uso;
- 2^a Lei - A cada leitor seu livro;
- 3^a Lei - A cada livro seu leitor;
- 4^a Lei - Economize o tempo do leitor;

5^a Lei - Uma biblioteca é um organismo em crescimento.

As Cinco Leis da Biblioteconomia são “[...] enunciados fortes e que podem atuar como norteadores das **práticas de toda e qualquer biblioteca** [...]” (Santos, 2015, p. 472, grifo nosso). Apesar do tempo da sua elaboração, ainda são vigentes. As leis almejam, segundo Fernandes, Pinto e Farias (2021, p. 112), “[...] uma aplicabilidade ainda maior que a contemplação filosófica adotada nas atividades, projetos, estruturação e organizações a serem planejadas dentro destes ambientes [...]”. Servem, no dizer de Campos, como orientação para todo o fazer do bibliotecário, em destaque as seguintes tarefas: “[...] seleção e aquisição; **administração de bibliotecas**; recuperação de informação; classificação e indexação; atendimentos aos usuários etc.” (1999, grifo nosso).

[...] são elas que permitem que o profissional possa compreender de uma forma mais abrangente a função de sua profissão, dentro de um contexto social, que permite definir critérios e princípios de ação que vão desde o posicionamento ético deste profissional até a escolha de métodos e técnicas para o seu fazer diário (Campos, 1999).

Segundo Panage (2008), a filosofia básica da Biblioteconomia e da profissão do bibliotecário é fundamentada pelas ideias do Ranganathan, e o conceito de biblioteca digital pode ser adequadamente compreendido ao aplicarmos suas leis.

Já se vê como as bibliotecas tradicionais são estruturadas, moldadas, constituídas e funcionam de acordo com os princípios do Dr. S.R. Ranganathan. A filosofia do Dr. Ranganathan é tão perspicaz e exaustiva em seu escopo que até mesmo novas tendências e técnicas são assimiladas ou acomodadas e explicadas com sua filosofia. Hoje em dia todo mundo fala muito sobre as novas tendências em Biblioteconomia, a saber, biblioteca digital e digitalização. É um esforço por parte do pesquisador assumir a tarefa do sistema de biblioteca digital no contexto da filosofia do Dr. Ranganathan (Panage, 2008, p. 274, tradução nossa).

A citação acima sugere que as bibliotecas tradicionais são fortemente influenciadas pelos princípios e pela filosofia de Ranganathan – tão abrangente que até mesmo as novas tendências e técnicas, como bibliotecas digitais e digitalização, são incorporadas e explicadas por ela. Isso indica que a filosofia de Ranganathan é duradoura e ainda relevante no campo da Biblioteconomia, mesmo diante das mudanças tecnológicas e das novas abordagens na gestão da informação. Nesse sentido, Panage (2008) assume o desafio de integrar o conceito de biblioteca digital e digitalização dentro do quadro conceitual estabelecido por Ranganathan, sugerindo que sua filosofia sirva como uma teoria sólida para entender e adaptar as bibliotecas às demandas da era digital.

Para Sousa e Targino (2016), as Cinco Leis de Ranganathan se encadeiam e se complementam. Os autores sugerem que os preceitos do indiano permanecem como

fundamentais para aqueles que enxergam, na Biblioteconomia, uma oportunidade única de exercer a cidadania e lutar pelo acesso universal à informação. Isso possibilita oferecer aos cidadãos informações adequadas às suas necessidades informacionais, em um momento histórico marcado pela presença marcante e irreversível das inovações tecnológicas.

Em 1984, Rajagopalan e Rajan conseguiram fazer uma adaptação das Leis de Ranganathan para a CI. É uma interpretação atualizada que está de acordo como o desenvolvimento de modernos serviços e sistemas de informação, cujo foco centraliza-se no uso e no usuário da informação, conforme pode ser observado:

1. a informação é para o uso; 2. a cada usuário sua informação; 3. cada informação a seu usuário; 4. economize o tempo do usuário – e o seu corolário: economize o tempo dos cientistas da informação; 5. um sistema de informação é um organismo em crescimento (*apud* Figueiredo, 1992, p. 189).

Nota-se que a nova versão substitui a palavra “livro”, do original de Ranganathan, por “informação”, ou seja, o conhecimento registrado, independentemente do seu suporte, impresso ou digital (livro, *e-book*, PDF, periódico eletrônico, CD-ROM e etc.).

Assim, pode-se argumentar que as Cinco Leis são um guia para o profissional da informação, especialmente o bibliotecário, na gestão da unidade/sistema de informação (tradicional ou digital).

A partir do que foi mencionado, questiona-se: as leis podem servir como princípios operacionais para a gestão de um RD? Então, conforme já foi dito na seção 2 da presente dissertação, por Rio Riande (2018), o RD é apontado como um objeto de fronteira entre a CI e as HD. Considerando que as leis de Ranganathan são para “qualquer” tipo de biblioteca, é possível fundamentar a sua aplicação como parâmetro para a gestão do RD. Dessa forma, o presente estudo propõe uma reinterpretação das Cinco Leis de Ranganathan com o intuito de encontrar diretrizes para a gestão ideal do RD.

Para analisar a primeira lei – os “livros são para serem usados” – no primeiro momento, é necessário pensar no período em que ela foi elaborada (início do século XX), quando as bibliotecas não eram organizações focalizadas na promoção do uso do livro, e sim na sua preservação (Ranganathan, 2009). A informação não estava disponível para todos, e o pensador indiano tentou mudar essa realidade, promovendo o acesso ao conhecimento:

[...] Ranganathan discute questões que irão envolver a democratização da informação, pois o que faz com que a instituição biblioteca exista é o fato do homem, ao desvendar o mundo, ao trocar experiências sobre suas descobertas e ao comunicar estas descobertas e avanços para possibilitar a transmissão de conhecimento, elabora registros, inscrições. Estes devem estar organizados, armazenados e preservados para propiciar a transmissão de conhecimento para a geração futura (Campos, 1999).

Na visão de Rajagopalan e Rajan (1984 *apud* Figueiredo, 1992, p. 189), a primeira lei – “a informação é para o uso” – trata das etapas do processo de transferência de informação, considerando que o seu teor é um recurso básico. “[...] Sendo um recurso de tal importância, deve ser acoplado, explorado e utilizado a fim de dar origem a benefícios plenos. O papel da informação no desenvolvimento nacional é percebido nesta lei [...].”

Segundo Henning *et al.* (2019, p. 390), “[...] o conhecimento científico é patrimônio da humanidade e, portanto, deve ser disponibilizado livremente para que todos os cidadãos possam acessá-lo livremente, sem quaisquer restrições econômicas ou legais.” Destaca-se, então, o papel do bibliotecário como um grande dinamizador, que vai possibilitar que os livros/documentos/informações sejam usados, não apenas armazenados (Campos, 1999). Corroborando essa posição, Fernandes, Pinto e Farias (2021) também destacam a atuação desse profissional no cumprimento da lei, na efetivação do uso da informação pelo usuário. Ou seja, seu papel no tratamento dos registros de conhecimento para promover o seu uso.

Observa-se, portanto, a necessidade de considerar todos os meios necessários para efetivar o uso de documentos eletrônicos no RD, como, por exemplo, implementar a "curadoria digital" como alternativa para organizar os itens na base de dados, a fim de facilitar a busca de informações.

A curadoria digital se fortalece a cada dia como ferramenta aliada às tecnologias para uma nova maneira de gerir informação estruturada que descreve, explica, localiza ou facilita a recuperação. Observa-se que sua metodologia já é adotada em diversos países para organizar e preservar os acervos de memória e cultura, como também em sites de naturezas diversas (Martins, 2021, p. 82).

Seguindo para segunda lei – “a cada leitor seu livro” (ou “a cada usuário sua informação”) –, é possível entender que após falar do uso da informação, Ranganathan destaca o usuário.

O pai da Biblioteconomia era um grande observador das necessidades humanas e entendia que nem todos precisam das mesmas informações, que o princípio da diferença é o que permite que as pessoas estejam em diversas cadeias de desenvolvimento ou em diferentes estados socioculturais. Portanto, possibilitar que cada usuário obtenha a sua informação é garantir que todo homem deve ter acesso ao conhecimento. “[...] O importante aqui é permitir a acessibilidade à informação - naquele momento, o livro -, a cada leitor/usuário, seja ele real ou virtual [...]”. (Campos, 1999). “[...] Assim, a revolução causada pela Segunda Lei é de natureza mais avançada e aproxima mais a humanidade de seu objetivo” (Ranganathan, 2009, p. 50).

Dessa maneira, a segunda lei “[...] evidencia a necessidade das bibliotecas conhecerem o público alvo, a comunidade de usuários que envolvem cada tipo de biblioteca [...]” (Fernandes; Pinto; Farias, 2021, p. 113). Considerando as unidades informações no século XXI, Sousa e Targino (2016, p 18) têm a seguinte observação:

Verdade que as bibliotecas do século XXI vivenciam um ambiente de informação distinto da época da formalização das Leis da Biblioteconomia / Leis de Ranganathan. Nos dias de hoje, a informação é mediada pelo uso das TIC, o que altera o cenário das bibliotecas em geral, sem, porém, transformar a essência dos preceitos, porque cada leitor mantém demandas específicas [...].

Na atualidade, os pesquisadores preferem buscar as informações no ambiente virtual, ao invés de procurar a unidade informacional. Eles consideram

[...] a comodidade oferecida pelas inovações tecnológicas aliada a uma série de fatores inerentes à sociedade contemporânea, em especial, nas grandes cidades: insegurança das ruas, trânsito caótico, acúmulo de atividades, redução de tempo, custo de deslocamento, etc. (Sousa; Targino, 2016, p. 19).

Mas, tanto em ambientes físicos quanto virtuais, a segunda lei enfatiza a utilidade de conhecer seu público-alvo para melhor atender às suas necessidades. Esta não é uma tarefa fácil, conforme pode ser observado na citação abaixo:

Atualmente existe uma dificuldade em identificar singularmente os usuários finais, já que são os mesmos produtores, mediadores, organizadores e agentes dos diferentes usos que se dão nos processos de apropriação. Por exemplo, no âmbito da produção do conhecimento, os usuários recorrem a diferentes fontes de informação para subsidiar as suas pesquisas e, ao mesmo tempo, atuam como mediadores no tratamento da informação, elaborando resumos, determinando os tópicos e fixando as palavras-chave (Nhacuongue; Dutra, 2016, p. 143).

No RD, o usuário pode se registrar, ou seja, criar o seu perfil na base de dados. O cadastro beneficia tanto a instituição, que passa a conhecer o seu público, quanto o usuário, que recebe em troca um atendimento personalizado. Portanto, os pesquisadores devem ser incentivados a se registrar no RD. No caso da BDTD/UFRRJ, o campo "Meu espaço" é destinado ao cadastro do usuário e acesso aos serviços oferecidos pelo sistema.

A terceira lei – “a cada livro seu leitor” (ou “cada informação a seu usuário”) – refere-se à divulgação do conhecimento. Trata-se do complemento da lei anterior e enfatiza que é necessário “[...] tornar os recursos informacionais conhecidos dos usuários o mais rápido possível [...]” (Sousa; Targino, 2016, p. 21). Em vista disso, o serviço de Disseminação Seletiva da Informação (DSI) pode funcionar como um perfeito método para fazer a mediação entre o documento digital, disponível no repositório, e o usuário que necessita dele, direcionando a

informação de acordo com o perfil previamente cadastrado na base, conforme dito anteriormente.

Para Lunh (1961, p. 132 *apud* Eirão, Cunha, 2012, p. 64), a DSI significa

[...] aquele serviço dentro de uma organização que se refere à canalização de novos itens de informação, vindos de quaisquer fontes, para aqueles pontos onde a probabilidade de utilização, em conexão com o interesse corrente do usuário, seja alta.

Ou seja, é um serviço personalizado que direciona informações ao pesquisador e funciona como um alerta sobre novidades de seu interesse. O usuário cadastrado no RD recebe atualizações por *e-mail* sobre os documentos cadastrados na base que sejam do seu interesse. Assim, um serviço que, de acordo com a quarta lei, também economiza o tempo do usuário.

Prosseguindo para a quarta lei – “Economize o tempo do leitor” (ou “economize o tempo do usuário”, e o seu corolário “economize o tempo dos cientistas da informação”) –, coloca-se em evidência a questão “tempo” nas pesquisas. De acordo com Ranganathan, “[...] o homem atual precisa estar informado sobre a evolução do conhecimento de sua área de interesse em menor tempo possível” (Campos, 1999). Isso significa, segundo Fernandes, Pinto e Farias (2021, p. 115), que “[...] a biblioteca não deve economizar esforços para que a informação seja disponibilizada quando e onde o usuário necessitar, pois caso contrário, a suposta oferta deixa de atender o usuário e, conseqüentemente, perde seu sentido”.

Ranganathan (2009, p. 211) expõe que

[...] talvez o método mais conveniente de estudar as conseqüências desta lei seja acompanhar um leitor desde quando ele entra na biblioteca até o momento em que sai, examinando criticamente cada processo pelo qual ele passa, prestando atenção na economia de tempo [...].

No caso do RD, quanto tempo o usuário permanece no ambiente virtual até encontrar a informação desejada?

Como, então, poupar o tempo do usuário no momento da pesquisa? Em Espírito Santo (2014, p. 100) encontra-se o seguinte esclarecimento:

A coleção deve ser organizada visando às possibilidades de recuperação. O ponto central desta lei é a organização/recuperação do acervo. Ela trata dos recursos e ferramentas apropriados para tratar, disponibilizar e acessar cada tipo de acervo. Esta é a principal tarefa da gestão e organização da informação: criar elementos que possam fazer com que a informação encontre o seu usuário no menor tempo possível. [...] Trazendo a Lei para o cenário atual, implica em constante atualização na formação do profissional de informação, pois com a evolução das Tecnologias de Informação e Comunicação, há necessidades de se avaliar qual o instrumento mais apropriado para tratar o acervo, de modo a oferecer um atendimento mais eficaz a cada tipo de público.

Para conseguir presteza no atendimento das demandas dos usuários é necessário retornar para o assunto “organização/recuperação”, conforme disse Espírito Santo (2014). Pensar na arquitetura da informação no RD, como, por exemplo, o uso de taxonomias para diminuir a quantidade de cliques do mouse para encontrar a informação, ou seja, possibilitar a busca pela informação de forma mais rápida e eficiente. A autora também defende a necessidade de capacitar os profissionais da informação para acompanhar a evolução das tecnologias. Como a questão é o tempo, é imprescindível se manter atualizado sobre as inovações no processamento da informação que podem ser aplicadas no RD.

Avançando para a última lei – “uma biblioteca é um organismo em crescimento” (ou “um sistema de informação é um organismo em crescimento”) –, tem-se o entendimento de que uma unidade/sistema informacional estará sempre em desenvolvimento. Para Ranganathan (2009, p. 241), sobre o crescimento da biblioteca, “[...] é um fato biológico indiscutível que somente o organismo que se desenvolve é o que sobrevive. Um organismo que pare de se desenvolver acabará por se paralisar e perecer [...]”. Ou seja, tanto a biblioteca tradicional como a digital receberão sempre novos depósitos de documentos em sua base. É um processo que não tem fim.

Por que a biblioteca/RD é um organismo em crescimento? Em Campos (1999) encontra-se a seguinte justificativa: “[...] a produção de conhecimento é um ato contínuo e dinâmico do ser humano [...]. Isto exige a todo momento, um repensar sobre as práticas e instrumentos utilizados e sobre as atividades realizadas”. Ou seja, o profissional da informação e a biblioteca (física ou digital) devem estar preparados para acompanhar esse desenvolvimento.

A quinta lei se conecta totalmente aos processos de gestão, pois passa a ideia de que a biblioteca não pode ser um espaço estagnado. Muito pelo contrário, deve ser um ambiente que acompanhe as mudanças nas sociedades, buscando sempre inovações, principalmente com o advento das novas tecnologias [...] (Fernandes; Pinto; Farias, 2021, p. 115).

Os autores também fazem uma analogia com a situação do mercado: “[...] assim como o mercado sofre transformações, as bibliotecas devem a cada dia se adaptar as novas tendências contemporâneas essenciais para a promoção do acesso” (Fernandes; Pinto; Farias, 2021, p. 116). Portanto, o sistema de informação deve explorar novas tecnologias digitais e buscar uma gestão otimizada. Entende-se que um RD atualizado, organizado e padronizado, gerenciado por profissionais capacitados, certamente estará pronto para ser ampliado.

Levando em consideração as observações dos autores, não há dúvida de que os princípios das Leis de Ranganathan são relevantes para orientar as ações realizadas em um RD. Sua essência permanece inalterada, mesmo quase um século após sua formulação, e seguir esses

princípios na gestão de um RD pode ser a chave para o seu sucesso. No entanto, é possível adaptar esses princípios para atender às necessidades das HD? Como ficaria? O documento digital é para ser baixado/ visualizado? A cada pesquisador a sua informação eletrônica? A cada informação eletrônica o seu pesquisador? Economize o tempo do pesquisador? O RD é uma base de dados em crescimento? Em resumo, ao aplicar os princípios ao repositório, é possível construir um ambiente virtual eficaz para atender às demandas da comunidade acadêmica e científica.

5.3.2 Princípios FAIR

O movimento FAIR iniciou-se em 2014, e dois anos depois, seus princípios foram publicados. No período, houve um aumento considerável no interesse global em melhorar o processamento de dados e outros objetos digitais, para torná-los mais facilmente encontráveis, acessíveis, interoperáveis e reusáveis (Santos, 2021).

Segundo Wilkinson *et al.*(2016), os princípios FAIR (acrônimo para Findable, Accessible, Interoperable e Reusable) têm como objetivo principal aperfeiçoar a capacidade das máquinas de localizar e usar dados automaticamente, como também proporcionar a sua reutilização pelos indivíduos.

Os princípios enfatizam a capacidade de ação da máquina (ou seja, a capacidade dos sistemas computacionais de encontrar, acessar, interoperar e reutilizar dados com nenhuma ou mínima intervenção humana) porque os humanos dependem cada vez mais do suporte computacional para lidar com dados como resultado do aumento no volume, complexidade e velocidade de criação de dados. [...] Os princípios referem-se a três tipos de entidades: dados (ou qualquer objeto digital), metadados (informações sobre esse objeto digital) e infraestrutura [...] (Go-Fair, [ca. 2016], tradução nossa).

Os princípios FAIR foram originalmente criados como uma guia para a gestão de dados de pesquisa, mas suas premissas podem ser adotadas em outros tipos de repositórios digitais, visto que apresentam elementos norteadores para a melhoria dos serviços prestados em qualquer tipo de base de dados. Sales (2021a, p. 92) descreve bem essa ideia quando fala que os “princípios FAIR devem ser aplicados não apenas a dados de pesquisa, mas a repositórios e outros objetos digitais”.

As orientações do FAIR têm sido amplamente divulgadas pela comunidade internacional nos últimos anos e chegaram recentemente ao Brasil como mais um campo de pesquisa e aplicação por gestores de dados. As discussões ocorreram para garantir o máximo reaproveitamento de informações científicas, “a partir da adoção de padrões, metadados,

vocabulários controlados, ontologias e identificadores persistentes que proporcionam significado preciso aos dados e aos demais objetos a eles vinculados” (Henning *et al.*, 2018, p. 4).

Segue o quadro com os princípios FAIR:

Quadro 7 - Princípios FAIR

| | | |
|-----------------------|---|--|
| Encontráveis | O primeiro passo para (re)usar dados é encontrá-los. Metadados e dados devem ser fáceis de encontrar tanto para humanos quanto para computadores. Os metadados legíveis por máquina são essenciais para a descoberta automática de conjuntos de dados e serviços, portanto, esse é um componente essencial do processo <i>FAIRification</i> . | <p>F1 – (Meta) dados são atribuídos a um identificador globalmente único e persistente;</p> <p>F2 – Os dados são descritos com metadados avançados (definidos por R1 abaixo);</p> <p>F3 – Os metadados incluem clara e explicitamente o identificador dos dados que descrevem;</p> <p>F4 – (Meta)dados são registrados ou indexados em um recurso pesquisável.</p> |
| Acessíveis | Uma vez que o usuário encontra os dados necessários, é necessário saber como eles podem ser acessados, possivelmente incluindo autenticação e autorização. | <p>A1 – (Meta)dados são recuperáveis por seu identificador usando um protocolo de comunicação padronizado;</p> <p>A1.1 – O protocolo é aberto, gratuito e universalmente implementável;</p> <p>A1.2 – O protocolo permite um procedimento de autenticação e autorização, quando necessário;</p> <p>A2 – Os metadados são acessíveis, mesmo quando os dados não estão mais disponíveis.</p> |
| Interoperáveis | Dados geralmente precisam ser integrados com outros dados. Além disso, os dados precisam interagir com aplicativos ou fluxos de trabalho para análise, armazenamento e processamento. | <p>I1 – (Meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representação do conhecimento;</p> <p>I2 – (Meta)dados usam vocabulários que seguem os princípios FAIR;</p> <p>I3 – (Meta)dados incluem referências qualificadas a outros (meta)dados.</p> |
| Reutilizáveis | O objetivo final do FAIR é otimizar a reutilização de dados. Para conseguir isso, metadados e dados devem ser bem descritos para que possam ser replicados e/ou combinados em diferentes configurações. | <p>R1 – (Meta)dados são ricamente descritos com uma pluralidade de atributos precisos e relevantes;</p> <p>R1.1 – (Meta) dados são liberados com uma licença de uso de dados clara e acessível;</p> <p>R1.2 – (Meta)dados são associados com proveniência detalhada;</p> <p>R1.3 – (Meta) dados atendem aos padrões da comunidade relevantes para o domínio.</p> |

Fonte: Go-Fair, [ca. 2016], tradução nossa.

No primeiro, *Findable* (encontráveis), é importante ressaltar o papel dos identificadores persistentes únicos para facilitar a recuperação de dados e informações no ambiente virtual. Exemplos desses recursos são o Identificador de Objeto Digital (DOI) e o ORCID, que já estão sendo amplamente utilizados no meio acadêmico e por diversas editoras de revistas eletrônicas, tanto no Brasil quanto no mundo.

A ideia por trás de um sistema de identificadores persistentes (sistema PID) é oferecer uma referência duradoura a uma entidade (física, digital ou abstrata), por exemplo, um documento digital, site web, pessoa ou instituição. Alguns sistemas PID bem conhecidos são *Archival Resource Key* (ARK), *Digital Object Identifier* (DOI), *Handle system*, *Persistent Uniform Resource Locator* (PURL), *Uniform Resource Name* (URN) e *Open Researcher and Contributor ID* (ORCID iD), sendo esse último exclusivamente para pessoas (Marín-Arraiza; Heredia, 2021, p. [29]).

Dessa forma, enquanto o ORCID funciona como padrão internacional para identificação persistente dos autores (Marín-Arraiza; Heredia, 2021), os identificadores de documentos, como o DOI, o International Standard Book Number (ISBN) e o International Standard Serial Number (ISSN), são ferramentas importantes tanto para o controle bibliográfico como também para legitimação da recuperação da informação (Coelho; Santos, 2016).

Prosseguindo para o segundo princípio, *Accessible* (acessível), após encontrar um dado/informação, o próximo passo é a sua obtenção. Segue o posicionamento dos autores Sayão e Sales, sobre a questão da acessibilidade:

O princípio da acessibilidade preconiza que os objetos de pesquisa sejam acessíveis preferencialmente por meios da implementação, quando apropriados, de protocolos automatizados de recuperação de dados; estipula também que os dados estejam disponíveis de acordo com procedimentos e condições claras e bem definidas. Estas condições envolvem o estabelecimento de processos de autenticação e autorização que estejam alinhados às políticas da organização e à cultura disciplinar, e ainda, às especificidades dos dados – por exemplo, nível de sensibilidade. Como mantra FAIR, que deve ser trabalhado pelos serviços de gestão de dados, temos que os metadados devem estar incondicionalmente acessíveis mesmo que os dados não estejam, ou deixaram de estar disponíveis (Sayão; Sales, 2021, p. 228).

Sobre o tema *Interoperable* (interoperável), que já foi tratado na seção 5.2, Sales (2021a, p. 40) ressalta que “os dados e os metadados devem estar em conformidade com formatos e padrões reconhecidos para permitir que sejam combinados e trocados entre sistemas”. Segundo a autora, “os dados só atingirão seu potencial máximo de reuso, gerando inovação, se puderem transitar entre plataformas” (2021b).

Para que se possa representar dados e metadados devem ser adotadas linguagens de representação do conhecimento que sejam padronizadas, acessíveis e amplamente aplicáveis. (Ex: RDF, XML, DICOM, etc.) [...] Dados e metadados devem possuir

referências a vocabulários e/ou ontologias que os descrevem. Devemos garantir que esses também sigam os princípios FAIR (Sales, 2021b).⁵

Para assegurar a capacidade de pesquisa automática e a interoperabilidade dos conjuntos de dados, é necessário usar

[...] (1) vocabulários controlados, ontologias e tesouros, fazendo uso de identificadores globais únicos, persistentes e identificáveis [...]; e (2) um bom modelo de dados e uma estrutura bem definida para descrever e estruturar os dados e metadados (Henning *et al.*, 2019, p. 402).

Quanto ao último, *Reusable* (reutilizável), o objetivo final do FAIR depende do cumprimento dos primeiros princípios para efetivar o reuso dos dados/informações. Por outro lado, “os dados devem estar em conformidade com as normas da comunidade e ter uma licença clara para que outros saibam quais tipos de reuso são permitidos” (Sales, 2021b).

Para Curty (2019, p. 180),

[...] diferentes modelos buscam representar o ciclo de vida de dados de pesquisa e preveem o reuso como uma ação desejada resultante dos esforços de curadoria e de gestão desses ativos científicos, assim como uma etapa que retroalimenta novas ideias e processos de pesquisa.

Dessa forma, de acordo com Campos, Dias e Sousa (2023, p.18-19),

[...] a encontrabilidade subsidia a acessibilidade, a inserção de metadados contribuindo para a interoperabilidade e ao uso e ao reuso dos dados mediante aos recursos de representação mediante os instrumentos de controle terminológico.

Consequentemente, os repositórios digitais podem ser considerados FAIR, segundo Sales (2021), quando se ajustam aos seus quatro princípios

[...] e abrigam dados, dados de pesquisa ou outros objetos digitais a partir de práticas de gestão que tornam seus objetos localizáveis, acessíveis, interoperáveis e reusáveis respeitando as questões éticas, legais de compartilhamento.

Dessa forma, os Princípios FAIR desempenham um papel importante na gestão de dados no RD, e têm o potencial de oferecer resultados satisfatórios no serviço prestado pelo sistema de informação. Ressalta-se, assim, a importância de adicionar Identificadores Persistentes (Persistent Identifier – PID) e atenção especial ao uso de metadados para viabilizar a recuperação de informações no ambiente virtual. Portanto, é oportuno se esforçar para alcançar a "FAIRificação" do RD.

⁵ DICOM: Digital Imaging and Communications in Medicine; RDF: Resource Description Framework; XML: Extensible Markup Language.

A próxima seção discorre sobre a BDTD da UFRRJ, abordando a coleção de teses e dissertações que em breve integrará o RIMA. Será apresentada a implementação de alguns conhecimentos previamente mencionados nesta dissertação, como as orientações do FAIR, evidenciando a aplicação prática dessas diretrizes.

6 BDTD DA UFRRJ: PESQUISA-AÇÃO

A UFRRJ⁶ tem uma trajetória que remonta à criação da Escola Superior de Agricultura e Veterinária (ESAMV) em 20 de outubro de 1910, o que significa que possui uma história de mais de um século. Embora tenha se estabelecido com foco no ensino agropecuário, a instituição tem se empenhado em diversificar seu perfil, oferecendo atualmente cursos em todas as áreas do conhecimento. A UFRRJ é vinculada à Secretaria de Educação Superior do Ministério da Educação e faz parte do sistema das Instituições Federais de Ensino Superior (IFES).

Segundo seu Catálogo Institucional⁷ de 2021, a universidade tem o seguinte quadro: são quatro *campi* (Seropédica, Nova Iguaçu, Três Rios e Campos dos Goytacazes), 56 cursos de graduação, dois cursos de graduação a distância, 29 cursos de mestrado acadêmico, oito cursos de mestrado profissional e 17 cursos de doutorado. Atualmente, a UFRRJ possui oficialmente as seguintes unidades informacionais: Biblioteca Central (Seropédica); Biblioteca Instituto Multidisciplinar de Nova Iguaçu (IMNI), em Nova Iguaçu; Biblioteca Instituto Três Rios (ITR), em Três Rios; Biblioteca Ivan de Otero Ribeiro, do Programa de Pós-Graduação de Ciências Sociais em Desenvolvimento, Agricultura e Sociedade (CPDA/RJ); e Biblioteca Campos dos Goytacazes (Campos dos Goytacazes).

A SPT, localizada na BC, é a responsável pela gestão da Biblioteca Digital de Teses e Dissertações da instituição, e também é encarregada pelo tratamento do acervo físico, que inclui livros, revistas, dissertações, teses, documentos multimídia e outros materiais. Em resumo, a missão da SPT é processar os documentos recebidos em sua unidade, independentemente de seu formato físico, a fim de disponibilizá-los para os usuários, tanto internos quanto externos.

A SPT desempenha um papel estratégico no gerenciamento do acervo da biblioteca, garantindo que os documentos sejam tratados de forma adequada e estejam prontos para uso. Isso envolve a catalogação, classificação, indexação e armazenamento dos materiais físicos, bem como o processamento dos documentos digitais. Dessa forma, a Seção assegura que os usuários tenham acesso eficiente aos recursos disponíveis na instituição.

As teses e dissertações da UFRRJ, quando estavam disponíveis apenas no seu formato impresso, apesar do seu acesso ser livre, ficavam limitadas ao espaço da BC, já que não é permitido o empréstimo desse tipo de material bibliográfico. Essa realidade mudou com o

⁶ Disponível em: <https://institucional.ufrj.br/ccs/historia-da-ufrj/>. Acesso em: 9 set. 2023.

⁷ Disponível em: <https://institucional.ufrj.br/ccs/catalogo-institucional-da-ufrj-edicoes-2021-e-2019/>. Acesso em: 9 set. 2023.

avanço das tecnologias digitais das últimas décadas. Atualmente, a divulgação dos conhecimentos produzidos pela universidade tem sido ampliada com o auxílio de novas ferramentas de processamento da informação.

Segundo Assis (2015, p. 21),

O avanço das tecnologias de informação possibilitou que as teses e dissertações possuam versão eletrônica e que sejam gerenciadas em sistemas de informação de acesso aberto, passando a ser mais amplamente disseminadas e acessíveis à comunidade científica.

Para a autora, quando as teses e dissertações estão disponíveis apenas na versão impressa, ficam com a circulação limitada, e o seu impacto é reduzido.

Na medida em que a tecnologia passou a ser utilizada na produção e, especialmente, na disseminação de resultados de pesquisa, as teses e dissertações tornaram-se canais de comunicação que veiculam, antes mesmo da literatura científica publicada, resultados de pesquisa. Isto foi possível graças ao surgimento de iniciativas de sistemas de informação digitais na internet para a obtenção, organização, armazenamento, preservação, recuperação e disseminação de teses e dissertações em todo o mundo. Em maior parte, esses sistemas de informação diziam respeito às bibliotecas digitais de teses e dissertações (Assis, 2015, p. 19).

No Brasil, um grande avanço para a disseminação desse tipo de material bibliográfico foi a criação da BDTD pelo IBICT. Baseado no modelo operacional da Networked Digital Library of Theses and Dissertations – NDLTD (banco de dados que reúne teses e dissertações eletrônicas de diversas instituições ao redor do mundo) e utilizando o protocolo Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a BDTD foi iniciada em 2002. A plataforma reúne as dissertações e teses eletrônicas defendidas nas instituições de ensino e pesquisa brasileiras e por brasileiros no exterior, disponibilizando gratuitamente ao público o acesso a seus textos completos para *download* (Assis, 2015). A Biblioteca Digital funciona da seguinte forma:

A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) integra e dissemina, em um só portal de busca, os textos completos das teses e dissertações defendidas nas instituições brasileiras de ensino e pesquisa. O acesso a essa produção científica é livre de quaisquer custos. [...] A Biblioteca Digital Brasileira de Teses e Dissertações segue os preceitos da Iniciativa de Arquivos Abertos (OAI), adotando assim, o modelo baseado em padrões de interoperabilidade. Assim, as instituições de ensino e pesquisa atuam como provedoras de dados e o Ibict/MCTI opera como agregador: coleta os metadados das teses e dissertações dos provedores (instituições), fornece serviços de informação sobre esses metadados e os expõem para coleta para outros provedores de serviços. Importante destacar que a BDTD é também coletada pela *Networked Digital Library of Theses and Dissertations* (NDLTD) que é a rede internacional que reúne teses e dissertações de todo o mundo (Biblioteca Digital, 2020).

Já o TEDE⁸ é um programa elaborado pelo IBICT e distribuído gratuitamente para instituições de ensino que desejam utilizar a solução para o gerenciamento de suas dissertações e teses na BDTD. Em 2014, foi disponibilizada a mais recente edição do TEDE2, que é uma customização do programa DSpace.

Conforme já foi mencionado na seção introdutória, a UFRRJ aderiu ao BDTD em 2005, e passou a depositar os arquivos PDF dos trabalhos finais dos programas *stricto sensu* da sua instituição em sua plataforma.

O processo de implantação da Biblioteca Digital de Teses e Dissertações no âmbito ruralino teve seu início por meio do edital Funcate/ Ibict nº 001/2005, de dezembro de 2005. A partir da realização de um workshop oferecido nos dias 10 a 12 de abril nas instalações do instituto, foram distribuídos kits tecnológicos para implantação da biblioteca digital e ministrado curso de capacitação para uso do TEDE (Biblioteca Digital, 2006 *apud* Oliveira, 2021, p. 86).

No momento, a sua plataforma disponibiliza 6.655 documentos para *download* (5.358 dissertações e 1.297 teses), do período de 1967 a 2022. São trabalhos de conclusão dos cursos de mestrado e doutorado da instituição, de diversas áreas do conhecimento, inclusive Ciências Humanas e Sociais.

A busca por documentos na BDTD da UFRRJ é facilitada por meio de diversos critérios, como título, autor, orientador, membros da banca, áreas do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), tipo de documento e ano de defesa, entre outros. O sistema de busca está organizado em "Comunidades" e "Coleções". A principal comunidade, denominada *Campus Seropédica*, é subdividida em subcomunidades correspondentes aos programas de pós-graduação existentes, cada uma com suas respectivas coleções. Essas coleções, por sua vez, são fragmentadas em programas de mestrado (acadêmico e profissional) e doutorado.

A BDTD/UFRRJ integra o movimento de acesso aberto da produção científica, permitindo que qualquer pessoa possa consultar e fazer *download* das teses e dissertações disponíveis em seu acervo.

Segue a página principal do repositório:

⁸ Disponível em: <https://bdttd.ibict.br/vufind/Content/tede>. Acesso em: 11 set. 2023.

Figura 2 – BDTD/UFRRJ



Fonte: *site* do TEDE (<https://tede.ufrj.br/>).

Para contemplar outros documentos digitais de interesse da instituição, como o estudo sobre a história do trabalho no Brasil, a Rural desenvolveu um importante RD, a saber: o RIMA. Sob a coordenação do docente Alexandre Fortes, o projeto tem como objetivo a “[...] difusão de acervos reunidos sobre a temática e compostos pelos mais variados registros. Sejam eles textuais, sonoros, bibliografia comentada” (Oliveira, 2021, p. 87). Com a ampliação do seu acervo e inclusão de novas comunidades, inclusive a coleção de teses e dissertações, o RIMA recebeu um novo nome, sem alterar a sua sigla: Repositório Institucional de Múltiplos Acervos. O repositório da UFRRJ está em constante aperfeiçoamento dos serviços prestados.

Na atualidade, a Universidade Rural está preparando a fusão do banco de dados da BDTD/UFRRJ para o RIMA, o que significa que, em breve, a instituição terá apenas o RI. Portanto, é imprescindível atuar desde já em busca de ajustes no tratamento das teses e dissertações na base, antes da sua migração para o RI.

Contudo, qual é a diferenciação entre esses dois sistemas de informações? De acordo com Assis (2015), enquanto os repositórios digitais são espaços que têm como objetivo o armazenamento da produção científica como um todo, sobretudo os artigos científicos, as bibliotecas digitais de teses e dissertações são voltadas para a administração exclusiva desses documentos. Os repositórios institucionais “[...] podem, com facilidade, assumir as funções das bibliotecas de T&D, enquanto estas, na maior parte dos casos, estão tecnologicamente impossibilitadas de se transformarem em RI” (Leite, 2011a).

Em outras palavras, não é possível inserir outros tipos de material bibliográfico na BDTD/UFRRJ, mas os dois sistemas possuem o mesmo formato e compartilham dos mesmos objetivos.

Tendo em vista que tanto os RI quanto as bibliotecas digitais de T&D constituem elementos de um sistema de gestão e comunicação da informação científica que não estão restritos às instituições ou países, a compreensão de seus problemas e as respostas a eles deve ocorrer sob uma perspectiva sistêmica. Por esta razão, nenhum esforço para tal deve considerar como parâmetro único as condições locais, sobretudo aquelas relacionadas com as particularidades referentes às infraestruturas tecnológicas, organizacionais e mesmo políticas das instituições (Leite, 2011a, n.p.).

De acordo com Leite (2011a, grifo nosso), as teses e dissertações “constituem os primeiros documentos a serem depositados em RI e são o **carro-chefe de seu povoamento**”.

Teses e dissertações constituem um dos primeiros relatos de pesquisa de doutorado e mestrado. Tais como outros produtos de informação científica, necessitam ser gerenciadas apropriadamente, isto é, reunidas, organizadas, armazenadas, preservadas, recuperadas e disseminadas amplamente, de modo que possam alimentar novos processos de produção do conhecimento científico [...] (Assis, 2015, p. 21).

Observa-se que esse tipo de coleção merece atenção especial devido às suas particularidades. Há, portanto, a necessidade de se estar constantemente atento à qualidade dos metadados, especialmente quanto aos registros dos autores, orientadores, títulos e palavras-chave, para efetivar a recuperação da informação.

6.1 Resultados e discussão

Em virtude dos fatos mencionados, a presente seção abordará a experiência em andamento da autora, na gestão da BDTD/UFRRJ, apresentando as ações implementadas durante o curso Mestrado em Humanidades Digitais (período de 2021-2023), da UFRRJ, conforme pode ser observado no quadro 8.

Quadro 8 - BDTD/UFRRJ – Ações x Resultados

| AÇÕES | RESULTADOS ALCANÇADOS |
|--|---|
| Elaboração do manual da BDTD/UFRRJ. | Padronização dos registros dos metadados. |
| Busca pelos documentos pendentes em diversos <i>sites</i> , como páginas de cursos da UFRRJ e banco de dados Sucupira. | Recuperação dos documentos pendentes (aproximadamente 80%). |
| Prioridade para o processamento do acervo digital (antes a seção processava o material | O processo dos registros das teses e dissertações no sistema se tornou mais ágil. |

| | |
|--|---|
| impresso e depois as teses e dissertações digitais). ⁹ | |
| Inclusão de novos dados (ORCID e o <i>link</i> do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) para o autor e os membros da banca. | Novas formas de recuperar a informação. |
| Padronização da referência bibliográfica da tese/dissertação (conforme normas da Associação Brasileira de Normas Técnicas – ABNT). | O pesquisador encontra a referência bibliográfica pronta, poupando o seu tempo, como também representa uma forma de recuperar a informação. |
| Correções de diversas falhas no registro dos metadados (retificação da nuvem de palavras, erros de digitação, Cadastro de Pessoas Físicas – CPF – sem padronização, não normalização dos nomes dos autores, orientadores e demais membros da banca e outras desconformidades). | Melhoria na apresentação do <i>site</i> , como também aperfeiçoamento na organização e recuperação da informação. |

Fonte: Elaborado pela autora (2023).

Inicialmente, foi observado que os três técnicos administrativos encarregados do registro de dissertações e teses na BDTD/UFRRJ trabalhavam sem seguir um padrão uniforme, resultando na falta de padronização no registro dos documentos digitais. Diante dessa constatação, foi desenvolvido um manual (documento em anexo) com o objetivo de estabelecer diretrizes claras para o preenchimento adequado dos metadados, buscando uma abordagem alinhada, porém sem desconsiderar a experiência desses servidores. Sua utilização contribuiu para a eficiência na recuperação de informações.

Para Borges *et al.* (2019, p. 255), o manual para o RD tem as seguintes funções:

Um manual, em aspecto amplo, estabelece os procedimentos e parâmetros de funcionamento de um repositório. Nele, está descrita as inter-relações possíveis no repositório de A a Z. Em linhas gerais, o manual estabelece informações completas, bem elaboradas e exaustivas sobre um determinado objeto [...].

No caso da BDTD/UFRRJ, o manual também cumpre o mesmo propósito, já que indica os metadados utilizados e os requisitos descritivos para cada elemento, conforme pode ser observado no documento. A padronização dos registros das teses e dissertações foi alcançada


⁹ A Portaria nº 2/2023, do dia 03 de abril de 2023, da Pró-Reitoria de Pesquisa e Pós-Graduação (PROPPG), suspendeu a obrigatoriedade da entrega do exemplar impresso de dissertações e teses para a concessão do diploma aos concluintes. Portanto, a BC receberá apenas a versão digital dos documentos da PROPPG, através do *drive* (disco virtual da Google), dos trabalhos defendidos a partir da data da publicação da portaria.

com sucesso, o que representa um passo importante para a gestão ideal do sistema de informação.

Após a elaboração do manual, procedeu-se a um levantamento dos documentos pendentes na BDTD/UFRRJ, ou seja, dissertações e teses que estavam disponíveis apenas em formato impresso na BC da UFRRJ. Esse trabalho foi realizado utilizando-se o relatório do Pergamum, a ferramenta de gestão de informações do acervo físico da BC. Com base nesse documento, os três técnicos responsáveis conduziram buscas em diferentes *sites*, como as páginas dos cursos da UFRRJ e o banco de dados Sucupira, a fim de localizar as versões digitais das teses e dissertações. No momento, ainda há 387 documentos que não foram localizados.

Segue uma página do relatório Pergamum, com o cadastro das teses e dissertações físicas:

Figura 3 – Relatório do Pergamum

| | | |
|--|---|-----------------------------------|
|  | UFRRJ Pergamum - Sistema Integrado de Bibliotecas ESTATÍSTICA DE PRODUÇÃO DOS USUÁRIOS - MODO ANALÍTICO Período : 01/01/1900 a 17/08/2023 Est-Usuarios-Produção (31) | Pag. 23 17/08/2023 17:39:34 |
| | <p>4700 AGNESE, MARCIO OCTAVIO; ESCOLA NACIONAL DE VETERINARIA (BRASIL). Anatomia da arvore bronquial em cães: Estudo dos bronquios segmentares. Rio de Janeiro, 1956. 59p. Classificação : 636.089123 A272a T Ac.4700</p> <p>4240 ALMEIDA, MARIA APARECIDA NOBREGA DE; ROSEMBERG, JOSE AUGUSTO; UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Liberação de fibras de linho por bacterias selecionadas. Rio de Janeiro, 1978. 64f. Classificação : 677.1125 A4471 T Ac.4240</p> <p>13680 JESUS, Eli Lino de; ALMEIDA, Dejáir Lopes de. Avaliação da sustentabilidade agrícola: uma abordagem conceitual e metodológica. 2003. 205f. Classificação : 631.584 J58a T (BC) Ac.13680</p> <p>4306 OLIVEIRA, Carlos Marcos Barcellos de. Biologia, flutuação populacional e patologia da Cochliomyia hominivorax (Coquerel, 1858) (Diptera: Calliphoridae). Rio de Janeiro, 1980. ix, 92f. Classificação : 595.77041 O48b T (BC) Ac.4306</p> <p>4459 ROBBS, Charles Frederick.. Estudo comparativo de algumas variantes de Corynebacterium michiganense (Smith) Jensen. Itaguaí, RJ, 1967. 43f. Classificação : 632.3 R637e T (BC) Ac.4459</p> <p>4622 SECCHIN, Haroldo. Estudo clínico-sanitario de bovinos destinados ao consumo no Grande Rio. Itaguaí, RJ, 1976. vi, 44f. Classificação : 636.20896 S444e T (BC) Ac.4622</p> <p>4313 HUNGRIA, Mariangela; NEVES, Maria Cristina P.. Fisiologia da fixação biologica do nitrogenio em Phaseolus vulgaris L.. Itaguaí, RJ, 1985. xxviii, 292f. Classificação : 631.847 H936f T (BC) (BUI TR) (BIM) (CPDA) (BLM) Ac.4313</p> <p>12338 SOUZA, Solange São Paulo de; CRUZ, Fernando Zanotta da. Ocorrência e bioecologia de aleirodídeos (Homoptera; Aleyrodidae) em plantas cítricas no estado do Rio Grande do Sul. 1999, xvi, 165f. Classificação : 634.35 S729o T (BC) Ac.12338</p> <p>12051 SILVA, Dácia Ibiapina da. Memórias da Guerrilha do Araguaia: relatos de moradores da Palestina do Pará. 2002, xii,290.[6]f. Classificação : 355.425 S586m T (BC) Ac.12051</p> | |

Fonte: Site Pergamum (<http://pergamum.ufrj.br:8080/pergamumweb/RelatoriosExibeResultadoServlet>).

Abaixo, a página do Programa de Pós-Graduação em Educação Agrícola (PPGEA), da UFRRJ, com as dissertações do curso:

Figura 4 – Página da PPGEA da UFRRJ



Fonte: Site da PPGEA/UFRRJ (<https://cursos.ufrrj.br/posgraduacao/ppgea/dissertacoes-concluidas/>).

Uma abordagem mais eficaz seria incorporar os *links* para as dissertações cadastradas na BDTD/UFRRJ na página do respectivo curso.

Borges *et al.* (2019, p. 261) têm o seguinte posicionamento sobre esse problema:

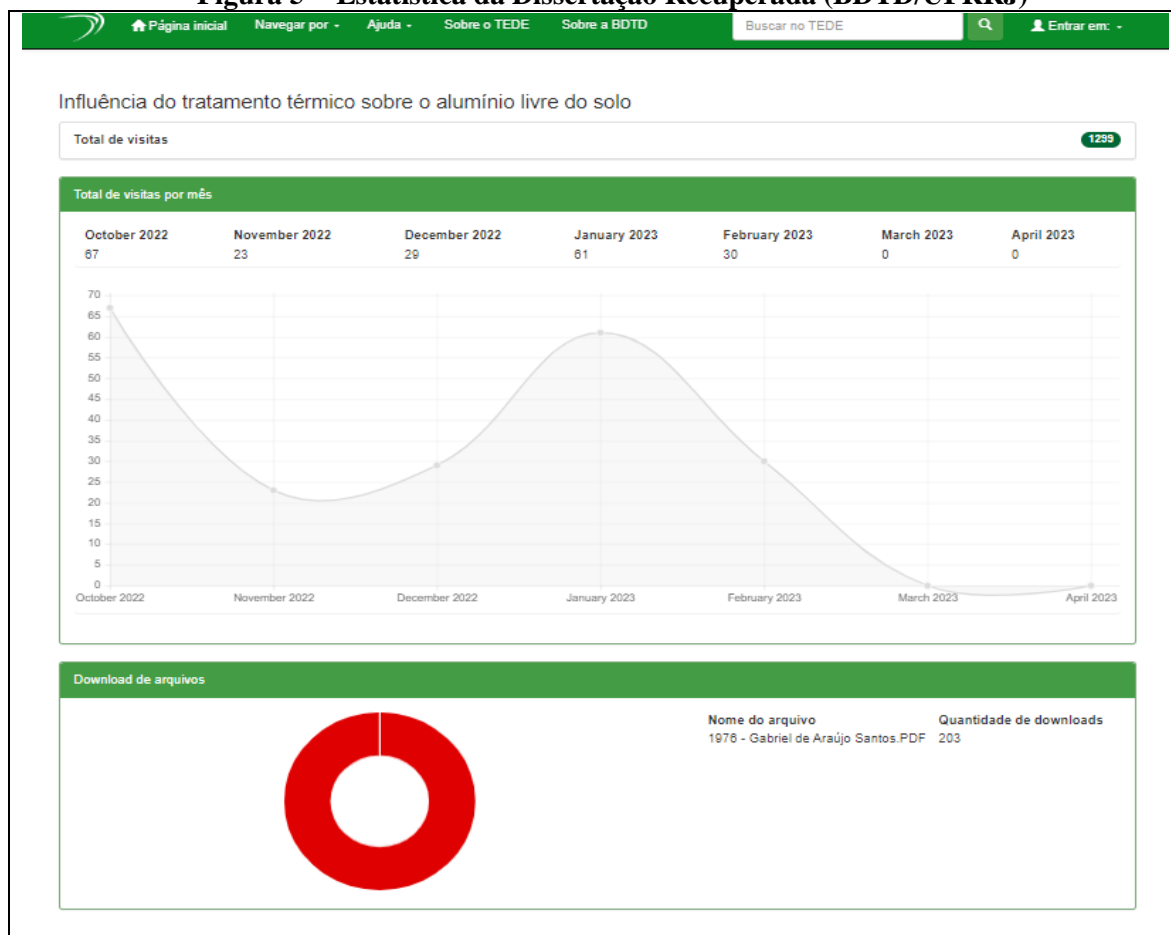
[...] outro ponto considerado está nas dispersões de documentos importantes sendo alocado em sites, que não seja o próprio repositório, algo que não poupa o tempo do usuário, como orienta uma das cinco leis da biblioteconomia.

Logo, é imprescindível canalizar os trabalhos de conclusão dos cursos de pós-graduação da instituição na BDTD/UFRRJ (futuramente, no RIMA).

Foram resgatadas pesquisas realizadas nas décadas de 60, 70 e 90, assim como estudos mais recentes. Os resultados foram altamente positivos, com vários documentos anteriormente pendentes alcançando uma elevada taxa de visualização por parte dos usuários. Um caso em destaque é a dissertação de Gabriel de Araújo Santos, defendida em 1976, a qual despertou considerável interesse entre os usuários.

Segue, abaixo, a figura 5 com a estatística da dissertação recuperada.

Figura 5 – Estatística da Dissertação Recuperada (BDTD/UFRRJ)



Fonte: site do TEDE (<https://tede.ufrrj.br/jspui/handle/jspui/1636/statistics>).

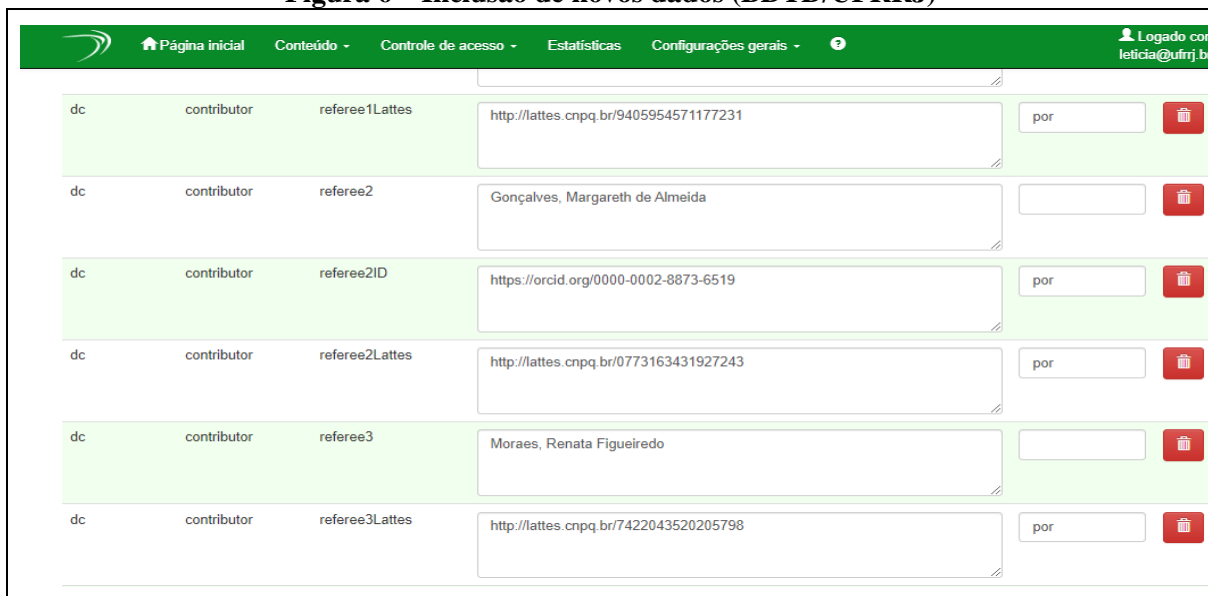
A dissertação recebeu um total de 1.299 visualizações e foi baixada 203 vezes em menos de três anos. Esses números refletem a relevância do trabalho acadêmico, mesmo após várias décadas de sua defesa. Ao tornar esses documentos acessíveis, a BDTD/UFRRJ está contribuindo para o acesso ao conhecimento produzido pela universidade ao longo do tempo.

Já a decisão de priorizar o processamento do acervo digital foi embasada na análise da demanda por parte dos usuários. Raramente, a BC recebe solicitações de acesso a dissertações e teses em formato físico, ao passo que os documentos eletrônicos são frequentemente consultados, conforme atestam as estatísticas dos objetos digitais disponíveis na BDTD/UFRRJ.

A quarta ação, a inclusão do ORCID e o *link* do currículo da plataforma Lattes para os autores, orientadores e demais membros da banca, tem como objetivo atender ao primeiro Princípio FAIR, *Findable* (encontráveis). Ao disponibilizar esses dados, a intenção é contribuir para uma melhor rastreabilidade das informações (por humanos e máquinas), como também facilitar a identificação desses indivíduos. Essas informações adicionais podem ser valiosas

para pesquisadores que desejam conhecer mais sobre a trajetória acadêmica deles. No entanto, é importante ressaltar que esses dados são recuperados, no momento, exclusivamente dentro do próprio ambiente da BDTD/UFRRJ.

Figura 6 – Inclusão de novos dados (BDTD/UFRRJ)



The screenshot displays a web interface for editing contributor data. At the top, there is a green navigation bar with links for 'Página inicial', 'Conteúdo', 'Controle de acesso', 'Estatísticas', and 'Configurações gerais'. On the right, it shows the user is logged in as 'leticia@ufrj.br'. Below the navigation bar, there is a table with six rows, each representing a contributor. Each row has a 'dc' identifier, a role of 'contributor', a name, a text input field for a URL or name, a dropdown menu with 'por' selected, and a red trash icon. The data in the rows is as follows:

| dc | contributor | referee | Input Field | Dropdown | Action |
|----|-------------|----------------|---|----------|--------|
| dc | contributor | referee1Lattes | <input type="text" value="http://lattes.cnpq.br/9405954571177231"/> | por | Trash |
| dc | contributor | referee2 | <input type="text" value="Gonçalves, Margareth de Almeida"/> | | Trash |
| dc | contributor | referee2ID | <input type="text" value="https://orcid.org/0000-0002-8873-6519"/> | por | Trash |
| dc | contributor | referee2Lattes | <input type="text" value="http://lattes.cnpq.br/0773163431927243"/> | por | Trash |
| dc | contributor | referee3 | <input type="text" value="Moraes, Renata Figueiredo"/> | | Trash |
| dc | contributor | referee3Lattes | <input type="text" value="http://lattes.cnpq.br/7422043520205798"/> | por | Trash |

Fonte: Site do TEDE (https://tede.ufrj.br/jsptools/edit-item?item_id=5797&submit=Editar).

Ao abordar a padronização da referência bibliográfica da tese/dissertação, de acordo com as normas da Associação Brasileira de Normas Técnicas (ABNT), dois objetivos principais são considerados. Primeiramente, busca-se economizar o tempo do pesquisador, fornecendo a informação pronta e correta para ser inserida em seu trabalho. Além disso, a inclusão do metadado “citação” (dc.identifier.citation, do Dublin Core) também facilita a recuperação do documento na rede mundial, o que pode ser observado na figura 7.

Citação:

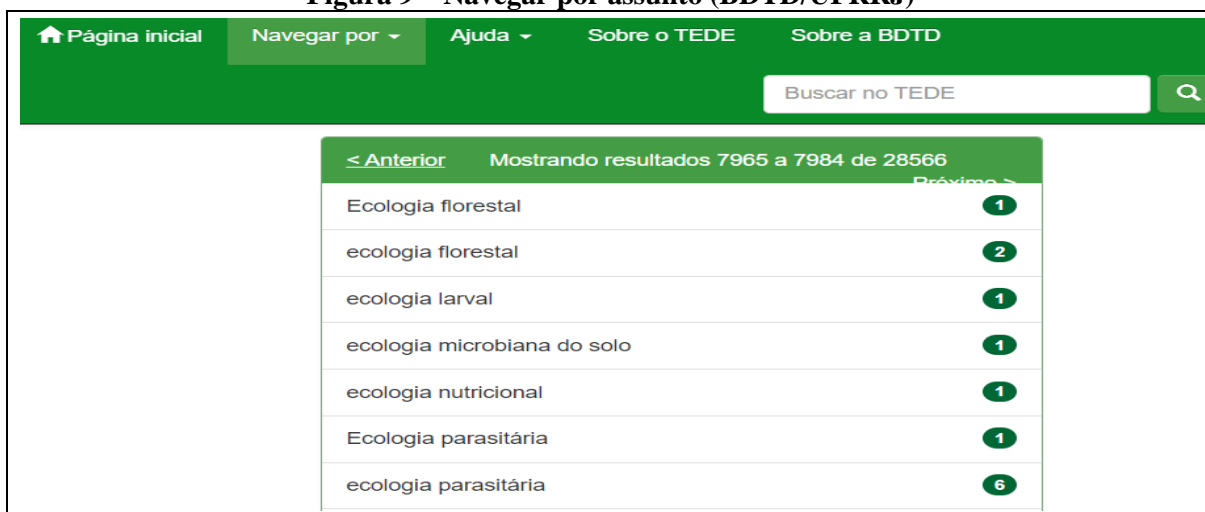
BESSA, Alessandra Castro Fiorini. **Princípios FAIR e Proveniência em repositórios de dados voltados para a Agricultura Digital:** um estudo exploratório à luz das Humanidades Digitais. 2021. 88 f. Dissertação (Mestrado em Humanidades Digitais) - Instituto Multidisciplinar de Nova Iguaçu, Universidade Federal Rural do Rio de Janeiro, Nova Iguaçu, RJ. 2021.

que apresentavam essa falha. Essa situação levanta a questão sobre como lidar com dados indisponíveis e seus metadados correspondentes. Nesse caso, a conclusão é que, quando um dado não está disponível, o metadado correspondente não deve ser preenchido. Assim, evita-se a inclusão de informações incorretas no sistema. Isso garante a integridade dos metadados e a precisão das informações apresentadas aos usuários.

Quanto aos erros de digitação e outras incorreções, há no manual a indicação para realizar a revisão completa dos dados antes de finalizar o depósito do documento na BDTD/UFRRJ. Contudo, é inevitável realizar uma inspeção regular em todo o sistema, com o objetivo de corrigir as falhas que prejudicam a recuperação da informação.

Um grande problema observado é que o sistema faz distinção entre letras maiúsculas e minúsculas, a presença e a ausência de acentos e outros detalhes, resultando em vários registros para um mesmo orientador, palavra-chave, curso e outros dados, como ilustrado nas figuras 9 e 10:

Figura 9 – Navegar por assunto (BDTD/UFRRJ)



The screenshot shows the TEDE website interface. At the top, there is a green navigation bar with links: 'Página inicial', 'Navegar por', 'Ajuda', 'Sobre o TEDE', and 'Sobre a BDTD'. A search bar labeled 'Buscar no TEDE' is also present. Below the navigation bar, a table displays search results for 'Ecologia florestal'. The table has two columns: the subject name and a count in a green circle. The results are as follows:

| Mostrando resultados 7965 a 7984 de 28566 | |
|---|---|
| Ecologia florestal | 1 |
| ecologia florestal | 2 |
| ecologia larval | 1 |
| ecologia microbiana do solo | 1 |
| ecologia nutricional | 1 |
| Ecologia parasitária | 1 |
| ecologia parasitária | 6 |

Fonte: Site do TEDE (<https://tede.ufrrj.br/jspui/browse?type=subject>).

Figura 10 – Navegar por orientador (BTD/UFRRJ)



The image shows a screenshot of the TEDE website. At the top, there is a green navigation bar with the following items: 'Página inicial' (with a home icon), 'Navegar por' (with a dropdown arrow), 'Ajuda' (with a dropdown arrow), 'Sobre o TEDE', and 'Sobre a BDT'. To the right of the navigation bar is a search box labeled 'Buscar no TEDE'. Below the navigation bar, there is a list of advisors with their names and the number of results for each name in a green circle:

| | |
|-----------------------------------|----|
| Alegrio, Leila Vilela | 4 |
| Alejo, José Luis Fernando Luque | 1 |
| Alejos, José Luis Fernando Luque | 27 |
| Alencar, Jeronimo Augusto Fonseca | 5 |
| Alencar, Jerônimo Augusto Fonseca | 1 |

Fonte: Site do TEDE (<https://tede.ufrrj.br/jspui/browse?type=advisor>)

Sem a revisão dos dados, o usuário empregará mais tempo em sua busca por informação na BDT/UFRRJ.

É importante destacar que a inconsistência entre letras maiúsculas e minúsculas no cadastro pode ser facilmente corrigida por meio da aplicação de um artefato computacional, como um programa ou script, que pode realizar automaticamente a conversão das entradas de texto para caixa alta ou caixa baixa conforme desejado, contribuindo significativamente para a manutenção da consistência e padronização nos dados inseridos no sistema.

Passando para o tópico “normalização dos nomes dos autores, orientadores e demais membros da banca”, observou-se que não era viável obter as informações necessárias para o cadastro no sistema a partir da folha de rosto do documento, uma vez que nem sempre estão corretos (nomes incompletos, erros de digitação, troca de letras, entre outros). Como solução para o problema, foi adotada a pesquisa nos currículos da Plataforma Lattes do CNPq, para o preenchimento adequado das informações. Segue imagem de um Currículo Lattes:

Figura 11 – Currículo Lattes

Alessandra Castro Fiorini Bessa

Endereço para acessar este CV: <http://lattes.cnpq.br/6511807481582383>

ID Lattes: 6511807481582383

Última atualização do currículo em 09/11/2022

Mestre em Humanidades Digitais pelo Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais da Universidade Federal Rural do Rio de Janeiro (PPGIHD/UFRRJ), especialista em Docência no Ensino Superior pelo Instituto A Vez do Mestre (UCAM) e bacharel em Biblioteconomia pela Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Atualmente trabalha como Bibliotecária-Documentalista, na Universidade Federal Rural do Rio de Janeiro (UFRRJ). Integrante de GT FAIR Agro BR. (Texto informado pelo autor)

Identificação

Nome **Alessandra Castro Fiorini Bessa**

Nome em citações bibliográficas Bessa, A. C. F.;Bessa, Alessandra C. F.

Lattes ID <http://lattes.cnpq.br/6511807481582383>

Orcid ID <https://orcid.org/0000-0002-7047-9294>

Fonte: Site Currículo Lattes (<https://buscatextual.cnpq.br/>).

É muito importante esse controle da autoridade “nome” no sistema, para a uniformidade dos dados. Em vista disso, a equipe cadastra na BDTD/UFRRJ o nome completo, conforme consta na plataforma, e segue as regras da catalogação, o Código de Catalogação Anglo-Americano – 2ª edição (*Anglo-American Cataloguing Rules – 2nd edition*) (AACR2), como no exemplo:

- BESSA, Alessandra Castro Fiorini (antes da padronização);
- Bessa, Alessandra Castro Fiorini (agora, conforme AACR2).

Outro ponto importante, é que não havia uma padronização no registro do Cadastro de Pessoas Físicas (CPF) dos autores e orientadores no sistema. A informação passou a ser cadastrada no formato válido, conforme segue: 000.000.000-00.

Após relatar quais foram as principais ações realizadas na BDTD/UFRRJ, é importante informar também sobre as questões que ainda não foram resolvidas no sistema, conforme pode ser observado no quadro a seguir:

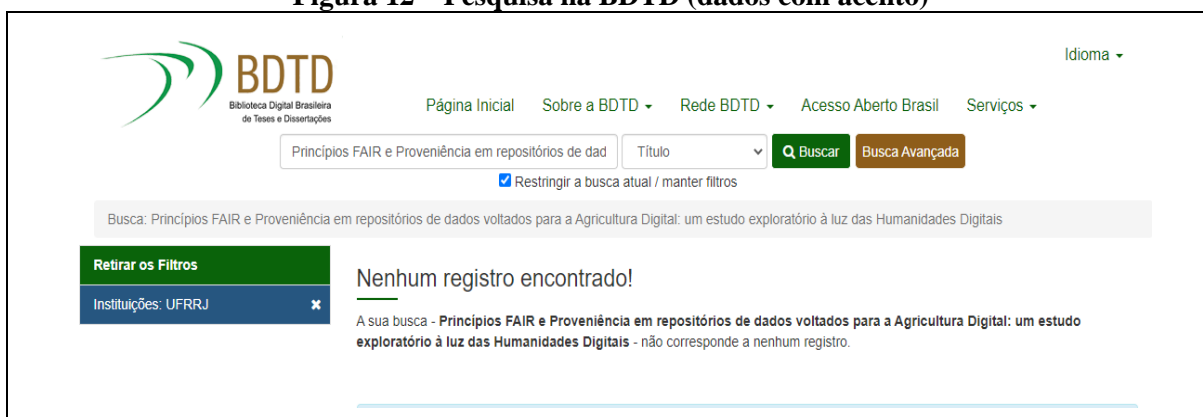
Quadro 9 – Questões não resolvidas X Soluções propostas (BDTD/UFRRJ)

| PROBLEMAS | SOLUÇÕES PROPOSTAS |
|--|------------------------------------|
| Os metadados registrados com acentos no sistema da UFRRJ não são recuperados no portal da BDTD. | Ajustes na configuração do DSpace. |
| Não é possível recuperar os documentos através do ORCID em busca na internet. | Ajustes na configuração do DSpace. |
| O link do CNPq só está disponível para o autor. A informação não aparece para os membros da banca. | Ajustes na configuração do DSpace. |

Fonte: Elaborado pela autora (2023).

Foi identificada uma questão grave na plataforma do BDTD, no portal do IBICT, em relação à recuperação de teses e dissertações com acento no título, autor ou em outro metadado. O documento não é encontrado na plataforma, conforme pode ser observado na figura 12.

Figura 12 – Pesquisa na BDTD (dados com acento)



Fonte: Site da BDTD (<http://www.bdttd.ibict.br/>).

Para encontrar esses registros no site da Biblioteca Digital é preciso remover da pesquisa as palavras com acento, como mostrado no exemplo abaixo:

Figura 13 – Pesquisa na BDTD (dados sem acento)

The screenshot shows the BDTD (Biblioteca Digital Brasileira de Teses e Dissertações) search interface. At the top, there is a navigation menu with links for 'Página Inicial', 'Sobre a BDTD', 'Rede BDTD', 'Acesso Aberto Brasil', and 'Serviços'. The search bar contains the text 'de dados voltados para a Agricultura Digital: um estudo' and a dropdown menu set to 'Título'. There are buttons for 'Buscar' and 'Busca Avançada'. Below the search bar, the search results are displayed. On the left, there is a 'Refinar a Busca' section with three categories: 'Instituições' (showing 'UFRRJ' with 1 result), 'Repositório' (showing 'Biblioteca Digital de Teses e Dissertações da UFRRJ' with 1 result), and 'Programa'. On the right, there is a 'Buscas alternativas' section with links for 'dados voltados > dados coletados, dados conectados'. Below this, it says 'A mostrar 1 - 1 resultados de 1 para a busca 'de dados voltados para a Agricultura Digital: um estudo', tempo de busca: 0.14s'. There is a dropdown for 'Ordenar' set to 'Relevância' and buttons for 'Ver Tudo' and 'Exportar'. The search results list shows one result: '1 Princípios FAIR e Proveniência em repositórios de dados voltados para a Agricultura Digital: um estudo exploratório? luz das Humanidades Digitais'.

Fonte: Site da BDTD (<http://www.bdttd.ibict.br/>).

Assim, a recuperação das teses e dissertações na BDTD encontra-se comprometida, sendo que a solução para esse problema dependerá da realização de ajustes no *software*. Nesse sentido, será necessário configurar o *charset* para LATIN (8-bit Unicode Transformation Format – UTF-8), Basic Latin (ASCII) a fim de corrigir o erro identificado.

Outro ponto relevante a ser considerado é sobre o registro do ORCID dos docentes e discentes no sistema. Infelizmente, ainda não é possível recuperar o dado na BDTD/UFRRJ, através da busca na internet, usando o Google, por exemplo, conforme pode ser observado na figura 14:

Figura 14 – Busca pelo ORCID no Google

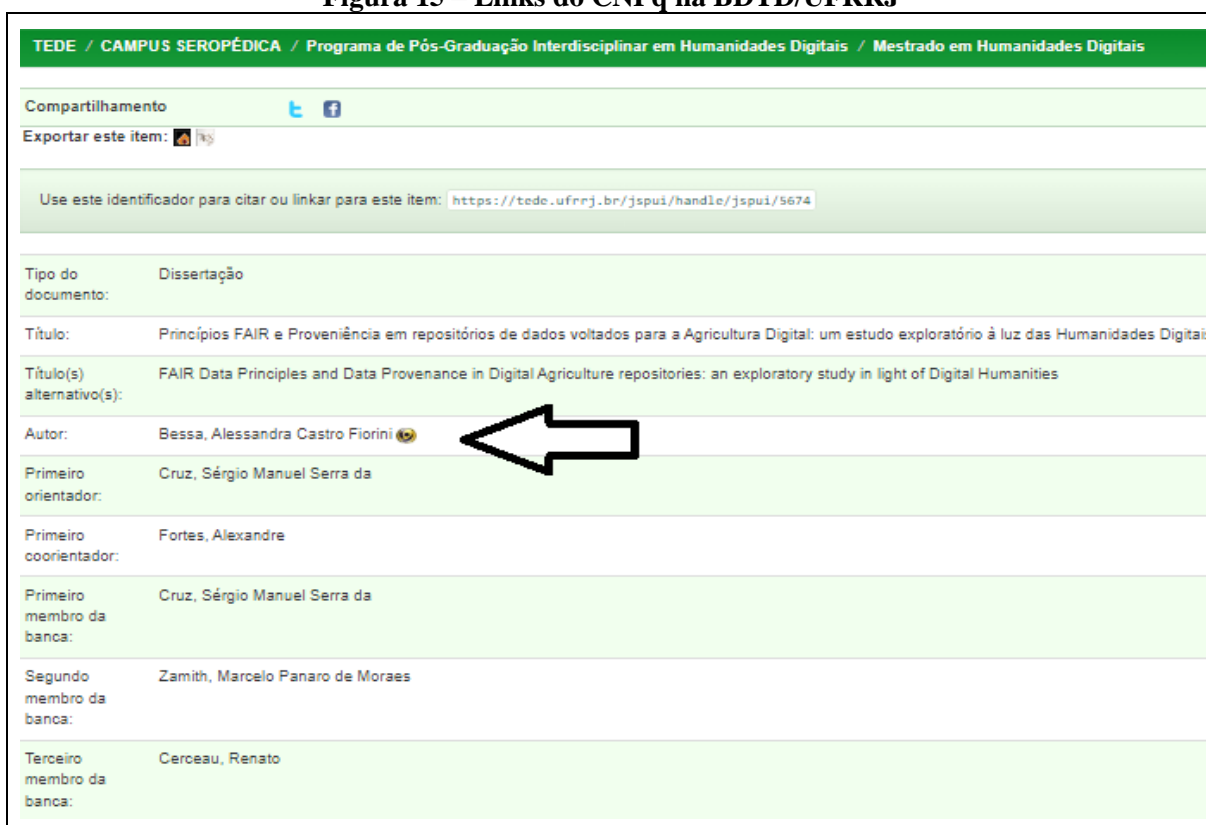


Fonte: Site Google (<http://www.google.com.br/>).



Em outras palavras, o dado é cadastrado na BDTD/UFRRJ, porém não é recuperado na rede mundial.



Quanto aos *links* do CNPq, o acesso fica disponível apenas para o autor, embora o dado seja inserido para todos os participantes da banca.

Figura 15 – Links do CNPq na BDTD/UFRRJ




TEDE / CAMPUS SEROPÉDICA / Programa de Pós-Graduação Interdisciplinar em Humanidades Digitais / Mestrado em Humanidades Digitais

Compartilhamento  

Exportar este item:  

Use este identificador para citar ou linkar para este item: <https://tede.ufrrj.br/jspui/handle/jspui/5674>

| | |
|---------------------------|--|
| Tipo do documento: | Dissertação |
| Título: | Princípios FAIR e Proveniência em repositórios de dados voltados para a Agricultura Digital: um estudo exploratório à luz das Humanidades Digitais |
| Título(s) alternativo(s): | FAIR Data Principles and Data Provenance in Digital Agriculture repositories: an exploratory study in light of Digital Humanities |
| Autor: | Bessa, Alessandra Castro Fiorini  |
| Primeiro orientador: | Cruz, Sérgio Manuel Serra da |
| Primeiro coorientador: | Fortes, Alexandre |
| Primeiro membro da banca: | Cruz, Sérgio Manuel Serra da |
| Segundo membro da banca: | Zamith, Marcelo Panaro de Moraes |
| Terceiro membro da banca: | Cerceau, Renato |

Fonte: Site do TEDE (<https://tede.ufrrj.br>).

Com base no exposto, ainda há alguns ajustes pendentes a serem realizados antes da migração do banco de dados para o RIMA da UFRRJ. No entanto, é evidente que a BDTD/UFRRJ está seguindo na direção correta para alcançar a gestão ideal do sistema. Isso permitirá aumentar a visibilidade de seu conteúdo e, conseqüentemente, promover um uso mais amplo por parte da comunidade acadêmica e científica.

6.2 O repositório em transição: propostas para o futuro

Considerando que, em breve, ocorrerá a migração da BDTD/UFRRJ para o RIMA e que os dados para o registro das teses e dissertações serão os mesmos, a presente pesquisa, que tem como objetivo aprimorar o tratamento dessa coleção no repositório, manterá sua relevância, uma vez que a estrutura dos metadados sofrerá alterações mínimas.

Como próximos passos para o RI da Rural, há algumas sugestões, como a adição de novos metadados, por exemplo, o campo "assunto". Atualmente, no cadastro da BDTD/UFRRJ, são utilizados os termos fornecidos pelos alunos, as palavras-chave. Essa representação temática do documento é uma informação muito importante para a recuperação da informação

no sistema. No entanto, nem sempre o autor consegue extrair as ideias principais do seu texto (dissertação ou tese) para incluir em seu trabalho.

Segue o posicionamento dos autores Tonello, Lunardelli e Almeida Júnior (2012) sobre o tópico palavras-chave:

[...] a representação da informação, no caso, por intermédio das palavras-chave consiste na substituição de uma entidade linguística (o texto do documento), por sua descrição abreviada (palavras-chave), requerendo, nesse processo, uma vez que contempla alguns aspectos específicos - a mediação do profissional da informação. As palavras que determinarão se o usuário acessará ou não a informação serão, reitere-se, extraídas do texto do documento, o que requer, para um resultado satisfatório, a qualificação do profissional da informação, pois todo esse processo de mediação do conteúdo informacional será feito por ele (Tonello; Lunardelli; Almeida Júnior, 2012, p. 29-30).

Dessa maneira, os autores sustentam a importância da intervenção do bibliotecário na análise e síntese das informações contidas nos documentos, visando a identificação das palavras-chave que melhor expressem seu conteúdo. Contudo, diante do expressivo volume de dissertações e teses na instituição, surge a indagação sobre como os profissionais da informação podem desempenhar essa tarefa de maneira eficaz. Qual seria a solução para esse desafio? Uma abordagem viável consiste em incluir o metadado "assunto" como uma alternativa adicional para a representação temática, superando a limitação do uso exclusivo do metadado "palavras-chave". Isso proporciona um novo caminho para a recuperação da informação. O procedimento inicial de registro das teses e dissertações no repositório digital pode ocorrer apenas com as palavras-chave, sendo posteriormente complementado, em um momento oportuno, com o cadastramento dos assuntos, empregando um vocabulário controlado. Diversas instituições no Brasil adotaram essas duas opções nos registros de seus recursos digitais, como o RI da Universidade Federal de Uberlândia (UFU) e o Repositório Institucional da Fiocruz (Arca). O levantamento dos assuntos contidos nos documentos poderia ser realizado pelo bibliotecário, com auxílio de ferramentas de mineração de texto, como, por exemplo, o Sobek, conforme já defendido na seção 4.2.

A seguir, seguem as demais recomendações destinadas ao tratamento da coleção das teses e dissertações da UFRRJ no RIMA:

- ✓ A adoção da atribuição de DOI em teses e dissertações da instituição, com base na experiência da UFU: a inclusão desse metadado (dc.identifier.doi) pode trazer benefícios significativos, como a garantia de uma identificação única e persistente para cada documento, facilitando a sua busca no ambiente virtual;

- ✓ Digitalização do acervo retroativo de teses e dissertações (são 387 documentos pendentes, que não foram recuperados durante a busca em diversos *sites*): o objetivo é tornar esses documentos disponíveis para acesso no RD, e assim a coleção digital ficará completa;
- ✓ Disponibilizar uma assistente virtual no repositório para tirar dúvidas dos usuários (baseado na experiência do Arca): é um *software* de inteligência artificial que pode auxiliar os usuários na navegação, busca e interação com o conteúdo do RD.

No futuro, a UFRRJ poderá implementar as recomendações do estudo, que visam melhorar ainda mais o funcionamento do seu RI. Agindo desta forma, pressupõe-se que o sistema de informação atingirá todo o seu potencial, oferecendo um serviço de alta qualidade, e contribuirá significativamente para o avanço da pesquisa na universidade.

7 CONSIDERAÇÕES FINAIS

Os profissionais da informação, que acompanharam de perto a evolução das tecnologias digitais nas últimas décadas, têm plena consciência da enorme transformação no tratamento dos registros de conhecimento. Um exemplo importante é a transição das bibliotecas, que usavam fichas catalográficas em papel, contendo os dados descritivos dos livros, para identificação e recuperação dos itens nas estantes. Hoje em dia, o cadastro dos exemplares ocorre em sistemas de informação digitais, utilizando metadados, e os cadastros ficam disponíveis para consulta em alcance global, por meio da internet. Essa mudança representa uma verdadeira revolução na maneira como a informação é acessada.

A principal finalidade de um sistema de informação (tradicional ou digital) é satisfazer as necessidades dos usuários em busca de conhecimento. Por essa razão, é necessário saber como aproveitar ao máximo os repositórios digitais. No contexto específico da BDTD, o estudo revelou que o processamento das teses e dissertações em sua base não se resume apenas à inserção de informações básicas, como autor, título e palavras-chave. É necessário adotar uma abordagem mais ampla e interdisciplinar no tratamento dessa coleção no repositório. Deve-se pensar em várias etapas, como a coleta, a padronização no registro dos dados no sistema e, principalmente, a representação temática dos objetos digitais. Dessa forma, será possível explorar todas as possibilidades oferecidas pela ferramenta, para garantir uma melhor experiência para os usuários em suas pesquisas.

No que diz respeito à abordagem metodológica adotada, a decisão de realizar a leitura dos textos selecionados revelou-se eficaz ao estabelecer uma base para as ações subsequentemente implementadas na BDTD. Portanto, a bibliografia permitiu uma imersão no contexto do tema de estudo. Isso não apenas proporcionou uma compreensão mais abrangente das melhores práticas na gestão de acervos digitais acadêmicos, mas também mostrou algumas lacunas na literatura, como a ausência de um documento oficial de normalização específico para o registro dos documentos em repositórios digitais.

Já a escolha de empregar o método de pesquisa-ação revelou-se uma opção correta para a efetivação das propostas da dissertação. Ao mesclar pesquisa e prática, realizou-se análise e aplicação, permitindo que as descobertas resultantes da leitura fossem traduzidas em intervenções visíveis no repositório. Dessa maneira, o método facilitou o aprendizado constante, abarcando tanto aspectos teóricos quanto práticos.

Voltando à questão norteadora da pesquisa, que trata do aumento do número de teses e dissertações eletrônicas da UFRRJ, o estudo mostrou que há caminhos para realizar a gestão

ideal dessa valiosa coleção da instituição. Em outras palavras, é possível se preparar para o crescimento desse acervo no RD e, assim, atender à quinta lei de Ranganathan.

Retomando o objetivo geral deste estudo, que visa aprimorar a forma como a coleção de teses e dissertações digitais da UFRRJ é tratada, por meio da implementação de uma abordagem sistemática em seu processamento na BDTD, é possível constatar, com base nos resultados da pesquisa, que um avanço significativo já foi alcançado na maneira como esse valioso acervo institucional está sendo gerenciado no ambiente virtual. Foram direcionados esforços para a melhoria da organização e a acessibilidade dessa coleção no RD.

Quanto aos objetivos específicos, ao investigar a inter-relação entre as HD e a CI, foi possível destacar a estreita conexão entre esses dois campos. Essa proximidade se manifesta tanto no emprego de tecnologias digitais, como também na utilização de estruturas semânticas, tais como taxonomias e tesauros.

A autora identificou, na disciplina Humanidades Digitais, métodos inovadores que podem contribuir para uma gestão ideal dos repositórios digitais. No entanto, é importante enfatizar que os conhecimentos tradicionais da CI não devem ser descartados, mas sim adaptados às tecnologias digitais. Um aspecto importante é a aplicabilidade das Leis de Ranganathan, que fornecem orientações valiosas para qualquer sistema de informação, conforme discutido no texto, e que em parte alinhados aos Princípios FAIR, podem fornecer um norte para a gestão ideal do RD. Lembrando que as questões do acesso e da usabilidade da informação são pontos muito importantes para ambos, conforme foi observado na dissertação.

No que diz respeito às abordagens para o tratamento informacional de teses e dissertações, a pesquisa destacou a relevância dos SOC, como a classificação, tesauro, ontologia e taxonomia. Esses sistemas desempenham uma função significativa na estruturação semântica dos dados dentro do RD, permitindo uma busca mais eficiente por objetos digitais.

O estudo também apresentou a aplicação do minerador de texto Sobek na tarefa de identificar os termos que representam os assuntos dos documentos digitais. A ferramenta digital, ao selecionar as palavras mais relevantes dos textos, consegue oferecer de forma automática alguns termos que podem ajudar na construção dos assuntos, ou seja, no preenchimento da representação temática das teses e dissertações na BDTD/UFRRJ. O Sobek tem várias funcionalidades, e a habilidade de indexar os tópicos pode ser aprimorada com o passar do tempo, assim como uma nova ferramenta de mineração de texto pode surgir no futuro para realizar o trabalho com mais precisão. Encontrar uma forma automática e eficaz para

realizar a descrição dos assuntos dos documentos eletrônicos representaria um grande avanço no processamento da informação.

A pesquisa também mostrou a importância da padronização dos dados registrados no RD, conforme visto na BDTD/UFRRJ. Essa uniformidade pode ter como base as regras da Biblioteconomia, como a AACR2. A normalização dos dados contribui para a organização da informação, poupando o tempo do usuário na busca por documentos no sistema de informação.

Quanto à ferramenta RD, a pesquisa mostrou uma análise de suas funcionalidades, fornecendo uma descrição sobre sua definição e explorando os diversos tipos em que se divide, não apenas temático e institucional. Além disso, explorou-se a importância dos metadados nesse contexto, destacando como eles desempenham um papel importante na recuperação dos documentos na base de dados. Outro ponto mencionado foi a interoperabilidade, examinando como essa característica permite a troca de informações com outras plataformas.

Em relação ao problema da dispersão das teses e dissertações, é importante lembrar que a BDTD (e, no futuro, o RIMA) deve reunir os trabalhos de conclusão dos cursos de pós-graduação da instituição (mestrado e doutorado). Com a transição para a versão exclusivamente digital desses documentos, é necessário estabelecer um meio que garanta a integridade dos itens no sistema de informação. Em outras palavras, é fundamental assegurar que todas as teses e dissertações da universidade estejam devidamente registradas no repositório.

As justificativas iniciais para abordar o tema da dissertação destacaram a relevância do estudo. A busca por melhorias em um sistema de informação, com o objetivo de proporcionar um serviço de qualidade para atender às necessidades tanto da comunidade acadêmica da UFRRJ quanto de usuários externos, resultou em benefícios concretos. A abordagem da preservação digital das teses e dissertações da UFRRJ também demonstrou ser eficaz, contribuindo para a conservação de longo prazo do conhecimento gerado pela instituição. Por fim, o presente documento servirá como uma base teórica valiosa para pesquisas futuras, especialmente para bibliotecários e profissionais da área de TI.

No que diz respeito às limitações da pesquisa, destaca-se como desafio principal o prazo restrito para a realização de uma investigação que demanda a análise aprofundada de diversos aspectos da gestão do RD. Apesar disso, considera-se que o estudo atingiu os objetivos inicialmente estabelecidos. Contudo, reconhece-se que há mais a percorrer nesse trajeto.

Uma sugestão relevante para estudos futuros envolve a viabilidade de incorporar vocabulários controlados nos repositórios digitais de teses e dissertações. Explorar mais a fundo esse tema em estudos posteriores seria de grande importância, visto que resultaria em

contribuições significativas para a gestão eficaz de informações acadêmicas nesse ambiente virtual. Além disso, outro aspecto a considerar é a análise da implementação de uma taxonomia facetada para a navegação, o que simplificaria a busca por documentos por meio de uma estrutura conceitual, proporcionando uma organização mais sistemática do conhecimento armazenado.

Essas ideias encontram respaldo no cenário atual, caracterizado pela ausência de padronização na busca por assuntos nessas plataformas. Essa falta de uniformidade tem se revelado um desafio para os usuários, que frequentemente enfrentam dificuldades na recuperação eficiente de informações em vários repositórios pelo Brasil.

Por fim, espera-se que a presente dissertação proporcione algumas reflexões sobre o tratamento das teses e dissertações eletrônicas tanto na BDTD, como também no RI.

REFERÊNCIAS

ABBOTT, Daisy. What is digital curation? **DCC Briefing Papers: Introduction to Curation**, Edinburgh, 2008, Identificador: 1842/3362. Disponível em:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>. Acesso em: 7 jul. 2023.

ALVES, Daniel. As Humanidades Digitais como uma comunidade de práticas dentro do formalismo acadêmico: dos exemplos internacionais ao caso português. **Ler História**, n. 69, p. 91-103, 2016. DOI: 10.4000/lerhistoria.2496. Disponível em:

<https://journals.openedition.org/lerhistoria/2496>. Acesso em: 08 dez. 2022.

ALVES, Rachel Cristina Vesú. **Metadados como elementos do processo de catalogação**. 2010. 134 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em:

[https://www.marilia.unesp.br/Home/Pos-](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves_rachel.pdf)

[Graduacao/CienciadaInformacao/Dissertacoes/alves_rachel.pdf](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves_rachel.pdf) . Acesso em: 16 ago. 2022.

ALVES, Rachel Cristina Vesú. **Web semântica**: uma análise focada no uso de metadados. 2005. 180 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2005. Disponível em: <http://hdl.handle.net/11449/93690>. Acesso em: 19 abr. 2023.

AMARAL, Marjorie Rosielle Silva do. **Disseminação Seletiva da Informação no contexto das Organizações Aprendentes**: proposta de um modelo digital integrado ao SIGAA na Universidade Federal do Rio Grande do Norte. 2014. 119 f. Dissertação (Mestrado) - Universidade Federal da Paraíba, 2014.

ANDRADE, Laura Mariane de. **Aproximações entre Humanidades Digitais e organização do conhecimento**. 2021. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de São Carlos, São Carlos, 2021. Disponível em:

<https://repositorio.ufscar.br/handle/ufscar/15930>. Acesso em: 15 dez. 2022.

ANDRADE, Laura Mariane de; DAL'EVEDOVE, Paula. Regina. Humanidades Digitais na Ciência da Informação brasileira: análise da produção científica. **Revista Ibero-Americana de Ciência da Informação**, [s.l.], v. 13, n. 1, p. 439-451, 2020.

DOI: 10.26512/rici.v13.n1.2020.29582. Disponível em:

<https://periodicos.unb.br/index.php/RICI/article/view/29582/25421>. Acesso em: 12 abr. 2022.

ARAÚJO, Carlos Alberto Ávila. O que é Ciência da Informação? **Informação &**

Informação, Londrina, v. 19, n. 1, p. 01, 1 dez. 2013. DOI 10.5433/1981-

8920.2014v19n1p01. Disponível em:

<http://www.uel.br/revistas/uel/index.php/informacao/article/view/15958>. Acesso em: 29 mar. 2022.

ARAÚJO JÚNIOR, Rogério Henrique de; TARAPANOFF, Kira. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. **Ciência da Informação**, Brasília, v. 35, n. 3, p. 236-247, set./dez. 2006. DOI: 10.18225/ci.inf..v35i3.1130. Disponível em: <https://www.brappci.inf.br/index.php/res/v/20958>. Acesso em: 20 jul. 2023.

ASSIS, Tainá Batista de. **Gestão de teses e dissertações eletrônicas no contexto do acesso aberto à informação científica no Brasil**. 2015. 208 f., il. Dissertação (Mestrado em Ciência da Informação) — Universidade de Brasília, Brasília, 2015. Disponível em: <https://repositorio.unb.br/handle/10482/18621>. Acesso em: 22 mar. 2023.

BARRERA, Marcial Contreras. Minería de texto en la clasificación de documentos digitales. **Biblios: Journal of Librarianship and Information Science**, Peru, n. 64, p. 33-43, 2016. DOI: 10.5195/biblios.2016.309. Disponível em: <https://brapci.inf.br/index.php/res/v/70501>. Acesso em: 16 dez. 2022.

BARROSO, Pedro Augusto de Lima; SALES, Igor Vinícius Pinto de; MELO, Josemar Henrique de. “Os repositórios digitais: da sua origem a atualidade tecnológica dentro da arquivologia.” In: SEMINÁRIO DE SABERES ARQUIVISTICOS, 8., 2017, João Pessoa, **Anais[...]** João Pessoa: Universidade Federal da Paraíba, 2017, p. 155-167.

BASTOS, Flávia Maria. **Organização do conhecimento em bibliotecas digitais de teses e dissertações**: análise da aplicabilidade das teorias macroestruturais para categorização de áreas de assunto. 2005. 118 f. Dissertação (Mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2005. Disponível em: <http://hdl.handle.net/11449/93702>. Acesso em: 28 jul. 2023.

BATISTA, Lucineia da Silva; SILVA, Stephanie Cerqueira; JORENTE, Maria José Vicentini. Convergência do design da informação com a curadoria digital como recurso potencializador para a mediação da informação em ambientes digitais de museus. In: ENCONTRO DE PESQUISA EM INFORMAÇÃO E MEDIAÇÃO, 3., 2021. **Anais eletrônicos [...]**. Marília: [s.n.], 2021. p. 1-16. Disponível em: https://www.researchgate.net/publication/354862492_CONVERGENCIA_DO_DESIGN_DA_INFORMACAO_COM_A_CURADORIA_DIGITAL_COMO_RECURSO_POTENCIALIZADOR_PARA_A_MEDIACAO_DA_INFORMACAO_EM_AMBIENTES_DIGITAIS_DE_MUSEUS. Acesso em: 11 jul. 2023.

BIBLIOTECA DIGITAL Brasileira de Teses e Dissertações celebra 18 anos em 2020. **Instituto Brasileiro de Informação em Ciência e Tecnologia**, 2020. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2020/12/biblioteca-digital-brasileira-de-teses-e-dissertacoes-celebra-18-anos-em-2020>. Acesso em: 28 mar 2023.

BLATTMANN, Ursula; WEBER, Claudiane. DSpace como repositório digital na organização. **Revista ACB**, Florianópolis, v. 13, n. 2, p. 467-485, nov. 2008. ISSN 1414-0594. Disponível em: <https://revista.acbsc.org.br/racb/article/view/593/694>. Acesso em: 12 jan. 2023.

BORGES, Leandro da Conceição *et al.* Potencialidades dos repositórios institucionais das universidades federais brasileiras: apontamentos sobre software, equipe, manual, tutorial e política. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 9, n. 2, p. 245-265, 2019. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/42962/27460>. Acesso em: 08 ago. 2022.

BORKO, H. Information Science. What is it? **American Documentation**, [s.l.], v. 19, n. 1, p. 3-5, jan. 1968. (Tradução Livre)

BRAYNER, Aquiles Alencar. Humanidades Digitais e as instituições memoriais brasileiras: evolução em descompasso. In: PIMENTA, Ricardo; ALVES, Daniel (org.). **Humanidades digitais e o mundo lusófono**. FGV Editora, 2021. p. 234-248. Disponível em: <https://run.unl.pt/handle/10362/119335?mode=full>. Acesso em: 13 jul. 2023.

BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. **Aslib Proceedings: New Information Perspectives**, [s.l.], v. 58, n. 1/2, p. 49-72, 2006. Disponível em: https://www.academia.edu/74329038/The_need_for_a_faceted_classification_as_the_basis_of_all_methods_of_information_retrieval. Acesso em: 17 nov. 2022.

BUDAPEST OPEN ACCESS INITIATIVE (BOAI). Declaração de Budapeste. 2002. Disponível em: <http://www.budapestopenaccessinitiative.org/>. Acesso em: 23 ago. 2022.

BUSA, Roberto. The annals of humanities computing: The index Thomisticus. **Computers and the Humanities**, v. 14, p. 83-90. Disponível em: <https://www.jstor.org/stable/30207304>. Acesso em: 8 dez. 22.

CÂMARA JÚNIOR, Auto Tavares da. **Indexação automática de acórdãos por meio de processamento de linguagem natural**. 141 f. 2007. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2007. Disponível em: https://repositorio.unb.br/bitstream/10482/2493/1/2007_AutoTavaresdaCamaraJunior.PDF. Acesso em: 25 nov. 2022.

CÂMARA JÚNIOR, Auto Tavares da. **Processamento de linguagem natural para indexação automática semântico-ontológica**. Brasília, 2013. 180 f., il. Tese (Doutorado em Ciência da Informação) –Faculdade de Ciência da Informação, Universidade de Brasília, 2013. Disponível em: <https://repositorio.unb.br/handle/10482/13768>. Acesso em: 3 dez. 2021.

CAMPOS, Arthur Ferreira; DIAS, Guilherme Ataíde; SOUSA, Marckson Roberto Ferreira de. Princípios FAIR, Encontrabilidade e Dados: por um diálogo teórico e uma aplicação prática. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, [s.l.], v. 28, p. 1–21, 2023. DOI: 10.5007/1518-2924.2023.e88060. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/88060>. Acesso em: 17 fev. 2023.

CAMPOS, Maria Luiza de Almeida. **As Cinco Leis da Biblioteconomia e o exercício profissional**. [s.l.]: 1999. Disponível em: <http://www.conexaorio.com/bit/mluiza/index.htm>. Acesso em: 24 mar. 2022.

CAMPOS, Maria Luiza Almeida; GOMES, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. **Datagramazero: Revista de Ciência da Informação**, [s.l.], v. 9, n. 4, ago. 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/6615>. Acesso em: 04 nov. 2022.

CARLAN, Eliana; MEDEIROS, Marisa Brascher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 4, n. 2, p. 53–73, 2011. DOI: 10.26512/rici.v4.n2.2011.1675. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/1675>. Acesso em: 28 nov. 2022.

CERRAO, Natalia Gallo; CASTRO, Fabiano Ferreira de. Repositórios institucionais das universidades federais brasileiras: análise da representação da informação. **Informação & Tecnologia**, Marília/João Pessoa, n. 1, v. 5, p. 92-104, 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/110458>. Acesso em: 01 nov. 2022.

COELHO, Odete Máya Mesquita; SANTOS, Raimundo Nonato Ribeiro dos. Importância do ISBN, ISSN e DOI para o processo de legitimação e recuperação da informação científica. In: GAUDÊNCIO, Mário; ALBUQUERQUE, Maria Elizabeth Baltar Carneiro de (Orgs.). **Criação intelectual na comunicação científica: reflexões e orientações**. Mossoró: EdUFERSA, 2016. p. 67-83. Disponível: <http://www.repositorio.ufc.br/handle/riufc/18984>. Acesso em: 2 mar. 2023.

CURTY, Renata. Abordagens de reuso e a questão da reusabilidade dos dados científicos. **Liinc em Revista**, [s. l.], v. 15, n. 2, 2019. DOI: 10.18617/liinc.v15i2.4777. Disponível em: <https://revista.ibict.br/liinc/article/view/4777>. Acesso em: 1 mar. 2023.

DZIEKANIAK, Gisele Vasconcelos. A organização da informação e a comunicação científica: implicações para os profissionais e usuários da informação. **Em Questão**, Porto Alegre, v. 16, n. 1, p. 45-59, 2010. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/10748>. Acesso em: 19 out. 2022.

EIRÃO, Thiago Gomes; CUNHA, Murilo Bastos. A atualidade e utilidade da disseminação seletiva da informação e da tecnologia RSS. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 17, n. 33, p. 59-78, jan./abr., 2012. ISSN 1518-2924. DOI: 1518-2924.2011v17n33p59. Disponível em: <https://brapci.inf.br/index.php/res/v/39729>. Acesso em: 14 out. 2022.

EPSTEIN, Daniel. **Uso do minerador de textos Sobek como ferramenta de apoio à compreensão textual**. 2017. 170 f. Tese (Doutorado em Informática na Educação) – Universidade Federal do Rio Grande do Sul, 2017. Disponível em: <https://www.lume.ufrgs.br/handle/10183/178332>. Acesso em: 15 abr. 2023.

ESPÍRITO SANTO, Carmelita do. A importância das Cinco Leis da Biblioteconomia para a gestão por processos e transformação da visibilidade da biblioteca especializada. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.12, n.3, p.96-118, set./dez. 2014. ISSN 1678-765X. Disponível em: <https://www.alice.cnptia.embrapa.br/bitstream/doc/994595/1/AIIMPORTANCIADASCINCOLEISDABIBLIOTECONOMIAPARAAGESTAO.pdf>. Acesso em: 14 out. 2022.

FERNANDES, Júlio César; PINTO, Virgínia Bentes; FARIAS, Gabriela Belmont de. Percepção dos bibliotecários sobre a aplicabilidade das cinco leis de Ranganathan ao contexto da gestão da informação. **Informação em Pauta**, Fortaleza, v. 6, n. especial, p. 108-125, maio 2021. Disponível em: https://repositorio.ufc.br/bitstream/riufc/60907/1/2021_art_jcfernandesvbpintogbfarias.pdf. Acesso em: 7 fev. 2023.

FERREIRA, Ana Maria Jensen Ferreira da Costa *et al.* Serviços de informação em repositórios institucionais. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 7., 2017, Londrina. **Anais eletrônicos...** Londrina: Universidade Estadual de Londrina, 2017. p. 600-

615. Disponível em:

<http://www.uel.br/eventos/cinf/index.php/secin2017/secin2107/paper/view/462>. Acesso em: 31 jan. 2023.

FICHE, R. H.; MOURA, L. R. C.; MAIA, L. C. G. A utilização dos repositórios digitais encontrados nas organizações. **RACIn**, João Pessoa, v. 5, n. 1, p. 61-80, 2017. Disponível em: http://www.luiзмаia.com.br/docs/2017-racin_v5_n1_artigo04.pdf. Acesso em: 12 set. 2022.

FIGUEIREDO, Nice Menezes de. A modernidade das Cinco Leis de Ranganathan. **Ciência da Informação**, Brasília, v. 21, n.3, set/dez.,1992. Disponível em: <http://revista.ibict.br/ciinf/article/view/430/430>. Acesso em: 24 mar. 2022.

FONSECA, Diego Leonardo de Souza; GONÇALEZ, Paula Regina Ventura Amorim. Repositórios digitais como serviços estratégicos de informação: uma análise do repositório digital institucional do instituto federal do amazonas. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 18, p. 1-22, 2022. Disponível em: <https://brapci.inf.br/index.php/res/v/211717>. Acesso em: 24 jan. 2023.

FRANÇA, Fernanda Percia; ARAUJO, Denise Oliveira de; SILVA, Márcio Bezerra da. A ferramenta para repositórios institucionais DSpace: conceitos e características. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 13, n. 2, n. 2, p. 603-618, 2020. DOI: 10.26512/rici.v13.n2.2020.31160. Disponível em: <https://brapci.inf.br/index.php/res/v/141313>. Acesso em: 11 jan. 2023.

FUJITA, Mariângela Spotti Lopes. A biblioteca digital no contexto da gestão de bibliotecas universitárias: análise de aspectos conceituais e evolutivos para a organização da informação. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais...** Salvador: UFBA, 2005. Disponível em: http://cinform-anteriores.ufba.br/vi_anais/docs/MariangelaFujita.pdf. Acesso em: 25 jul. 2023.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia & Ciência da Informação**, Campinas, v. 1, n. 1, p. 60-90, 2003. DOI: 10.20396/rdbci.v1i1.2089. Disponível em: <https://www.brapci.inf.br/index.php/res/v/40149>. Acesso em: 20 jul. 2023.

FUJITA, Mariângela Spotti Lopes. Aspectos evolutivos das bibliotecas universitárias em ambiente digital na perspectiva da rede de bibliotecas da UNESP. **Informação & Sociedade: Estudos**, João Pessoa, v. 15 n.2 2005, n. 2, 2005. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/91372>. Acesso em: 22 jul. 2022.

FUJITA, Mariângela Spotti Lopes; TOLARE, Jessica Beatriz. Vocabulários controlados na representação e recuperação da informação em repositórios brasileiros. **Informação & Informação**, Londrina, v. 24, n. 2, p. 93-125, 2019. DOI: 10.5433/1981-8920.2019v24n2p9. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/37985>. Acesso em: 26 out. 2022.

GO-FAIR. **FAIR Principles**, [ca. 2016]. Disponível em: <https://www.go-FAIR.org/FAIR-principles/>. Acesso em: 16 fev. 2023.

GOLUB, Koraljka; KAMAL, Ahmad M.; VEKSELIUS, Johan. Knowledge organisation for digital humanities: An introduction. In: **Information and Knowledge Organisation in Digital Humanities**. Nova York: Routledge, 2022. p. 1-22. DOI: 10.4324/9781003131816.

GOMES, Hagar Espanha. Marcos históricos e teóricos da organização do conhecimento. **Informação & Informação**, Londrina, v. 22, n. 2, p. 33-66, 2017. DOI: 10.5433/1981-8920.2017v22n2p33. Disponível em: <https://brapci.inf.br/index.php/res/v/33713>. Acesso em: 20 out. 2022.

GONÇALVES, Pâmella Benevides *et al.* Preservação digital distribuída para teses e dissertações: uma proposta para as bibliotecas universitárias da UNESP e UFRN. **Revista Brasileira de Preservação Digital**, Campinas, v. 3, n. 2022, 2006. DOI: 10.20396/rebpred.v3i00.16580. Disponível em: <https://brapci.inf.br/index.php/res/v/198647>. Acesso em: 15 ago. 2023.

GRACIO, José Carlos Abbud. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 127 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2002. Disponível em: <http://hdl.handle.net/11449/93722>. Acesso em: 19 abr. 2023.

GUERREIRO, Dália; BORBINHA, José Luís. Humanidades Digitais: novos desafios e oportunidades. **Revista Internacional del Libro, Digitalización y Bibliotecas**, Portugal, v. 2, n. 2, 2014. Disponível em: https://dspace.uevora.pt/rdpc/bitstream/10174/12155/2/Dalia_Guerreiro_artigo_repositorio.pdf. Acesso em: 19 out. 2022.

HENNING, Patrícia Corrêa *et al.* Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR. **Tendências da Pesquisa Brasileira e Ciência da Informação, ANCIB**, [s.l.], v. 11, n. 1. 2018. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/458>. Acesso em: 15 fev. 2023.

HENNING, Patrícia Corrêa *et al.* GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. **Em Questão**, Porto Alegre, v. 25, n. 2, p. 389-412, maio/ago. 2019. Disponível em: <https://www.arca.fiocruz.br/handle/icict/55755>. Acesso em: 16 fev. 2023.

IBICT. **Sobre repositórios digitais**. Instituto Brasileiro de Informação em Ciência e Tecnologia. 2015. Disponível em: <http://sitehistorico.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais>. Acesso em: 26 de maio. 2022.

KIRSCHENBAUM, Matthew G. What is digital humanities and what's it doing in english departments? **ADE Bulletin**, [s. l.], v. 150, n. 7, p. 55-61, 2010. Disponível em: Disponível em: <https://www.ade.mla.org/content/download/7914/225677>. Acesso em: 13 jul. 2023.

KURAMOTO, Hélio. Implantação de repositórios institucionais em universidades e instituições de pesquisa do Brasil: do projeto ao processo. In: GOMES, Maria João; ROSA, Flávia (Org.). **Repositórios Institucionais: democratizando o acesso ao conhecimento**. Salvador: UFBA, 2010. Disponível em:

<https://repositorio.ufba.br/bitstream/ri/616/3/Repositorios%20institucionais.pdf>. Acesso em: 21 jul. 2022.

LEITE, Fernando César Lima. A relação ainda não discutida entre repositórios institucionais de acesso aberto e teses e dissertações eletrônicas no Brasil: um convite à reflexão. In: CONFERÊNCIA LUSO-BRASILEIRA SOBRE ACESSO ABERTO, 2., 2011, **Anais...** Rio de Janeiro: CONFOA, 2011. 6 p. Disponível em: http://eprints.rclis.org/24487/1/CONFOA_2011%20-%20RIXBDTD.pdf. Acesso em: 27 mar. 2023.

LEITE, Fernando César de Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009. Disponível em: http://repositorio.unb.br/bitstream/10482/4841/1/LIVRO_ComoAmpliareGerenciar.pdf. Acesso em: 24 fev. 2023.

MACULAN, Benildes Coura Moreira dos Santos. **Taxonomia facetada navegacional**: construção a partir de uma matriz categorial para trabalhos acadêmicos. 2011. 185 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2011. Disponível em: <http://hdl.handle.net/1843/ECID-8LAN5K>. Acesso em: 20 dez. de 2022.

MACULAN, Benildes Coura Moreira dos Santos; LIMA, Gercina Angela Borém de Oliveira; PENIDO, Patrícia. Taxonomia facetada como interface para facilitar o acesso à informação em bibliotecas digitais. **Revista ACB**, [s.l.], v. 16, n. 1, p. 234-249, set. 2010. ISSN 1414-0594. Disponível em: <https://revistaacb.emnuvens.com.br/racb/article/view/746>. Acesso em: 20 dez. 2022.

MAGALHÃES, Lúcia Helena de; SOUZA, Renato Rocha. Sistema de recuperação da informação: uma abordagem baseada em ontologias. **Ponto de Acesso**, Salvador, v. 13, n. 2, p. 63-85, 2019. DOI: 10.9771/rpa.v13i2.28452. Disponível em: <https://www.brapci.inf.br/index.php/res/v/125642>. Acesso em: 31 out. 2022.

MARANHÃO, Ana Maria Neves; QUEIROZ, Claudete Fernandes de; RODRIGUES, Raphael Belchior. Curadoria digital de dados no Arca - Repositório Institucional da Fiocruz: relato de experiência. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [s.l.], v. 11, 2017. DOI: 10.29397/reciis.v11i0.1388. Disponível em: <https://brapci.inf.br/index.php/res/v/131801>. Acesso em: 24 jan. 2023.

MARCONDES, Carlos Henrique; SAYÃO, Luís Fernando. À guisa de introdução: repositórios institucionais e livre acesso. In: SAYÃO, Luís Fernando *et al.* (Org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 9-21. Disponível em: <https://repositorio.ufba.br/handle/ufba/473>. Acesso em 01 fev. de 2023.

MARÍN-ARRAIZA, Paloma; HEREDIA, Ana. FAIR PIDs: O papel da ORCID no fortalecimento dos Princípios FAIR. In: SALES, Luana Farias *et al* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 23 -30. DOI: 10.22477/9786589167242.cap2. Disponível em:

https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gest%C3%A3o%20de%20dados%20de%20pesquisa_2021.pdf. Acesso em: 19 abr. 2023.

MARQUES, Clediane de Araújo Guedes. Gerenciamento de repositórios digitais: apontamentos práticos para o desenvolvimento dos repositórios institucionais. **Convergência em Ciência da Informação**, [s.l.], v. 3 n. 2, n. 2, p. 135-162, 2020. DOI: 10.33467/conci.v3i2.13438. Disponível: https://repositorio.ufrn.br/bitstream/123456789/31514/4/GerenciamentoRepositoriosDigitais_Marques_2020.pdf. Acesso em: 18 maio 2022.

MARTINS, D. L.; SILVA, M. F.; SIQUEIRA, J. Comparação entre sistemas para criação de acervos digitais: análise dos softwares livres DSpace, EPrints, Fedora, Greenstone e Islandora a partir de novas dimensões analíticas. **InCID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 9, n. 1, p. 52-71, 2018. DOI: 10.11606/issn.2178-2075.v9i1p52-71. Disponível em: <https://www.revistas.usp.br/incid/article/view/134333>. Acesso em: 13 set. 2022.

MARTINS, Maria Madalena Schmid. Curadoria digital nos acervos de cultura e memória: Coleção Digital Campanha Civilista da Fundação Casa de Rui Barbosa. **RBHD**, Rio de Janeiro, v. 1, n. 1, Dossiê Temático 3, p. 22-39, jan./jun., 2021. Disponível em: <https://brapci.inf.br/index.php/res/v/168949>. Acesso em: 31 jan. 2023.

MATOS, J. C. M.; JACINTHO, E. M. D. S. B.; ALVAREZ, E. B. Humanidades digitais e a simbiose entre humano e máquina: algumas reflexões comparativas entre a interpretação e a mineração de textos. **Logeion: filosofia da informação**, Rio de Janeiro, v. 6, n. 1, p. 126-145, 2019. DOI: 10.21728/logcion.2019v6n1.p126-145. Disponível: https://www.researchgate.net/publication/335832452_HUMANIDADES_DIGITAIS_E_A_SIMBIOSE_ENTRE_HUMANO_E_MAQUINA_ALGUMAS_REFLEXOES_COMPARATIVAS_ENTRE_A_INTERPRETACAO_E_A_MINERACAO_DE_TEXTOS. Acesso em: 07 nov. 2022.

MATTOS, Bruno Marques de. **A Organização do Conhecimento nas Humanidades Digitais e o Conceito de Leitura Distante de Franco Moretti**. 2019. 121 f. Dissertação (Mestrado em História da Ciência) - Programa de Estudos Pós-Graduados em História da Ciência, Pontifícia Universidade Católica de São Paulo, São Paulo, 2019. Disponível em: <https://tede2.pucsp.br/handle/handle/22765>. Acesso em: 21 out. 2022

MEDEIROS, Wagner Oliveira; PINHO, Fábio Assis; CORRÊA, Renato Fernandes. Aplicação de software de mineração de texto na representação da informação de obras artístico-pictóricas. **Logeion: filosofia da informação**, Rio de Janeiro, v. 6, n. 1, p. 146-170, 2019. DOI: 10.21728/logcion.2019v6n1.p146-170. Disponível em: https://www.researchgate.net/publication/335831858_APLICACAO_DE_SOFTWARE_DE_MINERACAO_DE_TEXTO_NA_REPRESENTACAO_DA_INFORMACAO_DE_OBRAS_ARTISTICO-PICTORICAS. Acesso em: 07 nov. 2022.

MONTEIRO, Fernanda de Souza. **Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca**. 2013. 189 f., il. Tese (Doutorado em Ciência da Informação) –Faculdade de Ciência da Informação, Universidade de Brasília. Disponível em:

https://repositorio.unb.br/bitstream/10482/13461/1/2013_FernandadeSouzaMonteiro.pdf. Acesso em: 20 nov. 2021.

MOURA, Maria Aparecida. Ciência da Informação e Humanidades Digitais: mediações, agência e compartilhamento de saberes. **Perspectivas em Ciência da Informação**, [s.l.], v. 24, p. 57–70, jan./mar.2019. DOI 10.1590/1981-5344/3893. Disponível em: <http://www.scielo.br/j/pci/a/GzKTfV7D4s7RLMM6qgpHkJF/?lang=pt>. Acesso em: 29 mar. 2022.

NASCIMENTO, Felipe Mozart de Santana; PINHO, Fabio Assis. Sistemas de organização do conhecimento: semelhanças e diferenças. **Convergência em Ciência da Informação**, [s.l.], v. 2 n. 3, n. 3, p. 104-122, 2019. DOI: 10.33467/conci.v2i3.13673. Disponível em: <https://brapci.inf.br/index.php/res/v/140503>. Acesso em: 14 fev. 2023.

NHACUONGUE, Januário Albino; DUTRA, Moisés Lima. De Paul Otlet à Web Semântica: aportes teóricos sobre a organização do conhecimento. **Informação & Tecnologia**, Marília/João Pessoa, v. 3, n. 1, p. 138-158, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41780>. Acesso em: 26 dez. 2022.

NISO. **Understanding Metadata**. Bethesda, MD: NISO Press, 2004. Disponível em: https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf. Acesso em: 24 ago. 2022.

OLIVEIRA, Kadidja Valéria Reginaldo de; KAYA, Gabriela Tyemi; RONCAGLIO, Cynthia. Ciências da informação e humanidades digitais: produção, consumo e materialidade da informação em plataformas digitais. **Acervo - Revista do Arquivo Nacional**, [s.l.], v. 35, n. 1, p. 1-13, 2022. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/193731>. Acesso em: 11 ago. 2022.

OLIVEIRA, Marlene de Oliveira. Origens e evolução de Ciência da Informação. In: OLIVEIRA, Marlene de (Coord.). **Ciência da Informação e Biblioteconomia: novos conteúdos e espaços de atuação**. Belo Horizonte: Ed. UFMG, 2005. Cap. 1, p. 9-28.

OLIVEIRA, Thierre Xavier de. **O papel dos repositórios digitais na construção de políticas de Ciência Aberta**: o processo de implantação do Repositório Institucional de Múltiplos Acervos da UFRRJ. 2021. 182 f. Dissertação (Mestrado em Humanidades Digitais) - Instituto Multidisciplinar, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2021. Disponível em: <https://tede.ufrj.br/jspui/handle/jspui/1285>. Acesso em: 17 mar. 2023.

PALETTA, Francisco Carlos. Ciência da Informação e Humanidades Digitais: uma reflexão. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 19., 2018, Londrina. **Anais ... Londrina**, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/103662>. Acesso em: 12 abr. 2022.

PALETTA, Francisco Carlos; MOREIRO-GONZÁLEZ, José Antonio. Modelos de busca, acesso, recuperação e apropriação da informação na web de dados. **Informação & Informação**, Londrina, v. 24, n. 2, p. 182-210, 2019. DOI: 10.5433/1981-8920.2019v24n2p182. Disponível em: <https://brapci.inf.br/index.php/res/v/125792>. Acesso em: 25 out. 2022.

PANAGE, B.M., Application of Dr. S.R. Ranganathan's five laws of library science to digital library, In: **Proceedings of the International Conference on Social Sciences**. Sri Lanka: University of Kelaniya, 2008. p. 274. Disponível em: <http://repository.kln.ac.lk/handle/123456789/6858>. Acesso em: 21 jul. 2023.

PAVÃO, Caterina Groposo da Costa. *et al.* Metadados e repositórios institucionais: uma relação indissociável para a qualidade da recuperação e visibilidade da informação. **Ponto de Acesso**, Salvador, v. 9, n. 3, p. 103-116, 2015. DOI: 10.9771/1981-6766rpa.v9i3.15163. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/15163/10328>. Acesso em: 11 ago. 2022.

PIMENTA, Ricardo Medeiros. Por que Humanidades Digitais na Ciência da Informação? Perspectivas pregressas e futuras de uma prática transdisciplinar comum. **Informação & Sociedade: Estudos**, [s.l.], v. 30, n. 2, 15 jun. 2020. DOI 10.22478/ufpb.1809-4783.2020v30n2.52122. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/52122>. Acesso em: 31 mar. 2022.

PONTES, Flavio Vieira.; LIMA, Gercina Ângela Borém de Oliveira. A organização do conhecimento em ambientes digitais: aplicação da teoria da classificação facetada. **Perspectivas em Ciência da Informação**, [s.l.], v. 17, n. 4, p. 18-40, 2012. DOI: <https://doi.org/10.1590/S1413-99362012000400003>. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/35754>. Acesso em: 16 nov. 2022.

PRODANOV, Cleber Cristiano; FREITAS, Ernani César de. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2. ed. Novo Hamburgo: Feevale, 2013.

QUEIROZ, Claudete Fernandes de; ALVES, Aline da Silva. **Repositórios: conceitos, abordagens e perspectivas**. Rio de Janeiro: FIOCRUZ/ICICT, 2018. 51 p. Trabalho apresentado no Curso de Acesso à Informação Científica e Tecnológica em Saúde. Modalidade: Qualificação. Disponível em: https://www.arca.fiocruz.br/bitstream/icict/29474/4/Aula_sobre_Reposit%C3%B3rios_Claudete_Fernandes_ICICT_2018.pdf. Acesso em: 18 de mar. 2022.

QUEIROZ, Claudete Fernandes de; RODRIGUES, Raphael Belchior. **Curso repositórios digitais**. Rio de Janeiro: RBRD - Rede Sudeste; Fiocruz/Icict, 2022. 72 p. Disponível em: https://www.arca.fiocruz.br/bitstream/icict/53382/2/Curso_Reposit%C3%B3rios_Digitais_Claudete_Raphael_RIAA-Sudeste_2022.pdf. Acesso em: 05 de ago. 2022.

QUEIROZ, D. G. de C.; MOURA, A. M. M. de. Ciência da Informação: história, conceitos e características. **Em Questão**, [s. l.], v. 21, n. 3, p. 26-42, 2015. DOI: 10.19132/1808-5245213.26-42. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/57516>. Acesso em: 7 abr. 2022.

RANGANATHAN, S. R. **As cinco leis da Biblioteconomia**. Briquet de Lemos: Brasília, DF, 2009. Tradução de Tarcisio Zandonade.

REZENDE, Laura Vilela Rodrigues; BARBOSA, Sonia. Using the DATAVERSE project to move towards FAIR principles. In: SALES, Luana Farias *et al* (org.). **Princípios FAIR**

aplicados à gestão de dados de pesquisa. Rio de Janeiro: Ibict, 2021. p. 31 -46. DOI: 10.22477/9786589167242.cap3. Disponível em: https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gestão%20de%20dados%20de%20pesquisa_2021.pdf. Acesso em: 19 abr. 2023.

RIO RIANDE, Gimena del. Humanidades Digitales CAICYT: Las Humanidades Digitales en la agenda de investigación de las Ciencias de la Información. In: JOBICYT DEL CONGRESO DE ABGRA, 8, 2018, Argentina, Buenos Aires. **Anais...** Disponível em: <https://doi.org/10.5281/zenodo.3233995>. Acesso em: 28 abr. 2022.

RODRIGUES, Alessandra Pereira *et al.* A interoperação com repositórios digitais: protocolos e exemplos. **Ciência da Informação**, Brasília, v. 40, n. 3, 2013. DOI: 10.18225/ci.inf.v40i3.1294. Disponível em: <https://revista.ibict.br/ciinf/article/view/1294/1472>. Acesso em: 5 ago. 2022.

ROLLO, Maria Fernanda. Desafios e responsabilidades das Humanidades Digitais: preservar a memória, valorizar o patrimônio, promover e disseminar o conhecimento: o Programa Memória Para Todos. **Estudos Históricos**, Rio de Janeiro, v. 33, n. 69, p. 19-44. DOI: <https://doi.org/10.1590/S2178-149420200001000003>. Disponível em: <https://www.semanticscholar.org/paper/DESAFIOS-E-RESPONSABILIDADES-DAS-HUMANIDADES-A-O-E-Rollo/3fdd21438f6f8ad9aeae3d144409a458617690f8#citing-papers>. Acesso em: 9 maio 2022.

SALCEDO, Diego; BEZERRA, Vinícius Cabral Accioly. A gênese do repositório filatélico brasileiro: uma experiência interdisciplinar nas Humanidades Digitais. **Informação & Sociedade: Estudos**, João Pessoa, v. 28, n. 3, 2018. DOI: 10.22478/ufpb.1809-4783.2018v28n3.38295. Disponível em: <https://brapci.inf.br/index.php/res/v/109121>. Acesso em: 06 jul. 2023.

SALES, Luana Farias. **Curso princípios FAIR aplicados à repositórios.** Rio de Janeiro: Sudeste/RIAA; IBICT/PPGCI, 2021a. 94 p. Disponível em: <https://www.arca.fiocruz.br/handle/icict/50571>. Acesso em: 15 fev. 2023.

SALES, Luana Farias. FAIRificação de repositórios! In: ENCONTRO DA REDE SUDESTE DE REPOSITÓRIOS INSTITUCIONAIS, 2., 2021b, Rio de Janeiro. **Anais...** Rio de Janeiro: Fiocruz/Icict, 2021b. 40 p. Disponível em: <https://www.arca.fiocruz.br/handle/icict/48374>. Acesso em: 16 fev. 2023.

SALES, Luana Farias; CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Ontologias de domínio: um estudo das relações conceituais e sua aplicação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 7., 2006, Salvador. **Anais...** Salvador: ANCIB, 2006. Disponível em: <https://www.scielo.br/j/pci/a/qNXG6bfyKvsmtmKTcn6MwCp/?lang=pt>. Acesso em: 07 nov. 2022.

SALES, Luana Farias; ROCHA, Lucas de Lima; CAVALCANTI, Márcia Teixeira. Desenvolvimento de um vocabulário controlado para o repositório institucional CarpediEN. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [s.l.], v. 11, 2017. Disponível em: <https://brapci.inf.br/index.php/res/v/133788>. Acesso em: 01. nov. 2022.

SALES, Luana Farias *et al.* Um panorama histórico da iniciativa GO FAIR: da Europa para o Brasil. In: SALES, Luana Farias *et al.* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 9 -22. DOI: 10.22477/9786589167242.cap1. Disponível em:

https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gestão%20de%20dados%20de%20pesquisa_2021.pdf. Acesso em: 19 abr. 2023.

SANTAREM SEGUNDO, José Eduardo; SILVA, Marcel Ferrante; MARTINS, Dalton Lopes. Revisitando a interoperabilidade no contexto dos acervos digitais. **Informação & Sociedade: Estudos**, João Pessoa, v. 29, n. 2, 2019. Disponível em:

<https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/38107>. Acesso em: 18 jan. 2023.

SANTOS, Ana Paula Lima dos. **Relações interdisciplinares entre a Ciência da Informação e a Biblioteconomia: limites e possibilidades**. 2012. 133 f.: il. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal Fluminense, Departamento de Ciência da Informação, Niterói, 2012. Disponível em:

<https://app.uff.br/riuff/handle/1/10323>. Acesso em: 13 dez. 2022.

SANTOS, Francisco Edvander Pires; LIMA, Juliana Soares; ARAÚJO, Irlana Mendes de. Representação e recuperação da informação em bibliotecas universitárias: o tesouro como elo entre acervo, indexação e usuários da área de física. **Ciência da Informação em Revista**, Maceió, v. 4, n. 1, p. 43-58, 2017. Disponível em:

<http://hdl.handle.net/20.500.11959/brapci/36430>. Acesso em: 13 fev. 2023.

SANTOS, Izabel Lima dos. A informação e suas leis: um paralelo entre o pensamento de Ranganathan e o de Moody e Walsh. **Revista Digital de Biblioteconomia & Ciência da Informação**, Campinas, v. 13, n. 3, p. 469-479, 2015. DOI: 10.20396/rdbci.v13i3.8635792. Disponível em: <https://brapci.inf.br/index.php/res/v/40031>. Acesso em: 29 set. 2022.

SANTOS, Luiz Olavo Bonino da Silva. Prefácio. In: SALES, Luana Farias *et al.* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 5 -6. DOI: 10.22477/9786589167242.pref. Disponível em:

https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gestão%20de%20dados%20de%20pesquisa_2021.pdf. Acesso em: 19 abr. 2023.

SANTOS, Raimunda Fernanda dos. Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. **Revista Informação na Sociedade Contemporânea**, Natal, n. Especial, p. 1-22, 2017. Disponível em:

<https://brapci.inf.br/index.php/res/v/106607>. Acesso em: 01-nov.-2022.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, 1996. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/37415>. Acesso em: 22 dez. 2022.

SAYÃO, Luis Fernando. Afinal, o que é biblioteca digital? **Revista USP**, São Paulo, n. 80, p. 6-17, 2009. DOI: 10.11606/issn.2316-9036.v0i80p6-17. Disponível em:

<https://www.revistas.usp.br/revusp/article/view/13709>. Acesso em: 17 nov. 2022.

SAYÃO, Luis Fernando. Padrões para bibliotecas digitais abertas e interoperáveis. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, n. esp.

1. sem., p. 18-47, 2007. DOI: 10.5007/1518-2924.2007v12nesp1p18. Disponível em: <https://brapci.inf.br/index.php/res/download/96157>. Acesso em: 15 ago. 2022.

SAYÃO, Luis Fernando; MARCONDES, Carlos Henrique. Software livres para repositórios institucionais: alguns subsídios para a seleção. In: SAYÃO, Luís Fernando *et al.* (Org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 23-54. Disponível em: <https://repositorio.ufba.br/handle/ufba/473>. Acesso em 01 fev. de 2023.

SAYÃO, Luis Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, [s.l.], v. 21, n. 2, p. 90-115, dez. 2016. ISSN 1981-8920. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27939/20122>. Acesso em: 18 maio 2022.

SAYÃO, Luís Fernando; SALES, Luana Farias. Um modelo de implementação para a internet de dados & serviços FAIR. In: SALES, Luana Farias *et al* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 215 - 242. DOI: 10.22477/9786589167242.cap16. Disponível em: https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gestão%20de%20dados%20de%20pesquisa_2021.pdf. Acesso em: 19 abr. 2023.

SHINTAKU, M.; MEIRELES, R. **Manual do DSpace administração de repositórios**. Salvador: EDUFBA, 2010. Disponível em: <https://repositorio.ufba.br/bitstream/ri/769/1/Manual%20do%20Dspace%282%29.pdf> . Acesso em: 26 maio. 2022.

SIEBRA, Sandra de Alburquerque; BORBA, Vildeane da Rocha; MIRANDA, Májory Karoline Fernandes de Oliveira. Curadoria digital: um termo interdisciplinar. **Informação & Tecnologia**, Marília/João Pessoa, v. 3, n. 2, p. 21-38, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41848>. Acesso em: 06 jul. 2023.

SILVA, Marcel Ferrante; MARTINS, Dalton Lopes; SIQUEIRA, Joyce. Web Semântica em repositórios: ontologia para representação de bibliotecas digitais. **Ciência da Informação em Revista**, Maceió, v. 6, n. 1, p. 99-113, 2019. DOI: 10.28998/cirev.2019v6n1f. Disponível: <https://www.seer.ufal.br/index.php/cir/article/view/5566/5872>. Acesso em: 15 set. 2022.

SOBRAL, Renato Machado; SANTOS, Cibele Araújo Camargo Marques dos. Repositórios institucionais digitais de informação científica: implementação com o software Dspace como solução técnica. **PRISMA.COM**, Porto, n. 35, p. 152-184, 2017. Disponível em: <https://www.eca.usp.br/acervo/producao-academica/002916776.pdf>. Acesso em: 13 jan. 2023.

SOUSA, Maria Eliziana Pereira de; TARGINO, Maria das Graças. Cinco leis da Biblioteconomia / Cinco Leis de Ranganathan: resistindo bravamente ao tempo. **Ciência da Informação em Revista**, Maceió, v. 3, n. 1, p. 11-29, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/35917>. Acesso em: 29 set. 2022.

SULA, Chris Alen. Digital humanities and libraries: a conceptual model. **Journal of Library Administration**, [s.l.], v. 53, n. 1, p. 10-26, 2013. DOI: 10.1080/01930826.2013.756680.

Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01930826.2013.756680>. Acesso em: 18 jul. 2023.

TAVARES, Aureliana Lopes de Lacerda. **Análise de risco e preservação digital**: uma abordagem sistêmica na rede memorial Pernambuco. 2014. Dissertação. (Mestrado em Ciência da Informação) - Universidade Federal de Pernambuco, Recife, 2014. Disponível em: <https://repositorio.ufpe.br/handle/123456789/26453>. Acesso em: 11 jul. 2023.

TERRA, José Cláudio C. *et al.* **Taxonomia**: elemento fundamental para a gestão do conhecimento. 2005. Disponível em: <https://pt.scribd.com/document/54921452/Taxonomia-Elemento-Fundamental-para-a-Gestao-do-Conhecimento-Terra-Forum-8p#>. Acesso em: 01 dez. 2022.

THATCAMP. **Manifesto das Humanidades Digitais**. Paris: ThatCamp [The Humanities and Technology Camp]. 2011. Disponível em: <https://tcp.hypotheses.org/318>. Acesso em: 12 abr. 2022.

THE DIGITAL Humanities Manifesto 2.0. Humanities Blast. [20--]. Disponível em: https://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf. Acesso em: 13 jul. 2023.

TONELLO, Izângela Maria Sansone; LUNARDELLI, Rosane Alvares; ALMEIDA JÚNIOR, Oswaldo Francisco de. Palavras-chave: possibilidades de mediação da informação. **Ponto de Acesso**, Salvador, v. 6, n. 2, p. 21-34, 2012. DOI: 10.9771/1981-6766rpa.v6i2.4524. Acesso em: 14 abr. 2023.

TRIPP, David. Pesquisa-ação: uma introdução metodológica. **Educação e Pesquisa**, São Paulo, v. 31, n. 3, p. 443-466, set./dez. 2005. Disponível em: <https://www.scielo.br/j/ep/a/3DkbXnqBQyyq5bV4TCL9NSH/?format=pdf&lang=pt>. Acesso em: 07 ago. 2023.

VARELA, Ainda; BARBOSA; Marilene Lobo Abreu, GUIMARÃES, Igor Baraúna. Dos processos analógicos às tecnologias digitais contemporâneas de recuperação da informação: caminhos cognitivos na mediação para o acesso ao conhecimento. In: SAYÃO, Luís Fernando *et al.* (Org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 123-161. Disponível em: <https://repositorio.ufba.br/handle/ufba/473>. Acesso em: 1 fev. de 2023.

VIGNOLI, Richele Grengé; SOUTO, Diana Vilas Boas; CERVANTES, Brígida Maria Nogueira. Sistemas de organização do conhecimento com foco em ontologias e taxonomias. **Informação & Sociedade: Estudos**, João Pessoa, v. 23, n. 2, 2013. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/91940>. Acesso em: 21 out. 2022.

WEITZEL, Simone da Rocha. O desenvolvimento de coleções e a organização do conhecimento: suas origens e desafios. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 7, n. 1, p. 61 - 67, jan./jun. 2002. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/414/227>. Acesso em: 20 nov. 2021.

WEITZEL, Simone da Rocha. O mapeamento dos repositórios institucionais brasileiros: perfil e desafios. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da

Informação, [s. l.], v. 24, n. 54, p. 105–123, 2019. DOI: 10.5007/1518-2924.2019v24n54p105. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2019v24n54p105>. Acesso em: 7 dez. 2022.

WILKINSON, Mark D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, [s.l.], v. 3, 160018, 2016. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em: 2 mar. 2023.

ANEXO A – Manual da BDTD da UFRRJ

Manual da BDTD/UFRRJ

Elaborado pela equipe técnica da SPT

1. Introdução

A BDTD (Biblioteca Digital de Teses e Dissertações) funciona como um repositório digital das dissertações/teses da Universidade Federal Rural do Rio de Janeiro. O repositório armazena, preserva e possibilita o acesso à produção científica da instituição.

O objetivo principal do manual é padronizar os procedimentos realizados no sistema de informação, relacionados ao preenchimento dos dados necessários ao registro de documentos, conforme segue.

2. Procedimentos para inserção dos exemplares na BDTD:

- Observar, no Termo de Autorização, se o documento tem alguma restrição (pode ser liberado para publicação no TEDE?). Quando o documento possuir alguma limitação, deve ser respeitado o prazo de guarda, e não realizar o depósito no repositório antes do período.
- Salvar os PDFs das teses/dissertações na pasta do computador.
- Abrir a página de busca de currículos na Plataforma Lattes (<http://buscatextual.cnpq.br/buscatextual/busca.do>); pesquisar autor e orientador (deixar as páginas abertas para copiar endereço do currículo deles)
- Obs.: Não esquecer de marcar na busca doutores e demais pesquisadores**
- Abrir o TEDE <https://tede.ufrjr.br/jspui/>
- Entrar em -> Meu espaço -> Fazer login
- Clicar em iniciar um novo depósito



Login

Entrar no repositório

Usuário novo? [Clique aqui para se registrar](#)

Digite seu endereço de e-mail e senha nos espaços abaixo

Endereço de e-mail: leticia@ufrj.br

Senha:

Entrar

[Esqueceu sua senha?](#)

Iniciar um novo depósito

TEDE

Meu espaço: Letícia Schettini

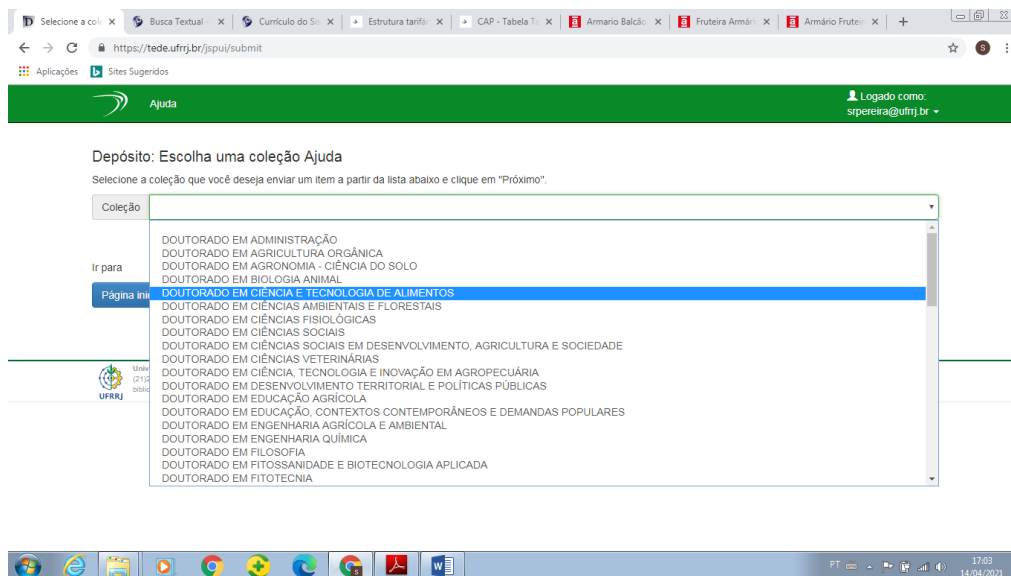
Iniciar um novo depósito

Ver depósito(s) aceito(s)

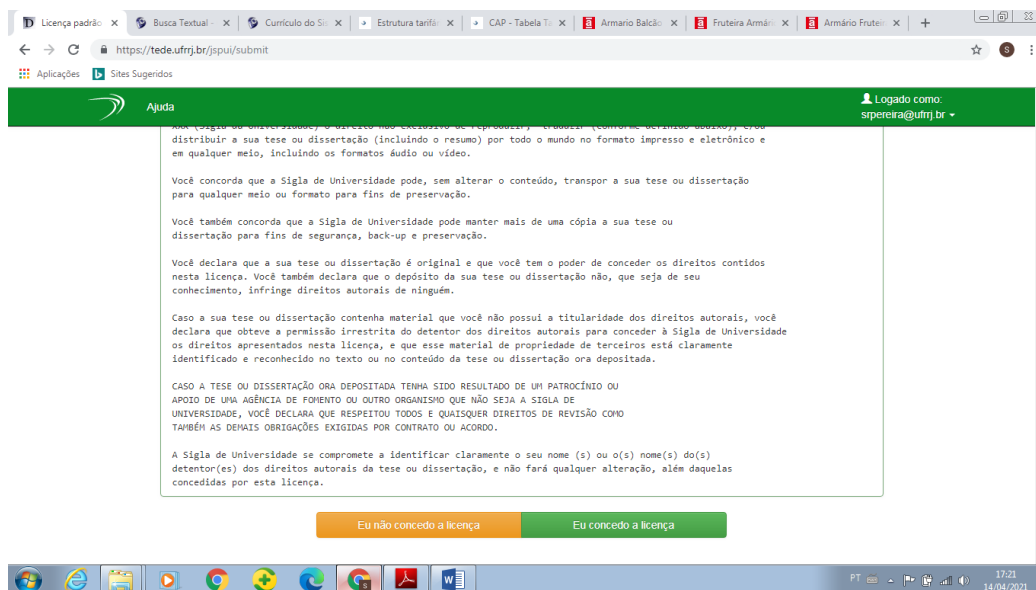
Universidade Federal Rural do Rio de Janeiro
(21) 2051-4530 / (21) 2051-4531
bibdig@ufrj.br

BDTD ibict

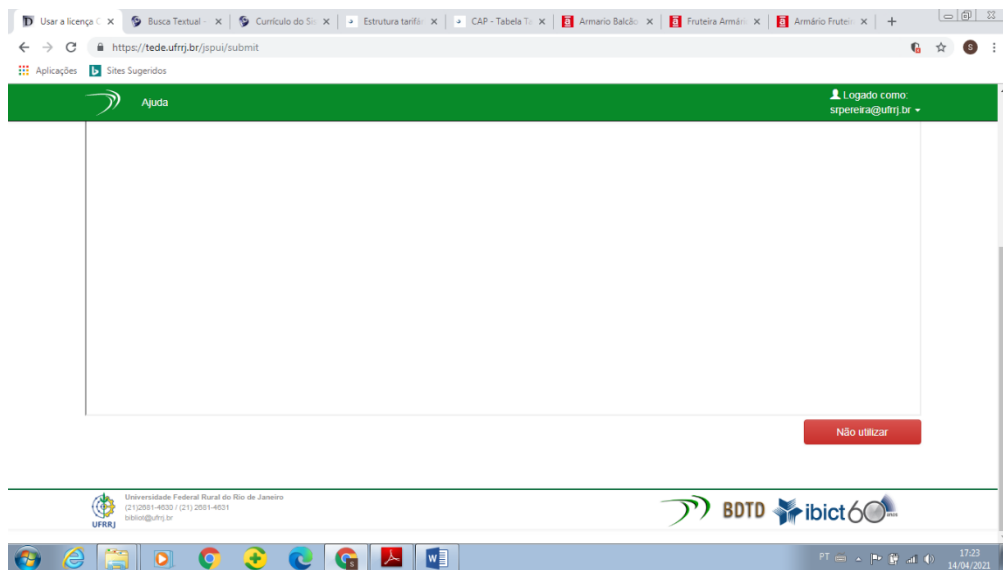
○ Escolha a coleção em que a tese ou dissertação pertence, e depois passe para a próxima página.



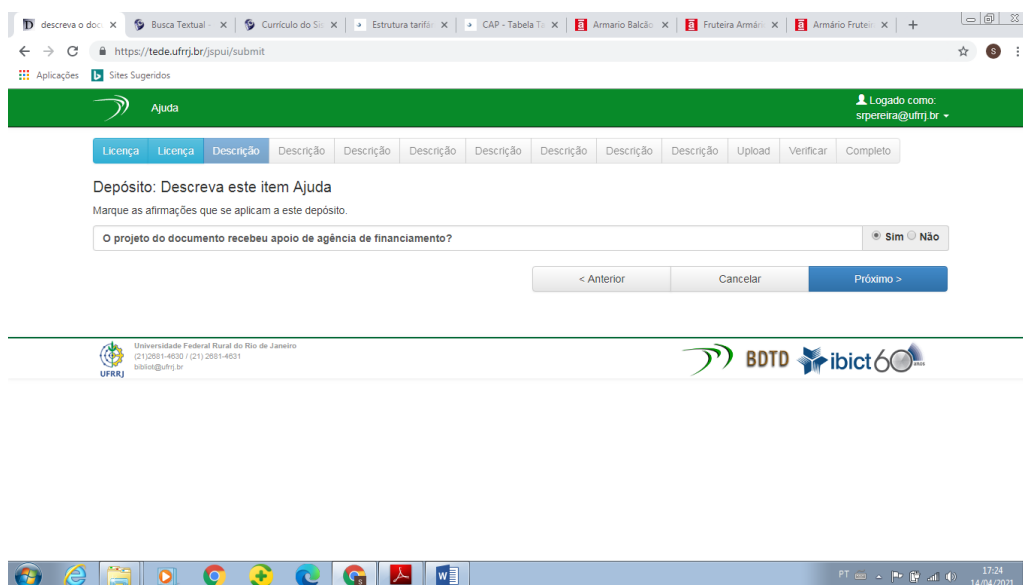
○ Selecionar: Eu concedo a licença



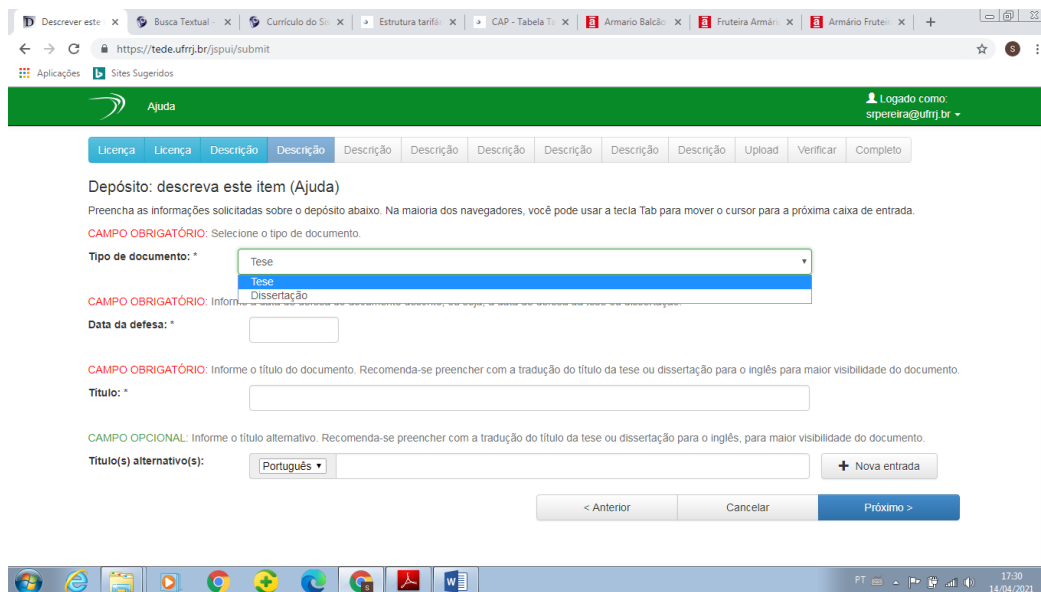
○ Clicar em “Não utilizar”



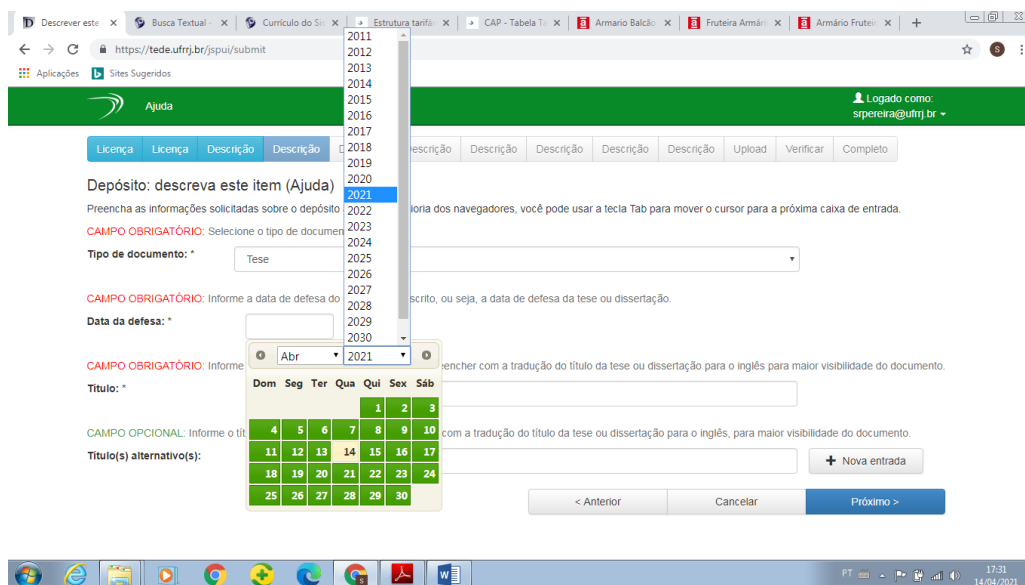
- Verificar na Plataforma Lattes do autor ou nos agradecimentos do documento se o trabalho teve ajuda financeira (exemplo: CAPES).
- Marcar Sim ou Não, e próximo.



- Seleciona o tipo de documento: Tese ou Dissertação



○ Inserir a data da defesa (informação disponível na página de aprovação do trabalho e no Termo de Autorização). Ordem: ano, mês e dia.



○ Inserir o título do trabalho, conforme aparece na folha de rosto, mas em caixa baixa (só a letra inicial do título ou nomes próprios podem entrar em caixa alta).
 ○ Para o título alternativo, utilizar a informação da referência no abstract, e substituir o Português para Inglês.

Obs.: Quando não encontrar o título em inglês no documento, pode usar o tradutor da Google para conseguir a informação.

Depósito: descreva este item (Ajuda)

Preencha as informações solicitadas sobre o depósito abaixo. Na maioria dos navegadores, você pode usar a tecla Tab para mover o cursor para a próxima caixa de entrada.

CAMPO OBRIGATÓRIO: Seleccione o tipo de documento.

Tipo de documento: *

CAMPO OBRIGATÓRIO: Informe a data de defesa do documento descrito, ou seja, a data de defesa da tese ou dissertação.

Data da defesa: *

CAMPO OBRIGATÓRIO: Informe o título do documento. Recomenda-se preencher com a tradução do título da tese ou dissertação para o inglês para maior visibilidade do documento.

Título: *

CAMPO OPCIONAL: Informe o título alternativo. Recomenda-se preencher com a tradução do título da tese ou dissertação para o inglês, para maior visibilidade do documento.

Título(s) alternativo(s):

- Inserir os dados do autor e do orientador. Copiar e colar o nome e o link (Lattes Id) do Currículo Lattes (plataforma do CNPq). **Coloque o nome completo, sem abreviação. É muito importante seguir as regras da citação bibliográfica (conforme documento em anexo).**
- Os CPFs estão disponíveis no Termo de Autorização.

Site do CNPq

Nome Felipe Dilelis de Resende Sousa

ID Lattes: 601316005724364

Última atualização do currículo em 21/06/2021

Passou graduação em Zootecnia pela Universidade Federal Rural do Rio de Janeiro(2010), mestrado em Zootecnia pela Universidade Federal Rural do Rio de Janeiro(2013) e doutorado em Zootecnia pela Universidade Federal Rural do Rio de Janeiro(2019). Atualmente é Assistente de Laboratório da Universidade Federal Rural do Rio de Janeiro. (Texto gerado automaticamente pela aplicação CVLattes)

Identificação

Nome Felipe Dilelis de Resende Sousa

Nome em citações bibliográficas SOUSA, F. D. R.; SOUSA, FELIPE DILELIS DE RESENDE; DILELIS, F.; SOUSA, FDR; DILELIS, FELIPE; DILELIS, F. DE RESENDE SOUSA, FELIPE DILELIS

Lattes ID <http://lattes.cnpq.br/601316005724364>

Endereço

AUTOR/ORIENTADOR

Descrição do **AUTOR** Exemplos

CAMPO OBRIGATÓRIO: Informe o nome do autor, conforme o formato de citação.

Autor:

CAMPO OPCIONAL: Informe a URL do currículo lattes do autor.

URL do Currículo Lattes do autor:

CAMPO OPCIONAL: Informe o identificador único do autor (CPF, ID de autor privativa da instituição ou passaporte).

Identificador único do autor:

Descrição do **PRIMEIRO ORIENTADOR** Exemplos

CAMPO OBRIGATÓRIO: Informe o nome do primeiro orientador, conforme o formato de citação.

Primeiro orientador:

CAMPO OPCIONAL: Informe a URL do currículo lattes do primeiro orientador.

- Na próxima página, só preenchemos se tiver co-orientador. Devemos incluir as informações disponíveis (CPF, URL do Lattes).

Licença Licença Descrição Descrição Descrição Descrição Descrição Descrição Descrição Descrição Upload Verificar Completo

Depósito: descreva este item (Ajuda)

Se houver, descrição do **PRIMEIRO CO-ORIENTADOR** Exemplos

CAMPO OPCIONAL: Informe o nome do primeiro coorientador, conforme o formato de citação.

Primeiro coorientador:

CAMPO OPCIONAL: Informe a URL do currículo lattes do primeiro coorientador.

URL do Currículo Lattes do primeiro coorientador:

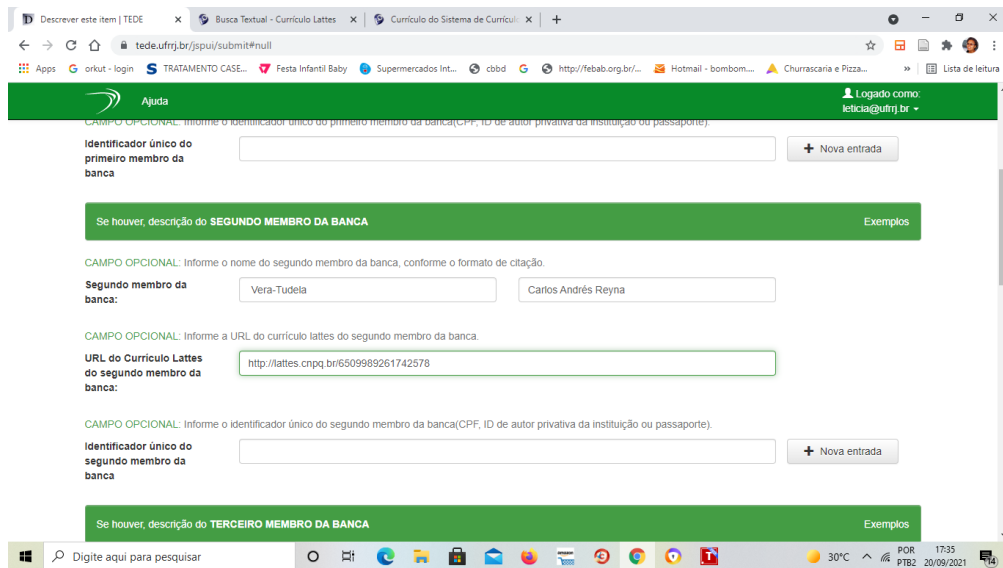
CAMPO OPCIONAL: Informe o identificador único do primeiro coorientador(CPF, ID de autor privativa da instituição ou passaporte).

Identificador único do primeiro coorientador:

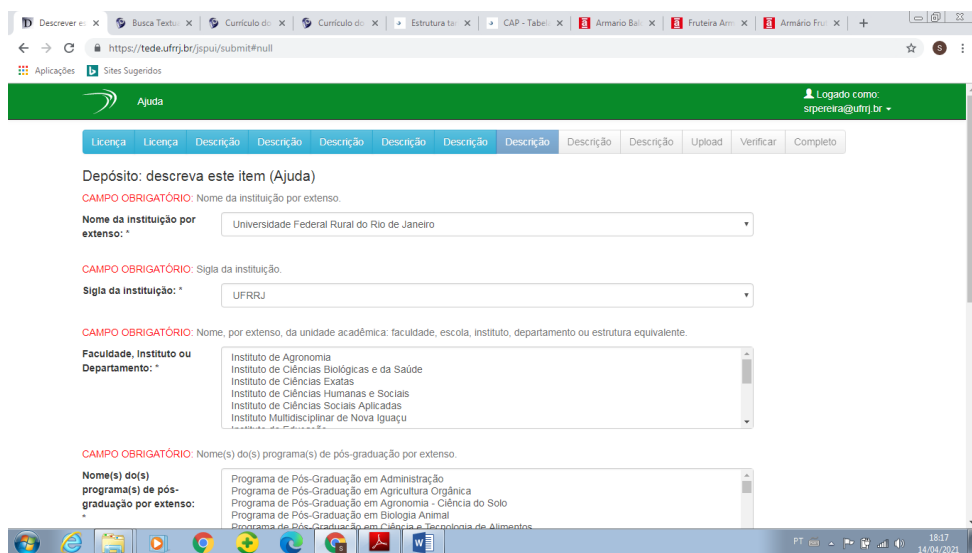
Se houver, descrição do **SEGUNDO CO-ORIENTADOR** Exemplos

CAMPO OPCIONAL: Informe o nome do segundo coorientador, conforme o formato de citação.

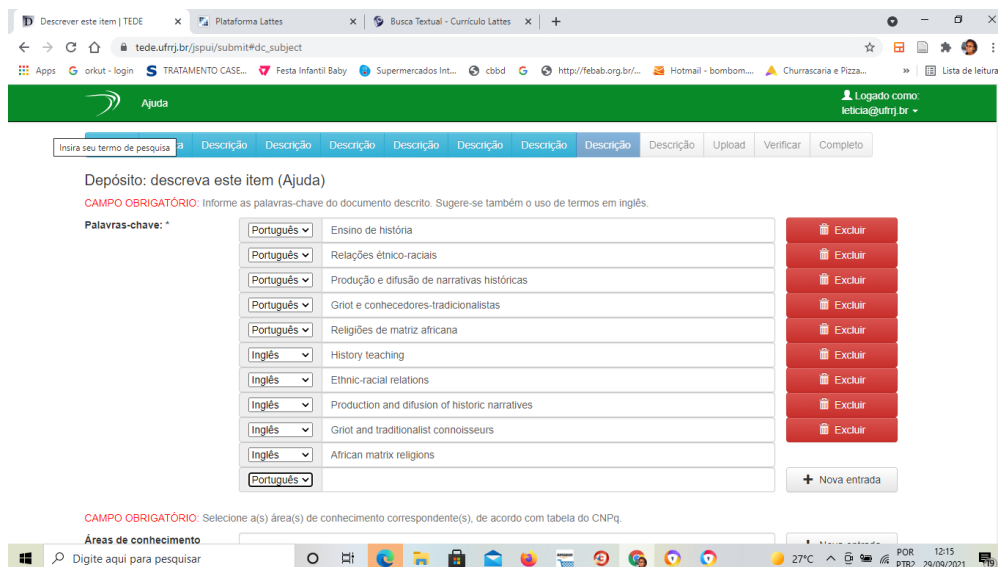
- Na próxima página, os membros da banca (colocar todos, conforme a ordem que aparece no PDF do documento).



- Nas próximas páginas, consultar as informações do PDF, para preenchimento dos campos. O nome da instituição por extenso, a sigla, o país e o idioma já estão cadastrados.
- Selecionar o nome do instituto e o seu programa;
- Copiar as referências bibliográficas do pdf (todas as páginas) e colar no campo;
- Não precisa preencher o identificador persistente.



- As palavras-chaves estão após o resumo, no PDF do documento. Após inserir a primeira, clicar em nova entrada. Colocar todas as palavras chaves da tese/dissertação (português e depois em inglês, do abstract).



- Inserir área de conhecimento (de acordo com a tabela CNPQ, em anexo);
- Copiar e colar resumo e logo abaixo o abstract (não esquecer de seleccionar a língua);
- Inserir agência de fomento (órgão que financiou a pesquisa):

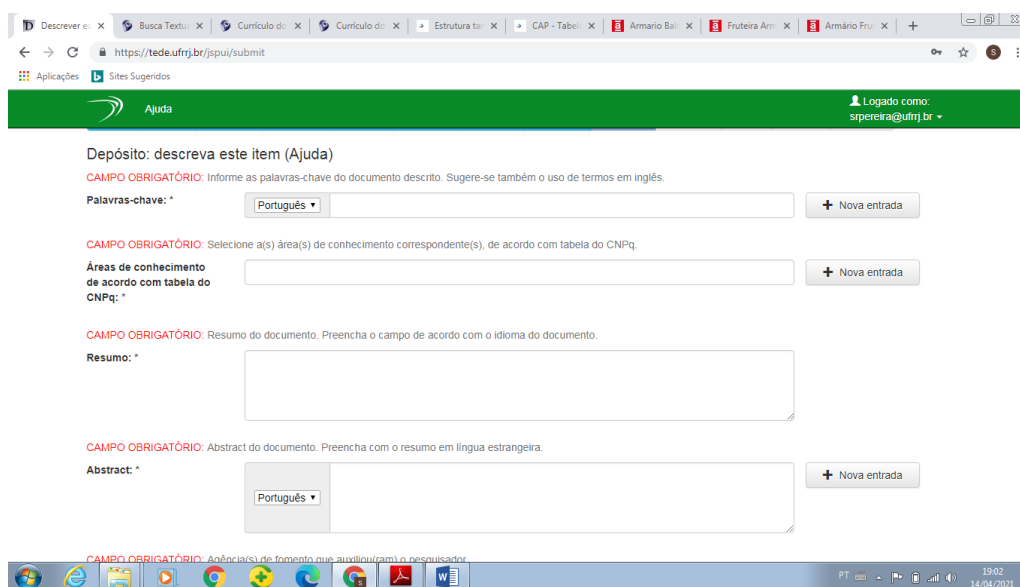
Exemplos:

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

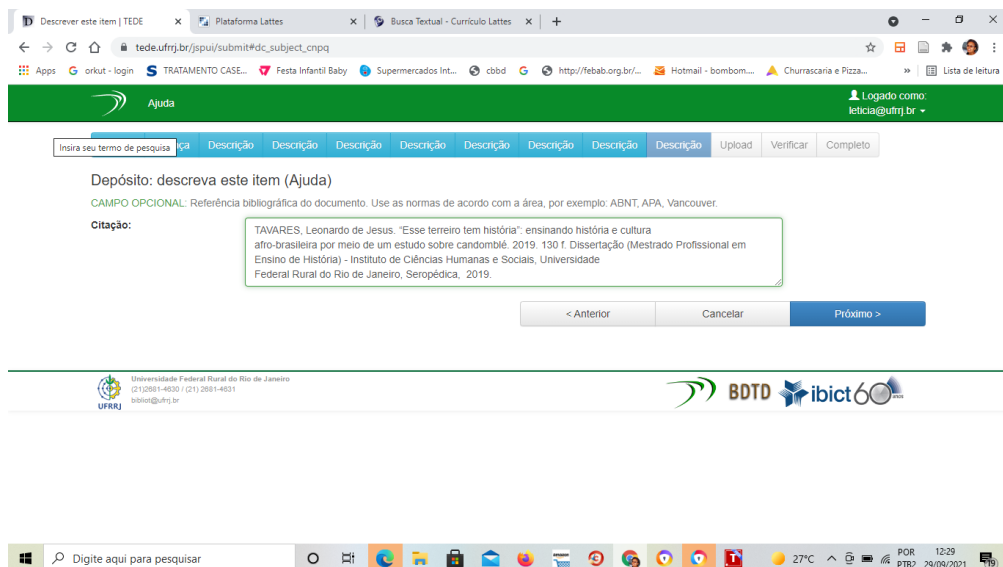
CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

FAPERJ - Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

- Não preenche a Relação

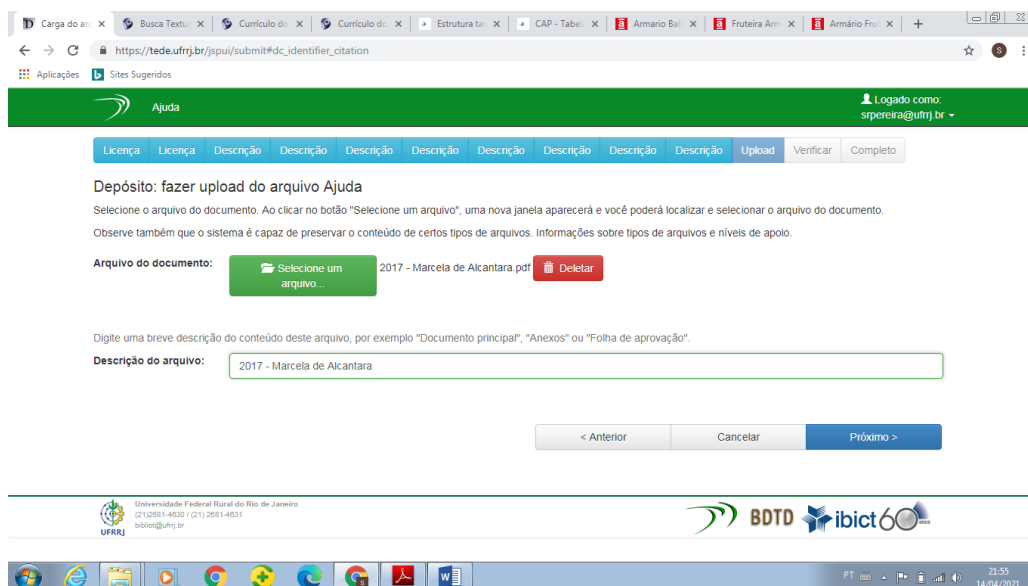


- Citação: copiar a referência bibliográfica do resumo e colar a informação no campo.

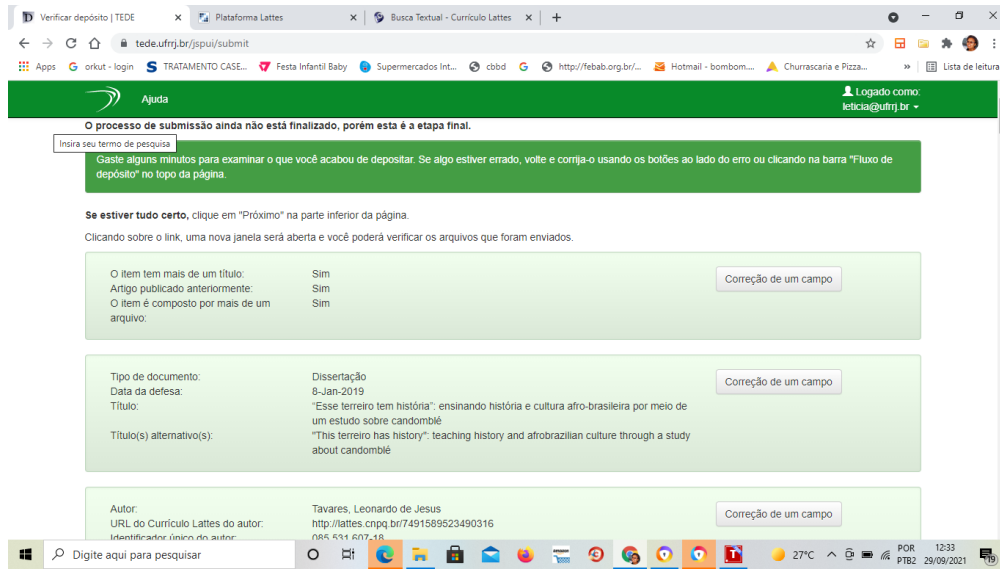


Realizar os acertos necessários da citação, conforme o modelo que segue:
 FARIA, Ada Guagliardi. Desenvolvimento do INOVA - um modelo para a adoção do Escritório de Processos como mecanismo de inovação organizacional. 2019. 126 f. Dissertação (Mestrado Profissional em Gestão e Estratégia) - Instituto de Ciências Sociais Aplicadas, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2019.

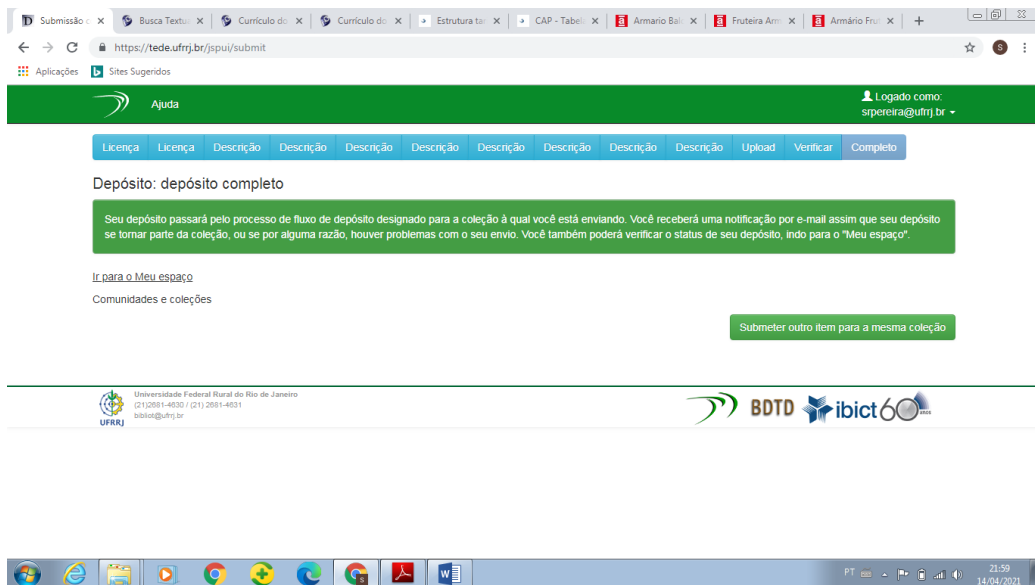
- Clicar em selecionar arquivo e procurar na pasta do seu computador o PDF da tese/dissertação.
- Quando o documento estiver carregado, copiar conforme abaixo na Descrição do arquivo



- As próximas telas são só para checagem de erros. Conferir as informações (corrigir os campos, se for necessário) e seguir para próxima página.



○ Depósito finalizado:



4. Programas de Pós-Graduação da UFRRJ e as sugestões de áreas do CNPq:

| Programa | Áreas do CNPq (INDICAÇÕES) |
|---|--|
| <i>Programa de Pós-Graduação em Administração</i> | Administração |
| <i>Programa de Pós-Graduação em Agricultura Orgânica</i> | Agronomia Engenharia Agrícola |
| <i>Programa de Pós-Graduação em Agronomia - Ciência do Solo</i> | Agronomia |
| <i>Programa de Pós-Graduação em Biologia Animal</i> | Biologia Geral Ecologia Genética |
| <i>Programa de Pós-Graduação em Ciência e Tecnologia de Alimentos</i> | Ciência e Tecnologia de Alimentos |
| <i>Programa de Pós-Graduação em Ciências Ambientais e Florestais</i> | Recursos Florestais e Engenharia Florestal Ecologia |
| <i>Programa de Pós-Graduação em Ciências Fisiológicas</i> | Fisiologia Medicina |

| | |
|--|---|
| <i>Programa de Pós-Graduação em Ciências Sociais</i> | <p>Ciências Sociais</p> <p>Sociologia</p> <p>Ciência Política</p> |
| <i>Programa de Pós-Graduação em Ciências Sociais em Desenvolvimento, Agricultura e Sociedade</i> | <p>Agronomia</p> <p>Engenharia Agrícola</p> <p>Ciências Sociais</p> <p>Planejamento Urbano e Regional</p> |
| <i>Programa de Pós-Graduação em Ciências Veterinárias</i> | <p>Medicina Veterinária</p> <p>Parasitologia</p> <p>Genética</p> <p>Microbiologia</p> <p>Zoologia</p> |
| <i>Programa de Pós-Graduação em Ciência, Tecnologia e Inovação em Agropecuária</i> | <p>Engenharia Agrícola</p> |
| <i>Programa de Pós-Graduação em Desenvolvimento Territorial e Políticas Públicas</i> | <p>Planejamento Urbano e Regional</p> <p>Economia</p> <p>Ciência Política</p> |
| <i>Programa de Pós-Graduação em Educação Agrícola</i> | <p>Agronomia</p> <p>Engenharia Agrícola</p> <p>Educação</p> |
| <i>Programa de Pós-Graduação em Educação, Contextos Contemporâneos e Demandas Populares</i> | <p>Educação</p> <p>Ciências Sociais</p> <p>Sociologia</p> |

| | |
|--|--|
| | <p>Ciência Política</p> <p>Planejamento Urbano e Regional</p> |
| <i>Programa de Pós-Graduação em Educação em Ciências e Matemática</i> | <p>Matemática</p> <p>Educação</p> |
| <i>Programa de Pós-Graduação em Engenharia Agrícola e Ambiental</i> | <p>Engenharia Agrícola</p> <p>Engenharia Sanitária</p> <p>Recursos Florestais e Engenharia Florestal</p> <p>Ecologia</p> |
| <i>Programa de Pós-Graduação em Engenharia Química</i> | <p>Química</p> <p>Engenharia Química</p> |
| <i>Programa de Pós-Graduação em Filosofia</i> | <p>Filosofia</p> |
| <i>Programa de Pós-Graduação em Fitossanidade e Biotecnologia Aplicada</i> | <p>Engenharia agrícola</p> <p>Biologia Geral</p> <p>Microbiologia</p> <p>Recursos Florestais e Engenharia Florestal</p> |
| <i>Programa de Pós-Graduação em Fitotecnia</i> | <p>Agronomia</p> <p>Engenharia agrícola</p> <p>Ecologia</p> |
| <i>Programa de Pós-Graduação em Geografia</i> | <p>Geografia</p> <p>Educação</p> |

| | |
|--|--|
| | Planejamento Urbano e Regional Ciência Política |
| <i>Programa de Pós-Graduação em Gestão e Estratégia</i> | Administração |
| <i>Programa de Pós-Graduação em Ensino de História</i> | História Educação |
| <i>Programa de Pós-Graduação em História</i> | História Educação Ciências Sociais Ciência Política |
| <i>Programa de Pós-Graduação em Letras</i> | Letras Linguística Educação |
| <i>Programa de Pós-Graduação em Matemática em Rede Nacional</i> | Matemática Educação |
| <i>Programa de Pós-Graduação em Medicina Veterinária (Patologia e Ciências Clínicas)</i> | Medicina Veterinária Parasitologia Farmacologia |
| <i>Programa de Pós-Graduação em Microbiologia Veterinária</i> | Medicina Veterinária Microbiologia |

| | |
|---|---|
| <i>Programa de Pós-Graduação em Modelagem Matemática e Computacional</i> | Matemática Ciências da Computação Probabilidade e Estatística |
| <i>Programa de Pós-Graduação em Patrimônio, Cultura e Sociedade</i> | Arquitetura e Urbanismo Sociologia História Antropologia |
| <i>Programa de Pós-Graduação em Práticas em Desenvolvimento Sustentável</i> | Recursos Florestais e Engenharia Florestal Economia |
| <i>Programa de Pós-Graduação em Psicologia</i> | Psicologia Ciências Sociais |
| <i>Programa de Pós-Graduação em Química</i> | Química Engenharia Química |
| <i>Programa de Pós-Graduação em Química em Rede Nacional</i> | Química Engenharia Química Educação |
| <i>Programa de Pós-Graduação em Zootecnia</i> | Zootecnia Medicina Veterinária |
| <i>Programa Multicêntrico de Pós-Graduação em Ciências Fisiológicas</i> | Fisiologia |

○ **Como cadastrar os nomes no repositório:**

- **Registra-se o nome a partir do último elemento do sobrenome.**

○ Exemplo: Ovídio Saraiva de Carvalho Silva

Registra-se: Silva, Ovídio Saraiva de Carvalho

- **Se o nome inclui palavras que indicam relação de parentesco, tais como Filho, Júnior, Neto, ou Sobrinho estas deverão ser consideradas como parte do sobrenome.**

Exemplo: Antonio Ribeiro de Castro Sobrinho

Registra-se: Castro Sobrinho, Antonio Ribeiro de

- **Nomes no idioma Espanhol:**

- **Autores com dois sobrenomes, faz-se entrada pelo primeiro destes.**

Exemplo: Eduardo Gonzáles Rivera

Registra-se: Gonzáles Rivera, Eduardo

- **Se o sobrenome inicia com um artigo, entra-se por este.**

Exemplo: Manuel Antonio Las Heras

Registra-se: Las Heras, Manuel Antonio

- **Alguns sobrenomes espanhóis são precedidos da partícula "de" (mulheres casadas). Neste caso, entra-se pelo sobrenome de solteira, seguido do sobrenome de casada.**

Exemplo: Antonia Murillo de Nogueira

Registra-se: Murillo de Nogueira, Antonia

- **Sobrenomes unidos pela letra "y": entra-se como se fossem compostos.**

Exemplos:

Emílio Cotarelo y Mori

Registra-se: Cotarelo y Mori, Emílio

Antonio Gonzáles y Gonzáles

Registra-se: Gonzáles y Gonzáles, Antonio

Fonte: [https://wiki.bireme.org/pt/index.php/Ap%C3%AAndice A -
Regras para entrada de autores](https://wiki.bireme.org/pt/index.php/Ap%C3%AAndice_A_-_Regras_para_entrada_de_autores)

ANEXO B – MTD3-BR (Metadados para Teses e Dissertações Digitais Brasileiras)



PADRÃO BRASILEIRO DE METADADOS DA BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES (MTD3-BR)

| | REPETITIVO | METADADO | VALOR |
|-----|------------|----------------------------------|-----------------------------|
| OBR | Não | dc.creator | Autor |
| OPC | Sim | dc.creator.ID | ID do autor |
| OPC | Não | dc.creator.Lattes | Currículo Lattes do autor |
| OBR | Sim | dc.contributor.advisor1 | Orientador1 |
| OPC | Sim | dc.contributor.advisor1ID | ID do orientador1 |
| OPC | Não | dc.contributor.advisor1Lattes | ID Lattes do orientador1 |
| OPC | Sim | dc.contributor.advisor2 | Orientador2 |
| OPC | Sim | dc.contributor.advisor2ID | ID do orientador2 |
| OPC | Não | dc.contributor.advisor2Lattes | ID Lattes do orientador2 |
| OPC | Não | dc.contributor.advisor-co1 | Co-orientador1 |
| OPC | Sim | dc.contributor.advisor-co1ID | ID do co-orientador1 |
| OPC | Não | dc.contributor.advisor-co1Lattes | ID Lattes do co-orientador1 |
| OPC | Não | dc.contributor.advisor-co2 | Co-orientador2 |
| OPC | Sim | dc.contributor.advisor-co2ID | ID do co-orientador2 |
| OPC | Não | dc.contributor.advisor-co2Lattes | ID Lattes do co-orientador2 |
| OPC | Sim | dc.contributor.referee1 | Membro da banca1 |
| OPC | Sim | dc.contributor.referee1ID | ID do membro da banca1 |
| OPC | Não | dc.contributor.referee1Lattes | Lattes do membro da banca1 |
| OPC | Sim | dc.contributor.referee2 | Membro da banca2 |
| OPC | Sim | dc.contributor.referee2ID | ID do membro da banca2 |
| OPC | Não | dc.contributor.referee2Lattes | Lattes do membro da banca2 |
| OPC | Sim | dc.contributor.referee3 | Membro da banca3 |
| OPC | Sim | dc.contributor.referee3ID | ID do membro da banca3 |
| OPC | Não | dc.contributor.referee3Lattes | Lattes do membro da banca3 |
| OPC | Sim | dc.contributor.referee4 | Membro da banca4 |
| OPC | Sim | dc.contributor.referee4ID | ID do membro da banca4 |
| OPC | Não | dc.contributor.referee4Lattes | Lattes do membro da banca4 |

| | REPETITIVO | METADADO | VALOR |
|------------|------------|-------------------------------|--|
| OPC | Sim | dc.contributor.referee5 | Membro da banca5 |
| OPC | Sim | dc.contributor.referee5ID | ID do membro da banca5 |
| OPC | Não | dc.contributor.referee5Lattes | Lattes do membro da banca5 |
| OPC | Não | dc.identifier.doi | Identificador permanente - DOI |
| OBR | Sim | dc.description.resumo | Resumo |
| OBR | Sim | dc.description.abstract | Abstract |
| OBR | Sim | dc.publisher | Nome da instituição por extenso |
| OBR | Não | dc.publisher.country | País da instituição |
| OBR | Sim | dc.publisher.department | Nome da faculdade, instituto ou departamento por extenso |
| OBR | Sim | dc.publisher.program | Nome do Programa de Pós-Graduação por extenso |
| OBR | Sim | dc.publisher.initials | Sigla da instituição |
| OPC | Não | dc.identifier.uri | Identificador persistente |
| OBR | Sim | dc.subject | Palavras-chave que descrevem o assunto do documento |
| OBR | Sim | dc.subject.cnpq | Área(s) do conhecimento do documento (Tabela CNPq) |
| OPC | Sim | dc.thumbnail.url | URL Arquivo (imagem) |
| OBR | Sim | dc.description.sponsorship | Agência de fomento |
| OBR | Não | dc.type | Tipo do documento: tese ou dissertação |
| OBR | Não | dc.title | O título e subtítulo da tese ou dissertação na língua do documento |
| OPC | Sim | dc.title.alternative | Título alternativo do documento e/ou traduzido em outro idioma |
| OBR | Não | dc.rights | Direitos de acesso ao documento |
| OBR | Não | dc.date.issued | Data de defesa |
| OPC | Não | dc.date.available | Data de embargo |
| OBR | Não | dc.language | Idioma |
| OPC | Não | dc.relation.references | Referências citadas na tese ou dissertação |
| OPC | Não | dc.identifier.citation | Citação do documento |
| OPC | Sim | dc.format | Formato do documento |
| OPC | Sim | dc.relation | Relação, quando for o caso, com outros materiais produzidos para a elaboração da tese/dissertação ou comentários do próprio autor sobre a tese/dissertação |

OBR – Obrigatório | OPC - Opcional

Fonte: https://bdt.d.ibict.br/vufind/Padrao_de_Metadados_da_BDTD_MTD3-BR_v_2017.pdf