

Izabel Lima dos Santos
Kalline Yasmin Soares Feitosa
Thalita Natasha Ferreira Damasceno
Organizadoras

Bibliotecas Universitárias estudos e experiências



Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

B477 Bibliotecas universitárias: estudos e experiências.

Bibliotecas universitárias [livro eletrônico] : estudos e experiências. / Izabel Lima dos Santos, Kalline Yasmin Soares Feitosa, Thalita Natasha Ferreira Damasceno (Organização); Jackson Sousa Serra (Diagramação). – Fortaleza, CE, 2022.
375 f. : il. color.

ISBN: 978-65-00-56657-4.

Coletânea elaborada no âmbito do Sistema de Bibliotecas da Universidade Federal do Ceará (UFC).

1. Bibliotecas Universitárias – Brasil. 2. Bibliotecas Universitárias – Administração. 3. Bibliotecas Universitárias – Serviços. 4. Bibliotecas Universitárias – Produtos. I. Santos, Izabel Lima dos. II. Feitosa, Kalline Yasmin Soares. III. Damasceno, Thalita Natasha Ferreira. IV. Serra, Jackson Sousa.

CDD 027.781

Ficha catalográfica elaborada por Izabel Lima dos Santos (CRB 3- 1339)

4 Indexação no Repositório Institucional da UFC

diagnóstico e soluções

María Marlene Rocha de Sousa¹⁰

Margareth de Figueiredo Nogueira Mesquita¹¹

Neiliane Alves Bezerra¹²

¹⁰ Bibliotecária na Universidade Federal do Ceará. Especialista em Tecnologias da Informação e Comunicação para o Gerenciamento da Informação pela Universidade Federal do Ceará (UFC).

Currículo Lattes: <http://lattes.cnpq.br/9000499717041322>

¹¹ Bibliotecária na Universidade Federal do Ceará. Diretora da Biblioteca de Pós-Graduação em Economia Agrícola. Mestra em Políticas Públicas e Gestão da Educação Superior pela Universidade Federal do Ceará.

Currículo Lattes: <http://lattes.cnpq.br/3941085605089505>

¹² Bibliotecária na Universidade Federal do Ceará. Diretora da Biblioteca do Curso de Arquitetura. Mestra em Políticas Públicas e Gestão da Educação Superior pela Universidade Federal do Ceará.

Currículo Lattes: <http://lattes.cnpq.br/5184095171001259>

1 Introdução

Os Repositórios Institucionais (RI) surgiram como uma consequência natural da explosão informacional gerada, com as funções de reunião, organização, tratamento, representação, armazenamento, recuperação, preservação e disseminação de coleções digitais referentes à produção intelectual dos membros de uma instituição, no sentido de agregar valor aos conteúdos intelectuais produzidos e ao incremento da visibilidade da produção científica, extrapolando, assim, o ambiente presencial das instituições. O RI foi desenvolvido, como ferramenta, para reunir toda a produção científica da comunidade acadêmica, quer sejam docentes, discentes e técnico-administrativos que integram a instituição, reunindo num único local a literatura produzida e os metadados a ela associados.

Nesse cenário, as Instituições de Ensino Superior (IES) começaram a utilizar os Repositórios como uma ferramenta que aceleraria as mudanças na produção científica e tecnológica, na divulgação de novos conteúdos e resultados de pesquisas realizadas no âmbito das universidades. Como consequência imediata, aumentou também a visibilidade da produção das unidades acadêmicas, contribuindo como um indicativo da

qualidade institucional. (COSTA; SOUSA, 2018). Além de constituírem uma forma de promover o acesso livre, eles contribuem também para aumentar o impacto da investigação nas diversas áreas do conhecimento, incrementando a sua visibilidade e acesso, facilitando a gestão da informação sobre a sua produção científica e preservando a memória intelectual em suporte digital (LEITE, 2009). Para Nunes, Marcondes e Weitzel (2012, p. 3), “Os Repositórios Institucionais reforçam a ideia de que todos os materiais de pesquisa devem estar disponibilizados publicamente na internet, sem restrições de acesso [...]”.

O acesso livre (*Open Access*) a resultados de pesquisas impacta diretamente na produção de novas pesquisas, passando a gerar novos conhecimentos, como também aponta tendências de áreas ou assuntos que estão surgindo, além de fornecer informações sobre o que é relevante ou não, direcionando os pesquisadores ao conteúdo das produções científicas, sem esquecer as outras formas de comunicação, como os relatórios de experiências e de pesquisas, os mais diversos recursos de aprendizagem, a literatura cinzenta, entre outras que podem fazer parte dos RI.

Para facilitar o acesso aos documentos, o RI é estruturado no *software DSpace*, que, segundo Bilhalva Cardoso (2019, p. 132), “faz parte de uma tecnologia aprimorada que se sujeita a variados e

existência se complementa com os repositórios que operam com o acesso aberto à Informação”.

O *DSpace* foi adquirido pela Universidade Federal do Ceará (UFC) por meio de um edital em parceria com o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Universidade de Brasília. A Biblioteca Universitária (BU) foi contemplada com um kit tecnológico em comodato, composto de um servidor pré-formatado e configurado com o sistema operacional baseado na plataforma *Unix/Linux*; os *softwares* básicos *Apache*, *MySQL* e *PHP* e ainda o *software DSpace*, para implantação do RI (BOLETIM, 2010 citado por MESQUITA, 2016, p. 47).

O RI da Universidade Federal do Ceará (RI-UFC) é gerenciado pela BU, ordenado em comunidades e subcomunidades, assim como em outras IES, e segue o modelo de departamentalização da Universidade. Como comunidades representadas no RI-UFC, temos: Biblioteca Universitária; Centros; Campi; COVID-19; Documentos em formato acessível; Encontros Universitários; Faculdades; Institutos; Museu de Arte da UFC; Memorial; PróReitorias; Secretarias; Trabalhos de Conclusão de Cursos de Graduação e de Especialização e seus diversos tipos de documentos: artigos de periódicos e eventos, dissertações, teses, monografias, resumos, capítulos de livro, livros, catálogos, relatórios, guias, vídeos, materiais acessíveis, documentos

administrativos e folhetos. São indexados os trabalhos de docentes, discentes e técnicos vinculados à instituição (REPOSITÓRIO, 2020).

Os RI possuem um papel destacado por aumentar o acesso e a visibilidade da produção científica das instituições e propiciar a possibilidade de criação de projetos colaborativos entre universidades distintas. Isso demonstra claramente a importância dos RI em auxiliar as universidades no processo de internacionalização e no alcance de melhores posições no *Ranking Web of Universities*, que mede a presença das instituições acadêmicas no ambiente *web*.

Na edição de setembro de 2020 do *Ranking Web of Repositories*, o RI-UFC está nas seguintes colocações:

- 66º lugar entre todos os tipos de repositórios do mundo;
- 56º lugar entre os repositórios institucionais do mundo;
- 7º lugar entre os repositórios institucionais das Instituições de Ensino Superior do Brasil; e em
- 1º lugar entre os repositórios institucionais da Região Nordeste.

Para que os usuários desses repositórios possam ter acesso ao texto completo, fazer *download* de arquivos e extrair destes as informações relevantes para seus estudos ou pesquisas, torna-se necessário

melhorar a gestão da indexação de conteúdo e da integração dos dados inseridos.

A escolha pelo tema gestão da indexação no RI-UFC partiu da atuação profissional como Bibliotecárias e pela percepção, contundente, de que a recuperação do conteúdo de toda a produção de uma instituição é uma tarefa grandiosa. Aliado ao fato de que o conteúdo recuperado pelos RI gera novos conhecimentos de forma colaborativa e novas pesquisas, contribuindo para alavancar o conhecimento humano e revelar o impacto na ampliação da visibilidade das produções de pesquisadores.

Diante do exposto e da visibilidade e colocação no *Ranking Web of Repositories*, que o RI-UFC alcançou, temos a seguinte inquietação e questionamento: Que fatores interferem na qualidade da indexação e, conseqüentemente, na recuperação de documentos relevantes para os usuários nas subcomunidades dos Departamentos de Economia Agrícola do Centro de Ciências Agrárias (DEA/CCA) e de Arquitetura e Urbanismo do Centro de Tecnologia (DAU/CT)?

Logo, o objetivo geral deste trabalho é elevar a qualidade da indexação de autoridades autor e assunto nas subcomunidades dos Departamentos de Economia Agrícola (DEA) e de Arquitetura e Urbanismo (DAU).

Como objetivos específicos, pretende-se:

- a) Verificar a recuperação documental por autores e por assuntos de forma comparada entre linguagem natural e linguagem controlada observando o que diz a literatura sobre o assunto;
- b) Identificar a ocorrência da variação de um mesmo termo de autoridades autor e assunto indexados; e
- c) Corrigir e completar os dados das autoridades de nomes pessoais e de assuntos na base *Pergamum*.

O tema em questão relaciona-se à participação da BU na disseminação da informação, atendendo a comunidade interna e externa da Universidade, conforme descrito em sua missão:

Organizar, preservar e disseminar a informação para a produção do conhecimento, dando suporte às atividades educacionais, científicas, tecnológicas e culturais da Universidade Federal do Ceará, possibilitando o crescimento e o desenvolvimento da Instituição e da sociedade (UFC, 2021).

A relevância do trabalho decorre da necessidade de detectar as inconsistências da indexação de autor pessoal e assunto no RI-UFC (situação verificada em várias bibliotecas do sistema), que poderá resultar em um número incorreto no quantitativo de produção de um autor. A duplicidade

de entrada da mesma autoria impacta diretamente no ranking do autor, que pode ter sua produção diminuída no RI-UFC, dado que o mesmo não utiliza uma ferramenta de indexação de remissivas.

A duplicidade pode ocorrer com o uso do nome dos autores com maiúsculo/minúsculo, plural/singular, agnomes, hífen, espaços em branco, entre outros. Além disso, a inclusão de palavras-chave utilizadas livremente pelos autores pode resultar em indexações diferentes para um mesmo assunto, impactando na recuperação dos documentos.

2 A Indexação da autoridade de nomes pessoais e sua relação com os RI

A padronização do nome de autores requer pesquisa e cuidado na etapa de descrição, uma vez que os próprios autores podem se identificar de maneira incompleta ou diferente em algumas ocasiões, assim como podem haver acréscimos ou retirada de nomes do meio, alterações no sobrenome de autoras após mudança de estado civil, sobrenomes latinos e ambiguidades (especialmente nos sobrenomes orientais) e finalmente os homônimos. (FERREIRA, 2018). Conseqüentemente, no trabalho cotidiano das bibliotecas e indexação nas bases de dados, essa inconsistência de autoridades gera recuperação da informação ineficiente ou equivocada.

Visto que os RI permitem o aumento da visibilidade da produção científica e tecnológica de uma IES, que essa produção gera novos conhecimentos e que os estudos mostram tendências de novas áreas de interesse científico a forma como um autor é indexado determina o sucesso, ou não, do resultado das buscas. Portanto, é importantíssimo definir fontes de informações confiáveis, antes da indexação, visando diferenciar os homônimos, formas não autorizadas do nome e a identificação segura da autoria de uma publicação.

Neste contexto, em um estudo sobre o RI da Universidade Federal do Rio Grande, Corrêa *et al.* (2012, p. 33) haviam discutido sobre a variação dos nomes e a possibilidade de padronizá-los no formato de catálogo, segundo elas, "Detectou-se que um nome de autor poderia ter diversas entradas".

Além do catálogo da Rede Pergamum, a Biblioteca Universitária utiliza, entre outras, as seguintes fontes de Informação sobre os currículos e nomes de autores acadêmicos:

- a) Plataforma Lattes: iniciativa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), gratuita, permite os autores registrarem, atualizarem seus currículos abrindo espaço também para a consulta aos mesmos, sendo considerado o padrão brasileiro de avaliação, pois apresenta o histórico das atividades

científicas, acadêmicas e profissionais dos pesquisadores.

- b) Open Researcher and Contributor ID (ORCID): plataforma de acesso aberto à busca, reúne uma variedade de nomes confusos e os identifica de maneira que predomine cada nome ligado à sua verdadeira identidade. O ORCID possibilita a desambiguação de nomes que, porventura, tenham sido registrados de diferentes formas, duplicados ou com acento, sem acento, etc. (SOARES, 2016).
- c) Catálogo de autoridade da Biblioteca Nacional (BN): inteiramente disponível à consulta livre através de mecanismos de busca, que facilitam as pesquisas realizadas por qualquer pessoa. “Utiliza a padronização estabelecida pelo AACR2 e também pelos exemplos extraídos do catálogo da Library of Congress (LC), o que a torna modelo em nomes de autoria pessoal padronizada”. (BILHALVA CARDOSO, 2019, p. 135).
- d) Arquivo da Autoridade Internacional Virtual (VIAF): plataforma de colaboração para identificação de nomes de autores em esfera mundial.

Nesta perspectiva, percebeu-se a necessidade de considerar estes e outros meios que facilitem a

identificação dos nomes dos autores, visando a uma padronização.

3 A indexação de assuntos e a sua conexão com os RI

No âmbito da organização da informação a grande preocupação é tornar as informações dos documentos acessíveis para quem delas necessitar. Para isso, o seu tratamento e a sua representação constituem condição sine qua non para a sua recuperação ou acesso.

Destacamos a diferença entre acesso ao documento e acesso ao seu conteúdo. Sob esse entendimento, Santos (2017, p. 5) verifica

[...] que o documento deve ser tratado em seu todo levando em consideração tanto as operações que contemplam a descrição física (tratamento descritivo ou representação descritiva), como a descrição do conteúdo temático (tratamento temático ou representação temática).

Para Fujita (2003, p. 61), a indexação “é uma combinação metodológica altamente estratégica entre o tratamento do conteúdo de documentos e sua recuperação por um usuário”, relacionando, de forma direta, o processo e a finalidade da uma indexação. Logo, a indexação descreve um conteúdo por meio de descritores. Nesse sentido, a indexação é a ligação

entre o documento e o usuário final e é concretizada quando o indexador analisa o conteúdo documental e realiza a sua tradução. Essa tradução consiste em representar os conteúdos dos documentos em termos oriundos dos instrumentos de indexação (linguagens de indexação), e, após a tradução, os termos são designados como descritores ou cabeçalhos de assunto.

Santos (2017, p. 6) afirma que

[...] em um Repositório Digital, os termos são pontos de acesso para os documentos, haja vista que é possível recuperar esses objetos informacionais sob qualquer uma das suas palavras-chave (que foram empregadas no sistema no ato da sua representação) ou até mesmo por meio da combinação delas. Em outras palavras, esses termos são atribuídos pelo indexador (pessoa responsável por realizar a atividade de indexação) e exercem a função de pontos de acesso mediante os quais um documento é identificado e recuperado.

No âmbito dos Repositórios, os termos presentes no título, resumo e palavras-chave dos documentos são permitidos como forma de indexação, visando à representação da informação e sua posterior recuperação. Na prática, esses termos podem despertar implicações negativas, pois são provenientes da linguagem natural do autor sem a ajuda de instrumentos de representação da informação. Logo, alguns termos podem fazer sentido para alguns, mas não para todas as pessoas que vão buscá-los.

Noriz (2007 citado por Santos, 2017, p. 9) menciona algumas implicações dessas práticas, entre outras:

- ausência de padronização em relação à flexão de número singular ou plural;
- emprego de palavras distintas, mas com mesmo significado; e
- descritores que não representam o objeto em sua totalidade.

Além dessas consequências, fica claro que problemas de revocação e precisão (medidas de desempenho) podem repercutir na busca pelo usuário final nos Repositórios. Com efeito, Assumpção (2012, p.17) considera que:

[...] o controle de autoridade pode ser definido como a consistência dos pontos de acesso (nomes) utilizados para identificar as entidades em um catálogo, tal controle tem como finalidade fazer com que o usuário do catálogo encontre os recursos informacionais associados à entidade por ele desejada independentemente do nome ou forma do nome da entidade que ele utiliza ao realizar sua busca.

Neste cenário, observou-se a importância da padronização dos nomes de autoridades para se manter a uniformidade do catálogo.

3.1 A linguagem natural

Para Lancaster (2004 citado por Santos, 2017, p. 7, grifo do autor), “na *indexação por extração* as palavras ou expressões empregadas pelo autor no documento analisado são selecionadas pelo indexador para representar o seu conteúdo”. Ou seja, linguagem natural utilizada pelo próprio autor do documento (palavras contidas no título, nas palavras-chave, resumo ou na própria estrutura textual do documento), onde os termos usados para descrever suas ideias são indexados para descrever seu conteúdo. A linguagem natural ou livre caracteriza-se por ser empregada na escrita e na fala de forma usual e comum, podem trazer ambiguidades, já que podem usar expressões regionais, siglas, abreviações, homógrafos, conter erros ortográficos, entre outros.

Quadro 1 – Linguagem natural: vantagens e desvantagens

VANTAGENS	DESVANTAGENS
<p>Permite o imediato registro da informação em um repositório digital, sem necessidade de consulta a uma linguagem de controle.</p>	<p>Os usuários da informação, no processo de busca, precisam fazer um esforço intelectual maior para identificar os sinônimos, as grafias alternativas, os homônimos etc.</p>

VANTAGENS	DESVANTAGENS
O processo de busca é facilitado com a ausência de treinamentos específicos no uso de uma linguagem de controle.	Haverá alta incidência de respostas negativas ou de relações incorretas entre os termos usados na busca (por ausência de padronização).
Os termos de entrada de dados são extraídos diretamente dos documentos que vão constituir o Repositório Digital.	Os custos de acesso tendem a aumentar com a entrada de termos de busca aleatórios
Os temas específicos citados nos documentos podem ser encontrados	Uma estratégia de busca que arrole todos os principais conceitos e seus sinônimos deve ser elaborada para ferramenta (ex.: nomes comerciais de substâncias químicas não ocorrem no <i>Chemical Abstracts</i>).
Elimina os conflitos de comunicação entre os indexadores e os usuários, pois ambos terão acesso aos mesmos termos.	Perda de confiança do usuário em uma possível resposta negativa

Fonte: Lopes (2002, p. 48).

3.2 A linguagem controlada

A indexação de atribuição ou controlada utiliza a atribuição de descritores ao documento, lançando mão de uma linguagem originária de uma fonte que não seja o próprio documento indexado. Sendo assim, são determinados os cabeçalhos a serem empregados objetivando-se controlar sinônimos, diferenciar homógrafos e ligar termos por meio de relações hierárquicas, associativas e de equivalência, utilizando-se uma linguagem artificial/linguagem controlada.

Para Lopes (2002, p. 47), o vocabulário controlado:

[...] pode ser definido como um conjunto de termos organizados de forma hierarquizada e/ou alfabética, com o objetivo de possibilitar a recuperação de informações temáticas, reduzindo substancialmente a diversidade de terminologia.

De acordo com Lancaster (2004, p. 19), "a indexação por atribuição envolve o esforço de representar a substância da análise conceitual mediante o emprego de termos extraídos de alguma forma de vocabulário controlado".

Quadro 2 – Linguagem controlada: vantagens e desvantagens

VANTAGENS	DESVANTAGENS
<p>Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuários.</p>	<p>Custos: a produção e manutenção da base de dados terá despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoal especializado na atualização do thesaurus.</p>
<p>Com o uso de um thesaurus e suas respectivas notas de escopo, os indexadores podem assinalar mais corretamente os conceitos dos documentos.</p>	<p>O vocabulário controlado poderá não refletir adequadamente os objetivos do produtor da base, caso esteja desatualizado.</p>
<p>Se bem constituído, o vocabulário controlado poderá oferecer alta recuperação e relevância e, também, ampliar a confiança do usuário diante de um possível resultado negativo.</p>	<p>Um vocabulário controlado poderá se distanciar dos conceitos adequados para a representação das necessidades de informação dos usuários.</p>

VANTAGENS	DESVANTAGENS
As relações hierárquicas e as remissivas do vocabulário controlado auxiliam tanto o indexador, quanto o usuário na identificação de conceitos relacionados.	Necessidade de treinamento no uso dos vocabulários controlados tanto para os intermediários, quanto para os usuários finais.
Redução no tempo de consulta à base, pois a estratégia de busca será mais bem elaborada com o uso do thesaurus.	Desatualização do vocabulário controlado poderá conduzir a falsos resultados.

Fonte: Lopes (2002, p. 47)

4 A indexação no RI-UFC

Atualmente, no RI-UFC, foram constatadas situações que podem impactar negativamente na recuperação de informações em virtude da inclusão, por parte dos indexadores (bibliotecários e outros), de palavras-chave utilizadas livremente pelos autores em suas publicações, ou pelo nome dos autores retirado diretamente dos documentos. O resultado dessas ações são indexações diferentes para o mesmo assunto, impactando de forma direta na recuperação dos documentos, ou ainda nomes que não identificam

uma autoria, uma abreviatura, por exemplo. A situação foi apresentada em novembro de 2018, em um encontro sobre o RI, realizado no auditório da Biblioteca Central do *Campus* do Pici.

Após a análise de todas as entradas do catálogo de autoridade pessoal e de assuntos, no intervalo de tempo de dois meses e utilizando a empiria do dia a dia de trabalho, constatou-se diversificação de registros de nomes de um mesmo autor com: nomes abreviados, com acentuação inexistente, com escrita equivocada, uso de letras maiúsculas/minúsculas, uso de espaço em branco indevido ou hífen, desconsideração de nomes compostos e de grafia de autores estrangeiros. Essas ocorrências geram dúvidas tanto em catalogadores, como no usuário que se deparar com essa situação, pois não saberão se os nomes representam a mesma pessoa, ou se existem pessoas diferentes com o mesmo sobrenome. Essa situação é observada com frequência em várias bibliotecas do sistema.

Quanto aos cabeçalhos de assunto, temos vários exemplos que demonstram a variedade de grafias para um mesmo termo e de diferentes termos para um mesmo conceito. Certamente, temos como respostas às buscas por assunto um quantitativo que não reúne a totalidade de publicações indexadas. Todas as ocorrências poderiam ser unificadas recorrendo ao catálogo de autoridades do sistema *Pergamum*, entre outros.

5 Metodologia

A pesquisa é um estudo de caso sobre a indexação de nomes pessoais e de assuntos, cujo objeto é a lista contida no catálogo do RI-UFC, mais especificamente, nas subcomunidades DEA e DAU. Gil (2010, p. 37) afirma que o estudo de caso “consiste no estudo profundo e exaustivo de um ou mais objetos, de maneira que permita seu amplo e detalhado conhecimento”.

Trata-se de uma pesquisa quanti-qualitativa e aplicada, pois adquire conhecimentos que se convertem para uma aplicação específica (GIL, 2010, p.27).

O universo da pesquisa compreende o Catálogo de autores e assuntos do RI-UFC nas subcomunidades do DEA e DAU e, para a organização, tratamento e análise de dados, valeuse de planilha eletrônica, com o objetivo de listar os nomes dos autores e assuntos, visando à sua correção ou unificação. Em uma coluna, foram listados os termos livres (nomes pessoais e assuntos), e em outra os respectivos termos controlados.

A pesquisa foi realizada no intervalo entre os meses de agosto de 2020 e janeiro de 2021, iniciando-se com a busca por autor, no catálogo do RI, de todos os professores do DEA e do DAU. Foram consultadas as diversas bases de dados com a

finalidade de verificar o controle das autoridades, como: Plataforma *Lattes*, *ORCID*, *VIAF* e *BN*. Em seguida, verificouse como é feita a citação do nome, observando-se o uso ou não de acentuação, homônimos, nomes latinos, entre outros. O passo seguinte foi realizar a consulta na base do *Pergamum* com a finalidade de corrigir ou completar os dados e atualizá-los.

Em um segundo momento, buscaram-se os assuntos duplicados, sinônimos ou que estavam no singular/plural, a fim de unificar os termos nas fontes e bases de dados especializadas. Novamente, consultou-se o catálogo *Pergamum*, com o objetivo de completar as informações não constantes na base.

6 Resultados

O estudo, que ainda está em andamento, foi dividido em dois momentos especificados a seguir:

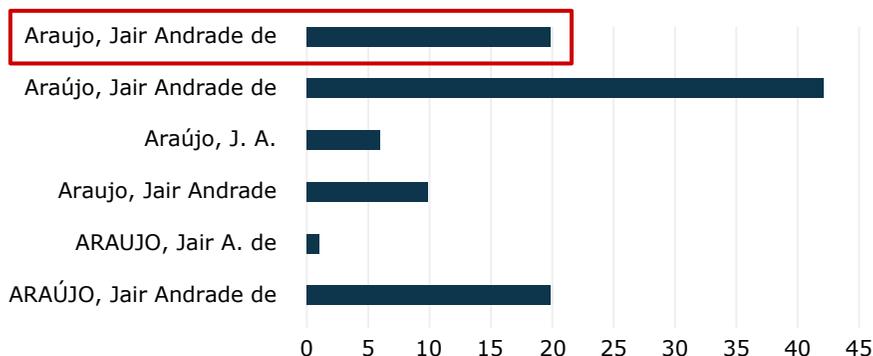
- a) **Busca de autoridade de nomes pessoais do DEA e do DAU**, onde procuramos identificar as diversas formas de citar a mesma autoria. Na busca facetada, encontramos 739 registros de autoria pessoal vinculados ao DEA e 154 vinculados ao DAU. (REPOSITÓRIO, 2021). Essa busca revelou a ausência de um padrão de indexação, o que gera dificuldades na identificação do autor e interfere no resultado da consulta.

Segundo Mugnaini *et al.* (2012, p. 264), o problema tem origem

[...] no próprio documento quando o autor se autodenomina de distintas maneiras em diferentes momentos (nome por extenso, abreviado, de casado, fantasia ou outras derivações). Essa falta de normalização gera um encadeamento de problemas de ambiguidade nas bases de produção científica, tendo em vista que o indexador normalmente vai entrar os nomes dos autores de acordo com o documento que tem em mãos.

Foram encontradas 245 autoridades do DEA e 18 do DAU, que estavam em duplicidade e precisaram ser unificadas. Elas se referiam ao uso de acentuação ou espaços em branco; a inserção de hífen entre o último sobrenome e o designativo de parentesco; a omissão de preposições; à abreviatura de prenomes; à grafia equivocada e à grafia em maiúscula do último sobrenome. Desse total, os 16 professores do DEA (UNIVERSIDADE, 2020) tiveram as ocorrências de citações coletadas, para seus respectivos nomes, bem como a quantidade de citações encontradas para cada autor. Logo abaixo, segue um exemplo (Gráfico 1):

Gráfico 1 - Quantidade de resultados associados às diferentes formas do nome de um mesmo autor



Fonte: Dados de pesquisa.

Legenda: O nome em destaque vermelho é o autorizado.

Na figura 1, a seguir, verificam-se as autoridades Santiago, Daiane Felix (1 ocorrência) e Santiago, Daiane Félix (3 ocorrências) como exemplo do uso de acentuação influenciando na quantidade de publicações encontradas pelo usuário.

Figura 1 - Consulta de autoridade no RI

Mostrando resultados 568 a 577 de 739	
Sampaio, Luciano Menezes Bezerra	1
Sampaio, Raquel Menezes Bezerra	1
Sancho, Daniel de Oliveira	1
Santana, Antônio Cordeiro de	2
Santana, Denise Moreira Araújo de	1
Santana, Geidson Ulson Seixas	1
Santana, Luiz Sérgio Lopes	1
Santiago, Daiane Felix	1
Santiago, Daiane Félix	3

Fonte: Dados de pesquisa.

Legenda: O nome em destaque vermelho é o autorizado.

Foi encontrado, ainda, exemplo de uso de abreviaturas, que não possibilita o usuário identificar a autoria quando a busca for realizada por autor. Ex. Reis, Ana P.; L. dos Rabelo e Janiquelli da S.

- b) **Nas autoridades de assuntos**, foram identificados 1645 cabeçalhos de assuntos da subcomunidade do DEA e 429 cabeçalhos de assuntos do DAU. Desse total, foram identificadas 178 autoridades do DEA, que precisavam ser unificadas e 53 do DAU. Após serem analisadas, foram alteradas, utilizando o catálogo online da BU, a base de dados da Rede Pergamum, BN, LC e bases de dados especializadas (Thesagro, ThesBio, BDPA, SIBI-USP).

Quadro 3 - Exemplos de termos livres e autorizados

TERMO LIVRE	TERMO AUTORIZADO
Crédito rural	Crédito agrícola
Desempenho escolar	Rendimento escolar
Carne bovina	Carne de boi
Crescimento econômico	Desenvolvimento econômico
Abelha (palavra do título)	Abelhas

Fonte: Dados de pesquisa.

Como os objetivos específicos visam verificar a recuperação documental por assuntos e por autores de forma comparada entre linguagem natural e linguagem controlada, identificar a ocorrência da variação de um mesmo termo de autoridades autor e assunto indexados e corrigir e completar os dados das autoridades de nomes pessoais e de assuntos na base do Pergamum, contempla-se o objetivo geral que é elevar a qualidade da indexação de autoridades autor e assunto nas subcomunidades DEA e DAU.

7 Considerações finais

A representação temática da informação, embora aparente ser uma atividade de execução simples, carrega alta complexidade, tendo em vista se tratar de uma tarefa de análise, síntese e tradução dos conteúdos de extrema profundidade. A problemática enfrentada pelas bibliotecas diz respeito à organização, ao tratamento das informações e à sua posterior recuperação pelos usuários.

A tarefa do indexador é traduzir os termos encontrados nos documentos considerados de linguagem natural para uma linguagem controlada extraída de lista de termos autorizados, pois do contrário haverá inconsistências derivadas da natureza dessas informações. Para tanto, existem instrumentos que auxiliam nesse processo, como o thesaurus, catálogos da BN, LC, entre outros.

Vale lembrar que a indexação está estritamente ligada à recuperação da informação. Neste contexto, fica evidente a necessidade de elaboração de manual ou guia de indexação para aprimorar a qualidade dos dados inseridos no RI-UFC, evitando, assim, as inconsistências detectadas nesta pesquisa. Isso eliminará muito a dispersão dos resultados, visto que essa é uma situação enfrentada por várias bibliotecas do sistema. É necessário também discutir com a equipe de trabalho e enviar sugestões para minimizar a ocorrência de falhas na indexação e, além disso, a capacitação constante voltada para o aperfeiçoamento dessa prática.

Os tutoriais também representam uma excelente ferramenta de ajuda para o bom funcionamento dos RI e podem ser disponibilizados para depósito, para validação de arquivo e para a validação dos metadados. Podem ser apresentados em forma de vídeos com demonstração dos passos a serem executados.

Como sugestão, para melhorar a interação entre pesquisas/buscas realizadas e o sistema, enfatizamos a conveniência de capacitações constantes dos nossos usuários que possam apresentar outras ferramentas de pesquisas que não sejam só termos retiradas do título ou de partes do trabalho, como também, as pesquisas avançadas que trazem mais precisão aos resultados.

Assim, a metodologia da parte empírica deste trabalho poderá apresentar-se como proposta para o controle de autoridades pessoal e de assunto e poderá servir como modelo para tomada de decisão quanto à política de indexação implementada no Sistema de Bibliotecas da UFC.

As limitações e dificuldades encontradas são decorrentes da ausência de desenvolvimento do *DSpace* (ainda não possui ferramenta de controle de autoridade na UFC) que permita o emprego de remissivas e hierarquização dos dados indexados. Os nomes dos autores e os termos livres vão se acumulando e se duplicando. Como resultado, longe de localizar a totalidade das pesquisas sobre determinado assunto, ou dos produzidos por determinado autor, o resultado apresentado nem sempre corresponde à realidade. Portanto, a indexação da autoria pessoal e de assunto interfere no resultado da busca executada no RI-UFC, impactando na recuperação dos documentos.

Outras pesquisas que sugerimos para o avanço dos estudos no tema em questão deveriam ter como foco os Repositórios de outras instituições, incluindo as estrangeiras, a fim de verificar o nível de controle da indexação de autoridades e a sua recuperação.

Referências

ASSUMPÇÃO, F. S.; SANTOS, P. L. V. A. C. A importância do controle de autoridade: uma abordagem baseada nos objetivos e nas funções dos catálogos. *In: ENCONTRO NACIONAL DE CATALOGADORES; ENCONTRO DE ESTUDOS E PESQUISAS EM CATALOGAÇÃO*, 1.; 3., 2012, Rio de Janeiro, 2012. **Anais...** Rio de Janeiro: Biblioteca Nacional, 2012. Disponível em: <http://eprints.rclis.org/18843/>. Acesso em: 03 fev. 2021.

BILHALVA CARDOSO, C. I.; MIRANDA, A. C. D. Modelo de Controle da Autoridade de nomes pessoais: uma proposta para o Repositório Institucional da Universidade Federal do Rio Grande. **ÁGORA: Arquivologia em debate**, [S. l.], v. 30, n. 60, p. 130-147, 2019. Disponível em: <https://agora.emnuvens.com.br/ra/article/view/894>. Acesso em: 18 jan. 2021.

BOLETIM Informativo do Sistema de Bibliotecas da UFC. **Biblionotícias**, Fortaleza, v. 17, n. 10-12, out-dez. 2010.

CORRÊA, T. P. P. *et al.* Implementação do RI FURG: uma visão através do catálogo decisório de autores. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 17, n. 1, p. 27-41, 2012. Disponível em: <http://www.brapci.inf.br/index.php/article/download/20089>. Acesso em: 20 jan. 2021.

COSTA, R. M.; SOUSA, M. M. R. de. Normalização de autoridade pessoal em repositório institucional. *In*: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 20., 2018, Salvador. **Anais...** Salvador: UFBA, 2018. p. 603-610. Disponível em: <http://repositorio.ufba.br/ri/handle/ri/27708>. Acesso em: 13 jan. 2021.

FERREIRA, A. G. C. **Padronizadores de autores:** uso em base de dados. *In*: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 6., 2018, p. 780-785. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/117330>. Acesso em: 18 jan. 2021.

FUJITA, M. S. L. 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, jul./dez. 2003. Disponível em: http://server01.bc.unicamp.br/revbib/artigos/art_5.pdf. Acesso em: 15 jan. 2021.

GIL, A. C. **Como elaborar projetos de pesquisa.** São Paulo: Atlas, 2010.

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

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

LOPES, I. L. Uso das linguagens controlada e natural em base de dados: revisão de literatura. **Ci. Inf.**, Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002. Disponível em: http://eprints.rclis.org/14449/1/linguagens_controladas.pdf. Acesso em: 12 dez. 2020.

MESQUITA, M. F. N. **Avaliação do impacto do Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), no Sistema de Bibliotecas Universitárias da Universidade Federal do Ceará.** Fortaleza, 2016. 117f. Dissertação (mestrado) - Universidade Federal do Ceará, Pró-Reitoria de Pesquisa e Pós-Graduação, Mestrado Profissional em Políticas Públicas e Gestão da Educação Superior - POLEDUC. Fortaleza, 2016. Disponível em: <http://www.repositoriobib.ufc.br/000026/00002648.pdf>. Acesso em: 03 fev. 2021.

MUGNAINI, R. *et al.* Normalização de nomes de autores em fontes de informação institucionais: proposta de um método automático de verificação de erros. **Em Questão**, Porto Alegre, v. 18, Edição Especial, p. 267-279, dez. 2012. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/33265>. Acesso em: 01 fev. 2021.

NUNES, R. R; MARCONDES, C. H; WEITZEL, S. R. Diretrizes para formulação de políticas mandatórias para consolidação dos repositórios institucionais brasileiros. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 23., 2012, Rio de Janeiro. **Anais...** Recife: ANCIB, 2012. Disponível em: <http://enancib.ibict.br/index.php/enancib/xiiienancib/paper/view/3782/2905>. Acesso em: 03 fev. 2021.

REPOSITÓRIO INSTITUCIONAL DA UNIVERSIDADE FEDERAL DO CEARÁ.

Disponível em: <http://www.repositorio.ufc.br>. Acesso em: 14 ago. 2020.

SANTOS, R. F. Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. **Rev. Inf. na Soc. Contemp.**, Natal, RN, v.1, p.1-22, 11 jun. 2017.

SOARES, S. B. C. **O que é ORCID? E porque usá-lo?** 17 maio 2016. 48 slides. [material de apoio para apresentação na 1º Semana de Escrita Científica]. Disponível em: <https://pt.slideshare.net/suelybcs/orcid-o-que-e-porque-uslo-62348608>. Acesso em: 15 jan. 2021.

UNIVERSIDADE FEDERAL DO CEARÁ. Biblioteca Universitária. **Missão, visão e valores**. Disponível em: <http://www.biblioteca.ufc.br/pt/sobre-a-bibliotecauniversitaria/missao-visao-e-valores/>. Acesso em: 14 jan. 2021.

UNIVERSIDADE FEDERAL DO CEARÁ. Departamento de Economia Agrícola. **Corpo docente**. Disponível em: <https://dea.ufc.br/pt/corpo-docente/>. Acesso em: 01 set. 2020