

Avaliação de vocabulário controlado para a representação e recuperação de teses e dissertações em repositório institucional

Lizandra de Souza Santos Alves

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, SP, Brasil
lizandra17alves@gmail.com

Roberta Cristina Dal'Evedove Tartarotti

Universidade Estadual de Campinas, Biblioteca Central, Campinas, SP, Brasil
roberta_tartarotti@yahoo.com.br

Mariângela Spotti Lopes Fujita

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Programa de Pós-Graduação em
Ciência da Informação, Marília, SP, Brasil
mariangela.fujita@unesp.br

DOI: <https://doi.org/10.26512/rici.v15.n1.2022.42496>

Recebido/Recibido/Received: 2022-01-04

Aceitado/Aceptado/Accepted: 2022-04-02

ARTIGOS

Resumo

Objetivo: A pesquisa teve como objetivo realizar a avaliação do controle de vocabulário de teses e dissertações no Repositório Institucional da Universidade Estadual Paulista (UNESP), cujos procedimentos metodológicos foram delineados em pesquisa exploratório-explicativo com estudo de caso. **Metodologia:** Para isso, foram aplicados critérios avaliativos em duas categorias de análise: avaliação intrínseca, contendo critérios voltados para o repositório propriamente dito e a avaliação extrínseca, com critérios voltados para a análise da representação e da recuperação da informação. **Resultados:** Os resultados demonstraram que o Repositório Institucional da Universidade Estadual Paulista (UNESP) necessita de maior controle terminológico dos termos correlacionados às pesquisas de teses e dissertações para maior precisão na recuperação dos recursos informacionais. **Conclusão:** A falta de padronização na representação acarreta falhas na recuperação, com inconsistências na interação do repositório institucional com seus usuários.

Palavras-Chave: Repositórios institucionais. Vocabulários controlados. Avaliação da indexação. Recuperação por assuntos.

Evaluación del vocabulario controlado para la representación y recuperación de tesis y disertaciones en repositorio institucional

Resumen

Objetivo: La investigación tuvo como objetivo realizar una evaluación del control de vocabulario de tesis y disertaciones en el Repositorio Institucional de la Universidad Estadual Paulista (UNESP), cuyos procedimientos metodológicos fueron esbozados en una investigación exploratoria-explicativa con estudio de caso. **Metodología:** Para ello, se aplicaron criterios de evaluación en dos categorías de análisis: evaluación intrínseca, conteniendo criterios dirigidos al propio repositorio y evaluación extrínseca, con criterios dirigidos al análisis de la representación y recuperación de información. **Resultados:** Los resultados mostraron que el Repositorio Institucional de la Universidad Estadual Paulista (UNESP) necesita mayor control terminológico de los términos correlacionados a la investigación de tesis y disertaciones para mayor precisión en la recuperación de recursos informativos. **Conclusión:** La falta de estandarización

en la representación provoca fallas en la recuperación, con inconsistencias en la interacción del repositorio institucional con sus usuarios.

Palabras-clave: Repositorios institucionales. Vocabularios controlados. Evaluación de indexación. Recuperación por tema.

Evaluation of controlled vocabulary for the representation and recovery of theses and dissertations in institutional repository

Abstract

Objective: The research aimed to carry out an evaluation of the theses and dissertations vocabulary control in the Institutional Repository of the São Paulo State University (UNESP), whose methodological procedures were outlined in an exploratory-explanatory research with a case study. **Methodology:** For this, evaluation criteria were applied in two categories of analysis: intrinsic evaluation, containing criteria aimed at the repository itself and extrinsic evaluation, with criteria aimed at the analysis of representation and information retrieval. **Results:** The results showed that the Institutional Repository of the São Paulo State University (UNESP) needs greater terminological control of the terms correlated to the research of theses and dissertations for greater precision in the retrieval of informational resources. **Conclusion:** The lack of standardization in the representation causes failures in the recovery, with inconsistencies in the interaction of the institutional repository with its users.

Keywords: Institutional repositories. Controlled vocabularies. Indexing evaluation. Subject retrieval.

1 Introdução

Tendo em vista a evolução e a utilização das Tecnologias da Informação e Comunicação (TICs), houve também um grande aumento da produção documental em formato digital e a necessidade de sua disseminação em sistemas de recuperação da informação no âmbito digital. Medeiros (2017, p. 8) apresenta que “isso tem proporcionado um novo ciclo de comunicação científica em todos os campos do conhecimento, haja vista a importância do acesso aberto no novo universo”.

Um dos sistemas de recuperação desenvolvidos e que tem grande importância no âmbito acadêmico são os repositórios institucionais, sendo sistemas de recuperação voltados para a guarda, preservação e disseminação da produção informacional científica de uma universidade. Em uma biblioteca tradicional, o acesso aos exemplares impressos de teses e dissertações pode ser realizado tanto via Sistema de Comutação Bibliográfica (COMUT), pouco utilizado, mas ainda em uso por algumas bibliotecas quanto pelo próprio usuário no catálogo online da biblioteca, por meio dos códigos de representação e realizar a recuperação nas estantes (ALVES, 2009, p. 32). Com o surgimento de repositórios institucionais, o usuário pode realizar a recuperação destes recursos informacionais já em formato digital, com maior rapidez, para além das fronteiras físicas, contanto que tenha acesso à internet.

Considerando a diferença entre repositório institucional e um catálogo online de uma biblioteca universitária,

[...] na prática, a visibilidade dos recursos informacionais produzidos pela comunidade acadêmica é possibilitada nos repositórios institucionais pelos processos de busca e recuperação dos registros dos recursos informacionais, permitindo conhecer os títulos, autores ou assuntos que compõem a coleção

e reunindo todos os recursos informacionais de um autor ou sobre um determinado assunto, além de outros pontos de acesso. Os repositórios institucionais diferenciam-se dos catálogos online, que, no geral, possibilitam a recuperação dos registros dos recursos informacionais físicos ou digitais e indicam a localização dos mesmos no acervo físico da biblioteca universitária, cuja coleção abarca diversos tipos de recursos informacionais, produzidos ou não pela comunidade acadêmica (TARTAROTTI, 2020, p. 4).

Pensando nas atividades de organização, representação, recuperação e disseminação da informação nesses repositórios institucionais, a Ciência da Informação é campo de estudo interdisciplinar junto à Organização do Conhecimento. Dahlberg (2006) acentua que a Organização do Conhecimento é uma ciência que cuidará de organizar e sistematizar os conceitos, sempre se preocupando com os valores dos conteúdos dos objetos ou assuntos que estão sendo organizados.

Guimarães (2008, p. 84) identifica que a Organização e Representação do Conhecimento apresenta três aspectos: **a) Processos:** abarca a análise, condensação e a representação da informação que será tratada; **b) Produtos:** produtos gerados serão os índices e os resumos que facilitarão o acesso aos recursos informacionais; **c) Instrumentos:** são ferramentas da linguagem documental para a padronização da representação temática das informações contidas nos recursos informacionais. Sousa (2019) ressalta que

O conhecimento sobre os objetos, ações e fenômenos é resultado de diversas percepções cognoscíveis do ser humano, por meio do espaço e do tempo, nisto é possível elencar atributos que possam representá-lo em determinados domínios. Para tanto, é natural recorrer aos métodos de organização e representação do conhecimento para sistematizar estes atributos, que passam por um tratamento técnico (SOUSA, 2019, p. 22).

Para que haja eficácia na representação e disseminação desses recursos informacionais, tais atributos dos recursos informacionais passam por um tratamento técnico a partir de métodos de organização e representação. Um dos métodos utilizados para auxiliar na representação pode ser o vocabulário controlado, que será utilizado pelo profissional indexador no momento da atribuição de descritores que representem o assunto do recurso informacional. No entanto, a inserção e a representação dos recursos informacionais em um sistema de recuperação da informação podem ser realizados tanto pelo profissional da biblioteca responsável pela atividade quanto pelos usuários, tendo em vista que os repositórios institucionais trouxeram autonomia aos usuários para essas atividades, sendo realizada pelos próprios autores das teses e dissertações pelo método de autoarquivamento, onde o autor irá inserir o recurso informacional no sistema e ele mesmo irá atribuir palavras-chave para a sua representação.

Em repositórios institucionais, o autoarquivamento possibilita autonomia para que os recursos informacionais como teses e dissertações possam ser inseridos pelos próprios autores, com atribuição de palavras-chave que irão representar o assunto do recurso informacional, visando sua representação e recuperação por assuntos. Entretanto, considerando-se que o uso da linguagem natural na representação pode causar inconsistências na recuperação, a utilização de um vocabulário controlado pelo profissional indexador ou pelo usuário é de suma importância.

Os termos que serão utilizados para descrever e recuperar os recursos informacionais são unidades terminológicas. Portanto, o controle de vocabulário é um recurso essencial para se realizar o controle terminológico no momento da indexação. Pensando nisso, os vocabulários controlados foram criados para auxiliar na organização e representação do acervo das bibliotecas, para que tivesse maior precisão no momento da recuperação da informação que o usuário estivesse buscando.

O vocabulário controlado é um instrumento dinâmico e que precisa estar em consonância com o conhecimento científico que vem se modificando e se atualizando constantemente, principalmente no âmbito acadêmico. Kobashi (2008) salienta que o vocabulário controlado é uma linguagem artificial construída de termos que irão representar um ou vários campos do conhecimento. A fim de garantir a precisão e eficácia na comunicação entre os usuários e os sistemas de informações,

[...] um vocabulário controlado surge para tentar solucionar problemas causados por características próprias da linguagem natural como ambiguidades, polissemia, sinonímia e homonímia, que quando não padronizados afetam os resultados dos processos de recuperação da informação quanto à precisão e relevância (SILVA, 2015, p. 21).

Todavia, para esse fim é importante a aplicação de avaliações nos vocabulários controlados utilizados pelos repositórios institucionais para que se tenha maior controle da padronização dos termos e assim os objetivos sejam alcançados de forma consistente, pois, como um vocabulário controlado é composto por termos e o conhecimento está sempre em movimento, ele está sujeito a ficar desatualizado. Essas avaliações precisam ser aplicadas pensando sempre no vocabulário controlado propriamente dito, onde irá avaliar se os elementos de sua estrutura estão cumprindo a função de auxiliar no momento da representação de maneira consistente; e avaliar como está sendo o seu impacto na recuperação da informação.

No âmbito da Ciência da Informação o controle de vocabulário utiliza métodos de avaliação com diferentes abordagens visando à qualidade de representação e recuperação. Entretanto, o uso dos vocabulários controlados para controle de vocabulário depende dos profissionais que administram esses sistemas de organização e recuperação do conhecimento.

O controle de vocabulário pode ser qualitativamente inferior ao que deveria ser, porque o vocabulário controlado:

[...] deve estar de acordo com as políticas de indexação definidas pelo sistema, intermediando o acesso à recuperação da informação e este, por outro lado, deve considerar a instituição onde se desenvolve; as expectativas e necessidades do usuário; as características do assunto tratado; os recursos humanos, físicos e financeiros; os produtos e serviços visados e a relação custo/desempenho (BOCCATO; FUJITA, 2006, p. 268).

A atualização e a avaliação do vocabulário controlado em repositórios institucionais visam uma maior consistência na representação e na recuperação por assuntos dos recursos informacionais. Porém, não só o vocabulário propriamente dito precisa ser avaliado, mas também o seu uso diretamente nos repositórios institucionais, pois alguns repositórios utilizam vocabulários controlados advindos de outras instituições e não vocabulários próprios que podem ter controle terminológico direto pelos profissionais da instituição que sejam responsáveis pela sua atualização.

Embora a avaliação do vocabulário controlado em repositórios institucionais seja um aspecto relevante, verifica-se a ausência na literatura de procedimentos metodológicos voltados para a avaliação do controle de vocabulários nestes sistemas de recuperação da informação. No contexto específico da Universidade Estadual Paulista (UNESP), verifica-se o uso do Tesauro UNESP no catálogo online Athena como uma linguagem documental padronizada, enquanto que no Repositório Institucional não há a adoção de uma linguagem controlada para a representação e recuperação por assuntos dos recursos informacionais, gerando inconsistências nestes processos.

Nesse sentido, a pesquisa teve como objetivo realizar avaliação do controle de vocabulário de teses e dissertações no Repositório Institucional da Universidade Estadual Paulista (UNESP), visando contribuir com a melhoria do controle de vocabulário neste sistema de recuperação da informação e com elementos de avaliação da indexação no campo de pesquisa teórico e prático da Organização e Representação do Conhecimento.

2 Procedimentos metodológicos

Os procedimentos metodológicos foram realizados com pesquisa exploratório-explicativa com estudo de caso, sendo aplicados critérios avaliativos em duas categorias de análise: *avaliação intrínseca*, contendo critérios voltados para o repositório propriamente dito e a *avaliação extrínseca*, com critérios voltados para a análise da representação e da recuperação da informação no repositório.

O universo de pesquisa foi o Repositório Institucional da Universidade Estadual Paulista (UNESP)¹ (Figura 1), cuja missão é promover a instituição em âmbito nacional e internacional, por meio do acesso aberto, da disseminação, da preservação, da valorização e do reconhecimento de sua produção científica, acadêmica, artística, técnica e administrativa. O Repositório Institucional da Universidade Estadual Paulista (UNESP) é dividido por comunidades e subcomunidades, tendo suas criações e alterações de comunidades e de subcomunidades em consonância com a criação ou alteração de denominação de Unidades universitárias, Unidades complementares; Unidades auxiliares, Centros interdepartamentais, Departamentos, Programas de pós-graduação, e Seções técnicas administrativas, de modo que o Repositório Institucional da Universidade Estadual Paulista (UNESP) reflita a organização da Universidade (UNESP, 2021)

Figura 1: Página inicial do Repositório Institucional da UNESP

The screenshot shows the homepage of the UNESP Institutional Repository. At the top, there is a search bar and navigation links for 'Busca Integrada', 'Entrar', 'Chat', and 'Sobre'. Below the search bar, there are three main sections: 'Apresentação', 'Autoarquivamento de Dissertações e Teses', and 'Dúvidas'. The 'Apresentação' section describes the repository's mission. The 'Autoarquivamento' section includes a login button and a link for more information. The 'Dúvidas' section offers an online service and a form to send questions.

Fonte: Dados da pesquisa.

Para que o recurso informacional possa ser arquivado no repositório, deverá seguir critérios gerais pré-estabelecidos no regulamento, sendo eles: possuir entre seus criadores ao menos uma pessoa vinculada à instituição; estar em formato digital; ser resultado de atividade científica, acadêmica, artística, técnica ou administrativa realizada nas dependências da universidade ou em seu nome; estar completo, pronto para publicação ou já ter sido publicado e com a cessão de direitos patrimoniais, quando for o caso; estar categorizado como: a) artigo publicado em periódico científico, em sua versão publicada ou aceita para a publicação; b) trabalho completo ou resumo publicado em anais de congresso, conferência, encontro, simpósio ou qualquer outra reunião de natureza técnico científica; c) livro ou capítulo de livro;

¹ Disponível em: <https://repositorio.unesp.br>.

d) tese de doutorado, tese de livre-docência, dissertação de mestrado ou trabalho de conclusão de curso de graduação ou de especialização; e) notação musical; f) imagem; g) material cartográfico; h) gravação de som ou gravação de vídeo; i) software; j) patente; k) outra categoria a ser estabelecida pelo Grupo Gestor do Repositório Institucional da Universidade Estadual Paulista (UNESP) e o de não violar direitos autorais. Fora os critérios gerais, deverão ser cumpridos critérios específicos para cada tipo de recurso informacional (UNESP, 2021)

O *Repositório Institucional da Universidade Estadual Paulista (UNESP)* tem o total de 174211 registros² dentre os variados tipos de recursos informacionais arquivados em seu sistema, que vão desde teses e dissertações a *podcasts*, sendo possível navegar e recuperar por tipos de produção, data do recurso informacional, autor, título e palavra-chave.

Gil Leiva (2008, p. 213) aponta duas metodologias de avaliação: uma tomando como objeto de análise o próprio vocabulário controlado, chamada de *avaliação intrínseca*, e a outra que estuda o papel do vocabulário controlado na representação e recuperação de informações, denominada de *avaliação extrínseca*. A partir dessas duas metodologias propostas pelo referido autor, foram criadas duas categorias de análise: avaliação intrínseca e avaliação extrínseca. Cada categoria possui critérios avaliativos diferentes tendo em vista aspectos internos e externos ao repositório institucional conforme Alves (2021):

Avaliação intrínseca:

a) Disponibilidade de link para o vocabulário controlado: corresponde ao fornecimento pelo repositório institucional, em sua interface, de um link que direcione o usuário prontamente para o vocabulário controlado.

b) Vocabulário correspondente ao domínio: se o vocabulário controlado empregado confere ao domínio contemplado pelo repositório institucional.

c) Autoarquivamento: disponibilização do autoarquivamento. Em caso positivo, analisar se é utilizada a linguagem natural ou se há o direcionamento para o vocabulário controlado para realizar a escolha de termos advindos da linguagem padronizada correspondente ao recurso informacional. Explorar no repositório institucional se está disponível algum manual de autoarquivamento.

Avaliação extrínseca: os critérios avaliativos serão pautados exclusivamente à recuperação da informação nos repositórios institucionais:

a) Controle terminológico: voltado para a análise de imprecisões advindas de ambiguidade, sinonímia, polissemia e homonímia na recuperação dos recursos informacionais.

² Verificado no dia 22 de março de 2022 em: [Repositório Institucional UNESP](#)

b) Forma dos termos: corresponde à forma com que os termos foram apresentados na recuperação. Se aparecem em plural ou singular, se apresenta entrada direta ou inversa, o idioma adotado, se os termos são simples ou compostos e se são adotadas abreviaturas (SOUZA; CAMPOS, 2007).

c) Grau de especificidade: análise do nível de especificidade terminológica no momento da recuperação dos recursos informacionais, isto é, verificar se a especificidade da recuperação está de acordo com os usuários potenciais (GIL LEIVA, 2008).

3 Resultados e discussão

Os resultados da aplicação dos critérios avaliativos do controle de vocabulário no Repositório Institucional da Universidade Estadual Paulista (UNESP) foram divididos em duas categorias de análise: *avaliação intrínseca* e *avaliação extrínseca*, conforme delineadas nos procedimentos metodológicos.

3.1 Resultados da *avaliação intrínseca*

a) Disponibilidade de link para o vocabulário controlado

Nesse critério, a busca pelo resultado foi feita por toda a interface do repositório institucional e em seus menus. Observou-se que o Repositório Institucional da Universidade Estadual Paulista (UNESP) não possui nenhum *link* em sua interface que direcione o usuário para um vocabulário controlado ou apresenta alguma informação do uso de um vocabulário para a representação e recuperação dos recursos informacionais, mesmo a instituição possuindo um tesauro próprio (Tesauro UNESP).

b) Vocabulário correspondente ao domínio

Para esse critério, foi analisado o domínio do vocabulário controlado, caso utilizado pelo repositório, a fim de verificar se é correspondente ao domínio abrangente. Conforme mencionado, não há nenhuma informação que relate o uso de um vocabulário controlado para realizar o controle dos termos utilizados no momento do autoarquivamento das teses e dissertações. Portanto, nesse critério não foi possível analisar se o vocabulário controlado corresponde ao domínio em que as teses e dissertações estão inseridas.

c) Autoarquivamento

Os resultados desse critério também foram obtidos a partir da análise da interface do repositório institucional e de seus menus. No Repositório Institucional da Universidade Estadual Paulista (UNESP) é realizado o autoarquivamento de teses e dissertações pelos autores-alunos da pós-graduação das diversas áreas oferecidas pela universidade. O autoarquivamento pode ser realizado em até 60 dias após a defesa e é obrigatório. Ao concluir a submissão, o arquivo passa

por uma conferência e mediante alguma inconsistência, o aluno é notificado por e-mail. Ao receber o aceite final, o texto será disponibilizado para acesso aberto no repositório institucional e o autor-aluno poderá entregar o exemplar impresso para a Seção Técnica de Pós-Graduação, onde será disponibilizado para acesso na Biblioteca da unidade correspondente ao mestrado ou doutorado.

Na interface do repositório há um *link* com a disponibilização de informações sobre autoarquivamento, contendo um breve conceito do que é esse processo, prazos para o arquivamento, informações sobre a versão impressa e informações de como realizá-lo. Essas informações estão disponíveis em um tutorial textual, um tutorial em vídeo, uma portaria de 2015, um folder explicativo e um fluxograma de autoarquivamento elaborado pela Biblioteca da UNESP de Rio Claro.

Após realizar login no *link* disponível exclusivamente para o autoarquivamento de teses e dissertações, o autor-aluno realiza a inserção dos metadados que correspondem ao seu trabalho, sendo compostos por: título e subtítulo (em português e em inglês), nome completo, nome do orientador e coorientador (se houver), data da defesa, resumo (em português e em inglês), palavras-chave, idioma, tipo de trabalho (dissertação ou tese), unidade da UNESP, programa de pós-graduação, área de concentração e linha de pesquisa. Em seguida, a escolha da versão do trabalho (parcial ou integral) para o *upload* de trabalho em formato PDF.

A disponibilização de um vocabulário controlado para a inserção de termos que correspondam ao trabalho é inexistente, fazendo com que os trabalhos sejam representados no repositório institucional pela linguagem natural, isto é, pelas palavras-chaves dos autores-alunos.

3.2 Resultados da avaliação extrínseca

Os critérios dessa categoria abordam o impacto do vocabulário controlado na representação e recuperação dos recursos informacionais. Foi escolhido um termo que correspondesse ao domínio do repositório para realizar uma busca simples e analisar o título das teses e dissertações que sejam apresentadas nos dez primeiros recursos informacionais recuperados. Nesse caso, o termo escolhido foi “Aquisição”, utilizado, por exemplo, nas teses e dissertações pertencentes ao campo da Ciência da Informação.

a) Controle terminológico

Observou-se como o termo de busca foi apresentado na recuperação em questão terminológica nos títulos dos recursos informacionais, buscando ambiguidades, sinonímia, polissemia, homonímia ou se foi recuperado exatamente como posto na busca. Para este resultado, foi necessário realizar a busca diretamente por tipo de material, pois, na busca simples pelo campo

disponível na interface não foi recuperado nenhuma tese ou dissertação nos dez primeiros resultados.

Na busca realizada em “dissertação de mestrado”, o termo “Aquisição” foi apresentado no título de seis dos dez recursos informacionais recuperados. Mesmo havendo o termo no título, só quatro dos seis recursos informacionais realmente eram referentes à aquisição de algo e entre os quatro que não apresentaram o termo no título não correspondiam ao assunto tratado. Já em relação ao tipo “tese de doutorado” foram recuperados seis dos dez primeiros recursos informacionais com o termo “Aquisição” no título. Desses seis, cinco correspondiam ao assunto tratado e entre os quatro que não apresentaram o termo no título não tinham correlação direta em suas pesquisas.

b) Forma dos termos

Analisou-se se nos títulos o termo escolhido apresentava pluralidade ou singularidade, o idioma adotado, se as palavras-chave são compostas ou simples e se são adotadas abreviaturas. Em “dissertações de mestrado” foram recuperados termos apenas no singular, em português e com a apresentação simples e composta, como, “aquisição de dados”, e não houve nenhuma abreviatura. No tipo “tese de doutorado”, foram recuperados recursos informacionais que apresentavam o termo no título apenas no singular e em português, havendo apresentação simples e composta, como, “aquisição volumétrica” e não apresentou abreviações.

c) Grau de especificidade

Para esse critério, observou-se se os recursos informacionais recuperados estavam de acordo com os usuários potenciais do repositório institucional. No tipo “dissertações de mestrado”, a busca recuperou recursos informacionais voltados para as áreas que a UNESP abrange. Um facilitador para isso é a seleção de critérios para o arquivamento e a disponibilização de autoarquivamento. O mesmo corresponde para as recuperações no tipo “teses de doutorado”.

4 Considerações finais

De forma geral, os resultados obtidos a partir dos critérios das categorias de análise permitiram observar um descontrole da linguagem no momento da recuperação dos recursos informacionais, demonstrando falha na representação dos mesmos. Por conta da falta de controle de vocabulário no momento da representação, há excessos de recursos informacionais na recuperação, apresentando pesquisas que não são correlacionadas com o termo buscado.

Esse fator pode estar relacionado com o uso da linguagem natural no momento da inserção dos metadados no autoarquivamento das teses e dissertações pelos autores-alunos. O fato de não haver indexação dos termos que representem esses materiais de forma padronizada com o domínio e que sejam prontamente correspondentes à pesquisa ocasionam

inconsistências na recuperação dos recursos informacionais no Repositório Institucional da Universidade Estadual Paulista (UNESP).

É possível salientar que pelo fato de a referida universidade possuir o seu próprio vocabulário controlado, o mesmo poderia estar integrado ao sistema do repositório institucional da Universidade Estadual Paulista (UNESP) de forma que estivesse disponível para conhecimento dos usuários gerais e implementado o Tesouro UNESP no momento do autoarquivamento para os autores-alunos, visando melhor padronização na representação e recuperação por assuntos nestes sistemas de recuperação da informação. Porém, essa questão é complexa, pois, ao mesmo tempo em que poderia facilitar, causaria dificuldades, visto que muitos autores-alunos não possuem conhecimento da existência do Tesouro UNESP e não sabem utilizá-lo.

Para resolver tal lacuna, seria necessária a disponibilidade de tutoriais de seu uso e treinamentos, preferencialmente no início do mestrado/doutorado. Além disso, a inserção de um novo campo de metadados direcionado para os bibliotecários que realizam a validação dos metadados inseridos pelos autores-alunos possibilitaria um melhor controle e padronização da linguagem, tanto na representação quanto na recuperação por assuntos.

A aplicação de critérios avaliativos no Repositório Institucional da Universidade Estadual Paulista (UNESP) ressalta a importância da utilização de um vocabulário controlado na representação dos recursos informacionais para melhorar a precisão na recuperação por assuntos no momento da busca pelos usuários e da integração ao repositório institucional de um vocabulário controlado que integre termos de seu domínio.

Referências

ALVES, J. C. **Organização e representação da informação na biblioteca digital de teses e dissertações da Universidade do Estado de Santa Catarina - UDESC: uma análise focada em metadados sob a luz do padrão MTD-br.** 2009. 161 p. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Santa Catarina, Florianópolis, 2009.

ALVES, L. S. S. **Avaliação de controle de vocabulários em repositórios Institucionais: uma proposta de aplicação metodológica Intrínseca e extrínseca.** 2021. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2021.

BOCCATO, V. R. C; FUJITA, M. S. L. Estudos de avaliação quantitativa e qualitativa de linguagens documentárias: uma síntese bibliográfica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 267-281, 2006.

DAHLBERG, I. Knowledge Organization: a new science? **Knowledge Organization**, v. 33, n. 1, p. 11-19, 2006.

GIL LEIVA, I. **Manual de indización: teoría y práctica**. Gijón: Trea, 2008.

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, n. 1, p. 77-99, 2008.

KOBASHI, N. Y. **Vocabulário controlado: estrutura e utilização**. Brasília: Escola Nacional de Administração Pública, 2008.

MEDEIROS, R. Prefácio. *In*: VECHIATO, F. *et al.* (org.). **Repositórios digitais: teoria e prática**. Curitiba: EDUTFPR, 2017.

SILVA, B. **Vocabulário de termos livres e controlados para a coleção de Teses e Dissertações da Universidade Federal de Viçosa**. 2015. 355 p. Dissertação (Mestrado em Biblioteconomia)- Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2015.

SOUSA, J. L. **Avaliação do padrão *simple Knowledge organization system (SKOS)* para a representação de vocabulários controlados**. 2019. 100 p. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de São Carlos, São Carlos, 2019.

SOUZA, J. C. C. E.; CAMPOS, M. L. A. Avaliação de linguagem de indexação aplicada à informação jornalística: estudo de caso. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8, 2007, Salvador. **Anais [...]**. Salvador, 2007. p. 1-17.

TARTAROTTI, R. C. D. Representação temática em repositórios institucionais de bibliotecas universitárias: a percepção de gestores e bibliotecários catalogadores-indexadores da USP, UNESP e UNICAMP. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 18, p. 1-25, 2020.

UNIVERSIDADE ESTADUAL PAULISTA. Repositório Institucional da Universidade Estadual Paulista (UNESP). Disponível em: <https://repositorio.unesp.br/> Acesso em: 01 jan. 2021.