



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

Maria Carolina Andrade e Cruz

GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DE LINGUAGENS DE
INDEXAÇÃO EM BIBLIOTECAS UNIVERSITÁRIAS

**Marília - SP
2024**

Maria Carolina Andrade e Cruz

Gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação como parte das exigências para a obtenção do título de Doutora em Ciência da Informação pela Faculdade de Filosofia e Ciências, Universidade Estadual Paulista (UNESP), Campus de Marília.

Área de Concentração: Informação, Tecnologia e Conhecimento

Orientadora: Prof^ª. Dra. Mariângela Spotti Lopes Fujita

Coorientadora: Prof^ª Dra. Franciele Marques Redigolo

**Marília - SP
2024**

C957g Cruz, Maria Carolina Andrade e
Gerenciamento da construção e manutenção de linguagens de
indexação em bibliotecas universitárias / Maria Carolina
Andrade e Cruz. -- Marília, 2024
184 p. : il.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),
Faculdade de Filosofia e Ciências, Marília
Orientadora: Mariangela Spotti Lopes Fujita
Coorientadora: Franciele Marques Redigolo

1. Linguagem de indexação. 2. Gerenciamento da linguagem
de indexação. 3. Bibliotecas universitárias. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da
Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília.
Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE MARIA CAROLINA ANDRADE CRUZ E SANTOS, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, DA FACULDADE DE FILOSOFIA E CIÊNCIAS - CÂMPUS DE MARÍLIA.

Aos 21 dias do mês de junho do ano de 2024, às 14:00 horas, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de MARIA CAROLINA ANDRADE CRUZ E SANTOS, intitulada **Gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof(a). Dr(a). MARIANGELA SPOTTI LOPES FUJITA (Orientador(a) - Participação Virtual) do(a) Programa de PósGraduação em Ciência da Informação / Unesp Faculdade de Filosofia e Ciências Marília, Prof(a). Dr(a). WALTER MOREIRA (Participação Virtual) do(a) Departamento de Ciência da Informação / Unesp, Faculdade de Filosofia e Ciências, Marília, Prof(a). Dr(a). CIBELE ARAÚJO CAMARGO MARQUES DOS SANTOS (Participação Virtual) do(a) Departamento de Informação e Cultura / Universidade de São Paulo, Escola de Comunicações e Artes, Prof(a). Dr(a). ISIDORO GIL LEIVA (Participação Virtual) do(a) Gestión de la Información y de la Comunicación en las Organizaciones / Universidad de Murcia, Facultad de Comunicacion y Documentacion, Prof(a). Dr(a). MARIA LUIZA DE ALMEIDA CAMPOS (Participação Virtual) do(a) Programa de Pós-Graduação em Ciência da Informação / Universidade Federal da Bahia. Após a exposição pela doutoranda e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, a discente recebeu o conceito final: Aprovada _ _ _ _ . Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Prof(a). Dr(a). MARIANGELA SPOTTI LOPES FUJITA



Documento assinado digitalmente

MARIANGELA SPOTTI LOPES FUJITA

Data: 24/06/2024 13:41:31-0300

Verifique em <https://validar.iti.gov.br>

Maria Carolina Andrade e Cruz

Gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias

Tese apresentada ao Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp), como requisito para a obtenção do título de Doutor em Ciência da Informação.

Área de concentração: Informação, Tecnologia e Conhecimento

Linha de pesquisa: Produção e Organização da Informação

Banca Examinadora

Prof^a. Dr^a. Mariângela Spotti Lopes Fujita

Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília.
Orientador

Prof. Dr. Walter Moreira

Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília.

Prof^a. Dr^a. Cibele Araújo Camargo Marques dos Santos

Universidade de São Paulo (USP)

Prof. Dr. Isidoro Gil Leiva

Universidad de Murcia

Prof^a. Dr^a. Maria Luiza De Almeida Campos

Universidade Federal Fluminense/Universidade Federal da Bahia

Marília, 21 de junho de 2024.

À minha família, bênção de Deus em minha vida.

AGRADECIMENTOS

Esperei muito esse momento em que finalmente terminaria a minha Tese e escreveria os agradecimentos às pessoas especiais que fizeram parte dessa jornada comigo.

Como sempre, primeiramente agradeço à Deus e peço: “inspirai, Senhor, as nossas ações e ajudai-nos a realizá-las, para que em Vós comece e para Vós termine tudo aquilo que fizemos.”

Pela minha família que tanto me deu suporte para enfrentar os desafios cotidianos:

Ao meu amado marido Kauê Cruz e Santos que me apoiou e incentivou por todos esses anos, fazendo todas as outras coisas que demandam tempo para que eu pudesse me dedicar a pesquisa e aos estudos, não me deixando desistir. Hoje, como fruto do nosso amor, estamos esperando nossa primeira filha, Catarina.

Aos meus pais, Maria Inês Andrade e Cruz e Nilson Augusto Correa da Cruz, meus eternos amores, que foram compreensivos com as minhas ausências e sempre deram todo o suporte para que eu pudesse atingir meus objetivos. Ao meu irmão Davi que nos ensina a cada dia (moleque arteiro que tanto amo).

Aos meus avós maternos (*in memoriam*) e paternos que fizeram ser exemplos de seres humanos fortes e capazes de grandes feitos a partir do amor, dedicação e humildade.

Às minhas amigas/irmãs, Ana Laura Dezen, Carolina Soares, Juliana Borges e Brunna Patrício que estiveram presentes nos meus momentos de alegrias e tristezas, independente da distância.

Às minhas queridas orientadoras, Prof^a Dr^a Mariângela Spotti Lopes Fujita e Prof^a Dr^a Franciele Marques Redigolo para sempre minha admiração e gratidão pelo apoio e ensinamentos passados.

À banca, por contribuir com a minha pesquisa.

À todas as amigas bibliotecárias e assistentes da Divisão Técnica de Biblioteca e Documentação, campus de Botucatu, em especial a Rosangela Lobo e a Luciana Pizzani, pelo incentivo, compreensão e apoio; a Rosemary pela correção das referências e doces palavras de encorajamento.

Aos docentes do Programa de Pós-Graduação em Ciência da Informação, campus de Marília.

Aos bibliotecários e profissionais da informação.

Minha profunda gratidão.

*“Aquele que trabalha duro pode superar um gênio.
Mas, de nada adianta trabalho duro se você não
confia em você mesmo.”*

Might Guy para Rock Lee

RESUMO

A construção de linguagens de indexação corresponde aos procedimentos estabelecidos por meio da literatura normativa, considerando os aspectos e particularidades da instituição, como a política de indexação e as perspectivas culturais. Em pesquisas anteriores, verificou-se que há bibliotecas universitárias que possuem uma linguagem de indexação própria ou estão com projetos para a sua elaboração. Pela importância de seus estudos e para a viabilização da pesquisa, foram selecionadas a USP e Unesp que possuem procedimentos documentados e definidos sobre o gerenciamento da construção e a manutenção das linguagens de indexação para responder à questão de pesquisa: como as bibliotecas universitárias da USP e da Unesp fazem o gerenciamento da construção e a manutenção das linguagens de indexação? A hipótese é de que outras bibliotecas universitárias não possuem procedimentos sólidos, no que diz respeito à fundamentação dos procedimentos metodológicos para o gerenciamento da construção e manutenção de uma linguagem de indexação, especialmente com base na literatura normativa e em uma política de indexação. Por isso, a tese apresentada é de que a experiência de sistematização dos procedimentos que implicam a construção e o gerenciamento de linguagens de indexação das duas instituições, USP e Unesp, contribuirá com as bibliotecas universitárias e/ou unidades de informação que desejam desenvolver uma ferramenta própria para o controle e representação de seus recursos informacionais. A pesquisa propõe investigar o gerenciamento da construção e manutenção das linguagens de indexação das bibliotecas universitárias da USP e Unesp. O objetivo geral desta pesquisa é contribuir com as bibliotecas universitárias que buscam o desenvolvimento de sua própria linguagem de indexação e também com os estudos da área de organização e representação da informação em Ciência da Informação. Os objetivos específicos se definem em: 1) realizar estudo teórico e metodológico a partir da revisão de literatura nacional e internacional sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação no âmbito da Ciência da Informação; 2) desenvolver estudo analítico descritivo sobre gerenciamento da construção e manutenção de linguagens de indexação pelas bibliotecas universitárias da USP e UNESP; 3) elaborar orientações sobre o gerenciamento da construção e manutenção de linguagens de indexação na política de indexação de bibliotecas universitárias, tendo como parâmetro a literatura técnico-científica e a *práxis* das bibliotecas universitárias. Na metodologia utilizou-se a pesquisa bibliográfica e a descritiva que permitiram a estruturação da revisão de literatura, análise categorial e elaboração de questões para entrevista com bibliotecárias gestoras e indexadores das duas universidades participantes. A partir das análises, pode-se reunir aspectos fundamentais na perspectiva da literatura, como as práticas utilizadas por outras instituições, a fundamentação dos processos a partir da literatura normativa e delimitação da trajetória e decisões da USP e Unesp no desenvolvimento de suas linguagens de indexação. A análise da *práxis* cotidiana das bibliotecas universitárias ocorreu através de entrevista com bibliotecárias gestoras e indexadoras reunindo relatos e experiências sobre a gestão da ferramenta e seu uso. Conclui-se que esta tese pôde sistematizar e reunir aspectos fundamentais para as bibliotecas universitárias que desejam fazer o gerenciamento da construção e manutenção de suas linguagens de indexação.

Palavras-chave: Gerenciamento da linguagem de indexação. Manutenção da linguagem de indexação. Construção de linguagem de indexação. Linguagem de indexação. Bibliotecas universitárias. ISO 25964.

ABSTRACT

The construction of indexing languages corresponds to the procedures established through normative literature, considering the aspects and particularities of the institution, such as the indexing policy and cultural perspectives. Previous research has shown that some university libraries have their own indexing language or are planning to develop one. Due to the importance of these studies and to facilitate the research, USP and UNESP were selected because they have documented and defined procedures for managing the construction and maintenance of indexing languages to address the research question: how do the university libraries of USP and UNESP manage the construction and maintenance of indexing languages? The hypothesis is that other university libraries do not have solid procedures regarding the methodological foundations for managing the construction and maintenance of an indexing language, especially based on normative literature and an indexing policy. Therefore, the thesis presented is that the experience of systematizing the procedures involved in the construction and management of indexing languages at the two institutions, USP and UNESP, will contribute to university libraries and/or information units that wish to develop their own tool for controlling and representing their informational resources. The research aims to investigate the management of the construction and maintenance of indexing languages at the university libraries of USP and UNESP. The general objective of this research is to contribute to university libraries seeking to develop their own indexing language and also to the studies in the field of organization and representation of information in Information Science. The specific objectives are defined as: 1) conducting a theoretical and methodological study based on a review of national and international literature on the management of the construction and maintenance of indexing languages from the perspective of indexing policy within Information Science; 2) developing a descriptive analytical study on the management of the construction and maintenance of indexing languages by the university libraries of USP and UNESP; 3) preparing guidelines on the management of the construction and maintenance of indexing languages in the indexing policy of university libraries, based on technical-scientific literature and the practices of university libraries. The methodology used was bibliographic and descriptive research, which allowed for the structuring of the literature review, categorical analysis, and the development of questions for interviews with librarian managers and indexers from the two participating universities. From the analyses, fundamental aspects could be gathered from the perspective of the literature, such as practices used by other institutions, the foundation of processes based on normative literature, and the trajectory and decisions of USP and UNESP in developing their indexing languages. The analysis of the everyday practices of university libraries was conducted through interviews with librarian managers and indexers, gathering reports and experiences on the management of the tool and its use. It is concluded that this thesis was able to systematize and gather fundamental aspects for university libraries that wish to manage the construction and maintenance of their indexing languages.

Keywords: Indexing language management. Indexing language maintenance. Indexing language construction. Indexing language. University libraries. ISO 25964.

LISTA DE ILUSTRAÇÕES

Quadro 1 -Sistematização da Pesquisa.....	20
Quadro 2 - Articulação das categorias com os textos adotados.....	24
Quadro 3 - Símbolos e abreviações.....	57
Quadro 4 - Termos compostos em uma única palavra.....	62
Quadro 5 - Considerar ou não a divisão dos termos.....	63
Quadro 6 - Sistematização das etapas de planejamento do tesauro segundo a ISO 25964 (2011).....	67
Quadro 7 - Resultados da BRAPCI.....	76
Quadro 8 - Resultado de busca nas bases de dados Scopus e <i>Web of Science</i>	76
Quadro 9 - 1ª estratégia de busca e trabalhos recuperados.....	77
Quadro 10 - 2ª estratégia de busca e trabalhos recuperados.....	78
Quadro 11 - Textos para análise descritiva.....	80
Quadro 12 - Questões articuladas com a fundamentação teórica.....	83
Quadro 13 - Pontos de destaques da entrevista.....	125

SUMÁRIO

1 INTRODUÇÃO	10
1.1 Delimitação do problema	17
1.2 Justificativa	18
1.3 Organização da tese articulada com os objetivos	20
2 REVISÃO DE LITERATURA	24
2.1 Procedimentos de implementação da linguagem de indexação: relatos de experiências.....	26
2.2 O uso de linguagens de indexação por bibliotecas universitárias.....	33
2.3 Capacidade de interoperabilidade nas linguagens de indexação	38
2.4 Recuperação da informação a partir das linguagens de indexação	42
3 GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DE LINGUAGENS DE INDEXAÇÃO EM BIBLIOTECAS UNIVERSITÁRIAS.....	46
3.1 Linguagem de indexação na perspectiva da política de indexação em bibliotecas.....	51
3.2 Diretrizes normativas no contexto das linguagens de indexação.....	53
3.2.1 Diretrizes iniciais da International Organization for Standardization (25964, 2011)	56
3.2.2 Símbolos e siglas.....	57
3.2.3 Relação entre termos equivalentes.....	59
3.2.4 Orientações quanto à gramática	61
3.2.5 Relações entre conceitos	64
3.3 Gerenciamento da construção e manutenção de linguagens de indexação (ISO 25964, 2011).....	67
4 METODOLOGIA.....	74
4.1 Pesquisa bibliográfica.....	74
4.2 Pesquisa descritiva	79
4.2.1 Análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva da literatura.....	80
4.2.2 Análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação na perspectiva dos gestores e dos bibliotecários da USP e UNESP	81
4.2.2.1 Entrevista	82
5 RESULTADOS E DISCUSSÃO DA ANÁLISE DESCRITIVA DO GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DAS LINGUAGENS DE INDEXAÇÃO DA USP E UNESP NA PERSPECTIVA DA LITERATURA	90
5.1 Considerações a respeito do VocaUSP e Tesouro Unesp	101
6 RESULTADOS E DISCUSSÃO DA ANÁLISE DESCRITIVA DO GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DAS LINGUAGENS DE INDEXAÇÃO DA USP E UNESP NA PERSPECTIVA DOS GESTORES E BIBLIOTECÁRIOS.....	103
7 RECOMENDAÇÕES SOBRE O GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DE LINGUAGENS DE INDEXAÇÃO	128
8 CONSIDERAÇÕES FINAIS	134
REFERÊNCIAS.....	139
ANEXO 1 - E-mail para as Bibliotecas	148
ANEXO 2 - Transcrição da entrevista com G1	149
ANEXO 3 – Transcrição da entrevista com G2.....	158
ANEXO 4 - Transcrição da entrevista com In1.....	168
ANEXO 5 – Transcrição da entrevista com In2	174

1 INTRODUÇÃO

No Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP), a linha de pesquisa que este trabalho se insere é de produção e organização da informação, como o tema delimitado no gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias.

Nesta linha de pesquisa, a produção da informação refere-se à produção científica e documental; e a organização da informação abrange estudos sobre os processos de “análise, síntese, condensação, representação, e recuperação do conteúdo informacional, bem como das competências e comportamentos informacionais do usuário inerentes a tais processos” (Programa de Pós-Graduação em Ciência da Informação - PPGCI, 2022). Os processos que estão sob a organização da informação estão relacionados ao tratamento temático da informação, tendo em vista o objetivo de identificar elementos do documento, assegurar sua representação e recuperação.

Guimarães (2008) aponta que o tratamento temático da informação é um conceito que busca reunir os elementos das operações documentais e de processamento técnico da informação e, portanto, o tratamento temático “estabelece intermediação entre a coleta e a difusão de documentos”. As operações documentais para o processamento técnico, dessa forma, realizam atividades “tanto sob a ótica do suporte material - tratamento descritivo - quanto do conteúdo – tratamento temático” (Guimarães, 2008, p. 79).

Braschër e Guimarães (2018) explicam que na área da Ciência da Informação, há distinção entre a descrição física e a temática, porém muitas vezes no cotidiano esses processos são únicos e realizados de maneira simultânea. Considera-se que na prática, esses processos são realizados de forma conjunta pelo bibliotecário. Contudo, são processos que possuem suas particularidades e direcionam-se para um objetivo em comum: o catálogo.

Por isso, pensando na abordagem prática e no cunho empírico desta pesquisa, entende-se que o tratamento temático da informação é capaz de traduzir os objetivos aqui propostos, ressaltando o direcionamento para o assunto/tema de recursos documentais.

A indexação constitui-se de processos influenciados por contextos socioculturais e que devem ser considerados tanto na prática quanto na teoria da área, cooperando para o estabelecimento de critérios pragmáticos (Hjørland, 2018). Sobre isso, no processo de indexação, realiza-se a análise de assunto com a finalidade de identificar o conteúdo do documento e traduzi-los por meio de uma linguagem de indexação com uso de controle de vocabulário (Dias; Naves, 2007; Souza; Hillesheim, 2014). A linguagem de indexação é

instrumento de controle de vocabulário utilizado no processo de indexação e a sua composição, ou seja, o modo como é constituído recebe influência do meio social e cultural que está inserido. Desta forma, é fundamental relacionar as contribuições teóricas e normativas com as práticas e necessidades organizacionais que utilizam a linguagem.

O gerenciamento da construção e manutenção de linguagens de indexação é um conceito presente na norma ISO 25964 (International Standardization Organization, 2011) que se refere ao planejamento para a elaboração de tesauros. Há orientações para gerenciar o desenvolvimento dessa linguagem de indexação e como os responsáveis devem mantê-la (Clarke, Zeng, 2012). Sua manutenção envolve atualização periódica e avaliação dos descritores presentes em sua estrutura. Apesar de abranger orientações a respeito dos tesauros, contém informações importantes para o desenvolvimento de outros tipos de linguagem de indexação que podem ser aproveitadas.

A linguagem de indexação é uma ferramenta utilizada para promover o controle de vocabulário e a organização e representação da informação. Segundo Lancaster (2004, p. 19) as linguagens de indexação atuam como vocabulários controlados e constituem, em sua essência, uma lista autorizada de termos que inclui uma “forma de estrutura de semântica”. Ressalta-se que na literatura existem termos como linguagem de indexação, vocabulário controlado, linguagem documentária e sistemas de organização do conhecimento, por exemplo, que podem ser considerados equivalentes (Barité, 2011; Fujita; Santos; Alves, 2018).

Segundo Barité (2011) nos últimos tempos muitas nomenclaturas surgiram para designar ferramentas na representação e recuperação temática, sendo influenciadas por suas correntes teóricas; termos como linguagem documentária ou linguagem documental, são frutos da corrente francesa; termos representados por uma corrente anglo-saxônica utiliza as expressões linguagem de indexação e vocabulário controlado que são utilizados para identificar ferramentas específicas de indexação. O termo sistema de organização do conhecimento surgiu como uma alternativa para contemplar todos os esquemas para organização da informação e sua gestão, incluindo sistemas de classificação, cabeçalho de assuntos, autoridade de assuntos, redes semânticas e ontologias (Hodge, 2000). Percebe-se que o termo “guarda-chuva”, sistema de organização do conhecimento, abrange além de ferramentas de indexação, mas também aquelas utilizadas na classificação de assuntos. Fujita, Santos e Alves (2018) concluem que os termos linguagem documentários, linguagem de indexação e sistema de organização do conhecimento podem ser considerados equivalentes e assim como Barité (2011) demonstram a influência da vertente teórica de cada conceito.

Em seu livro, Van Slype (1991) aponta que a linguagem de indexação trata especificamente da representação do conteúdo dos documentos e da busca documental para viabilizar a recuperação da informação. Com isso, o termo definido para essa pesquisa será linguagem de indexação, considerando a delimitação do termo coerente com a linha desta pesquisa, isto é, compreende-se que o termo é direcionado ao processo de indexação, busca e recuperação em ambientes de informação, tal qual as bibliotecas universitárias. Mesmo assim, não serão excluídos trabalhos que utilizam outros termos para se referirem ao mesmo conceito.

A construção de linguagens de indexação corresponde aos procedimentos estabelecidos por meio da literatura normativa, considerando os aspectos e particularidades da instituição, como perspectivas culturais (Biscalchin; Moreira, 2020). Segundo Rio-Branco (2020) a atualização das linguagens de indexação deve ser constante, pois são instrumentos vivos que derivam da linguagem natural, ou seja, são ferramentas responsáveis pela representação das palavras do autor sob uma linguagem técnica, passível de controle sintático e de recuperação.

A autora também enfatiza que todo o processo que envolve a construção e a manutenção das linguagens de indexação é complexo, pois para ser possível a abrangência significativa de termos que representem o acervo de uma biblioteca universitária, é preciso a linguagem constituir-se de um vocabulário estruturado, consistente e amplo nas áreas do conhecimento, dotado de exaustividade e especificidade.

O gerenciamento dessas atividades por uma equipe de pessoas especializadas e comprometidas é fundamental para o cumprimento dos objetivos inicialmente delimitados, além de sua permanente atualização.

A norma mais atualizada que envolve a construção de tesouros para a recuperação da informação e a capacidade de interação com outras ferramentas de controle de vocabulário utilizando a interoperabilidade é a ISO 25964, “*Information and documentation: Thesauri and interoperability with other vocabularies*”. Esta norma divide-se em duas partes, a primeira parte “*Thesauri for information retrieval*”, publicada em 2011; e a segunda parte, o “*Interoperability with other vocabularies*”, publicada em 2013.

Em especial, na primeira parte da publicação da ISO 25964-1, no capítulo 13, é abordado o gerenciamento da construção e manutenção de tesouros, seguindo especificações como o planejamento, primeiros estágios da compilação, a construção em si, a introdução da ferramenta, disseminação e atualização. Será tratado adiante maiores desdobramentos a respeito dessa norma constituída de orientações para o desenvolvimento de uma linguagem de indexação.

Além da base normativa para a construção de uma linguagem de indexação, outro componente preciso a ser contemplado neste processo é a política de indexação. A política de indexação é um documento formalizado e elaborado pela biblioteca ou sistema de informação para definir elementos e procedimentos que devem ser norteadores na indexação. Destaca-se a importância da política de indexação para as bibliotecas universitárias, tendo em vista uma característica muito comum entre as bibliotecas universitárias brasileiras, apresentarem várias unidades contempladas por um mesmo catálogo online ou repositório institucional. Isso significa que o tratamento da informação é realizado por diferentes indexadores. Através de um manual contendo a política de indexação da instituição é possível buscar um alinhamento nas tomadas de decisão que envolvem a indexação. Além da formalização dos procedimentos adotados na indexação de assuntos em uma política de indexação registrada em um manual, considera-se que a capacitação contínua dos bibliotecários atuantes na indexação seja fundamental para esse alinhamento e deve ser realizado com certa periodicidade ou sempre que houver necessidade, como com a presença de novos indexadores. Com isso, a uniformidade dos procedimentos contribuirá para a padronização e consistência dos catálogos e, consequentemente, para a recuperação da informação.

Destaca-se que na política de indexação são definidas variáveis como níveis de exaustividade, especificidade, revocação e a linguagem de indexação. São aspectos qualitativos da política que devem ser estabelecidos de acordo com as demandas e determinações do sistema de informação (Fujita, 2016).

As linguagens de indexação representam uma parte fundamental da política de indexação, além do papel mediador na organização e recuperação da informação (Fujita; Cruz; Patrício; Rio-Branco, 2019). Isso pois, a linguagem de indexação é responsável por representar toda a política de indexação, tendo em vista que é a partir das definições previamente estabelecidas que afetará a maneira de trabalhar dos indexadores, na quantidade de termos para a representação dos documentos, a especificidade a ser aplicada, os procedimentos realizados quando não são encontrados os termos na linguagem utilizada, e principalmente, a definição de qual utilizar.

Agora, é importante descrever as conclusões e reflexões destacadas a partir das pesquisas que foram desenvolvidas anteriormente e que têm relação com a problemática desta tese. Desde a Iniciação Científica, no decorrer da graduação em Biblioteconomia (Cruz, 2015, 2017, Cruz; Santos; Fujita, 2016, Fujita; Cruz; Patrício, 2017, Cruz; Santos; Fujita, 2017, Fujita; Cruz; Patrício; Rio-Branco, 2019), até a Dissertação de Mestrado (Cruz, 2019, Cruz; Fujita, 2021), as pesquisas desenvolvidas estiveram inseridas no âmbito de dois projetos financiados

pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), inicialmente o “Política de indexação para bibliotecas” (Fujita, 2010) e, posteriormente, o “Linguagem de indexação para bibliotecas na perspectiva da política de indexação” (Fujita, 2015), ambos sob coordenação da Professora Doutora Mariângela Spotti Lopes Fujita (Fujita; Moreira; Santos; Cruz; Ribas, 2018, Fujita; Cruz; Patrício; Rio-Branco, 2019, Cruz; FERNEDA; Fujita, 2022, Tolare; Cruz; Fujita, 2022).

O início das investigações envolvendo o modo do emprego das linguagens de indexação pelas bibliotecas universitárias, teve como campo de estudo as regiões Sul e Sudeste do Brasil, para um diagnóstico preliminar. Contou com a aplicação de questionário e a participação de 60 bibliotecas. A partir dos resultados das análises realizadas, pode-se destacar as seguintes características nestas regiões brasileiras: várias bibliotecas não utilizam linguagem de indexação no processo de indexação, apenas linguagem natural, ou seja, sem controle de vocabulário; ou mesmo o uso combinado de linguagem de indexação com a linguagem natural. É recorrente o uso de mais de uma linguagem de indexação para representação dos recursos informacionais do local. Tanto a prática de combinação de linguagem natural com linguagem controlada ou o uso de várias linguagens de indexação podem levar a inconsistência no catálogo da biblioteca pela falta de controle de vocabulário (Cruz, 2019).

Das respostas analisadas, foi possível observar que todas as bibliotecas realizam algum tipo de procedimento quando o termo preferido não é encontrado na linguagem de indexação utilizada. Porém, o método de inclusão dos termos na descrição do conteúdo dos documentos não apresenta procedimentos padronizados, considerando a falta de pesquisa, critérios e análise para a inclusão, evidenciam-se os seguintes procedimentos, de acordo com a incidência de casos:

- 1) realizam pesquisas antes de adicionar os termos novos (19 bibliotecas);
- 2) adicionam novos termos por meio da linguagem natural, sem controle de vocabulário (17 bibliotecas);
- 3) adequam a termos gerais já existentes (14 bibliotecas);
- 4) adequam e adicionam termos novos por meio de estudos realizados em grupo (11 bibliotecas) (Fujita; Cruz; Patrício; Rio-Branco, 2019).

Em sequência, constatou-se que a maioria das participantes relata apresentar um manual de política de indexação, sendo que o documento aponta a linguagem de indexação utilizada na representação de seus documentos. Vinte bibliotecas que relataram utilizar uma linguagem de indexação construída pela própria instituição, descreveram recursos que não compreendem

softwares, como linguagem de programação, programas para gestão de bibliotecas, rede de catalogação cooperativa e linguagens de indexação.

Para finalizar esta síntese dos estudos realizados nas regiões Sul e Sudeste, destaca-se que há bibliotecas buscando aprimoramentos no tratamento temático da informação ao apresentar projeto para a elaboração de uma linguagem de indexação própria, alguns em andamento, além da política de indexação como norteadora no processo de indexação, a linguagem utilizada. Todavia, há falta de compreensão dos bibliotecários considerando a forma de uso e de como proceder com a atualização das linguagens de indexação, e o desconhecimento da função da ferramenta para controle de vocabulário (Fujita; Cruz; Patrício; Rio-Branco, 2019).

A respeito disso, em uma outra perspectiva, Sunny e Angadi (2018) expõem um estudo realizado sobre avaliação da efetividade do tesouro em sistemas digitais de recuperação da informação. Apesar de não ser no contexto universitário, o objetivo dos sistemas de recuperação da informação pode ser considerado essencialmente similares ao de uma biblioteca universitária. Ao realizarem a avaliação em 15 sistemas, os autores consideraram pertinentes o uso de tesouros na indexação, busca e navegação de sistemas de recuperação de informações digitais. Tendo em vista que um tesouro pode melhorar significativamente a eficácia da recuperação de documentos e também no processo de busca realizada pelo usuário ao utilizar os termos de um tesouro, desde que sejam devidamente orientados sobre seu uso.

Na Dissertação de Mestrado foi ampliado o escopo de pesquisa para bibliotecas universitárias de todo o território nacional, resultando em um estudo analítico sobre o uso de linguagens de indexação no Brasil. Com o mesmo método anteriormente utilizado, através da aplicação de questionário, obteve-se 46 respostas que foram examinadas através de categorias de análise elaboradas a partir dos objetivos específicos estabelecidos na pesquisa e das respostas obtidas pelos participantes. Um componente fundamental para a melhora no desempenho da recuperação da informação que foi inserido nesse estudo, foi a disponibilização da linguagem no catálogo *online* para o usuário (Cruz, 2019).

Desta forma, realça-se os seguintes resultados da Dissertação de Mestrado, apresentados no artigo Cruz e Fujita (2021), a grande parcela das bibliotecas universitárias brasileiras utiliza linguagem de indexação, sendo que metade delas não disponibilizam a linguagem à disposição do usuário, o que pode causar imprecisão no momento da busca, por incompatibilidade entre as linguagens utilizadas, a do sistema e a do usuário.

No artigo de Cruz, Fernalda e Fujita (2022), a questão da disponibilização da linguagem no catálogo *online* ao usuário foi aprofundada. Com isso, verificou-se 18 bibliotecas que

afirmam fazer essa disponibilização da ferramenta, foi identificada a presença do recurso em 16 catálogos *online*, sendo que em duas delas não foi encontrado.

Em sequência, pela natureza das bibliotecas universitárias que são compostas de acervos diversificados com vários cursos e especialidades, observou-se que mais da metade utiliza mais de uma linguagem no processo de indexação. Contudo, isso dificultaria a busca e consulta do usuário, pois seria necessário consultar diferentes linguagens de indexação para recuperação da informação. Ainda, grande parte das bibliotecas utiliza linguagem natural quando não são encontrados os termos correspondentes na linguagem de indexação utilizada, de forma mais detalhada, os seguintes resultados foram obtidos em relação a essa realidade:

- 1) adicionam novos termos por meio da linguagem natural, sem controle de vocabulário (18 bibliotecas);
- 2) realizam pesquisas antes de adicionar os termos novos, em outras linguagens (13 bibliotecas);
- 3) adequam a termos gerais já existentes (7 bibliotecas);
- 4) adequam e adicionam termos novos por meio de estudos realizados em grupo (6 bibliotecas) (Cruz; Fujita, 2021).

Com as sete bibliotecas que já possuem linguagem de indexação própria, foi visto que quatro delas souberam relatar o *software* para a gestão e manutenção de suas linguagens, o TemaTres.

Tendo em vista, a relação dos procedimentos adotados quando não são encontrados os termos na linguagem de indexação utilizada, o uso de linguagem natural, estar em uma rede ou sistema de bibliotecas, o uso de várias linguagens de indexação e também não disponibilizar a linguagem ao usuário, estima-se que muitas bibliotecas se beneficiariam se houvesse a possibilidade de usufruírem de uma linguagem de indexação própria e condizente com a filosofia institucional a qual pertencem. Além disso, podemos citar como benefícios:

- Disponibilizar uma estrutura de termos condizentes com as necessidades de informação, ou seja, uma linguagem especializada que abrange os assuntos tratados pela coleção da biblioteca e/ou pesquisados na instituição.
- Contribuir com o trabalho do bibliotecário indexador, trazendo elucidação sobre qual caminho seguir na atribuição de descritores;
- Ter controle sobre os termos que são inseridos e excluídos (a partir de justificativa e embasamento);
- Linguagem atualizada;

- Recuperação da informação a partir da busca inteligente com termos autorizados.

1.1 Delimitação do problema

A partir do levantamento dos resultados de pesquisas anteriores, destacam-se os seguintes pontos que constituem nossa problemática:

- As bibliotecas universitárias, em sua maioria, não possuem procedimentos pré-definidos por uma convenção ou grupo específico, em relação às decisões a serem tomadas pelos indexadores quando não são encontrados os termos específicos na ferramenta utilizada, estima-se que haja uma lacuna no manual de política de indexação sobre essa questão.
- Desconhecimento do *software* utilizado para o gerenciamento da construção e manutenção das bibliotecas universitárias que possuem linguagem de indexação própria ou daquelas que estão com projetos para a elaboração. Isso reflete que as bibliotecas não realizam a manutenção estratégica da ferramenta, pois desconhecem seu programa.
- A não disponibilização da linguagem ao usuário, podendo ser pelo uso de diversas linguagens para o tratamento da informação, contribui para a falta de compatibilidade entre a linguagem do usuário e da utilizada para a representação dos documentos.

Ao reunir os resultados de Fujita, Cruz, Patrício e Rio-Branco (2019) e Cruz e Fujita (2021) conclui-se que vinte e seis bibliotecas responderam apresentar uma linguagem de indexação própria e sete bibliotecas estão com projetos de elaboração. Este resultado é importante para essa pesquisa, pois com esse cenário é possível perceber que as bibliotecas universitárias brasileiras buscaram estratégias e metodologias para desenvolverem essas linguagens controladas. Contudo, para viabilizar o desenvolvimento da pesquisa foram selecionadas duas instituições, das vinte e seis anteriormente participantes, que possuem procedimentos documentados e definidos sobre o gerenciamento da construção e a manutenção das linguagens de indexação: Universidade de São Paulo (USP), desde 2001 (Lima *et al.* 2006), e Universidade Estadual Paulista (UNESP), desde 2013 (Fujita, 2021).

Por isso, tendo em vista que as bibliotecas universitárias estão buscando aprimoramentos no tratamento da informação, especificamente com a elaboração de linguagens de indexação, configura-se como **questão de pesquisa**: como as bibliotecas universitárias da USP e da UNESP fazem o gerenciamento da construção e a manutenção das linguagens de indexação?

A **hipótese** desta pesquisa é de que as bibliotecas universitárias não possuem procedimentos sólidos, no que diz respeito à fundamentação dos procedimentos metodológicos para o gerenciamento da construção e manutenção de uma linguagem de indexação, especialmente com base na literatura normativa e em uma política de indexação.

A **tese** apresentada é de que a experiência de sistematização dos procedimentos que implicam a construção e o gerenciamento de linguagens de indexação das duas instituições, USP e UNESP, contribuirá para com as bibliotecas universitárias e/ou unidades de informação que desejam desenvolver uma ferramenta própria para o controle e representação de seus recursos informacionais, tendo como direcionamento a revisão de literatura e de normativas sobre o tema, reunindo experiências e metodologias já aplicadas pelas bibliotecas universitárias da USP e da UNESP.

A partir disso, propõe-se investigar o gerenciamento da construção e manutenção das linguagens de indexação das bibliotecas universitárias brasileiras, resgatando duas instituições participantes da primeira fase da pesquisa relacionadas nos artigos de Fujita, Cruz, Patrício e Rio-Branco (2019) e Cruz e Fujita (2021): USP e UNESP.

O **objetivo geral** desta pesquisa é contribuir com as bibliotecas universitárias que buscam o desenvolvimento de sua própria linguagem de indexação e também com os estudos acerca da área de organização e representação da informação em Ciência da Informação.

Em relação aos **objetivos específicos**, busca-se:

1. Realizar estudo teórico e metodológico a partir da revisão de literatura nacional e internacional sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação no âmbito da Ciência da Informação;
2. Desenvolver estudo analítico descritivo sobre gerenciamento da construção e manutenção de linguagens de indexação pelas bibliotecas universitárias da USP e UNESP.
3. Elaborar orientações sobre o gerenciamento da construção e manutenção de linguagens de indexação na política de indexação de bibliotecas universitárias, tendo como parâmetro a literatura técnico-científica e a *práxis* das bibliotecas universitárias.

1.2 Justificativa

Este trabalho final de Doutorado é uma continuação dos demais estudos realizados no âmbito da política de indexação e das linguagens de indexação em bibliotecas universitárias

brasileiras que vem percorrendo uma trajetória desde 2016 até atingir as discussões e reflexões aqui descritas.

É de conhecimento geral que novas tecnologias como a inteligência artificial estão sendo apresentadas para o nosso cotidiano, trazendo novos desafios sobre o seu uso e discussões sobre ética nas produções acadêmicas, plágio, direitos autorais e até mesmo a pauta sobre o fim de algumas profissões. O pensamento aqui presente é que teremos sempre que nos reinventar para utilizar essas ferramentas de forma ética e inteligente em nossas atividades profissionais. Em relação ao gerenciamento da construção e manutenção de linguagem de indexação por ser um trabalho árduo para os responsáveis, percebe-se que utilizar de meios que podem favorecer o processo de seleção de termos por meio da automação, por exemplo, não exime o fator humano de organização e validação dos resultados pelos usuários e utilizadores da linguagem controlada.

O gerenciamento da construção e a manutenção de uma linguagem de indexação, responsável por representar as necessidades e particularidades descritas em uma política de indexação, é uma tarefa complexa e exige grande planejamento, tempo, estudo, equipe responsável pelo desenvolvimento e, posteriormente, manutenção da ferramenta. Isso faz com que as bibliotecas possam ter dificuldades na hora de compilar todas as informações necessárias ao elaborar um projeto e mantê-lo consistente com o propósito inicial da biblioteca.

Segundo Fujita, Cruz, Patrício e Rio Branco (2019) consideram que haja desafios a serem enfrentados na representação e recuperação da informação com foco na compreensão do profissional responsável pelo uso da linguagem no processo de indexação.

Estudos sobre uso de linguagens de indexação em bibliotecas não são usuais na literatura porque bibliotecas estão mais associadas ao tratamento descritivo pela catalogação e ao tratamento temático da catalogação de assuntos com uso de listas de cabeçalhos de assunto (Fujita; Cruz; Patrício; Rio-Branco, 2019, p. 219).

Contudo as autoras ressaltam ainda que o cenário das linguagens de indexação em bibliotecas universitárias além dos avanços que estão presentes, abrange também problemas e lacunas desafiadores aos profissionais e pesquisadores da área de Organização e Representação do Conhecimento (Fujita; Cruz; Patrício; Rio-Branco, 2019).

Por isso, percebe-se necessidade de continuar as investigações a respeito das linguagens de indexação na perspectiva da política de indexação, em particular direcionada às especificidades contidas em uma biblioteca universitária, como por exemplo, a quantidade de áreas do conhecimento que o acervo abrange, diferentes unidades com diversos indexadores, os usuários que são atendidos e atualização recorrente da ferramenta, tendo em vista os avanços

da tecnologia e das ciências, bem como a indexação realizada tanto para os recursos informacionais dispostos em repositórios institucionais para como os acervos físicos e digitais.

Por esses motivos, esta pesquisa justifica-se pela intenção de prosseguir com os estudos sobre o tema de linguagens de indexação e viabilizar a árdua tarefa das bibliotecas universitárias em sistematizar os procedimentos necessários para a construção de uma ferramenta consistente e que de fato contribua para a indexação de assuntos e recuperação da informação.

1.3 Organização da tese articulada com os objetivos

Para explicitar a sistematização da pesquisa, apresenta-se o Quadro 1, com as definições dos capítulos articulados com os objetivos traçados.

Quadro 1 – Sistematização da Pesquisa

ESTRUTURA DA TESE	DELIMITAÇÃO
Questão da pesquisa	Como as bibliotecas universitárias da USP e da UNESP fazem o gerenciamento da construção e a manutenção das linguagens de indexação?
Proposta	Investigar o gerenciamento da construção e manutenção das linguagens de indexação das bibliotecas universitárias brasileiras, resgatando duas instituições participantes da primeira fase da pesquisa relacionadas nos artigos de Fujita; Cruz; Patrício; Rio-Branco (2019) e Cruz e Fujita (2021): USP e UNESP.
Objetivo Geral	Contribuir com as bibliotecas universitárias que buscam o desenvolvimento de sua própria linguagem de indexação e também com os estudos acerca da área de organização e representação da informação em Ciência da Informação.
Objetivos Específicos articulados às seções	Objetivo específico 1: Realizar estudo teórico e metodológico a partir da revisão de literatura nacional e internacional, incluindo as normativas sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação. Seção 2 - Revisão de literatura Seção 3 - Gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias Seção 4: Metodologia

	Objetivo específico 2: Desenvolver estudo analítico descritivo sobre gerenciamento da construção e manutenção de linguagens de indexação pelas bibliotecas universitárias da USP e UNESP.
	Seção 4: Metodologia
	Seção 5: Resultados e discussão da análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva da literatura
	Seção 6: Resultados e discussão da análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva dos gestores e bibliotecários.
	Objetivo específico 3: Elaborar instruções sobre o gerenciamento de construção e manutenção de linguagens de indexação na política de indexação de bibliotecas universitárias, tendo como parâmetro a literatura técnico-científica e a <i>práxis</i> das bibliotecas universitárias.
	Seção 7: Recomendações às bibliotecas sobre o gerenciamento da construção e manutenção de linguagens de indexação.
	Seção 8: Considerações finais

Fonte: elaborado pela autora.

A partir dessa introdução e resgate de pesquisas anteriores, a segunda seção desta tese apresenta a revisão de literatura como resultado da pesquisa bibliográfica que está descrita na seção de metodologia. Essa seção estabeleceu categorias baseadas na análise sistemática dos textos recuperados em bases de dados nacionais e internacionais.

Na terceira seção são descritos os referenciais teóricos a respeito do “Gerenciamento da construção e manutenção em linguagens de indexação em bibliotecas universitárias”, que abrange discussões sobre novo contexto das bibliotecas em relação à indexação de assuntos e enfatiza as mudanças que ocorreram nos últimos tempos em relação ao papel do bibliotecário e as novas demandas que surgiram com o tempo. Nessa mesma seção, discute-se sobre a visibilidade que o tratamento da informação pode oferecer a um determinado tema e a relevância do uso de linguagens de indexação em um cenário de ciência aberta com periódicos e repositórios institucionais e a atribuição de palavras-chave em linguagem natural, além da indexação automática por extração de linguagem natural e por atribuição com uso de linguagens controladas e quais as suas implicações no contexto da biblioteca universitária. Como parte complementar, a subseção “3.1 Linguagem de indexação na perspectiva da política de indexação em bibliotecas” demonstra a importância de se documentar a tomada de decisões sobre os procedimentos norteadores para o trabalho do indexador e a função da linguagem de

indexação na política. Em sequência, a subseção “3.2 Diretrizes normativas no contexto das linguagens de indexação” contextualiza a trajetória histórica e evolutiva das normas internacionais existentes sobre os tesauros, linguagem de indexação enfatizada pelas diretrizes. Sobre isso, enfatiza-se a *International Organization for Standardization* (ISO 25964, 2011) com as diretrizes basilares para o desenvolvimento e manutenção de tesauros monolíngues e multilíngues, bem como a interoperabilidade semântica, voltados à recuperação da informação. A próxima subseção, “3.3 Gerenciamento da construção e manutenção de linguagens de indexação (ISO 25964, 2011)” aborda os elementos considerados essenciais da norma para estruturar o projeto e desenvolvimento de uma linguagem de indexação.

Na metodologia, inserida na quarta seção desta pesquisa, discute-se especialmente: a pesquisa bibliográfica, com as estratégias de buscas utilizadas nas bases de dados científicas, sendo uma delas nacional e as outras internacionais. Os resultados nas bases de dados foram evidenciados por meio de quadros para explicitar as estratégias utilizadas e os textos recuperados em cada base de dados; a pesquisa descritiva, em que se buscou descrever como foi a análise descritiva proposta sobre o gerenciamento da construção e manutenção das linguagens de indexação da USP e da UNESP, na perspectiva da literatura, dos gestores e bibliotecários. Logo após, destaca-se a técnica de entrevista utilizada para averiguar a experiência dos bibliotecários que já trabalharam com a construção e manutenção de linguagens de indexação ou que utilizam a linguagem construída no seu cotidiano para a descrição dos assuntos de seus recursos informacionais.

Na quinta seção intitulada “Resultados e discussão da análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva da literatura” e na sexta, “Resultados e discussão da análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva do gestor e do bibliotecário” são discutidas de forma categorial os resultados obtidos a partir da análise da literatura e entrevistas realizadas, respectivamente. São apresentados os objetivos de cada categoria, a exposição das informações levantadas através da análise dos textos publicados sobre o tema e por meio de trechos retirados das transcrições das entrevistas realizadas que complementam as discussões, respectivamente.

A sétima seção cumpre o último objetivo específico da pesquisa em elaborar “Recomendações às bibliotecas sobre o gerenciamento da construção e manutenção de linguagens de indexação”, onde reúne os principais aspectos considerados fundamentais que auxiliam as bibliotecas universitárias para executar suas linguagens de indexação. Na última

seção, “Considerações finais”, foram apresentadas as principais conclusões levantadas pelos resultados obtidos, restrições e êxitos constatados na pesquisa.

2 REVISÃO DE LITERATURA

A pesquisa bibliográfica desta tese foi sistematicamente organizada em categorias de acordo com os assuntos identificados nos textos recuperados por meio da estratégia de busca, que está descrita na seção de metodologia. Em suma, essas categorias foram criadas a partir da revisão e análise da literatura de obras estrangeiras e nacionais.

Ressalta-se que há trabalhos que são discutidos em mais de uma categoria na busca de aproveitar a visão desses autores nos demais assuntos que são discutidos em suas publicações.

Abaixo são apresentadas essas categorias que se configuram em: 1) Procedimentos de implementação da linguagem de indexação: relatos de experiências; 2) O uso de linguagens de indexação por bibliotecas universitárias; 3) Capacidade de interoperabilidade das linguagens de indexação; 4) Recuperação da informação a partir das linguagens de indexação.

Quadro 2 - Articulação das categorias com os textos adotados

Categoria	Textos utilizados
1) Procedimentos de implementação da linguagem de indexação: relatos de experiências	FOSTER, A. K. A Controlled vocabulary for an electronic resources problem reporting system creation, implementation and assessment. Library resources & technical services, Chicago, v. 65, n. 1, 2021. SANTOS, L. B. P.; FUJITA, M. S. L. A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. Scire: representación y organización del conocimiento, Zaragoza, v. 22, n. 2, 2016. FUJITA, M. S. L.; MOREIRA, W.; SANTOS, L. B. P.; CRUZ, M. C. A.; RIBAS, R. R. B. Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: An experience of the Sao Paulo state university. Knowledge Organization, Wurzburg, v. 45, n. 3, p. 220-231, 2018. CHO, J. H. Cataloging for a celebration: metadata for an institutional repository from the ground up. Cataloging and Classification Quarterly, Philadelphia, v. 60, n. 2, p. 141– 163, 2022. KLOSE, A. C; GOLDSTEIN, S.; LEVY, M. S. Numismatics & Bibliographic Description: how Rutgers University Libraries described coins with MODS. Journal of Library Metadata, Philadelphia, v. 22, n. 1-2, p. 75-104, 2022. HO, J.; STOKES, C. Core metadata element recommendations for institutional repositories at Texas A&M university libraries. Journal of Library Metadata, Philadelphia, v. 19, n. 3-4, p. 187-214. THENAGLIA, G. Diseño de tesauros como estrategia didáctica para fortalecer su comprensión. Palabra Clave, Barcelona, v. 12, n. 1, 2022.
2) O uso de linguagens de indexação por bibliotecas universitárias;	ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. Knowledge Organization, Wurzburg, v. 43, n. 2, p. 107–112, 2016.

	CAMPBELL, H. M.; DIECKMAN, C. S.; TEAL, W.; WINTERMUTE, H. E. Improving subject headings for iowa indigenous peoples. Library Resources and Technical Services, Chicago, v. 66, n. 1, p. 48-59, 2022. CRUZ, M. C. A.; FUJITA, M. S. L.; SANTOS, L. B. P. Linguagem de indexação no contexto da política de indexação: estudo em bibliotecas universitárias. In: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). Memória, tecnologia e cultura na organização do conhecimento. Recife: UFPE, 2017. p. 217- 224. v. 4 CRUZ, M. C. A.; FUJITA, M. S. L. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. Informação & Informação, Londrina, v. 26, n. 1, 2021. FUJITA, M. S. L.; CRUZ, M. C. A.; PATRÍCIO, B. O. M.; RIO BRANCO, L. B. P. R. Linguagens de indexação em bibliotecas universitárias: estudo analítico. Informação & Informação, Londrina, v. 24, n. 1, p. 190-225, 2019. TARTAROTTI, R. C. D.; FUJITA, M. S. L. Subject cataloguing & indexing in university libraries: a verbal-protocol comparative study. Scire, Zaragoza, v. 23, n. 1, p. 57-66, 2017. TOLARE, J. B.; FUJITA, M. S. L. A utilização de linguagens de indexação por bibliotecas universitárias: revisão bibliográfica. Scire: representación y organización del conocimiento, Zaragoza, v. 28, 2022.
3) Capacidade de interoperabilidade de das linguagens de indexação;	ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. Knowledge Organization, Wurzburg, v. 43, n. 2, p. 107-112, 2016. CRUZ, M. C. A.; FUJITA, M. S. L. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. Informação & Informação, Londrina, v. 26, n. 1, 2021. ACOSTA BRAVO, A.; DÍAZ JIMÉNEZ, A.; LEIVA MEDEROS, A. A. Propuesta de procedimiento para la construcción semiautomática de tesauros en bibliotecas universitarias. Anales de Investigacion, Havana, v. 17, n. 2, 2021. CLARKE, S. G. D. The information retrieval thesaurus. Knowledge Organization, Wurzburg, v. 46, n. 6, p. 439-459, 2019. MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. The support of constructs in thesaurus tools from a Semantic Web perspective: Framework to assess standard conformance. Computer Standards and Interfaces, Amsterdam, v. 65, p.79-91, 2019. MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. semantic web methodological framework to evaluate the support of integrity in thesaurus tools. Journal of Information Science, [S.l.], v. 46, n. 3, p. 378- 391, 2020. DOI: 10.1177/0165551519837195 CHETI, A.; VITI, E. Functionality and merits of a faceted thesaurus: the case of the <i>Nuovo Soggettario</i>. Cataloging & Classification Quarterly, Philadelphia, v. 61, n. 5-6, p. 708-733, 2023. GARCÍA MARCO, F. J.; LÓPEZ DEL RAMO, J. Developing an interoperable taxonomy by mapping empiric taxonomies: a case study on the websites of the associations of friends of the Way of St James. Scire, Zaragoza, v. 28, n.1 , p. 55-73, 2022.

4) Recuperação da informação a partir das linguagens de indexação.	ALEXIEV, V.; ISAAC, A.; LINDENTHAL, J. On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). International Journal on Digital Libraries, New York, v. 17, n.1, p. 39-48, 2016. ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. Knowledge Organization, Wurzburg, v. 43, n. 2, p. 107-112, 2016. MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. The support of constructs in thesaurus tools from a Semantic Web perspective: Framework to assess standard conformance. Computer Standards and Interfaces, Amsterdam, v. 65, p.79- 91, 2019. CRUZ, M. C. A.; FUJITA, M. S. L. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. Informação & Informação, Londrina, v. 26, n. 1, 2021. CRUZ, M. C. A.; FERNEDA, E.; FUJITA, M. S. L. A disponibilização de vocabulário controlado aos usuários para a recuperação da informação. Revista Ibero-Americana de Ciência da Informação, Brasília, v. 15, n. 1, p. 266-282, 2022. DOI: 10.26512/rici.v15.n1.2022.42464.
--	--

Fonte: elaborado pela autora.

2.1 Procedimentos de implementação da linguagem de indexação: relatos de experiências

Nesta categoria são apresentados trabalhos que identificaram os procedimentos e decisões tomadas para a implementação de uma linguagem de indexação em seus sistemas de informação. Apesar de estarem inseridas no contexto universitário, algumas relatam suas experiências com a implementação e uso em repositórios institucionais ou em acervos especiais, não necessariamente no acervo geral ou físico da biblioteca universitária. Com isso, aqui são reunidos os trabalhos de: Fujita e Santos (2016), Fujita, Moreira, Santos, Cruz e Ribas (2018), Martínez-González e Alvite-Díez (2019), Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021), Foster (2021), Cho (2022), Klose, Goldstein e Levy (2022) e Cheti e Viti (2023).

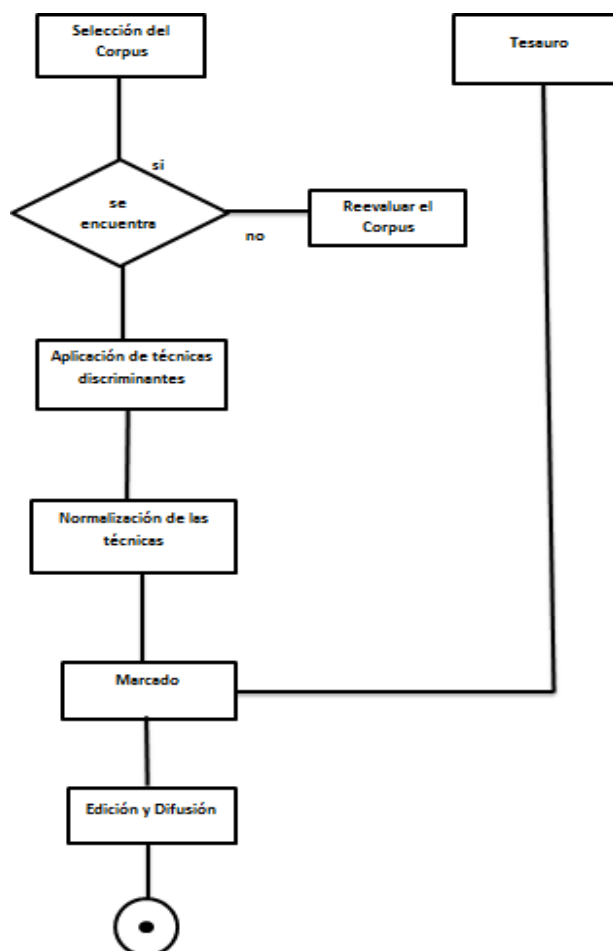
Os procedimentos de implementação de uma linguagem de indexação devem ser planejados e sistematicamente estruturados com a definição de seus objetivos. A exemplo dos estudos de Foster (2021) em que a linguagem de indexação desenvolvida possui objetivo mais específico e direcionado a uma questão em particular. O estudo se passa na Universidade Estadual de Ohio e conta com a Equipe de Recursos Eletrônicos e Seriais das Bibliotecas (ERMT), pessoal responsável pela gestão de recursos eletrônicos, também chamado de *e-resources*, onde desenvolveram uma linguagem de indexação para auxiliar no gerenciamento do “*ticketing system*”, sistema de tíquetes para relatar quaisquer problemas encontrados com os recursos eletrônicos, como livros digitais, periódicos online, bancos de dados digitais, etc. Esses

problemas podem ocorrer em várias áreas, como dificuldades de acesso aos recursos, erros ou imprecisões na catalogação (ou seja, na organização e descrição dos itens disponíveis) e falta de abrangência ou completude na cobertura dos acervos digitais. Por isso, foi desenvolvido uma linguagem de indexação, no formato de lista para descrever esses problemas específicos do ambiente de *e-resources*.

Para isso, a autora define os principais pontos de trabalho para implementação da ferramenta: Retirada dos termos a partir de relatórios (1), onde foi conduzida uma análise minuciosa, seguida da definição precisa de cada termo com base em notas e referências pertinentes (2). Em seguida, ocorreu o período de avaliação, em que o vocabulário foi disponibilizado para uso, sendo o primeiro ano de sua utilização sem alterações (3). Por fim, os termos foram submetidos à avaliação criteriosa por parte da equipe responsável e também por especialistas no campo, buscando garantir a qualidade e a adequação do vocabulário considerando as correções que identificaram nesse período (4).

O trabalho dos autores Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021) realizam uma proposta de construção semiautomática de tesauros para as bibliotecas universitárias associadas a Rede TIC do projeto VLIR, em Cuba, utilizando estruturas já existentes, linguagens de indexação já utilizadas pelos bibliotecários das unidades pesquisadas. Para isso o seguinte diagrama foi seguido:

Figura 1 - Diagrama do fluxo das etapas



Fonte: Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021).

As etapas foram detalhadas conforme as apresentadas no diagrama. Com base nisso, podemos destacar que na seleção do *corpus* (1) deve ser considerada principalmente sua representatividade em relação a sua área de domínio, esta é uma parte fundamental pois ela sustentará o tesauro e recomenda-se que seja escolhido um *corpus* já existente caso a ferramenta seja construída desde o início; em técnicas discriminantes (2) é tratado o controle de léxico, em que os autores apontaram o programa chamado *AntConc* para a “análise léxica, tratamento de palavras vazias, tratamento de termos flexionados, tratamento de palavras compostas e filtragem de termos (Acosta Bravo; Díaz Jiménez; Leiva Mederos, 2021, p. 67). Mesmo assim, na proposta de implementação de tesauro desenvolvida pelos autores, não foram usadas todas essas técnicas, mas foram utilizadas: na organização dos conceitos a lei de Zipf, para determinar a frequência das palavras no corpus; e na eliminação de palavras vazias, o *Term Frequency — Inverse Document Frequency* (TF-IDF) para obtenção de termos com melhor qualidade linguística.

A discriminação dos termos resultou na formação de *clusters* (conjuntos, em português). Utilizou-se o método de N-Grams para extrair o significado das estruturas linguísticas, dividindo os termos em grupos homogêneos com base em suas características morfológicas e linguísticas, facilitando a obtenção de frases compostas para o tesouro.

Na fase seguinte (3), de normalização das técnicas foram sugeridas algumas regras para padronizar os termos candidatos, visando facilitar o processo de edição, em que ocorre a normalização das entradas, formas gramaticais e seleção dos tipos de descritores. A penúltima fase nomeada de “mercado” (4), abrange duas subfases: a primeira refere-se a organização e adoção de esquema para realizar um conjunto de tarefas de inferência semântica e deduzir logicamente que o conhecimento está correto; a segunda subfase, “mercado para o tesouro” em que se apresentam os conceitos em SKOS dos termos que compõem um tesouro. Eles se distinguem pelas categorias de significado e sua interação com os objetos e propriedades correspondentes aos termos usados para rotulá-los. A partir disso se estabelecem as relações entre os termos, sendo de equivalência, associativas e hierárquicas. A última fase (5) sobre edição e difusão foi empreendido o VocBench

Por fim, Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021) concluem que esse estudo pode contribuir para projetos para a construção semiautomática de tesouros, destacando os mecanismos de controle e inferência semântica associados ao SKOS na eliminação de erros de sinonímia, polissemia, exaustividade e correção lexical. Além de estabelecer procedimentos para sanar, em certa medida, questões que antes prejudicavam a interoperabilidade semântica.

Em continuidade, nessa categoria de análise da literatura, para além dos procedimentos de implementação da linguagem de indexação, destacamos também a atualização/manutenção e como esse aspecto foi abordado nos textos recuperados. Sobre isso, Fujita e Santos (2016) destacam os procedimentos realizados na adaptação e compatibilização de três linguagens de indexação para o desenvolvimento do Tesouro Unesp, no texto ainda com o nome inicial de Linguagem Unesp. As linguagens aplicadas foram: a Terminologia de Assuntos da Biblioteca Nacional (TBN), o Library of Congress Subject Headings (LCSH) e o Medical Subject Headings (MeSH). Ressalta-se que essas linguagens apesar de distintas apresentam a mesma raiz terminológica.

Pensando na construção e atualização da ferramenta, tanto em Fujita e Santos (2016), como em Fujita, Moreira, Santos, Cruz e Ribas (2018) há a indicação que a sua atualização depende da organização da estrutura hierárquica para viabilizar classificação e inserção de termos novos.

Por isso, a proposta inicial de Fujita e Santos (2016) foi realizar um estudo nas macroestruturas da linguagem LCSH, Terminologia da Biblioteca Nacional, Linguagem do BIBLIODATA e no Vocabulário Controlado da USP, para as áreas de física e matemática.

Sendo assim, a partir da comparação entre essas macroestruturas, pode-se verificar que a linguagem do BIBLIODATA por não estar sendo atualizada e não apresentar em sua estrutura hierarquização dos termos, foi retirada das linguagens consideradas para a proposta de macroestrutura do Tesouro Unesp. O Vocabulário USP se destacou por sua estrutura hierárquica, havendo concordância que ela seria importante para a macroestrutura da Linguagem UNESP, especialmente por estar igualmente no contexto de bibliotecas universitárias. Também a LCSH, juntamente com a TBN (tendo em vista que esta é derivada da LCSH, mas não tão diversificada) pela “exaustividade e diversidade terminológica, oferecendo opções para o trabalho do indexador, de modo que este possa atender às necessidades mais atuais de interdisciplinaridade científica” (Fujita; Santos, 2016, p.45).

Essa tarefa de compatibilização de linguagens implica em prévios estudos por combinar diferentes estruturas e seus contextos que refletem tanto em suas estruturas conceituais como nas hierárquicas.

Em continuidade aos estudos iniciados em Fujita e Santos (2016), Fujita, Moreira, Santos, Cruz e Ribas (2018) apontam os próximos passos tomados para a construção e atualização do Tesouro Unesp, que foram definidas categorias e subcategorias com base nas linguagens mencionadas acima. Isso possibilitou a avaliação da estrutura das linguagens utilizadas na organização de termos superordenados e subordinados, o que contribuiu para a operacionalização de procedimentos do manual da linguagem de indexação para catálogos online em bibliotecas.

Esse mesmo esquema de categorização das macroestruturas foi utilizado por Cheti e Viti (2023) na implementação de um tesauro denominado “*Nuovo soggettario*”, coordenado pela Biblioteca Nacional Central de Florença, diferentemente dos trabalhos apontados anteriormente em que o ambiente da linguagem de indexação é universitário (Foster, 2021; Fujita; Santos, 2016; Fujita; Moreira; Santos; Cruz; Ribas, 2018). Sendo assim, a esquematização do tesauro em macroestrutura de forma hierarquizada promoveu o agrupamento de conceitos homogêneos. O Tesouro do *Nuovo Soggettario* seguiu as recomendações da ISO 25964 (2011, 2013), assim como apontado por Fujita, Moreira, Santos, Cruz e Ribas (2018) no Tesouro Unesp.

Sobre a norma citada, Martínez-González e Alvite-Díez (2019, p. 18) afirmam que a ISO 25964 de 2011 e 2013, representa um “renascimento” no que diz respeito às ferramentas

de controle de vocabulário, como os tesauros. Além de substituir a diretriz anterior, a ISO 2788, publicada em 1974, reúne novos elementos como a Web Semântica e o conceito de interoperabilidade semântica.

Outro trabalho recuperado, mas que não mostrou os procedimentos para a elaboração de uma linguagem de indexação e sim apontou o caminho para a tomada de decisões para definir uma linguagem já estabelecida foi Cho (2022) que apresenta o panorama das Bibliotecas da Universidade Adelphi, em seu Repositório Institucional, e que a partir de estudos e testes realizados decidiram adotar a *Library of Congress Subject Headings* (LCSH) como linguagem de indexação, devido a sua flexibilidade e universalidade. O trabalho de Klose, Goldstein e Levy (2022) evidencia o trabalho das Bibliotecas da Universidade Rutgers que desenvolveram uma coleção digital de moedas romanas antigas, no Repositório Comunitário da Universidade Rutgers, utilizando o Padrão de Descrição de Objetos de Metadados (MODS) para metadados descritivos.

Em relação à descrição desse material, foi criado um vocabulário controlado buscando consistência com os termos e a ortografia usados na comunidade numismática. Esse projeto começou com listas na documentação das melhores práticas do projeto, definindo os termos preferenciais e não preferenciais. Contudo não foi utilizado nenhum *software* de gestão de linguagem específico a essa tarefa, mas se estabeleceu uma equipe responsável por sua manutenção. Os autores apresentaram que o padrão de metadados bibliográficos é uma decisão importante para a descrição desejada dos recursos informacionais. Sobre isso, Ho e Strokes (2019) destacam a importância da compatibilidade entre os padrões de metadados para que seja possível a interoperabilidade entre linguagens de indexação que estão sendo desenvolvidas continuamente. O padrão de metadados citado como exemplo pelas autoras foi o Dublin Core.

Apesar de não ser um relato de implementação de uma linguagem de indexação em um sistema de informação, podemos destacar o trabalho interessante de Thenaglia (2022) que faz uma proposta de atividades aos seus estudantes de biblioteconomia, seguindo a norma ISO 25.964. Para isso, o autor inicialmente contextualiza que a ISO 25964 é uma versão atualizada que substitui a ISO 2.788 (1986) e a ISO 5.964 (1985). Expõe com mais ênfase a primeira parte da norma, a seção 13, que como já visto anteriormente se concentra na gestão da construção e manutenção de tesauros, dividindo-se em seis categorias: planejamento do tesouro, compilação de termos, construção por meio das hierarquias, introdução do tesouro, difusão e a atualização da ferramenta que compreende uma etapa contínua para o grupo gestor. Para atingir seu objetivo, o autor explica que o projeto almejado foi a elaboração de um tesouro tendo como um domínio escolhido por cada um, sendo: gírias, fotografia, confeitaria,

cartografia, basquete e ginecologia. A partir disso, os alunos selecionaram de 3 a 5 documentos específicos sobre essa temática, como material de referência ao trabalho proposto. O experimento em sala de aula contou com a participação de quinze alunos e teve uma duração de seis semanas, estipuladas no planejamento de atividades, além de seguir a seção 13 da ISO 25964 como guia para o desenvolvimento do tesouro. Como resultado observou-se que a familiaridade e preferência por determinados temas escolhidos pelos alunos, facilitou a extração e organização dos termos, tornando o processo mais intuitivo e assertivo. Isso ressalta o conhecimento prévio do domínio temático na eficácia do trabalho de projeto e confecção de tesouros. Além disso, a análise revelou que, embora os alunos já tivessem familiaridade teórica e prática com os tesouros, a construção de um tesouro consolidou e destacou fraquezas em alguns conceitos previamente adquiridos, como evidenciado pela recusa em usar termos não autorizados e a falta de relações de equivalência na estrutura do tesouro, em alguns casos. Isso sugere uma necessidade de melhor compreensão dos conceitos relacionados ao vocabulário e às relações de equivalência entre os termos apresentados nos tesouros por alguns alunos (Thenaglia, 2022).

Apesar de ser um experimento em sala de aula, podemos trazer esses fatos para uma linguagem de indexação desenvolvida no contexto da biblioteca universitária, tendo em vista a importância da presença de especialistas que possam auxiliar o trabalho do bibliotecário ou estarem presentes nos grupos de desenvolvimento e manutenção da linguagem de indexação para serem consultados. Treinamentos e capacitações da equipe responsável, como dito anteriormente, podem fazer com que o conhecimento prévio desses indivíduos sobre essas ferramentas seja consolidado e mais eficiente. Especialmente para servirem de referência sobre os procedimentos que devem ser adotados por demais colegas que não estão na linha de frente da gestão da linguagem, mas fazem seu uso no dia-a-dia.

Com base nas etapas elencadas pelas pesquisas, percebe-se que as bibliotecas desempenham a construção e elaboração de suas linguagens de indexação de forma diferente uma das outras. Mesmo assim, os procedimentos são realizados a partir de uma organização lógica. O estudo de Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021) buscou técnicas e programas diferenciados para a organização e validação dos termos. Mostrando que é possível utilizar diversos recursos, mesmo que sejam de uso mais recorrentes em outras áreas do conhecimento, para viabilizar a elaboração de uma linguagem de indexação. Destaca-se os trabalhos que estabeleceram macroestruturas através de categorias e subcategorias, como forma de operacionalização de procedimentos tanto para a gestão na construção da ferramenta, como

para a sua posterior atualização (Cheti; Viti, 2023; Fujita; Santos, 2016; Fujita; Moreira; Santos; Cruz; Ribas, 2018).

Apesar da categoria que discute a interoperabilidade de linguagens de indexação ser mais adiante, foi identificado também aqui a importância de serem estabelecidos sistemas inteligentes para o reconhecimento e compatibilização de padrões de metadados, visando a importação e exportação de dados através da interoperabilidade semântica (Ho; Strokes, 2019; Klose, Goldstein e Levy, 2022). Todas essas informações reforçam que é fundamental que haja um diagnóstico do processo de indexação e dos resultados obtidos através dele. Por meio disso, é possível identificar as lacunas presentes nesse âmbito e como isso pode ser resolvido ou aprimorado, assim como visto em alguns casos nesta categoria, que isso se deu pela elaboração de linguagens de indexação.

2.2 O uso de linguagens de indexação por bibliotecas universitárias

A descrição e discussão sobre o uso de linguagens de indexação no ambiente universitário foi encontrada em maior quantidade em artigos de autores brasileiros, mesmo assim nesta categoria reúnem-se trabalhos de Fujita, Boccato e Rubi (2010), Andrade e Lara (2016), Tartarotti e Fujita (2017), Cruz, Fujita e Santos (2017), Fujita, Cruz, Patrício e Rio Branco (2019), Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021), Cruz e Fujita (2021), Campbell, Dieckman, Teal, Wintermute (2022), Tolare e Fujita (2022).

Para iniciar identificar a necessidade da construção de uma linguagem de indexação própria, muitas vezes são identificados problemas no processo de indexação e na recuperação da informação. Esse foi o caso apresentado por Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021) sobre a Rede de Bibliotecas TIC do projeto VLIR, composta por diversas bibliotecas, cada uma com sua particularidade, mas sendo de natureza universitária. Os autores identificaram dificuldades na recuperação da informação, devido a inconsistências no processo de indexação, tendo em vista a falta de padronização das ações pelos bibliotecários da rede. A partir disso, identificou-se também obstáculos para a interoperabilidade semântica.

A pesquisa foi realizada por meio de entrevista e aplicação de questionário a quinze especialistas responsáveis pela indexação de sete bibliotecas universitárias da Rede TIC, de Cuba. A partir do diagnóstico realizado pelos autores, foi possível identificar concordância pelos entrevistados que a construção de tesouros enfrenta desafios significativos, devido à falta de políticas, abandono de boas práticas e lacunas na formação dos especialistas. Identificou-se também o uso de palavras-chave extraídas do texto do autor, ou seja, o uso de linguagem natural

para a indexação dos documentos, representando 29% de uso. Mesmo assim, foi visto que se priorizam as linguagens controladas, representando 71% das respostas obtidas. Outro ponto destacado são os baixos salários desses profissionais que podem se sentir não estimulados a realizar o processo de indexação com maior cuidado e atenção. Há predominância da linguagem controlada com estrutura principal alfabética nos tesauros e pouco uso de indicadores de relações hierárquicas, de equivalência e associativas. Embora os métodos tradicionais de indexação sejam amplamente utilizados, métodos mais inovadores não são comuns nas bibliotecas universitárias participantes dessa pesquisa. Também foi visto que não são empregadas todas as etapas estabelecidas pelas normas, especialmente no que diz respeito a introdução, disseminação e atualização da ferramenta.

Assim como no estudo exposto acima, Campbell, Dieckman, Teal e Wintermute (2022) também identificaram problemas na linguagem de indexação utilizada e propuseram mudanças na estrutura da linguagem. Os bibliotecários responsáveis pelos serviços de metadados da Biblioteca da Universidade Estadual de Iowa verificaram a necessidade e oportunidade de práticas de incluir em seus metadados diversidade, equidade e inclusão¹. Como resultado, foi formado o Grupo de Trabalho de Metadados DEI que identificaram lacunas na representação da informação no âmbito das coleções concernentes aos povos indígenas norte-americanos pela linguagem de indexação utilizada LCSH. Esse projeto de pesquisa relacionou alguns objetivos, como: “identificar termos inaceitáveis e os seus equivalentes culturalmente mais apropriados”; “complementar os termos antigos com os novos nos metadados da biblioteca local”; “construir relacionamentos bem-sucedidos com as nações indígenas americanas e compartilhar informações obtidas no projeto com outras bibliotecas na esperança de ajudar em trabalhos semelhantes” (Campbell; Dieckman; Teal; Wintermute, 2022, p. 49, traduzido pela autora).

Inicialmente o grupo se concentrou em grupos indígenas com ligação em Iowa e por conta da especificidade do tema, relacionou contatos de especialistas para validação de informações relacionadas ao tema, como representantes reconhecidos pelas tribos indígenas, além de apoio em recursos acadêmicos e sites oficiais indígenas. Ressalta-se que não poderia simplesmente substituir o campo da linguagem da LCSH, por isso, buscou-se criar cabeçalho de assuntos potenciais semelhantes aos encontrados no LCSH que seguiriam a mesma abordagem do termo existente, amplos e referenciados aos povos indígenas, não a entidades políticas específicas. A proposta de compilação desses termos se deu por como os povos originários se referem a si mesmos, seguidos do sufixo (Povos Indígenas da América do Norte)”

¹ Diversity, equity, and inclusion (DEI)

no lugar de “Índios”, por exemplo. A implementação de qualificadores ocorreu de forma automatizada e em lotes pelos recursos oferecidos pelo sistema Ex Libris Alma. O desafio enfrentado nesse projeto foi pela diversidade de nomes usados para se referir aos povos indígenas. Os termos da LCSH existentes não puderam ser diretamente relacionados aos nomes atualizados pela falta de correspondência direta entre os termos desatualizados e preferenciais. Além disso, alguns termos da LCSH podem se relacionar com mais de um nome preferencial, complicando ainda mais o mapeamento. A grafia dos povos indígenas também foi um desafio a ser enfrentado. Sendo assim, os registros que não puderam contar com a atualização de forma automática foram analisados e modificados individualmente, sempre com respaldo na pesquisa dos bibliotecários do grupo e, quando possível, com a orientação de especialistas. Com isso, os autores concluíram que as atualizações e melhorias são previstas e consideradas para o futuro, além de considerar que os objetivos propostos foram alcançados, pois identificaram os cabeçalhos de assuntos culturalmente mais apropriados que descreviam os povos indígenas de Iowa, desenvolveram e implementaram um processo automatizado para complementar os termos do LCSH com termos aprimorados nos metadados da biblioteca local, estabeleceram também relações colaborativas com várias nações indígenas americanas e começaram a compartilhar seu trabalho com outras bibliotecas.

Em sequência à exposição dos estudos de Campbell, Dieckman, Teal e Wintermute (2022), Tartarotti e Fujita (2017) enfatizam um outro aspecto das linguagens de indexação que deve ser considerado: o nível de consistência que ela traz ao processo de indexação. Sobre isso, as autoras afirmam que a linguagem de indexação para tradução de conceitos também é um dos fatores para a consistência da indexação, especialmente quando se dispõe de uma linguagem própria e única pela biblioteca ou rede de bibliotecas em questão. Ao considerar essa questão, Andrade e Lara (2016, p. 108) apontam os contextos em que as linguagens de indexação podem ser utilizadas, como na inserção no sistema a partir da indexação do documento, por meio de termos selecionados em campo bibliográfico específico, como o campo 650 no formato MARC21 (*Machine Readable Cataloging*); na saída do sistema de informação utilizado, como nos catálogos das bibliotecas em que especialistas e bibliotecários consultam a linguagem de indexação e selecionam termos para a recuperação de documentos; e na recuperação por hierarquias, sendo que algumas bases de dados permitem restrição na recuperação dos termos desejados e não os termos a ele subordinados (se houver).

Em um panorama brasileiro, os estudos sobre linguagem de indexação no contexto das universidades se intensificaram com o projeto de pesquisa de Fujita (2015) sobre “Linguagem de indexação para bibliotecas na perspectiva da política de indexação”. Investigações foram

realizadas para identificar o cenário do uso de linguagens de indexação como forma de controle de vocabulário em sistemas de informação com foco nas bibliotecas e que serão abordadas a seguir.

Algumas das pesquisas no espectro do projeto trouxeram uma perspectiva interessante das ferramentas de controle de vocabulário, como em Cruz, Fujita e Santos (2017) que através de um estudo preliminar realizado em cinco bibliotecas universitárias que verificaram problemas relacionados situações de quando não encontram o termo desejado na linguagem de indexação utilizada, ao uso de linguagem natural, ou seja, sem controle de vocabulário e à utilização de múltiplas linguagens de indexação. Esses problemas podem levar a inconsistências no catálogo e prejudicar a recuperação de informações por assunto devido à falta de padronização na linguagem utilizada.

Segundo Fujita, Cruz, Patrício e Rio Branco (2019) a linguagem natural é aquela utilizada no dia a dia, presente no discurso dos autores e pelos usuários quando desconhecem a linguagem de indexação para a estratégia de busca. As autoras realizaram um estudo analítico semelhante ao preliminar de Cruz, Fujita e Santos (2017) agora se ampliando para a região sudeste do Brasil com a participação de sessenta bibliotecas que responderam questões a respeito do uso da linguagem e política de indexação em seus ambientes de trabalho. Em seus resultados foi possível verificar a falta de compreensão dos bibliotecários sobre a função mediadora da linguagem de indexação, sendo que ao utilizar várias linguagens ou a linguagem natural com a controlada, não garante uma consistência do catálogo das bibliotecas. Sobre isso, as autoras reforçam que “Caso essa representação com várias linguagens não seja devidamente orientada por meio de diretrizes precisas e correspondentes, os profissionais não terão critérios definidos para o uso de linguagens de indexação” (Fujita; Cruz; Patrício; Rio-Branco, 2019).

Porém, duas observações que se destacaram neste estudo foram 20 bibliotecas, representando um terço do total da amostra, que afirmaram apresentar a própria linguagem de indexação, enquanto outras 10 tinham planos para fazê-lo. Esse dado significativo indica uma tendência na gestão e na atenção em relação ao tratamento temático no ambiente universitário. Em contrapartida nenhuma biblioteca apresentou linguagem de indexação na interface de busca para consulta dos usuários, ou seja, no catálogo online. Por isso, considera-se complexa e inviável a tarefa de fazer com que o usuário consulte mais de uma linguagem de indexação para sua busca. Sendo assim, evidencia-se a disposição das bibliotecas de construir sua própria linguagem de indexação, tendo em vista os problemas ocasionados pelo uso de várias linguagens e a necessidade de disponibilização de uma linguagem para o acesso do usuário.

A ampliação do estudo se deu mais uma vez, agora em território nacional brasileiro com o artigo de Cruz e Fujita (2021), que teve como objetivo uma perspectiva ampla do emprego das linguagens de indexação. Obtiveram a participação de quarenta e seis bibliotecas universitárias e os resultados alcançados foram semelhantes ao estudo realizado na região sudeste (Fujita; Cruz; Patrício; Rio-Branco, 2019) em relação a falta de compreensão dos bibliotecários sobre a função mediadora da linguagem, aos problemas sobre quando o bibliotecário não encontra os termos desejados na linguagem e em iniciativas que podem afetar a consistência do catálogo da biblioteca. Do mesmo modo que em Fujita, Bocatto e Rubi (2010) que a partir de sua pesquisa em nove bibliotecas universitárias brasileiras da Rede Unesp, identificaram a incompatibilidade da linguagem de indexação e problemas de recuperação por assunto do catálogo. Relatando que não era apenas um problema identificado pelos bibliotecários, mas também pelos usuários e seria necessária uma solução a partir da fundamentação e sistematização de processos no âmbito da indexação de assuntos. Evidenciou-se também em Cruz e Fujita (2021) o desconhecimento de softwares de gestão e manutenção de linguagens de indexação e a possibilidade de construção compartilhada da linguagem por meio da interoperabilidade. Um fato que se mostrou diferente foi que, desta vez, metade das bibliotecas participantes relatou disponibilizar a linguagem de indexação para consulta do usuário na interface de busca.

Tolare e Fujita (2022) realizaram um mapeamento sistemático da literatura em bases de dados nacionais e internacionais, a respeito das linguagens de indexação em bibliotecas universitárias. As autoras observaram que na maioria dos trabalhos analisados, a discussão é pautada pela influência do autor e do indexador no processo de indexação tendo em consideração a visão de mundo de cada um, assim como as diferentes interpretações de que pessoas podem ter sobre um determinado tema/assunto. Outro fator identificado que interfere no processo de indexação são as limitações ou lacunas do sistema, como falta de política de indexação e manuais de procedimentos (Fujita; Bocatto; Rubi, 2010). Destacaram os demais resultados como a importância da colaboração na indexação, a necessidade de treinamento tanto para os autores quanto para os indexadores, a fim de atribuir corretamente os termos conforme a linguagem de indexação, e destacaram a relevância da avaliação contínua e da atualização da linguagem utilizada.

A partir desses trabalhos é possível concordar com a premissa dita por Fujita, Bocatto e Rubi (2010) que há uma tendência muito grande, como visto em esfera nacional, de estabelecer um rigor científico e metodológico para alcançar especificidade e precisão tanto na análise e representação de assuntos quanto na recuperação. Tendo em vista que as bibliotecas

universitárias estão identificando os problemas de seus próprios sistemas e buscando alternativas para saná-los. Além disso, as autoras fazem um apontamento interessante ao dizerem que “Os catálogos das bibliotecas se assemelham a bases de dados, por isso o indexador deve apresentar uma postura compromissada semelhante à de um indexador que trabalha na produção dessas bases” (Fujita; Boccato; Rubi, 2010).

Sendo assim, podemos destacar alguns pontos considerados importantes da compilação desses estudos. A trajetória para a gestão da construção de uma linguagem de indexação requer um esforço preliminar em identificar as lacunas existentes no processo de indexação e consequentemente na recuperação da informação, como inconsistências na representação da informação. A partir do diagnóstico desenvolvido por meio dessa investigação, é possível fundamentar as demandas existentes para que haja um planejamento estratégico de solução para essas questões. A partir do planejamento, é possível verificar os investimentos necessários para desenvolvimento e continuidade da ferramenta. Investimentos estes referentes ao pessoal responsável, como coordenadores, especialistas para atuarem como consultores, equipe de tecnologia e etc., e eventuais valores de caixa para serem desembolsados como por exemplo para a contratação de softwares e suas atualizações. Partindo então para a elaboração de uma ferramenta sólida capaz de atender os profissionais bibliotecários e contribuir para a consistência da indexação.

2.3 Capacidade de interoperabilidade nas linguagens de indexação

A capacidade de interoperabilidade dos sistemas ou *softwares* que compõem as linguagens de indexação faz parte dos recursos existentes para a própria construção e manutenção da linguagem de indexação. Muito tem se falado sobre esse tema que se encontra nos trabalhos de Andrade e Lara (2016), Clarke (2019), Martínez-González e Alvite-Díez, (2019, 2020), Acosta Bravo, Díaz Jiménez, Mederos, (2021), Cruz e Fujita (2021), García Marco e López del Ramo (2022) e em Cheti e Vitti (2023).

Segundo Andrade e Lara (2016), os principais padrões de metadados estabelecem a interoperabilidade como uma característica fundamental das linguagens de indexação. As autoras apontam que a interoperabilidade se refere à capacidade de dois ou mais sistemas serem capazes de trocar e utilizar informações entre si, ou seja, são sistemas capazes de se comunicar e compartilhar dados. Além de usar informações que foram alteradas (Clarke, 2019). Isso é particularmente relevante em ambientes digitais dinâmicos ou mesmo em bibliotecas universitárias, nos quais os dados estão sujeitos a atualizações, revisões e modificações

constantes, tendo em vista as manutenções que devem ser realizadas nas linguagens de indexação periodicamente para que estas estejam atualizadas conforme os novos conhecimentos que surgem com as descobertas nas diversas áreas do conhecimentos e retratadas nos recursos informacionais do acervo eletrônico ou físico da biblioteca. Sendo assim, podemos dizer que a interoperabilidade promove a integração e o compartilhamento de recursos de informação.

Dessa forma, a interoperabilidade não apenas facilita a comunicação e o compartilhamento de dados entre sistemas, mas também promove a acessibilidade e a utilização eficiente de informações atualizadas e modificadas ao longo do tempo. Essa capacidade é essencial para a construção de ambientes informacionais mais integrados, colaborativos e adaptáveis às necessidades dos usuários e das organizações.

Em Clarke (2019) é relacionado principalmente às funcionalidades do tesauro, destacando aqui a capacidade de interoperabilidade dessa ferramenta, a autora destaca dois contextos principais em que ela ocorre: na integração vertical e na horizontal.

A integração vertical apresenta o tesauro embutido com *software* de indexação ou pesquisa, ou seja, onde o tesauro é incorporado a softwares de indexação ou pesquisa. Neste caso, o tesauro atua como uma ferramenta complementar, auxiliando na categorização e organização dos dados, o que facilita a recuperação de informações relevantes para os usuários. A interoperabilidade nesta perspectiva permite que o tesauro se comunique de forma eficaz com o software subjacente, garantindo uma indexação precisa e uma busca eficiente. No segundo contexto “engajamento horizontal” do tesauro com outro sistema de organização do conhecimento, normalmente sendo necessária a conversão da indexação e de expressões de pesquisa entre os idiomas presentes nas linguagens. Aqui, a interoperabilidade é permite a comunicação entre diferentes sistemas que podem ser baseados em diferentes idiomas ou estruturas conceituais. Muitas vezes, isso requer a conversão da indexação e expressões de pesquisa entre os idiomas presentes nas linguagens dos sistemas envolvidos. Essa capacidade de interoperabilidade horizontal amplia significativamente o alcance e a utilidade dos tesauros, permitindo que eles sejam utilizados em uma variedade de contextos e ambientes informacionais.

Por isso, a interoperabilidade dos tesauros desempenha um papel fundamental na facilitação da troca de informações e no aumento da eficiência na recuperação de dados. Ao permitir a integração vertical com softwares específicos e o engajamento horizontal com outros sistemas de organização do conhecimento, os tesauros se tornam ferramentas versáteis e

indispensáveis no gerenciamento e acesso à informação em diferentes contextos e ambientes digitais.

Em Cruz e Fujita (2021) referem-se ao conceito de interoperabilidade com a adição da palavra “semântica”, interoperabilidade semântica (Acosta Bravo, Díaz Jiménez e Leiva Mederos, 2021, Martínez-González e Alvite-Díez, 2019), sendo um recurso para a construção de linguagens de indexação que utiliza a fusão com outras linguagens existentes, tendo também a capacidade de sistemas autônomos de trocar ou compartilhar conteúdos conceituais, por meio de padrões semelhantes e de protocolos utilizado, tornando a organização e o acesso à informação em ambientes digitais cada vez mais complexos e interconectados.

Martínez-González e Alvite-Díez, (2019, p. 80) veem essa capacidade diretamente relacionada com a Web Semântica, considerando uma abordagem e pesquisa baseada em conceitos. Em outro estudo, reforçam também sobre as diretrizes normativas, especialmente a ISO 25964, que substitui o ISO 2788, sobre a sua evolução para tesouros baseados em conceitos e mais alinhados com a abordagem da Web Semântica (Martínez-González e Alvite-Díez, 2020).

No contexto da interoperabilidade semântica, as autoras Martínez-González e Alvite-Díez, (2019) também apontam que linguagens de indexação podem ser reutilizadas em contextos abertos e que a premissa de reutilização de dados é fortalecida. Isso significa que as linguagens de indexação podem ser compartilhadas e utilizadas por diferentes sistemas de informação, independentemente de suas especificidades técnicas ou de domínio, promovendo uma maior integração e interoperabilidade entre diferentes fontes de informação. Diante disso, pode-se dizer que as linguagens de indexação, como os tesouros, que fornecem uma estrutura semântica para a organização e a categorização de conceitos, contribuem para a recuperação mais precisa e contextualizada da informação. Em consonância, Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021) defendem o desenvolvimento de metodologias que usufruam de tecnologias atualizadas, especialmente no atual contexto da Web 3.0, para tornar os sistemas de informação capazes de comportar linguagens controladas eficientes e para que possam ter a competência de evolução cada vez mais presente.

Relacionado à essa questão, em seu relato de caso, Cheti e Viti (2023) destacam que o tesouro desenvolvido, “*Nuovo soggettario*”, está em conformidade com a IFLA *Library Reference Model* (IFLA LRM) e os princípios FAIR (*Findable, Accessible, Interoperable, Reusable*), sendo que a partir da estrutura do tesouro é possível estar em conformidade com esses princípios que são:

- Encontrabilidade (Findable): metadados são passíveis de leitura por máquinas.

- Acessíveis (Accessible): capacidade de acessar os dados e metadados por meio de consultas, aqui destacando o protocolo SPARQL e a linguagem RDF, que são padrões amplamente utilizados na recuperação de informações de bases de dados.

- Interoperáveis (Interoperable): os dados são compatíveis e podem ser integrados com outros sistemas, fontes de dados ou bancos de dados, permitindo a troca de dados de diferentes plataformas ou linguagens de indexação para facilitar a colaboração e o compartilhamento de recursos.

- Reutilizáveis (Reusable): Refere-se à qualidade dos metadados, ou seja, quando são bem descritos podem ser reutilizados ou replicados em diferentes contextos.

Considera-se que este é um diferencial da linguagem de indexação “*Nuovo soggettario*”. Não sendo pontuado em outros trabalhos a conformidade da linguagem de indexação com os princípios FAIR que são fundamentais para que os dados e metadados sejam úteis e eficazes em ambientes digitais. A respeito da interoperabilidade, seus dados abertos estão disponíveis nos formatos Zthes, MARC21 e SKOS e permitem interoperabilidade com outros sistemas de informação (Cheti, Viti, 2023).

Em outro estudo, García Marco e López del Ramo (2022) apresentam o seu trabalho que teve como objetivo apresentar o desenvolvimento de uma taxonomia interoperável para os sites das Associações de Amigos do Caminho de Santiago, construída por meio de categorias dos menus de navegação, hierarquicamente organizadas. Especialmente com foco na interoperabilidade, os autores destacam que essa funcionalidade foi desenvolvida a partir do mapeamento das categorias utilizadas nos menus de navegação, fazendo com que possam ser identificadas e organizadas para serem compreendidas e utilizadas em diferentes contextos ou sistemas. No âmbito apresentado pelos autores, a interoperabilidade é enfatizada pela sua capacidade de ser utilizada em uma variedade de funções de navegação e busca, bem como na gestão da construção de linguagens de indexação, no suporte ao design de arquiteturas web, funcionalidades de integração e agregação de páginas web. Portanto, considera-se um aspecto fundamental na construção de taxonomias web, permitindo que as informações sejam compartilhadas e utilizadas de forma eficiente em diferentes contextos e sistemas.

A partir dos estudos aqui relacionados, podemos destacar a importância da capacidade de interoperabilidade semântica no contexto das linguagens de indexação. Destaca-se também que é papel dos responsáveis pelos sistemas de informação em proporcionar uma estrutura semântica consistente às suas ferramentas de controle de vocabulário para que a interoperabilidade entre diferentes linguagens seja possível e um meio facilitador no processo de gestão e manutenção.

2.4 Recuperação da informação a partir das linguagens de indexação

A recuperação da informação é uma preocupação de longa data em bibliotecas e centros de informação, impulsionada pela necessidade de acessar e utilizar eficientemente o vasto volume de informações disponíveis. As linguagens de indexação favorecem o controle de vocabulário e são ferramentas, como já dito anteriormente, que são empregadas para viabilizar a recuperação da informação tendo em vista que através da manipulação de uma estrutura de termos podem organizá-los para uma melhor representação de recursos informacionais. Por isso, pode-se dizer que as linguagens de indexação permitem a organização sistemática de termos relacionados a um determinado domínio de conhecimento. Elas fornecem um conjunto padronizado de termos e relações semânticas, facilitando a busca e a recuperação de informações relevantes. Configuram-se como aliadas em conceitos chaves da Biblioteconomia e Ciência da Informação, como na representação, organização e recuperação da informação.

Com isso, os trabalhos que verificamos que tratam sobre a recuperação da informação no contexto das linguagens de indexação são de Alexiev, Isaac e Lindenthal (2016), Andrade e Lara (2016), Clarke (2019), Martínez-González e Alvite-Díez (2019), Cruz e Fujita (2021) e Cruz, Ferneda e Fujita (2022).

A linguagem é a matéria-prima de uma linguagem de indexação, sendo que através dela é possível a viabilização de pesquisas bibliográficas e tem o intuito de selecionar termos para a elaboração de uma estratégia de busca e recuperação da informação (Andrade, Lara 2016). Esse tipo de sistema é formado “por variações morfológicas (classes gramaticais), semântica (limites de significado), sintaxe (posição dos termos dentro do descritor, uso de pontuação) e tipografia” (Andrade, Lara, 2016, p. 107), por isso o gerenciamento da construção e manutenção de uma linguagem de indexação é uma tarefa minuciosa e que requer empenho dos responsáveis em mantê-la. Cabe ressaltar, que um dos propósitos da construção de um linguagens de indexação é a capacidade de auxiliar na recuperação da informação (Martínez-González, Alvite-Díez, 2019).

Ao aprofundar esse assunto, Clarke (2019) apresenta que o conceito de recuperação da informação remetido em sua pesquisa é da atividade de obter recursos de informação, ou seja, documentos presentes em uma determinada organização, relevantes para uma demanda de informação de uma ou mais coleções.

Em seu artigo, Clarke (2019) contextualiza a evolução dos tesauros relacionada à recuperação da informação desde seu uso em sistemas mais tradicionais, como agora aparados com diretrizes internacionais, como a ANSI/NISO Z39.19 ([2005], R2010) e a ISO 25964

(2013), sendo as mais recentes a tratar desse assunto. Segundo a autora, após a Segunda Guerra Mundial, a necessidade urgente de avanços científicos e industriais impulsionou a pesquisa e desenvolvimento, levando à criação do tesouro de recuperação de informação. Apesar das dúvidas sobre sua eficácia ao longo do tempo e da predominância dos motores de busca como o Google, o tesouro ainda desempenha um papel complementar na organização de informações, sendo ainda uma ferramenta aliada nos sistemas de informação. Destaca-se que a interoperabilidade é crucial para o sucesso atual do tesouro, abrindo caminho para aplicações em dados interconectados. Além disso, a demanda por recuperação de informação impulsiona o interesse em sistemas híbridos de organização do conhecimento que combinam elementos dos tesouros, ontologias e esquemas de classificação.

A autora aborda ainda que os tesouros são destinados a recuperação da informação e enfatiza o sistema de organização do conhecimento que utiliza de termos preferidos, ou seja, termos pré-definidos ao invés de códigos ou esquemas de caracteres representa fator facilitador àquelas pessoas que não estão familiarizadas com a ferramenta. Contudo, os tesouros para a recuperação da informação têm o desafio da ambiguidade presente na linguagem, pois um conceito pode ser expresso de diferentes maneiras e um caminho para solucionar esse desafio é uma linguagem de indexação responsável pelo controle de sinônimos, em que os homógrafos são separados e cada termo preferido pode ter apenas um significado (Clarke, 2019). Por assim dizer, ao estabelecer as remissivas em um determinado conceito, é possível atingir mais resultados e com maior especificidade.

Os autores Alexiev, Isaac, Lindenthal (2016, p. 40) abordam “expansão de consulta na recuperação de informações”, sendo um outro ponto na estrutura das linguagens de indexação que é a confiabilidade dos resultados de busca por meio dos relacionamentos hierárquicos, desde que priorizados com uma semântica refinada. Para facilitar o entendimento, demonstram com o seguinte exemplo: “Se Sofia BTP Bulgária BTP Europa, então Sofia BTPE Europa. Isto permite uma pesquisa de locais na Europa para encontrar também Sófia”.

Considera-se que isso pode favorecer a tarefa do bibliotecário indexador que teria uma melhor visão de onde as peças de um domínio se encaixam. Do mesmo modo, permite que o usuário também tenha maiores possibilidades em sua busca. Contudo, Alexiev, Isaac, Lindenthal (2016) questionam a viabilidade dessa possibilidade na interface do sistema de recuperação da informação, pois a expansão da consulta em hierarquias é algo significativo, mas que essa interface deve permitir ao usuário escolher se deseja incluir ou não elementos vinculados por relações de instância nos resultados da busca, a fim de evitar sobrecarregar o

usuário com informações indesejadas. Por isso, considera-se que o sistema de recuperação da informação possua recursos para fazer essa filtragem nos resultados.

Ressaltam ainda que a dificuldade surge devido à natureza complexa das hierarquias de tesauros, que frequentemente contêm um compilado de relações semânticas distintas. Isso compromete a precisão dos resultados de recuperação, especialmente quando a inferência automatizada de relações é aplicada. Esse desafio representa um obstáculo significativo para a representação refinada do conhecimento e a eficácia da expansão de consultas em sistemas de recuperação de informação (Alexiev; Isaac; Lindenthal, 2016). Por isso, é fundamental que tenhamos em mente a importância das tecnologias ao dispor dos profissionais da informação, especialmente em trabalhos que exigem grande empenho como a gestão na construção e manutenção de linguagens de indexação, porém que a validação da estrutura, aspectos semânticos e relações entre os termos sejam realizados por pessoas qualificadas e aparadas por diretrizes e políticas de procedimentos.

O artigo de Cruz, Ferneda e Fujita (2022) recupera os dados da pesquisa realizada em Cruz e Fujita (2021) especialmente sobre a disponibilização de vocabulários controlados, ou seja, das linguagens de indexação aos usuários para a recuperação da informação. Os autores relatam que o conhecimento e acesso a essas ferramentas fornece autonomia aos usuários no processo de recuperação da informação, além de que “a eficiência do processo de recuperação de informação é dependente da qualidade da representação dos documentos e das buscas realizadas pelos usuários” (Cruz; Ferneda; Fujita, 2022).

Em sequência, foi realizada uma pesquisa caracterizada como exploratória com análise de dados em sites dos catálogos online de bibliotecas universitárias brasileiras. Constatou-se que das quarenta e seis bibliotecas universitárias participantes, trinta e oito que utilizam linguagem de indexação no tratamento da informação e dezoito bibliotecas disponibilizam o recurso no catálogo e vinte, não. As tipologias das linguagens encontradas nessa amostra são: lista de autoridades (11 bibliotecas); tesauro (4) e Lista de cabeçalho de assuntos (1). Por meio dessa investigação, também ocorreu que a linguagem fica fora do campo de visão do usuário, sendo que apenas pessoas já familiarizadas com esse recurso conseguem visualizar o local de acesso para a consulta. Sobre isso, é enfatizado sobre a divulgação e capacitação desse recurso pelos bibliotecários para os usuários, a partir de treinamentos, tutoriais e manuais.

O sucesso da recuperação da informação está intrinsecamente ligado ao processo de indexação, como apontado por Acosta Bravo, Díaz Jiménez e Leiva Mederos (2021). A indexação envolve uma série de processos nos quais a sua eficácia depende de diversos fatores que influenciam a consistência dos resultados. Estabelecer uma política de indexação e,

principalmente, fornecer uma linguagem de indexação padronizada para os bibliotecários de uma rede pode ser uma estratégia eficaz para melhorar esse processo. Essa linguagem padronizada permite aos bibliotecários selecionar e traduzir termos de maneira uniforme e consistente, facilitando assim a recuperação precisa e eficiente das informações pelos usuários. Ao ter uma linguagem de indexação padronizada à disposição, os profissionais da informação podem tomar decisões mais embasadas, garantindo que os recursos informacionais sejam adequadamente representados e facilmente localizados pelos usuários, contribuindo assim para o sucesso da recuperação da informação.

Portanto, consideramos que as linguagens de indexação desempenham um papel essencial na viabilização da recuperação da informação, proporcionando uma estrutura organizada e padronizada para representar recursos informacionais e facilitar sua busca e recuperação. Essas ferramentas são fundamentais para garantir que as bibliotecas e centros de informação atendam eficazmente às demandas dos usuários por acesso rápido e preciso às informações necessárias.

3 GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DE LINGUAGENS DE INDEXAÇÃO EM BIBLIOTECAS UNIVERSITÁRIAS

Esta seção possui o objetivo de apresentar os estudos de cunho teórico e metodológico a partir da revisão de literatura nacional e internacional, incluindo também diretrizes normativas da ISO 25964-1 (2011), sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação. Destaca-se que esta seção contribuiu com a elaboração das questões da entrevista, tida como instrumento de coleta de dados desta pesquisa (Quadro 12). A revisão de literatura aqui realizada utilizou dos marcos teóricos e normativos da área de Biblioteconomia e Ciência da Informação.

O contexto das universidades e faculdades representa um espaço de dinamização na construção de conhecimento em prol da sociedade, pois o conhecimento gerado e os resultados alcançados através da pesquisa retornam para comunidade em geral de alguma forma. Considera-se que Saleh e Suleiman (2022) de forma assertiva consideram a biblioteca universitária como o coração de instituições de ensino, isso porque os bibliotecários atuam ou deveriam atuar, em conformidade com os aspectos de democracia, diversidade e equidade que são valores que os bibliotecários carregam consigo em sua atuação profissional cotidiana. Com isso, as bibliotecas universitárias concentram grande parte desse conhecimento, por meio do registro em recursos informacionais em diversos formatos, e têm como objetivo amparar e suprir as demandas dos pesquisadores, docentes, discentes e da própria universidade.

O papel da biblioteca universitária tem se expandido para diversas esferas nas universidades que exigem capacitação e atualização dos bibliotecários. Neste sentido, Silva (2023) coloca a biblioteca universitária atuante como um escritório de pesquisa, capaz de contribuir com produtos e serviços destinados a comunidade acadêmica e a comunicação científica. O autor elenca diversas frentes em que o bibliotecário pode atuar, como na gestão de dados de pesquisa, identificadores persistentes de autores, auxílio no processo de publicação, estudos de métricas, impacto da pesquisa de autores, questões que envolvem o direito autoral, além do incentivo às práticas de acesso aberto e à preservação e divulgação da produção científica e acadêmica da instituição, desempenhado pelos repositórios institucionais e bases de dados. É possível ainda citar a capacitação da comunidade, acadêmica ou não, por meio de *webinars*, cursos à distância, elaboração de tutoriais e entre outros, para que as barreiras físicas sejam rompidas e as pessoas sejam capazes de terem o suporte da biblioteca aonde quer que estejam.

Como se percebe, as atividades do bibliotecário se multiplicaram, tornando a profissão fundamental no meio em que a produção da informação e do conhecimento são constantes. No âmbito do tratamento da informação, especialmente no que se refere à indexação, há de se questionar como aprimorar as ferramentas de serviço habituais do bibliotecário para gerir toda essa demanda informacional. Muitos dos afazeres elencados por Silva (2023) passam inicialmente pelo tratamento da informação para posteriormente serem possíveis de acontecer. O rol de atividades do bibliotecário transpassa as paredes físicas dos prédios que abrigam os acervos e recursos físicos da biblioteca e passa a atuar também no ambiente digital e com a gestão da organização do conhecimento de recursos informacionais que podem ser acessados a qualquer momento (Seleh; Suleiman, 2022). Sobre isso, podemos citar os repositórios institucionais e as bases de dados que abrangem trabalhos, teses, dissertações, artigos, dentre outros, constantemente com originalidade ou atualizações sobre determinados temas, fazendo com que a manutenção de ferramentas de organização e representação da informação, tais como as linguagens de indexação, seja necessária. Assim como a divulgação científica, na qual só é possível com a catalogação e a indexação adequada dos recursos informacionais.

Em consonância com alguns pontos já elencados acima, a *Academic and Research Libraries* (ARL), uma das seções da *International Federation of Library Associations and Institutions* (IFLA), indicam os seguintes pontos que são destinados às novas atribuições do bibliotecário universitário, em especial ao apoio a aprendizagem, pesquisa e ensino: a transformação e incentivo às publicações em acesso aberto, o gerenciamento de dados abertos, recursos educacionais abertos, atuação em ambientes de código aberto e seguir políticas de ciência aberta. Aspectos de trabalho que são coerentes com a natureza e valores do bibliotecário e que agora ganham destaque pelo ambiente digital que se encontram, sendo que ainda devem ser voltados à democratização do conhecimento. Outros pontos elencados pela seção são: suporte para pesquisa em ambientes digitais, incluindo a alfabetização informacional de seus usuários, adquirir conhecimento em legislação de direitos autorais, além de novas formas disponíveis de gerenciamento de bibliotecas, sendo que para isso é necessário habilidades avançadas e treinamento contínuo dos profissionais que atuam em bibliotecas. Além dos bibliotecários, podemos citar também os demais funcionários que fazem parte da constituição de uma biblioteca universitária que também deve ter ciência de seu envolvimento no processo de construção do conhecimento.

Sobre o aspecto de incentivo à ciência aberta e a indexação desses documentos, Silva (2023, p. 43) aponta que “a ciência aberta promove uma nova dinâmica de pesquisa, como o acesso ilimitado ao conteúdo e a necessidade de sua indexação promove o desenvolvimento de

ações que visam dar visibilidade a documentos “invisíveis” em fontes de informação tradicionais”.

A visibilidade é um resultado alcançado pela representação e recuperação da informação; e pode ser entendida como trazer algo ou um tema à luz, ou seja, favorecer o conhecimento das pessoas que buscam a informação, especialmente em base de dados científicas. Com essa perspectiva, o estudo de Cruz, Tolare e Fujita (2023) apresenta o exemplo do tema da Covid-19 e como é retratado em diferentes linguagens de indexação, como a: Terminologia de Assuntos da Biblioteca Nacional; Tesauro Unesp e DeCS/MeSH. A pandemia de Covid-19 foi algo que surgiu rapidamente e todo o planeta uniu forças para combater essa ameaça à sobrevivência humana. Os profissionais da informação, como mediadores entre o conhecimento registrado e quem o almeja, tiveram que atuar rapidamente para disponibilizar os estudos recém lançados e atualizar as linguagens de indexação para fazer com que não houvesse disparidades na recuperação da informação. Esse é um caso em que a visibilidade do tema foi imprescindível, tanto para combater a desinformação, como para contribuir com os pesquisadores profissionais da saúde no rápido desenvolvimento científico e salvar vidas.

Em outro contexto, mas ainda sim sobre a visibilidade que é possível oferecer a um tema com a representação da informação, Achilles, Sousa e Sabbag (2019) relatam que “tanto a seleção quanto a representação temática são empregadas enquanto possibilidade de pensar um (ou mais) assunto(s) e podem fomentar a (in)visibilidade de uma coleção.” A presença de conceitos que abrangem representatividade social traz à tona temas que podem ser/estar invisibilizados por não serem contemplados na representação documental.

Contudo, a linguagem natural é uma realidade por ser utilizada por bases de dados que optam por termos que não tem controle de vocabulário. Strader (2009) enfatiza que esse é um debate antigo, principalmente diante de importantes ferramentas de busca, como o Google. A empresa Google, merece destaque devido a magnitude abrangente de produtos e serviços oferecidos, mas a sua principal característica é por ser um gigante buscador que dispõe de uma base de dados digital e um sistema de busca igualmente imponentes e complexos. A forma de funcionamento desse buscador é por meio de um mecanismo de pesquisa totalmente automatizado que usa softwares rastreadores da *web* e realiza a indexação com procedimentos de processamento e análise do conteúdo das páginas através de tags, como título, imagens, vídeos e entre outros (Central da Pesquisa Google, 2023).

Essa automação contribui significativamente para a área de indexação automática e para a organização em massa de documentos digitais presentes nas bibliotecas universitárias e em repositórios institucionais. Estudos sobre a indexação automática com o foco em bibliotecas

têm sido realizados (Gil Leiva; Fujita; Ortuño; Reis, 2020; Stachokas, 2014; Frederick, 2016), além de como o rol de atividades dos bibliotecários mudou e traz a indexação automática como uma ferramenta que pode auxiliar o trabalho do bibliotecário dentro da universidade devido a extensa gama de atividades que demandam sua atenção. Segundo Lancaster (2004) há a distinção de dois tipos de indexação, sendo esse processo automático ou não, que pode ocorrer por extração ou por atribuição. A indexação por extração, também conhecida como derivada, utiliza de termos presentes do próprio documento indexado. Na indexação por atribuição, utiliza-se o vocabulário controlado, ou seja, as linguagens de indexação na tradução dos conceitos. Corrêa e Fujita (2024) verificam o crescente uso da indexação automática por atribuição em documentos técnico-científicos no formato digital.

Porém, a dinâmica em bases de dados específicas e bibliotecas universitárias pode ser considerada diferente e requer um tratamento particularizado de seus recursos informacionais de forma mais padronizada e consistente tendo em vista as especificidades de cada instituição com a política organizacional.

Assim como veremos no decorrer desta tese, linguagens mais robustas que estabelecem relacionamentos entre os termos, como tesouros, por exemplo, precisam dispor de *softwares* para sua construção e manutenção. Com isso, um quesito fundamental para contribuir com esse processo é a criação de registros de autoridades.

Os registros de autoridades podem ser utilizados para padronizar nomes e assuntos. Fujita, Santos, Cruz e Moreira (2017, p. 65) explicam que quando um registro bibliográfico é criado é feita a descrição do recurso informacional pelo catalogador que utiliza a Base de Autoridades Assuntos ao descrever o assunto do item que está catalogando. Essa base contém os registros de autoridades que viabilizam o controle de vocabulário e estão no formato Marc 21. Com isso, destacam-se dois pontos fundamentais sobre o papel dos registros de autoridade: “têm o objetivo de assegurar a constante padronização da linguagem de indexação de forma automática e podem ser transferidos de uma base de dados para outra na formação dos catálogos de bibliotecas e rede de bibliotecas.” Por isso, compreende-se que os registros de autoridade auxiliam na construção e manutenção das linguagens de indexação, tendo em vista que atuam na correção e adequação no catálogo de registros bibliográficos no campo de assuntos. Além de que é possível a sua exportação para *softwares* de gestão de linguagens de indexação, por serem em formato do Marc 21.

Diante disso, é perceptível a complexidade que envolve a decisão da biblioteca sobre usar uma linguagem de indexação ao invés de termos livres/palavras-chave que envolvem alguns pontos que devem ser considerados. Ao construir uma ferramenta específica para as suas

necessidades organizacionais também é preciso estabelecer uma relação contundente entre os princípios da “garantia literária” e da “garantia de uso”. Segundo a norma ANSI/NISO Z39.19 (R2010), a garantia literária deve optar pelos termos que correspondam o mais próximo possível de termos utilizados na literatura. Já a garantia de uso está relacionada aos termos mais utilizados pelos usuários do sistema de informação. Esses dados de uso podem ser verificados por meio das pesquisas em bases de dados ou mesmo solicitações de materiais feitos diretamente aos colaboradores da biblioteca.

Barité, Fernández-Molina, Guimarães e Moraes (2010) fazem considerações a respeito da garantia literária ao dizerem que seu uso como um critério para a elaboração de um *corpus* na elaboração de um sistema de organização do conhecimento (SOC) pode não ser suficiente para ser utilizado como um respaldo ou justificativa. Contudo, os autores elencam outras perspectivas que a garantia literária pode contribuir para as linguagens de indexação, como utilizar a garantia literária como uma forma de avaliação de linguagens e a qualidade da terminologia utilizada, tendo em vista aspectos de “pertinência; reconhecimento, de especialistas; reconhecimento de usuários não especialistas; atualidade *versus* obsolescência; e representação idiomática”. Além da “extensão do teste de justificação em relação às relações conceituais”, tendo em vista a verificação e validação das relações hierárquicas e associativas entre conceitos, tanto das relações bem estabelecidas e consolidadas na literatura da área como aquelas que são mais “frágeis”, pois ainda estão ou não nesse estágio de validação pela comunidade científica-acadêmica. Esse processo pode envolver uma investigação mais aprofundada para determinar a validade e a relevância de certas relações conceituais, garantindo assim a precisão e a robustez da linguagem de indexação. Como última aplicação da garantia literária, os autores apresentam o seu uso para a “análise ou mapeamento de domínios” que se refere à reconstrução da estrutura conceitual de um determinado campo de conhecimento com base na identificação dos termos que representam o conteúdo dos documentos relacionados a esse campo. Isso possibilita entender melhor a organização do domínio e estrutura conceitual. Com isso, pode-se dizer que a garantia de uso e a garantia literária são princípios que podem contribuir para a gestão e manutenção de linguagens de indexação, basta definir em quais aspectos e para que serão utilizados.

A ciência de ambos os princípios, tanto a literária como a de uso, pode fazer com que as relações entre os termos dentro da estrutura de uma linguagem de indexação seja completa, no sentido de antecipar as variações terminológicas e harmonizá-las, por meio de remissivas, por exemplo.

Além disso, podem ser levantados quatro conceitos fundamentais sobre as particularidades que uma linguagem de indexação traz consigo como, a exaustividade e especificidade com relação à revocação e precisão. A exaustividade e a especificidade são formas com que a representação da informação é realizada, pois quanto maior for a exaustividade da indexação, ou seja, quanto maior for a cobertura de representação dos documentos, maior será a revocação das informações, porém com menor precisão/relevância nos resultados alcançados (Lancaster, 2004). Diferentemente da indexação que prioriza a especificidade nos termos escolhidos para a representação, atuando como um filtro para selecionar documentos com maior relevância diante da busca realizada (Given; Olson, 2003).

Esses elementos e outros aspectos são definidos na política de indexação, na qual representa o padrão de técnicas e procedimentos adotados no tratamento temático de documentos de uma biblioteca ou um sistema de informação e está relacionada com a cultura de processos organizacionais e administrativos pré-estabelecidos. Sobre isso, Rubi (2004) enfatiza que a política de indexação padroniza suas práticas, auxilia na capacitação de profissionais recém direcionados ao setor de indexação e aproxima a compatibilização de procedimentos adotados por diferentes indexadores da mesma instituição.

3.1 Linguagem de indexação na perspectiva da política de indexação em bibliotecas

A política de indexação é uma medida gerencial administrativa (Rubi, 2004) e não deve ser confundida com conhecimento tácito, ou seja, aquele que não é formalizado e difícil de ser transmitido a outras pessoas. No caso, ela deve estar oficializada em um documento que desempenhe a função padronizadora para ser seguido de forma sistemática pelos bibliotecários. Por isso, muitas vezes no cotidiano das bibliotecas esse documento pode ser chamado de manual de política de indexação, juntamente para transparecer que se trata de um guia para os procedimentos que envolvem a indexação.

De acordo com Fujita e Gil Leiva (2009, p. 156) a política de indexação apresenta-se “[...] de modo a compor-se como um conjunto de procedimentos, materiais, normas e técnicas orientadas por decisões que refletem a prática e princípios da cultura organizacional”.

Marília Vidigal Carneiro reúne em seu clássico artigo “Diretrizes para uma política de indexação” (1985) os elementos a serem considerados para uma política de indexação: a cobertura de assuntos do sistema/catálogo da biblioteca, a seleção e aquisição dos documentos-fonte, o processo de indexação, a estratégia de busca, a forma de saída, ou seja, como os resultados são apresentados e a avaliação do sistema. O processo da indexação na política exige

estabelecer as variáveis que irão influenciar em toda a representação e recuperação da informação, como:

- níveis de exaustividade;
- níveis de especificidade;
- escolha da linguagem;
- capacidade de revocação e precisão do sistema.

Estas são variáveis que quando definidas refletem a qualidade almejada da indexação. Especialmente a escolha da linguagem na política de indexação é crucial, pois influencia o desempenho do sistema de recuperação de informação. A qualidade da linguagem orienta o indexador na seleção precisa dos termos que representam o conteúdo do documento e guia o usuário na formulação de estratégias de busca (Carneiro, 1985; Rio-Branco, 2020).

Choi e Park (2020) apontam que essas medidas são consideradas as mais relevantes para a efetividade da indexação. Porém, Piovezan e Fujita (2014) destacam que as definições que abrangem esses dois níveis são tratados de formas distintas na literatura. A exaustividade está relacionada à cobertura dos termos/conceitos que são utilizados para representar um documento. Não se restringe ao assunto geral, mas busca abranger todos os assuntos abordados no documento (Carneiro, 1985, Gil Leiva, 1999; Olson; Boll, 2001). Na visão de Fosket (1973) a exaustividade está relacionada a análise exaustiva que o indexador realiza do documento a fim de conhecer ao máximo seu conteúdo temático. Já Choi e Park (2020) argumentam que a exaustividade se refere à medida em que várias partes de um assunto composto são reconhecidas na indexação, porém, com isso pode recorrer a linguagem de indexação que reúne todos os termos que representam um único conceito, por meio das remissivas. Desta forma, a questão de sinônimos e a dispersão na hora de recuperar documentos em uma base de dados, seria resolvida.

A especificidade por sua vez está relacionada à precisão que será utilizada na cobertura de assuntos de um documento, ou seja, é decidido o quão específico serão os termos utilizados no momento da indexação (Guimarães; Evangelista, 2015). De acordo com Kim (2006) o nível de especificidade é um critério para identificar termos de indexação ou descritores. Ressalta que a preocupação com este elemento surgiu a partir de uma questão prática para a catalogação de assuntos, especialmente na escolha de cabeçalho de assuntos.

De acordo com Fernandéz, Alonso e Montavez (2006) os níveis de exaustividade e a especificidade, trazem um indicador de qualidade à indexação. A partir da conceituação dessas variáveis da indexação, é nítida a percepção de que são decisões fundamentais para a

recuperação da informação. Além de serem decisões que devem ser pensadas e pré-estabelecidas para não haver divergências discrepantes nos sistemas de informação. Em concordância com Rio-Branco (2020) que resgata a explicação de Soergel (1994) em dizer que para que os níveis de especificidade e exaustividade estejam em sintonia é preciso considerar o contexto da indexação, definidos na política, como também o contexto da estrutura das linguagens de indexação que a biblioteca utiliza.

A linguagem de indexação é uma parte fundamental da política de indexação e apresenta papel mediador na organização e recuperação da informação (Fujita; Cruz; Patrício; Rio-Branco, 2019). A linguagem é uma variável determinante que representa toda a política de indexação, pois é a partir das definições previamente definidas no documento que o indexador irá escolher a quantidade de termos para a representação dos documentos, a especificidade definida a ser utilizada, os procedimentos realizados quando não são encontrados os termos no vocabulário utilizado, e principalmente a definição de qual linguagem de indexação utilizar.

Os autores Boccato, Fujita, Gil Leiva (2011) também destacam que a relação harmoniosa entre a política de indexação com a linguagem de indexação facilita o acesso e a recuperação eficaz de informações. Ponderam que o sistema de recuperação da instituição deve ser moldado considerando a natureza institucional, o que o usuário busca, as peculiaridades dos assuntos abordados no catálogo da biblioteca, os recursos disponíveis e a necessidade de otimizar o custo e a entrega de produtos e serviços oferecidos pelo sistema. Fujita, Moreira, Santos, Cruz e Ribas (2018) relatam sobre a elaboração de um manual de linguagem de indexação e a sistematização dos procedimentos que envolvem o catálogo online da biblioteca, como a diversidade e a especificidade de áreas do conhecimento para isso ser considerado na composição da estrutura hierárquica da linguagem de indexação. Esses mesmos autores consideram que esse manual de linguagem de indexação pode ser uma extensão da política de indexação, por todos esses elementos serem complementares uns dos outros em um mesmo ambiente organizacional.

3.2 Diretrizes normativas no contexto das linguagens de indexação

No âmbito das linguagens de indexação, o tesouro é o recurso que apresenta mais normas internacionais para o seu desenvolvimento. Acredita-se ser por conta da sua estrutura mais robusta e seu desempenho em sistemas de recuperação da informação. Por isso, antes de

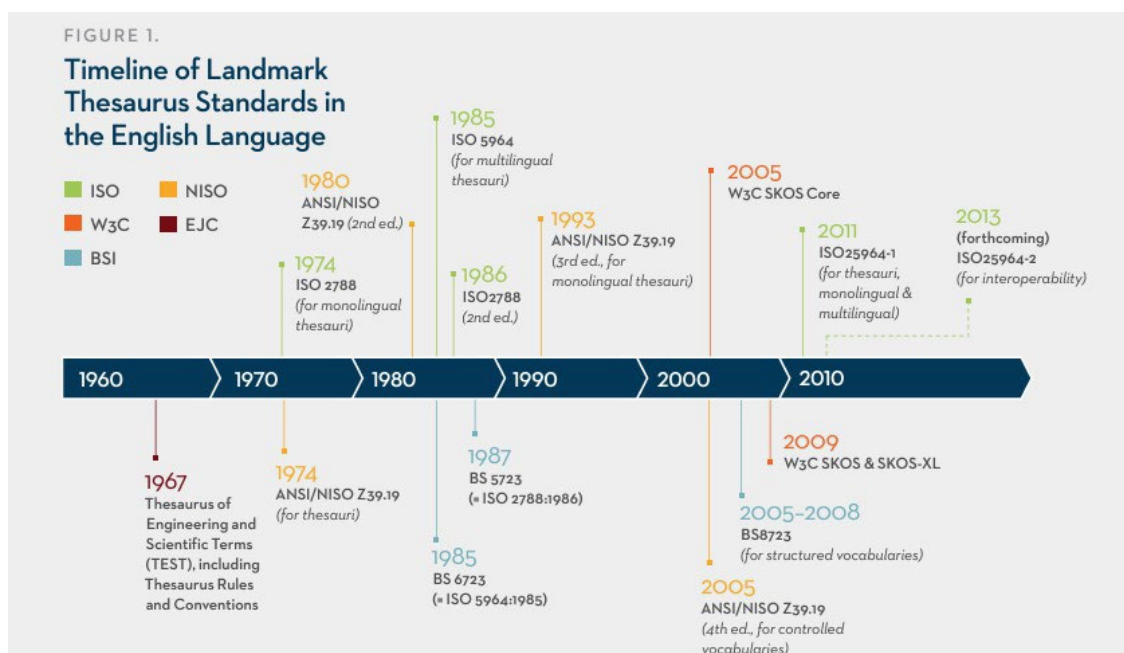
aprofundar mais na norma principal que rege esse trabalho, propõe-se apresentar um breve histórico das normas internacionais já existentes para a gestão de tesouros.

De acordo com Aitchison e Clarke (2004) o tesouro mais famoso foi desenvolvido por Peter Mark Roget, em 1852, com o nome de *Thesaurus of English Words and Phrases*, estruturado de acordo com os conceitos em que apresentava de forma sistemática e não em ordem alfabética. Mas foi em 1967 que foi publicada uma norma consolidada para tesouros, o *Thesaurus of Engineering and Scientific Terms (TEST)*. A título de curiosidade, esse tesouro, apesar da década em que foi elaborado, um pouco mais antiga no aspecto tecnológico, já dispunha de funcionalidades que utilizamos atualmente, como relacionamento de equivalência, hierárquica e associativa (Aitchison; Clarke, 2004).

Sendo assim, com o passar do tempo, evidenciando o aumento da produção de informação e o avanço científico-tecnológico, outras normas nacionais e internacionais foram sendo desenvolvidas para retratar a evolução que os tesouros estavam apresentando.

Na Figura 2, Clarke e Zeng (2012) apresentam uma linha do tempo na evolução das normas existentes sobre tesouros especificamente em língua inglesa.

Figura 2 - Linha do tempo com os marcos das normas sobre tesouros



Fonte: Clarke e Zeng (2012, p. 22).

Apesar dessa informação não constar na linha do tempo de Clarke e Zeng (2012), em 1970, a UNESCO publica *Guidelines for the Establishment and Development of Monolingual*

Scientific and Technical Thesauri, em vista que em 1986 esse padrão seria incorporado a ISO 2788, a primeira norma internacional para tesauros monolíngues.

Cervantes (2009), destaca outras duas normas, sendo elas de origem francesa e alemã, que se originaram também da norma da UNESCO, publicadas pela *Association Française de Normalization*, a AFNOR Z47-100, em 1973, e a *Deutsches Institut für Normung*, DIN 1463, de 1976. No âmbito americano, a *American National Standards Institute* publica a ANSI Z39.19 (1974) que contempla diretrizes para a construção de tesauros, em que seis anos mais tarde foi publicada sua segunda edição. Aliás, de acordo com a figura apresentada, as normas que foram atualizadas com o passar dos anos foram: a ANSI/NISO, chegando em sua quarta edição, em 2005, com foco no controle de vocabulário; e a ISO 2788, chegando a segunda edição no ano de 1986.

Em 1985, é publicada a norma internacional para tesauros multilíngues pela ISO 5964. Na esfera das normas de matriz britânica sobre os tesauros, são observadas a *British Standards*, a BS 6723 (1985) e a BS 5723 (1987), sendo consideradas idênticas a 2788 e ISO 5964 (Aitchison; Clarke, 2004).

Em 2005, um marco interessante se destaca, o foco das normas de tesauros no ambiente *web*, um novo padrão surge para contemplar a Web Semântica, elaborada pela *World Wide Web Consortium* (W3C), desenvolvido com recomendações para utilizar o *Simple Knowledge Organization Systems* (SKOS) e projetado para apoiar a publicação de linguagens de indexação, como tesauros, no ambiente web.

Para a normas mais recente no âmbito do gerenciamento da construção e manutenção das linguagens de indexação apresenta-se a norma ISO 25964, que foi publicada em duas partes: “Information and documentation: part 1: Thesauri and interoperability with other vocabularies”, em 2011, e a “Information and documentation: part 2: Thesauri and interoperability with other vocabularies”, de 2013 (Fujita; Cruz; Patrício, 2017).

. Segundo Martínez-González e Alvite-Díez (2020), a norma ISO 25964 (2011) substitui a ISO 2788 “*Documentation - Guidelines for the establishment and development of monolingual thesauri*” (1986) e avança nas discussões a respeito de tesauros baseados em conceitos, alinhados com a Web Semântica e a interoperabilidade com outras linguagens. Will (2012) relata que essa norma é a mais extensa já publicada em relação aos tesauros, pois busca não somente padronizar e esclarecer as interpretações existentes sobre entidades e relacionamentos, mas também busca auxiliar na implementação de forma consistente em sistemas automáticos.

A partir das diretrizes da ISO 25964 (2011, 2013), pretende-se sintetizar (e não esgotar) os pontos considerados basilares para o gerenciamento de linguagens de indexação, como: diretrizes iniciais para gerenciamento de linguagem de indexação; símbolos e siglas; relação entre termos equivalentes; orientações quanto à gramática; relação entre conceitos.

3.2.1 Diretrizes iniciais da International Organization for Standardization (25964, 2011)

Logo no início é indicado o desenvolvimento e manutenção de tesauros monolíngues e multilíngues, bem como a troca de dados tendo em vista a interoperabilidade, voltados à recuperação da informação. Esta etapa é inicialmente apresentada para depois mostrar como utilizar a linguagem na indexação. A abordagem da norma não exclui a gestão eletrônica dos documentos, pelo contrário, busca fazer a transição necessária para que documentos físicos e digitais possam ser geridos por um vocabulário controlado. Ressalta-se que a norma considera o fator humano na indexação, ou seja, presume que “o intelecto humano está geralmente envolvido na seleção de termos de indexação e na seleção de termos de indexação” (ISO 25964-1, 2011).

O tesouro costuma ser utilizado em sistemas de recuperação pós-coordenados, mas pode ser utilizado em outros sistemas, como o pré-coordenados, índices e diretórios hierárquicos. Visando a recuperação da informação, os tesauros podem ser utilizados para recuperar diversos tipos de recursos informacionais. Ressalta-se que são comumente mesclados com outras linguagens, ou seja, é normal que diferentes linguagens sejam utilizadas para compor somente uma. Por isso, a norma contempla não somente tesauros, mas também outros tipos de linguagens estruturadas e fornece orientações, que serão relacionadas adiante, para a interoperabilidade de vocabulários nas etapas do processamento técnico, armazenamento e recuperação da informação.

Frequentemente a palavra “termos” é tratada na norma, mas é especificado que isso é por questão de praticidade e o propósito em questão é manipular esses termos para contribuir para a recuperação de conceitos representados através dos termos ou mesmo “descritores”. A norma esclarece essa distinção, mas ressalta que o modelo apresentado para a estruturação de um tesouro é por meio de conceitos. Em Fujita, Moreira e Campos (2021) são definidos conceitos básicos sobre o tesouro, com base na ISO 25964 (2011) e em Dahlberg, (1978). Os autores explicam que “termo é a forma verbal de um conceito, o componente que, convenientemente, sintetiza e representa um conceito com o propósito de designá-lo e comunicá-lo” e os conceitos são designados como unidades de conhecimento representados por

termos. Destacam que “Cada termo incluído em um tesouro deve representar um conceito único”. Os conceitos variam de simples para complexos, sendo que geralmente termos compostos são utilizados para expressar conceitos mais complexos (Fujita, Moreira e Campos, 2021, p. 100).

Sobre isso, determinar um termo para representar um conceito específico é estabelecer o controle de vocabulário, pois um único conceito pode ser descrito de diferentes formas, por meio de diferentes termos (ISO 25964, 2011). As linguagens de indexação buscam esse controle para não haver disparidades e equívocos na descrição dos conceitos.

Um ponto interessante a respeito do controle de vocabulário é que o usuário precisa compreender e aceitar grau de “artificialidade” nos termos, pois é comum que o termo desejado não seja o termo autorizado/preferido, mas é aquele que reúne os documentos a respeito do tema em questão.

A elaboração de uma linguagem de indexação envolve várias nomenclaturas específicas, por isso a ISO 25964-1 traz um glossário com os principais termos e seus respectivos significados.

3.2.2 Símbolos e siglas

Além do glossário, as ferramentas apresentam símbolos e siglas para facilitar a esquematização de seus elementos. No quadro 2, são relacionados esses elementos, com o símbolo, *tag* (sigla) e significado de cada uma. Estes foram adaptados para a língua portuguesa, tendo em vista a norma que contempla essas informações no inglês, francês, alemão, dinamarquês, finlandês, dinamarquês, sueco, espanhol, chinês e russo.

Quadro 3 - Símbolos e abreviações

Descrição	Símbolos	Tag	Significado
Elementos descritivos		NE	Nota de escopo
		DEF	Definição
		NH	Nota histórica
Códigos		CA	Categoria de assunto; um código ou notação aplicada para um grupo de conceitos relevante para um determinado assunto.

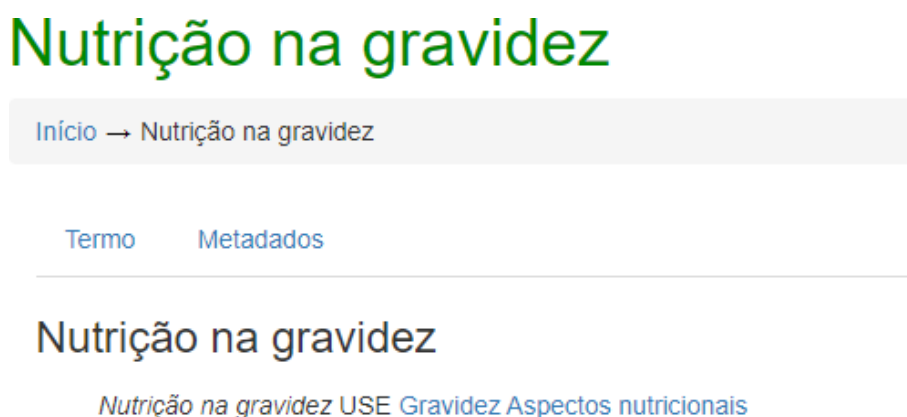
		CC	Código ou notação conceitual
Relações	->	USE	Use; o termo que segue a tag é o termo preferido que deve ser usado no lugar do termo não preferencial que precede a tag
	=	UP	Usar para ou Usado para; o termo que segue a tag não é preferencial termo para o qual o termo preferencial que precede a tag deve ser usado em vez de
		USE...+	Os dois ou mais termos preferidos após as tags devem ser usados juntos para representar o conceito indicado
		UP+	O termo não preferencial a seguir deve ser representado por uma combinação de termos preferenciais, incluindo o termo preferencial que precede a tag
		TT	Termo tópico; o termo preferido que segue a tag representa o conceito mais amplo em uma hierarquia à qual o conceito específico pertence
	<	TG	Termo genérico; o termo que segue a tag representa um conceito com um significado mais amplo
		TGM	Termo genérico (maior)
		TGI	Termo genérico (por instância)
	—<	TGP	Termo genérico (partitivo)
	>	TE	Termo específico: o termo que precede a tag refere a um conceito com um significado mais específico
		TEM	Termo específico (maior)
		TEI	Termo específico (por instância)
	>—	TEP	Termo específico (partitivo)
	—	TR	Termo relacionado: o termo que precede a tag é associado, mas não é um sinônimo, um quase-sinônimo, um termo genérico ou um termo específico.

Fonte: adaptado pela autora a partir da ISO 25964-1 (2011).

3.2.3 Relação entre termos equivalentes

O objetivo das linguagens de indexação são fundamentalmente guiar o indexador e o pesquisador para a escolha de um termo pré-estabelecido com o mesmo conceito. Para isso, é preciso que haja uma organização por meio de uma listagem de todos os termos que podem descrever o conceito (domínio). Por meio disso, é estabelecido o termo preferido com a *tag* “USE”, de forma a facilitar a identificação pela pessoa que está utilizando a ferramenta. O estabelecimento das relações e a apresentação dos termos de forma estruturada pode contribuir para essa clareza, como o exemplo abaixo:

Figura 3 - Exemplo da *tag* “USE”



Fonte: Tesauro Unesp.

Com o relacionamento entre os termos, é possível verificar que a pesquisa realizada utilizou o termo “Nutrição na gravidez” que não é o termo preferido, mas sim “Gravidez Aspectos nutricionais”, um exemplo de termo complexo com dois termos preferidos, usados juntos para representar um único conceito. Ao selecionar e abrir o termo, as seguintes indicações aparecem:

Figura 4 – Exemplo de dois termos preferidos

Gravidez Aspectos nutricionais

Início → Gravidez Aspectos nutricionais

Termo	Metadados
Gravidez Aspectos nutricionais	
Termos não preferidos	
<u>UP</u>	↔ Grávidas Nutrição
<u>UP</u>	↔ Nutrição na gravidez
<u>UP</u>	↔ Nutrition in pregnancy
<u>UP</u>	↔ Pregnant women Nutrition
Termos específicos	
<u>TE1</u>	↓ Desnutrição na gravidez ►
Termos relacionados	
<u>TR</u>	↔ Pregnancy Nutritional aspects

Fonte: Tesauro Unesp.

Fujita, Moreira e Campos (2021) definem que conceitos são unidades de conhecimento representados por termos, sendo que termos ou frases compostas são utilizados para conceitos complexos. O exemplo das Figuras 3 e 4 representa um termo de conceito complexo formado por um subcabeçalho que qualifica o termo tópico.

Na figura 2, são apresentados os termos não preferidos depois da *tag* que significa usado para - UP; um termo específico - TE1 relacionado ao assunto pesquisado e um termo relacionado - TR, no idioma inglês. Esclarecemos que na Norma ISO 25.964 (2011, p.44) “A relação de equivalência é a relação entre um termo preferido e seu(s) termo(s) não preferencial(is) correspondente(s) na mesma linguagem natural”. A relação de equivalência, dessa forma, não estabelece relação entre termos e, por isso, é explicitada separadamente das relações entre termos hierárquicos e associativos.

A partir dos relacionamentos entre os termos, os pesquisadores podem ter uma visão mais ampla de onde o termo está enquadrado na ferramenta. Com isso, é possível obter: Expansão na busca; Sugestão de termos de busca alternativos; Suporte para agrupamento ou outras formas de refinar uma busca; Identificação de erros comuns de digitação; Suporte para indexação automática (ISO 25694-1, 2011, tradução nossa)².

² Search expansion; suggestion of alternative search terms; support for clustering or other means of refining a search; identification of common spelling mistakes; support of automatic indexing.

Os conceitos e termos de uma linguagem de indexação são restritos a um escopo, perceptíveis através de uma cadeia hierárquica e/ou alfabética. Mesmo quando representados por uma macroestrutura que engloba diferentes áreas do conhecimento, geralmente presentes em bibliotecas universitárias, a partir da sua estrutura é possível distinguir um conceito de outro. Caso ainda seja passível de dúvidas, é possível colocar uma nota de escopo acompanhando o termo para melhor explicá-lo. De acordo com a norma, a nota de escopo deve restringir o conceito a um significado, “define ou esclarece os limites semânticos de um conceito conforme é usado no vocabulário estruturado”³ (ISO 25694-1, 2011).

Se a linguagem de indexação for multilíngue, deverá ser incluído um termo preferido e o escopo para cada idioma da ferramenta. Isso contribui não somente para as pessoas que utilizam esses idiomas como primeira língua, mas também nas buscas em bases de dados estrangeiras, facilitando assim a elaboração de estratégias de buscas e a recuperação da informação.

Os termos da linguagem podem ser únicos ou compostos para expressar um determinado conceito. Quando existem termos homógrafos são utilizados qualificadores entre parênteses para diferenciá-los, mesmo assim, é recomendado colocar uma nota explicativa.

Exemplos:

Mercúrio (planeta)

Mercúrio (elemento químico)

Mercúrio (mitologia)

3.2.4 Orientações quanto à gramática

Também são elencadas instruções acerca da forma gramatical dos termos e como eles devem ser apresentados, utilizados ou evitados. O uso de substantivos ou sintagma nominal, mas também pode ocorrer adjetivos para complementação das frases nominais. Os adjetivos não devem ocorrer sozinhos, pois podem gerar dispersão na recuperação da informação. Devem ser evitados: advérbios, tais como, advérbios de intensidade devem ser desconsiderados nas linguagens de indexação a não ser que tenham um significado fundamental para o termo, como por exemplo: “Problemas de muitos corpos” (TG - Física Nuclear); não utilizar artigos no início dos termos; verbos no infinitivo e particípio não devem ser utilizados sozinhos. Quando

³ Citação original: Scope note - note that defines or clarifies the semantic boundaries of a concept (2.11) as it is used in the structured vocabulary.

possível, é preferível transformar os verbos em substantivo, como “Escrever a mão” para “Escrita à mão”. Este exemplo orienta com uso da gramática para evitar flexão verbal.

Como mencionado anteriormente, a linguagem de indexação pode apresentar termos simples e compostos. A ISO 25694-1 (2011) discorre de forma detalhada sobre as decisões e implicações a respeito dos termos complexos. Porém, considera-se apresentar de forma resumida as especificidades apresentadas em relação a essa questão.

Sendo assim, os termos complexos podem ser compostos, dos quais utilizam a combinação de termos para representar um único conceito, ou pode ser uma única palavra que utiliza de prefixos e sufixos na sua composição, como “Biodegradável”.

Especialmente nos tesouros há a capacidade de fazer combinação de termos. Porém, é preciso tomar uma decisão em relação a esse aspecto. No exemplo dado pela Norma ISO 25694-1 (2011, p.37) “Biodegradável” não ficaria bem se fossem termos separados, “Bio” e “Degradável”, porém há termos que podem estar separados e posteriormente combinados, como “gestão + recursos humanos”. Com isso, eles são apresentados de forma consistente e satisfatória na estrutura da linguagem, evitando a dissolução do conceito. Adicionalmente, consideramos que o uso da terminologia é uma boa metodologia para esses casos mesmo que não conste da Norma ISO 25964-1 (2011)

Outro exemplo para analisar é “agroindústria” e “biossegurança” que compreende termos compostos em uma única palavra, por isso são considerados como complexos. Observa-se que esses termos apresentam:

Quadro 4 - Termos compostos em uma única palavra

Termo	Termo central	Diferença
Agroindústria	Indústria	Agro (derivado de Agricultura)
Biossegurança	Segurança	“Bio” - prefixo associado à Biologia

Fonte: ISO 25964-1 (2011).

Sendo assim, percebe-se que termos complexos, como os expostos no quadro acima, irão aparecer na linguagem. Sendo que resta decidir se os termos complexos que utilizam duas palavras ou mais será admitido ou não na linguagem. Nesse sentido, a norma apresenta cinco pontos a serem analisados em relação à conceitos que complexos na linguagem:

- 1) Simplificar o termo complexo, utilizando um único termo preferencial.

2) Representar o conceito complexo como uma combinação de dois ou mais termos, mas também incluir um termo não preferencial que aponta para a combinação dos termos preferenciais mais simples. Isso é chamado de "divisão" do conceito. Exemplo: Software de gestão de tesouros USE Gestão de tesouros + Software.

3) Assumir o conceito complexo e fazer uma avaliação depois de um período experimental.

4) Rejeitar o conceito complexo por considerá-lo desnecessário.

5) Não incluir diretamente o conceito complexo, contudo manter um conceito mais amplo e incluir um termo não preferencial direcionando para a forma autorizada. Exemplo: *Software* de gestão de tesouros USE *Software*.

Percebe-se que essa decisão está relacionada com o tipo de especificidade que a linguagem pode conter. No ponto 5, a norma aponta como termo preferencial o “Software”, porém se a área de concentração da ferramenta estiver no âmbito da Biblioteconomia ou Ciência da Informação, o termo preferencial provavelmente seria “Gestão de tesouros”. Na perspectiva da política de indexação, decisões como neste exemplo podem ser incluídos para a definição da especificidade e exaustividade.

Percebe-se que a decisão por meio da análise dos termos, especialmente dos termos complexos, pode ser trabalhosa, tendo em vista o uso comum e de como o termo é popularmente conhecido, fazendo com que se a expressão estiver utilizando termos separadamente dificultaria a compreensão ou mesmo a perda de significado e ambiguidade.

O quadro abaixo demonstra as circunstâncias que favorecem ou não a divisão dos conceitos. Ressalta-se que a análise individual de cada termo é fundamental para a tomada de decisões.

Quadro 5 - Considerar ou não a divisão dos termos

Dividir o termo quando	Não dividir o termo quando
Conceito bastante específico e está fora do escopo principal do tesouro. A inclusão de um grande número de termos periféricos aumenta o volume e a complexidade do vocabulário sem proporcionar muitos benefícios de recuperação.	Um termo/conceito familiar no uso comum, ou no campo abrangido pelo tesouro; sua expressão como elementos separados prejudicaria a compreensão.
Haver a probabilidade de que poucos documentos sejam indexados com o termo proposto, a inclusão como termo preferencial pode não valer a pena.	Dividir o conceito em suas partes levaria a perda de significado ou ambiguidade.
Quando houver dois focos diferentes em um	O conceito é representado por um termo

mesmo conceito, por exemplo, "câmeras de cinema subaquáticas" deve ser dividido em "câmeras subaquáticas" e "câmeras de cinema".	estabelecido que é ou incorpora um nome próprio.
Quando o foco representar uma propriedade, parte ou componente da diferença.	A diferença em um termo estabelecido perdeu seu significado original.
	O conceito é representado por um termo que contém uma diferença sugerindo uma semelhança, como uma metáfora, com algo ou evento não relacionado.
	O conceito é representado por um termo em que o foco tem outro significado na ausência da diferença.

Fonte: elaborado pela autora com base na ISO 25694-1 (2011).

Sendo assim, é possível perceber que não há uma resposta única para a inclusão de termos complexos e estes devem passar por uma análise individual levando em consideração as circunstâncias relacionadas na norma para que a melhor decisão seja empregada. Conceitos complexos são mais comuns na linguagem natural e depende de decisão tomada com base na terminologia da área de domínio.

Ao estabelecer o termo preferido, outra decisão deve ser estabelecida em relação à estrutura da linguagem, o relacionamento entre os conceitos, podendo ser uma relação hierárquica ou associativa entre os termos.

3.2.5 Relações entre conceitos

❖ Relação hierárquica

A relação hierárquica ocorre quando o escopo de um determinado conceito se enquadra completamente no escopo do outro. Se estabelece uma relação de superordenação e subordinação, na qual o conceito superordenado representa uma classe geral, e os conceitos subordinados referem-se aos seus membros. Moreira (2019, p. 14) explica que as relações hierárquicas são estabelecidas entre dois conceitos “quando o espectro semântico de um deles encaixa-se completamente no espectro semântico do outro”.

Há três tipos de relações hierárquicas que são logicamente distintas uma das outras, são elas:

a) Relação genérica: relacionamento entre uma classe/categoria e seus membros, por exemplo:

Frutas exóticas

Alguns membros da categoria 'frutas' são 'kiwis', mas apenas alguns kiwis são considerados 'frutas exóticas'."

b) Relação hierárquica todo-parte: esse tipo de relacionamento é uma forma específica de relação que se aplica em casos limitados, onde uma parte é inerente e exclusivamente parte de um todo particular. Segundo Moreira (2019), as relações hierárquicas genéricas ou partitivas são mais fáceis de serem percebidas do que as relações associativas, especialmente quando visualizadas em uma estrutura arborescente e com o uso de *tags* para diferenciá-las.

A relação hierárquica todo-parte é um tipo de relação que descreve a forma como as partes estão relacionadas a um todo. Nesse contexto, a relação todo-parte é definida como uma situação onde uma parte faz parte de um todo exclusivo. No entanto, é importante destacar que essa relação não é aplicável a todas as situações, mas sim a um conjunto específico de casos. Essa relação é geralmente aplicada a quatro classes principais de termos:

- Sistemas e órgãos do corpo;
- Localizações geográficas;
- Disciplinas ou campos do discurso (áreas do conhecimento e suas especialidades);
- Estrutura social hierárquica;

c) Relação de instância: relação entre os termos em que a conexão entre eles se dá através de um conceito amplo, que pode ser uma classe de objetos, fenômenos ou eventos, e outro específico ou exemplar dessa classe, geralmente representada por um nome próprio. Por exemplo, a classe "carros", na qual "carros" é um conceito geral que inclui todos os veículos automotores com as características comuns de um carro; uma instância dessa classe poderia ser "Mustang", que é um carro específico representado por um nome próprio.

Além da relação hierárquica, comum nos tesouros e em outras linguagens de indexação, pode-se estabelecer um relacionamento associativo.

❖ Relação associativa

A relação associativa entre os conceitos de uma linguagem não é relacionada hierarquicamente, mas sim são associados devido a semântica ou ao seu significado enquanto conceito. Por isso, a ligação entre os conceitos deve ser sempre explicada dentro da linguagem, tendo em vista que outros termos podem surgir ou serem modificados conforme a atualização da ferramenta. A relação entre os termos é indicada pela *tag* "TR" - termo relacionado.

Na relação associativa, ao usar um termo preferencial para representar um conceito, o outro termo associado deve ser implicitamente compreendido pelos usuários familiarizados com as referências expostas na ferramenta. O exemplo exposto na norma é de que muitas vezes, um dos termos está intrinsecamente ligado ao outro de tal forma que é fundamental para explicar ou definir o conceito, por exemplo, o termo "pássaros" é um componente necessário para compreender completamente o conceito de "ornitologia", que é o estudo científico das aves. Isso ilustra como certos termos estão interconectados e como a compreensão de um é vital para a compreensão do outro.

Exemplos:

Pássaros TR Ornitologia

Ornitologia TR Pássaros

As relações associativas também apresentam diferentes formas de fazer a ligação entre conceitos, como:

- a) Termos e conceitos com significados sobrepostos;
- b) Campo de estudo e os objetos/fenômenos estudados;
- c) Operação/processo e agente/instrumento;
- d) Ação e o produto da ação;
- e) Ação e seu destinatário/alvo;
- f) Objetos/materiais e suas propriedades definidoras;
- g) Artefato e suas partes (caso não se qualificarem para o relacionamento hierárquico todo-parte);
- h) Conceitos ligados por dependência causal;
- i) Objeto/processo e seu contra-agente;
- j) Conceito e sua unidade de medida;
- k) Termo composto e o substantivo que lhe dá foco;
- l) Organismo ou substância ou derivação de outro;

Essas são decisões técnicas que devem ser estabelecidas de acordo com o planejamento da ferramenta e do acervo físico e/ou digital a ser representado. Os relacionamentos dentro das linguagens de indexação funcionam como engrenagens que ligam um conceito a outro, fazendo com que haja uma compreensão lógica e sistemática da linguagem de indexação pelos atores responsáveis pela sua manutenção, pelos próprios usuários e indexadores.

3.3 Gerenciamento da construção e manutenção de linguagens de indexação (ISO 25964, 2011)

Nesta subseção, serão apresentados os pontos trazidos no décimo terceiro capítulo da norma ISO 25964-1 (2011) especificamente, por se tratar de informações fundamentais para os propósitos desta pesquisa, justamente por compreender o gerenciamento da construção e manutenção de tesouros.

Assim como em uma política de indexação, esta etapa é fundamental para a elaboração de uma linguagem de indexação, principalmente por estabelecer os objetivos, determinar o passo a passo das decisões a serem feitas e outros aspectos importantes que veremos a seguir. Desenvolver uma ferramenta que seja consistente, durável e manipulável requer esforços que podem levar anos e grandes investimentos por parte da instituição que a almeja.

Sendo assim, são apresentados estágios no decorrer do gerenciamento da construção e da manutenção das ferramentas de controle de vocabulário, especialmente os tesouros, ferramenta tratada pela norma ISO 25964-1. Esses estágios são fixos, ou seja, é possível utilizar esses estágios para o desenvolvimento de quaisquer linguagens de indexação tendo em conta que são sistematizadas as principais etapas do processo. São 6 etapas: a) planejamento do tesouro, b) estágios iniciais de compilação, c) construção; d) introdução ao tesouro; e) disseminação; e f) atualização.

a) Planejamento do tesouro

De acordo com a organização das etapas de planejamento da ISO 25964 (2011) é possível visualizar a sistematização no Quadro 6. Esse quadro foi desenvolvido com base na norma e com o objetivo de relacionar as etapas com elementos chaves de cada um.

Quadro 6 - Sistematização das etapas de planejamento do tesouro segundo a ISO 25964 (2011)

Planejamento do tesouro	
1 - Objetivos	Fator norteador Para que construir uma linguagem de indexação? Por quem será utilizado? Área(s) temática(s)
2 - Características do tesouro	Escopo Estilo de exibição Periodicidade nas atualizações

3 - Recursos	Recursos humanos Recursos financeiros Aporte tecnológico
4 - Estabelecer as responsabilidades	Coordenador(a) Grupo de revisão Equipe
5 - <i>Software</i> de gerenciamento do tesauro	Operacionalização da linguagem
6 - Compilação dos termos	Fonte dos termos: Seleção automática Índice de pesquisa Outras linguagens de indexação Referências tradicionais
7 – Construção	Organização dos termos Agrupamentos dos termos Relacionamento entre os termos
8 – Introdução	Apresentação da ferramenta Exibição das regras utilizadas Formas de utilização Contato dos responsáveis
9 – Disseminação	Divulgação Preservação
10 – Manutenção	Atualização da linguagem Frequência Revisões Critérios para as adequações

Fonte: Desenvolvido pela autora com base na ISO (25964).

No planejamento do tesauro é preciso inicialmente definir seus objetivos, pois é algo que servirá por muitos anos e irá refletir na representação e recuperação da informação de todos os recursos informacionais da instituição. Por isso, é necessário descrever:

- a) para que o tesauro será usado e por quem;
- b) se será limitado pelas restrições do software existente com o qual será usado;
- c) qual será o conhecimento dos usuários na área temática do tesauro e em seu uso.

Muitas pessoas estão envolvidas desde a gestão da ferramenta, como a equipe especializada nas áreas do conhecimento, bibliotecários, técnicos em informática, e quem fará o uso como instrumento de pesquisa e recuperação como, os usuários e indexadores. Por isso, é preciso alinhar os objetivos entre as partes para que esta se faça compreender.

Em seguida, características do tesauro como o escopo e o estilo devem ser definidos. Estabelecendo-se o formato da linguagem, esta será impressa, eletrônica ou ambos; o estilo de

exibição mais contundente; formatos especiais para integração com outros sistemas, como sistemas de busca ou indexação; e, como serão feitas as atualizações e com que frequência serão necessárias.

Como em todo projeto, os recursos utilizados devem ser também definidos, tais quais os recursos humanos; de quem virá o financiamento para as despesas, ferramentas de *software* e recursos da linguagem. Outro ponto destacado na norma é que se o tesouro for multilíngue, o orçamento precisará considerar as despesas gerais de comunicação, bem como o custo software especializado se necessário.

A partir disso, é preciso estabelecer as responsabilidades de cada integrante da equipe. Em um primeiro momento, um coordenador ou editor deve assumir a responsabilidade pela gestão do projeto em todas as suas etapas e possivelmente, em sua continuação e manutenção. Ao depender da magnitude do projeto, pode-se estabelecer um grupo de revisão editorial. A vantagem em combinar a responsabilidade sobre a ferramenta é o controle de qualidade da indexação temática de recursos.

Com relação aos recursos investidos, a escolha do *software* de gerenciamento do tesouro é um elemento a ser definido também, pois é a partir do software que poderá ser feita a operacionalização. O estudo dos *softwares* deve ser realizado para verificar se todos os requisitos desejados são atendidos. Caso nenhum software atenda a todos os requisitos, algumas adequações deverão ser feitas para a realidade da linguagem e do software.

b) Estágios iniciais de compilação

A partir disso, algumas recomendações gerais são oferecidas para como e quando começar a estruturação. Sabe-se que a linguagem de indexação tem como função apoiar a indexação e a busca de um sistema de informação, sendo assim, deve ser feita uma compilação dos termos de forma substancial para compor uma base de dados previamente.

Ressalta-se que alguns *softwares* disponíveis no mercado fazem o processo de análise do texto e extração de palavras e frases automaticamente de acordo com a frequência de incidência, até mesmo organizam esses termos de forma estruturada. Porém, é a qualidade de um processo realizado automaticamente não corresponde à qualidade do processo intelectual desenvolvido pela capacidade humana. A possibilidade de automatizar parte do processo pode contribuir para agilizar a compilação dos termos e posteriormente passar pela análise e validação intelectual.

Os termos de uma linguagem construída são representações de conceitos tratados nos recursos informacionais dos quais a ferramenta foi construída para atender. Tendo isso em vista,

a construção da linguagem é desenvolvida a partir da coleta e organização dos termos de forma que representem adequadamente os conceitos necessários da sua instituição.

A fonte para a compilação dos termos pode ocorrer a partir da literatura central, técnicas automáticas, outras linguagens de indexação, como tesouros e esquemas de classificação que apresentam o escopo parcialmente ou totalmente similar, termos separados pelos indexadores, relatórios que tratam dos índices de pesquisas de bases de dados, sites ou ferramentas de referências relevantes a partir da frequência de uso dos termos de busca, bem como obras de referências tradicionais, como dicionários e enciclopédias de especialidades que também podem ser úteis para a escolha, verificação e validação do termo.

c) Construção

A construção em si da linguagem consiste na entrada de termos/conceitos e na sequência de manutenção da ferramenta, através da alimentação com termos novos. Na entrada dos termos, deve-se agrupá-los em grupos ou hierarquicamente; do mesmo modo os sinônimos ou quase-sinônimos que devem estar agrupados e considerados em conjuntos para indicar as relações de equivalência. Nessa etapa é necessária atenção para não haver duplicidade de conceitos, sobreposições, omissões, erros de ortografia e empregar a especificidade necessária. As notas de escopo podem ser inseridas nesse primeiro momento também. Como visto, é possível utilizar, como ponto de partida nas seleções de termos, uma outra linguagem de indexação, sendo assim é possível importar uma parte deste outro vocabulário com auxílio do software de gerenciamento escolhido. A importação auxilia na economia de tempo, mas ressalta-se que é fundamental obter permissão do detentor dos direitos autorais.

A administração da construção da linguagem pode ser mais eficiente quando trabalhada em camadas, no agrupamento de termos, a partir das hierarquias, ou seja, ao estabelecer camadas superiores pode-se garantir a cobertura desejada e o equilíbrio adequado ao objetivo da linguagem. É interessante obter a contribuição de especialistas, docentes e outros usuários para organizar e validar a estrutura construída.

d) Introdução ao tesouro

Como a maioria dos projetos e pesquisas desenvolvidas, a linguagem deve apresentar uma introdução a fim de apresentar o escopo e os objetivos da ferramenta para o público usuário e para os potenciais usuários. Além disso, na introdução é possível esclarecer como deve ser utilizada a linguagem e em qual contexto pode ser empregada. Observa-se que os pontos tratados na introdução, segundo a norma ISO 25964-1 (2011), destacam grande parte do

arcabouço de estudos e decisões tomadas para a elaboração da linguagem. A partir da norma, destacam-se os seguintes pontos considerados importantes:

- os campos de assuntos abordados, identificação das áreas centrais e periféricas;
- os idiomas da linguagem;
- normas nacionais e/ou internacionais seguidas;
- tradução das convenções, abreviaturas e quaisquer sinais de pontuação utilizados;
- regras na seleção de conceitos;
- regras adotadas na seleção dos termos preferenciais, incluindo a referência a qualquer manual de estilo seguido, e as regras para decidir relações;
- declaração sobre a política de atualização, incluindo a frequência, datas e procedimentos, e o contato da agência ou pessoal responsável;
- referências e fontes utilizadas na compilação e revisão da ferramenta.

e) Disseminação

A disseminação da linguagem de indexação está atrelada a alguns procedimentos de divulgação, assim como de preservação. A usabilidade da linguagem no processo de indexação, navegação e busca em bases de dados faz com que seja necessário que a linguagem de indexação esteja integrada com um sistema eletrônico, passível de haver a exportação e a importação da ferramenta em um formato padrão, a fim de evitar dificuldades ao disponibilizá-la em outras aplicações, permitir compatibilidade e a sua transferência entre diferentes sistemas de indexação e busca. A publicação eletrônica, a distribuição de cópias impressas e até mesmo a publicação convencional representam outras formas de disseminar a ferramenta, destaca-se como exemplo o acesso à linguagem sem que haja a necessidade de estar conectado à internet nos casos em que há a impressão da linguagem. Assim como produções intelectuais devem ser depositadas na biblioteca nacional do país de origem, uma cópia da edição atualizada do tesouro desenvolvido e publicado também pode ser depositado em um órgão centralizador, tal qual bibliotecas nacionais. Por fim, os diretórios de sites podem contribuir para a disseminação da ferramenta, na qual o desenvolvedor responsável deve enviar detalhes completos da linguagem de indexação ao administrador de sites gerar e disponibilizar um link direto para ele, priorizando o estabelecimento de registro de metadados, como o *Dublin Core*.

f) Atualização

Para finalizar este tópico sobre gerenciamento da construção e manutenção de tesouros, destaca-se a importância e presença da atualização da linguagem, que se refere especialmente

a sua manutenção. Essa deve ser feita durante todo o período de vida útil da ferramenta. Segundo a norma ISO 25964-1 (2011), a necessidade de atualização começa no dia em que um tesouro é emitido, especialmente nesse começo, para que testes sejam realizados tanto pelos indexadores quanto pelos usuários na estratégia de busca em bases de dados. Estabelecer um procedimento de sugestão de inclusão ou alteração de termos, pode se tornar um mecanismo conveniente e um aliado na tarefa de atualizar a linguagem. Esse pode ser realizado através de formulários online ou mesmo impressos. Com isso, é possível realizar possíveis mudanças no formato e qualidade no desempenho de tratar e organizar a informação.

Com as mudanças realizadas na linguagem, é fundamental que seja realizado em períodos pré-estabelecidos a revisão dessas alterações para que seja verificado a coerência desses procedimentos realizados e ajustes necessários. O editor pode monitorar a frequência de uso dos termos preferidos na indexação, observando os termos mais utilizados e aqueles muito pouco que podem ser candidatos à exclusão ou modificação. Do mesmo modo, as consultas dos usuários na base de dados em questão e termos empregados em buscas mal sucedidas também podem ser critérios para se embasar a manutenção.

No início da implementação da linguagem é possível haver algum prejuízo no desempenho da recuperação para os recursos já indexados com outro vocabulário ou mesmo em linguagem natural. Sendo assim, as opções encontradas e escolhidas pela unidade de informação podem ser entre fazer a correção automática ou semiautomática de registros do acervo a partir do software de gestão do tesouro ou atender às novas demandas dos novos registros informacionais e realizar as correções de forma retrospectiva.

Considera-se que esses sejam os pontos principais da norma ISO 25694-1, especialmente no que se refere ao item 13 do documento. Para maiores detalhamentos a respeito do gerenciamento e manutenção de linguagens de indexação, a norma está disponível para a aquisição mediante compra online. Outros assuntos também são abordados como diretrizes para o gerenciamento do software de tesouros, modelo de dados, integração do tesouro com aplicações, intercâmbio de formatos e protocolos. Para concluir, a segunda parte da norma, ISO 25694-2 (2013) tratará a interoperabilidade com outros vocabulários, sendo tipos adicionais de linguagens controladas, tais como esquemas de classificação, listas de autoridade de nomes, ontologias, etc. Reúne diretrizes para a interoperação de linguagens em todos os estágios do processo de armazenamento e recuperação da informação.

A interoperabilidade é a capacidade de dois sistemas ou mais trocar e utilizar informações. A construção de linguagens de indexação pode utilizar a interoperabilidade para

o crescimento da ferramenta, especialmente no início de seu desenvolvimento (Bocato, Torquetti, 2012).

A próxima seção, conta com a apresentação das metodologias, dos procedimentos e instrumentos metodológicos que amparam e viabilizaram a investigação proposta nesta pesquisa.

4 METODOLOGIA

“A ciência se propõe a captar e entender a realidade, a metodologia se preocupa em estabelecer formas de como chegar até ela” (Demo, 1985). Para entender a realidade a partir da questão de pesquisa deste trabalho, a delimitação de caminhos e decisões foram estabelecidos.

A natureza desta pesquisa classifica-se como qualitativa, pois reúne e analisa os dados levantados cedidos pelas bibliotecas universitárias brasileiras. Também se caracteriza como pesquisa exploratória e descritiva, tendo em vista o propósito em responder à questão dessa pesquisa: Como as bibliotecas universitárias da USP e UNESP fazem o gerenciamento da construção e a manutenção de suas linguagens de indexação?

Melhor dizendo, define-se como pesquisa exploratória, pois, investiga e proporciona familiaridade com a questão a ser respondida, sendo que comumente se utiliza a pesquisa bibliográfica para esse objetivo (Gil, 2002). Segundo Selltiz *et al.* (1967 *apud* Gil, 2002, p. 41) em muitos casos, essas pesquisas envolvem: “(a) levantamento bibliográfico; (b) entrevistas com pessoas que tiveram experiências práticas com o problema pesquisado; e (c) análise de exemplos que “estimulem a compreensão””. Por isso, uma das etapas deste estudo utiliza a pesquisa bibliográfica, com a análise de práticas empregadas por outras bibliotecas universitárias, tanto no âmbito nacional como no internacional.

A pesquisa descritiva tem como objetivo descrever e estudar as principais características de determinada população (Gil, 2002), no caso, duas bibliotecas universitárias, da USP e da UNESP, com suas ferramentas desenvolvidas especificamente para o tratamento de seus recursos informacionais, com a análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação, na perspectiva da literatura e de entrevistas realizadas com os gestores e bibliotecários de ambas as instituições.

4.1 Pesquisa bibliográfica

Nesta etapa de pesquisa foram obtidos subsídios teóricos relatados, analisados e discutidos na Seção 2 para o suporte à discussão dos resultados alcançados por meio da aplicação da entrevista. A pesquisa bibliográfica atendeu o primeiro objetivo específico da pesquisa em realizar estudo teórico e metodológico a partir da revisão de literatura nacional e internacional, incluindo as normativas sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação.

A pesquisa se caracteriza como bibliográfica por procurar, explicar e discutir um assunto, tema ou problema com base em referências publicadas em diversos recursos informacionais. Por meio desse método, busca-se conhecer, analisar e explicar contribuições já desenvolvidas e publicadas por outros pesquisadores, sobre determinado assunto, tema ou problema (Martins; Theóphilo, 2016).

Para viabilizar a pesquisa bibliográfica utilizou-se do levantamento bibliográfico a partir da busca teórica e metodológica em bases de dados nacionais e internacionais. Para isso, foi utilizada a base científica “Dados Referenciais de Artigos de Periódicos em Ciência da Informação” (BRAPCI), por trazer trabalhos nacionais nos resultados de sua busca, fundamentais para o objetivo desta pesquisa que foca especialmente nas bibliotecas universitárias brasileiras; a *Scopus* e a *Web of Science*, por serem bases científicas conceituadas e de grande relevância internacional, nas quais os trabalhos recuperados são de periódicos que passaram por revisões criteriosas para poderem estar indexadas nelas. Além disso, foram realizadas revisões e estudos na literatura normativa a respeito do desenvolvimento de ferramentas de controle de vocabulário, como a elaboração de tesauros e a interoperabilidade dessas ferramentas. A norma apresentada mais recentemente é a ISO 25964-1 (2011) e a ISO 25964-2 (2013), mesma norma, publicada em duas partes.

Para as bases de dados da *Scopus* e *Web of Science*, foram utilizadas estratégias de buscas iguais. A primeira foi pensada a partir de variações de termos equivalentes encontradas para linguagem de indexação, com as citadas na introdução desta Tese e citadas a partir da sua corrente teórica: “linguagem documentária”, “vocabulário controlado” e “sistema de organização do conhecimento” e que tivessem relação com as bibliotecas universitárias. Na *Scopus* e na *Web of Science*, foram utilizados dois filtros: a restrição para trabalhos de Ciências Sociais, em um período de 10 anos (2013-2023), na busca por relacionar publicações mais atualizadas sobre o tema proposto. Na segunda estratégia, foi retirado o termo “*university libraries*” para incluir o termo referente à norma “ISO 25694”, na intenção de recuperar trabalhos que apresentassem métodos de elaboração de linguagens de indexação.

As estratégias mais complexas, como as utilizadas na *Scopus* e *Web of Science*, na BRAPCI não recuperaram nenhum documento, por isso foi necessário encontrar uma forma mais simplificada para obter resultados. Foram várias tentativas, seguindo as recomendações para a definição de termos em busca composta, com as variações dos termos e idioma, aplicando os operadores booleanos e tendo ciência das peculiaridades do sistema da base. A cada tentativa, a estratégia foi sendo modificada e facilitada. Sendo assim, a pesquisa na BRAPCI foi realizada com termos apenas em português “linguagem de indexação” e “bibliotecas universitárias”, na

modalidade de artigos, resultando em 6 trabalhos recuperados, como demonstrado no Quadro 7.

Quadro 7 - Resultados da BRAPCI

BRAPCI
FUJITA, M. S. L.; CRUZ, M. C. A.; PATRÍCIO, B. O. M.; RIO BRANCO, L. B. P. R. Linguagens de indexação em bibliotecas universitárias: estudo analítico. Informação & Informação , Londrina, v. 24, n. 1, p. 190-225, 2019.
CRUZ, M. C. A.; FUJITA, M. S. L.; SANTOS, L. B. P. Linguagem de indexação no contexto da política de indexação: estudo em bibliotecas universitárias. In: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). Memória, tecnologia e cultura na organização do conhecimento . Recife: UFPE, 2017. p. 217- 224.
CRUZ, M. C. A.; FUJITA, M. S. L. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. Informação & Informação , Londrina, v. 26, n. 1, 2021.
TOLARE, J. B.; FUJITA, M. S. L. A utilização de linguagens de indexação por bibliotecas universitárias: revisão bibliográfica. Scire: representación y organización del conocimiento , Zaragoza, v. 28, 2022.
FUJITA, M. S. L.; BOCCATO, V. R. C.; RUBI, M. P. O contexto de indexação para catalogação de livros usando uma abordagem sociocognitiva. Brazilian Journal of Information Science: research trends , Marília, SP, v. 4, n. 2, p. 22-40, 2010.
SANTOS, L. B. P.; FUJITA, M. S. L. A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. Scire: representación y organización del conocimiento , Zaragoza, v. 22, n. 2, 2016.

Fonte: elaborada pela autora.

O Quadro 8, demonstra as bases de dados e as quantidades de trabalhos recuperados da Scopus e *Web of Science*.

Quadro 8 – Resultado de busca nas bases de dados Scopus e *Web of Science*

Bases de dados	1ª estratégia	2ª estratégia
	("indexing language" OR "indexing languages" OR "controlled vocabulary" OR "controlled vocabularies" OR "knowledge organization system" OR "KOS" OR "documentary language") AND ("university libraries")	("indexing language" OR "indexing languages" OR "controlled vocabulary" OR "controlled vocabularies" OR "knowledge organization system" OR "KOS" OR "documentary language") AND ("ISO 25964")
Scopus	17 artigos	10 artigos
Web of Science	8 artigos	11 artigos

Fonte: elaborada pela autora.

O quadro abaixo identifica as bases de dados e os trabalhos recuperados, com a primeira estratégia de busca:

Quadro 9 - 1ª estratégia de busca e trabalhos recuperados

SCOPUS
TOLARE, J. B.; FUJITA, M. S. L. A utilização de linguagens de indexação por bibliotecas universitárias: revisão bibliográfica. Scire: representación y organización del conocimiento , Zaragoza, v. 28, 2022.
FOSTER, A. K. A Controlled vocabulary for an electronic resources problem reporting system creation, implementation and assessment. Library resources & technical services , Chicago, v. 65, n. 1, 2021.
KLOSE, A. C; GOLDSTEIN, S.; LEVY, M. S. Numismatics & Bibliographic Description: how Rutgers University Libraries described coins with MODS. Journal of Library Metadata , Philadelphia, v. 22, n. 1-2, p. 75-104, 2022.
CAMPBELL, H. M.; DIECKMAN, C. S.; TEAL, W.; WINTERMUTE, H. E. Improving subject headings for iowa indigenous peoples. Library Resources and Technical Services , Chicago, v. 66, n. 1, p. 48-59, 2022.
CHO, J. H. Cataloging for a celebration: metadata for an institutional repository from the ground up. Cataloging and Classification Quarterly , Philadelphia, v. 60, n. 2, p. 141-163, 2022.
HO, J.; STOKES, C. Core Metadata Element Recommendations for Institutional Repositories at Texas A&M University Libraries. Journal of Library Metadata , Philadelphia, v. 19, n. 3-4, p. 187-214.
FUJITA, M. S. L.; MOREIRA, W.; SANTOS, L. B. P.; CRUZ, M. C. A.; RIBAS, R. R. B. Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: An experience of the Sao Paulo state university. Knowledge Organization , Wurzburg, v. 45, n. 3, p. 220-231, 2018.
TARTAROTTI, R. C. D.; FUJITA, M. S. L. Subject cataloguing & indexing in university libraries: a verbal-protocol comparative study. Scire: representación y organización del conocimiento , Zaragoza, v. 23, n. 1, p. 57-66, 2017.
SANTOS, L. B. P.; FUJITA, M. S. L. A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. Scire: representación y organización del conocimiento , Zaragoza, v. 22, n. 2, 2016.
Jurić V.V., Olujić T.I. The importance of genre and physical characteristic headings in subject analysis of fine arts: The example of cataloguing fine arts materials in the graphic arts collection in the National and University Library in Zagreb. Vjesnik Bibliotekara Hrvatske , v. 57, n. 1-3, p. 377-390, 2014.
Web of Science
TOLARE, J. B.; FUJITA, M. S. L. A utilização de linguagens de indexação por bibliotecas universitárias: revisão bibliográfica. Scire: representación y organización del conocimiento , Zaragoza, v. 28, 2022.
CRUZ, M. C. A.; FERNEDA, E.; FUJITA, M. S. L. A disponibilização de vocabulário controlado aos usuários para a recuperação da informação. Revista Ibero-Americana de Ciência da Informação , Brasília, v. 15, n. 1, p. 266-282, 2022.
FOSTER, A. K. A Controlled vocabulary for an electronic resources problem reporting system creation, implementation and assessment. Library resources & technical services , Chicago, v. 65, n. 1, 2021.
ACOSTA BRAVO, A.; DÍAZ JIMÉNEZ, A.; LEIVA MEDEROS, A. A. Propuesta de procedimiento para la construcción semiautomática de tesauros en bibliotecas universitarias. Anales de Investigación , Havana, v. 17, n. 2, 2021.

CHO, J. H. Cataloging for a celebration: metadata for an institutional repository from the ground up. **Cataloging and Classification Quarterly**, Philadelphia, v. 60, n. 2, p. 141-163, 2022.

GUNJAL, B; URS, S; SHI, H. Australian Digital Libraries: an overview. In: World Congress On Engineering And Computer Science, 2008, São Francisco. **Proceedings [...]** São Francisco: WCECS, 2008.

Fonte: elaborada pela autora.

O artigo de Juric e Olujić foi desconsiderado da análise por estar em Bósnio. No Quadro 10, são identificados os trabalhos recuperados utilizando a 2ª estratégia de busca.

Quadro 10 - 2ª estratégia de busca e trabalhos recuperados

SCOPUS
ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. Knowledge Organization , Wurzburg, v. 43, n. 2, p. 107–112, 2016.
LARA, M.L.G. Documentary languages and knowledge organization systems in the context of the semantic web. Transinformação , Campinas, v. 25, n. 2, p. 145-150, 2013.
MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. The support of constructs in thesaurus tools from a Semantic Web perspective: Framework to assess standard conformance. Computer Standards and Interfaces , Amsterdam, v. 65, p.79-91, 2019.
CLARKE, S. G. D. The information retrieval thesaurus. Knowledge Organization , Wurzburg, v. 46, n. 6, p. 439- 459, 2019.
MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. semantic web methodological framework to evaluate the support of integrity in thesaurus tools. Journal of Information Science , [S.l.], v. 46, n. 3, p. 378- 391, 2020.
ALEXIEV V.; ISAAC, A.; LINDENTHAL, J. On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). International Journal on Digital Libraries , Nee York, v. 17, n.1, p. 39-48, 2016.
CHETI, A.; VITI, E. Functionality and merits of a faceted thesaurus: the case of the <i>Nuovo Soggettario</i> . Cataloging & Classification Quarterly , Philadelphia, v. 61, n. 5-6, p. 708-733, 2023.
THENAGLIA, G. Diseño de tesauros como estrategia didáctica para fortalecer su comprensión. Palabra Clave , Barcelona, v. 12, n. 1, 2022.
GARCÍA MARCO, F. J.; LÓPEZ DEL RAMO, J. Developing an interoperable taxonomy by mapping empiric taxonomies: a case study on the websites of the associations of friends of the Way of St James. Scire: representación y organización del conocimiento , Zaragoza, v. 28, n.1 , p. 55-73, 2022.
Web of Science
ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. Knowledge Organization , Wurzburg, v. 43, n. 2, p. 107–112, 2016.
LARA, M.L.G. Documentary languages and knowledge organization systems in the context of the semantic web. Transinformação , Campinas, v. 25, n. 2, p. 145-150, 2013.

MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. The support of constructs in thesaurus tools from a Semantic Web perspective: Framework to assess standard conformance. Computer Standards and Interfaces , Amsterdam, v. 65, p.79-91, 2019.
CLARKE, S. G. D. The information retrieval thesaurus. Knowledge Organization , Wurzburg, v. 46, n. 6, p. 439- 459, 2019.
MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. semantic web methodological framework to evaluate the support of integrity in thesaurus tools. Journal of Information Science , [S.l.], v. 46, n. 3, p. 378- 391, 2020.
ALEXIEV V.; ISAAC, A.; LINDENTHAL, J. On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). International Journal on Digital Libraries , New York, v. 17, n.1, p. 39-48, 2016.
CHETI, A.; VITI, E. Functionality and merits of a faceted thesaurus: the case of the <i>Nuovo Soggettario</i> . Cataloging & Classification Quarterly , Philadelphia, v. 61, n. 5-6, p. 708-733, 2023.
ALVITE-DÍEZ M.-L., MARTÍNEZ-GONZÁLEZ M.M. Thesauri representation in SKOS: trends and challenges for their integration in the Semantic Web. Scire: representación y organización del conocimiento , Zaragoza, v. 27, 2, p. 13-20, 2021.
THENAGLIA, G. Diseño de tesauros como estrategia didáctica para fortalecer su comprensión. Palabra Clave , Barcelona, v. 12, n. 1, 2022.
GARCÍA MARCO, F. J.; LÓPEZ DEL RAMO, J. Developing an interoperable taxonomy by mapping empiric taxonomies: a case study on the websites of the associations of friends of the Way of St James. Scire , Zaragoza, v. 28, n.1 , p. 55-73, 2022.

Fonte: elaborado pela autora.

Nas estratégias de busca utilizadas algumas bases de dados recuperaram os mesmos textos, quais sejam: Alexiev e Isaac (2016), Alvite-Díez e Martínez-González (2019, 2020, 2021), Andrade e Lara (2016), Chet e Viti (2023), Cho (2022), Clark (2019), Foster (2021), Fujita *et al.* (2018), García Marco e López Del Ramo (2022), Lara (2013), Tolare e Fujita (2022), Thenaglia (2022). O total de trabalhos com a exclusão de duplicados foi de 36 artigos.

4.2 Pesquisa descritiva

A pesquisa descritiva tem o intuito de verificar, analisar e correlacionar fatos ou fenômenos, além de suas variáveis, sem manipulá-los. A partir da pesquisa descritiva pretende-se identificar, com maior precisão possível, como ocorre determinado fato, a sua frequência, natureza, conexão com os outros e suas características (Cervo; Bervian; Da Silva, 2006).

Essa etapa foi desenvolvida a partir da análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e da UNESP na perspectiva da literatura (1) e dos gestores e bibliotecários (2). Em ambas as etapas foram desenvolvidas

categorias de análise para que viabilizasse a investigação e comparação das informações coletadas.

4.2.1 Análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e UNESP na perspectiva da literatura

A análise descritiva do gerenciamento na perspectiva da literatura teve como objetivo destacar o trabalho de duas universidades brasileiras, USP e UNESP, pois construíram as suas próprias linguagens de indexação e realizaram publicações sobre os processos realizados, os desafios encontrados e as decisões tomadas. Ademais, refletir sobre o papel das bibliotecas universitárias frente às demandas de representação e recuperação da informação nos ambientes físico e digital especialmente em relação ao tratamento da informação e o uso de linguagens de indexação são fundamentais pela experiência relatada e motivação para a busca de aprimoramento por outras bibliotecas.

A literatura analisada foi essencialmente de autores nacionais, mas algumas publicações com coautoria de pesquisadores estrangeiros atuantes como colaboradores e agregadores nos projetos de pesquisa conforme Quadro 11 de referências abaixo. Também se observou a divulgação desse tema em periódicos brasileiros e estrangeiros, além de trabalhos apresentados em congressos científicos e a publicação de livros.

Quadro 11 – Textos para análise descritiva

UNESP	Textos
	FUJITA, M. S. L. Política de indexação para bibliotecas: funções e finalidades. In: FUJITA, M. S. L. (org.). Política de indexação para bibliotecas: elaboração, avaliação e implantação . Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016.
	FUJITA, M. S. L. Linguagem de indexação para bibliotecas na perspectiva da política de indexação. In FUJITA, M. S. L. ; MOREIRA, W. (Org.). Manual do planejamento, construção e manutenção do Tesauro Unesp para bibliotecas: do conceitual a práxis . Marília: Oficina Universitária ; São Paulo : Cultura Acadêmica, 2021.
	FUJITA, M. S. L.; MOREIRA, W; CAMPOS, M. L. A. Aspectos conceituais do tesauro Unesp. In FUJITA, M. S. L. ; MOREIRA, W. (Org.). Manual do planejamento, construção e manutenção do Tesauro Unesp para bibliotecas: do conceitual a práxis . Marília : Oficina Universitária ; São Paulo : Cultura Acadêmica, 2021.
	FUJITA, M. S. L. Linguagem de indexação para bibliotecas na perspectiva da política de indexação. In FUJITA, M. S. L. ; MOREIRA, W. (Org.). Manual do planejamento, construção e manutenção do Tesauro Unesp para bibliotecas: do conceitual a práxis . Marília: Oficina Universitária ; São Paulo : Cultura Acadêmica, 2021.
	FUJITA, M. S. L.; SANTOS, L. B. P. ; CRUZ, M. C. A. e ; MOREIRA, W. . Avaliação das características do TemaTres e Multites para o controle de autoridades nas bibliotecas universitárias. Scire: representación y organización del conocimiento , Zaragoza, v. 23, p. 71-81, 2017.

	JESUS, S. A. Análise do Tesauro UNESP sob a perspectiva da abordagem facetada. 132 p. 2022. Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, Marília. SANTOS, J. C. F.; FUJITA, M. S. L.; MOREIRA, W. Tesauro Unesp: integração do registro de autoridade para o TemaTres. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. <i>Anais [...]</i>. Londrina: UEL; ANCIB, 2018.
USP	LIMA, V. M. A. O Vocabulário Controlado USP: elaboração, implantação e gerenciamento: uma experiência profissional, 2013. LIMA, V. M. A. <i>et al.</i> Estudos para implantação de ferramenta de apoio à gestão de linguagens documentárias: vocabulário controlado da USP. TransInformação, Campinas, v. 18, n. 1, p. 17-25, jan./abr., 2006. KOBASHI, N. Y.; LIMA, V. M. A.; LEME, M. Â. T. Manual de indexação de assuntos com uso do Vocabulário Controlado USP: versão preliminar. 1. ed. São Paulo: [s. n.], 2006. 72 p SANTOS, C. A. C. M. <i>et al.</i> Sistema de gestão para linguagem documentária: metadados e rede colaborativa no vocabulário controlado do SIBI/USP. 2010, Anais [...] Rio de Janeiro: UFRJ/SIBi, CRUESP, 2010.

Fonte: elaborado pela autora.

Dessa forma e, para melhor compreensão, foi decidido fazer a análise de cada linguagem de indexação a partir dessa literatura com base em sete categorias formuladas, conforme estudo normativo sobre construção e manutenção de linguagens de indexação (ISO 25964, 2011): *Background* das linguagens de indexação; Compatibilização de linguagens; Grupo gestor; Critérios e frequência para atualização; Software de gestão e manutenção; Disponibilização online para consulta; Diretriz normativa adotada. Essa etapa finaliza-se com as considerações sobre as duas linguagens na Seção 5.

4.2.2 Análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação na perspectiva dos gestores e dos bibliotecários da USP e UNESP

Para essa etapa de análise descritiva foi utilizada a técnica de entrevista aplicada com gestores e bibliotecárias, sendo duas bibliotecárias da UNESP e outras duas da USP. O critério de escolha desses profissionais fora que um tivesse participado ativamente da construção e manutenção da linguagem de indexação e outro que fosse da área de tratamento informação e que fizesse uso cotidianamente da ferramenta tendo em vista analisar, também, o *background* de elaboração da linguagem.

Destaca-se que a entrevista é uma técnica de interação entre o entrevistador e o entrevistado, com ela é possível aprender a opinião dos participantes, seus saberes e argumentos

(Severino, 2007), com isso é possível descobrir, não somente o resultado das práticas realizadas pelas bibliotecas na construção de suas linguagens, mas também as dificuldades e desafios do percurso percorrido.

A estruturação da entrevista ocorreu a partir das respostas obtidas no questionário dos estudos de Cruz (2019) e Fujita, Cruz, Patrício e Rio-Branco (2019), além do embasamento na revisão de literatura.

4.2.2.1 Entrevista

A entrevista é um instrumento para a coleta de dados voltada para a perspectiva do participante, tendo em vista que buscam a óptica do outro através de suas opiniões, avaliações, concepções e informações (Leitão, 2021). Esta técnica foi escolhida justamente por contribuir com o terceiro objetivo específico da pesquisa que busca parâmetros na literatura técnico-científica e a *práxis* das bibliotecas, ou seja, na prática e experiência dos bibliotecários que já trabalharam com a construção e manutenção de linguagens de indexação.

Segundo Leitão (2021) a entrevista é composta por três dimensões a partir de suas características: a dimensão temporal, a dimensão espacial e a dimensão estrutural. As dimensões elencadas pela autora definem-se como:

- A dimensão temporal trata-se de uma comunicação direta entre pesquisador e participante, caracterizada como um contato síncrono, ou seja, uma interação em tempo real que pode ocorrer pessoalmente ou não.
- A dimensão espacial concerne sobre a posição das pessoas envolvidas, pesquisador e entrevistado, em relação ao ambiente no qual a entrevista é realizada.
- A dimensão estrutural na entrevista refere-se à presença ou ausência de um roteiro que oriente a condução do pesquisador. Entrevistas podem ser não estruturadas, semiestruturadas ou estruturadas, com base no roteiro previamente definido.

Sendo assim, pode-se estabelecer que a entrevista foi realizada de forma síncrona, com duração aproximada de 45 minutos, realizada à distância através do Google *Meet*, o que facilitou o contato com os bibliotecários e até mesmo a transcrição da entrevista, pois a ferramenta possui a capacidade de gravação. A dimensão estrutural da entrevista foi de forma estruturada, pois contou com questões previamente estabelecidas diante da fundamentação teórica.

A elaboração da entrevista seguiu os procedimentos metodológicos expostos por Leitão (2021), seguindo as etapas de preparação; a construção do instrumento e a realização do estudo-piloto; a coleta de dados, com a condução da entrevista; e a análise- a transcrição e a análise dos depoimentos coletados.

- Preparação: Configura-se em uma fase preliminar de planejamento do *design* metodológico, na qual opta-se pela entrevista e a definição da amostra. Já discutidos acima na seção referente ao Universo de Pesquisa e escolha da entrevista por contribuir com os objetivos da pesquisa, especialmente o objetivo específico 3. É nessa fase também que aconselha-se a realizar “uma revisão consistente da literatura relacionada”.

- A construção do instrumento e a realização do estudo-piloto

Nesta pesquisa a entrevista teve um papel fundamental para analisar e realizar um panorama das bibliotecas universitárias que responderam às questões sobre a construção da linguagem de indexação, expondo as suas características e peculiaridades. A elaboração das questões foi baseada na revisão de literatura, na norma ISO 25964 e também na aplicação preliminar da entrevista, ou seja, o estudo piloto se concretizou com uma bibliotecária da Rede Unesp que foi a primeira entrevista para avaliar a clareza e pertinência das questões elaboradas.

Inicialmente, foram realizadas 28 perguntas para a bibliotecária da Rede Unesp e no decorrer da entrevista foram adicionadas outras duas questões que se mostraram pertinentes aos procedimentos necessários para a construção e atualização da linguagem de indexação. As questões foram previamente separadas em categorias para obedecer a uma sequência lógica. No Quadro 12 apresenta-se as questões da entrevista, juntamente com o aporte da revisão de literatura, além da contribuição adquirida pela aplicação piloto.

Ressalta-se que a organização das perguntas se deu pela divisão de blocos temáticos, também compreendidos aqui como blocos categóricos com base no conhecimento adquirido pela redação da fundamentação teórica, na qual buscou-se agrupar as perguntas de assuntos semelhantes e incentivar a linha de raciocínio do entrevistado ao aprofundar o tema em um percurso reflexivo e exploratório da entrevista (Leitão, 2021).

Informações sobre a biblioteca	
1. O catálogo da biblioteca é coletivo?	Segundo a norma ISO 25964-1 (2011) o

2. Quais são os principais assuntos que a coleção abrange?	planejamento de um tesauro ou de uma linguagem de indexação deve determinar a finalidade de uso da ferramenta e quem será o público-alvo. Por isso é fundamental conhecer o escopo da biblioteca universitária em questão.
3. A biblioteca está inserida em uma rede de bibliotecas?	
Política de indexação das Bibliotecas	
4. A biblioteca tem um documento com a política de indexação? Se sim, a política de indexação é local ou para todos da rede?	“A política de indexação é composta por decisões que devem contemplar os aspectos que versam sobre: a política e o processo de indexação, a linguagem de indexação e o sistema de recuperação da informação” (Rubi; Fujita; Bocato, 2012). Segundo Carneiro (1985) a linguagem de indexação é uma variável determinante que deve estar presente na política de indexação, pois é a partir dela que o indexador irá selecionar os descritores para representação da informação. Fujita e Rubi (2006) destaca que determinar a linguagem de indexação é fundamental, pois essa é uma decisão que afetará o desempenho do sistema de recuperação de informação, na estratégia de busca (expressão que o usuário busca no sistema) e na indexação (estabelece a precisão com que o indexador pode descrever o assunto do documento).
5. A linguagem de indexação é descrita na política de indexação?	
6. A forma como faz a gestão, o uso e a manutenção/atualização da linguagem é descrita em algum documento?	
Background das linguagens de indexação	
7. O que levou a construção de uma linguagem de indexação da própria instituição?	Contextualização da realidade da biblioteca.
8. Foram utilizadas outras linguagens de indexação no desenvolvimento da linguagem da sua instituição? (compatibilização de L.I) 8.1 Se sim, quais? 8.2 Se não, como foi realizada a seleção dos termos para compor a linguagem?	“Em um ambiente automatizado, a ausência de compatibilidade pode aumentar substancialmente os custos de tempo, inteligência e dinheiro [...]” (Zeng, p. 169, 1992, tradução nossa).
9. Foi utilizada alguma norma para guiar as decisões na elaboração da L.I.?	“Compartilhar registros bibliográficos e de autoridade é mais uma vantagem dos catálogos online das bibliotecas, cuja função é evitar a duplicação de esforços. Nesse sentido, normas e padrões desempenham um papel essencial na cooperação entre as bibliotecas.” (Fujita; Moreira; Santos; Cruz, Ribas, p. 222, 2018, tradução nossa).

10. Existe algum documento publicado ou não que tenha a trajetória da elaboração da linguagem de indexação, como um projeto, livro, relatório ou artigo?	“A elaboração de um manual de linguagem de indexação contém procedimentos sistematizados para catálogos online de bibliotecas” (Fujita; Moreira; Santos; Cruz, Ribas, p. 221, 2018, tradução nossa).
Aspectos técnicos da linguagem de indexação	
11. Qual é a linguagem de indexação da sua instituição?	Contextualização da realidade da biblioteca.
12. Quais são os campos de assuntos abordados na L.I.?	Contextualização da realidade da biblioteca.
13. A linguagem utiliza de siglas (como USE, UP, etc.) para esquematizar e organizar seus elementos?	A norma ISO 25964-1 (2011) reúne em uma tabela os símbolos, siglas e outros tipos de convenções, como as <i>tags</i> que são utilizadas em tesauros para indicar a relação entre os termos e a sua função.
14. A linguagem de indexação é multilíngue?	“Romper as barreiras linguísticas e culturais no acesso e representação da informação, via construção de vocabulários controlados multilíngues que sigam os conceitos da transculturalidade, da multiculturalidade e da garantia cultural é uma maneira de atender a esses propósitos.” (Biscalchin; Moreira, 2020).
15. A linguagem utiliza qualificadores para diferenciar um mesmo termo com diferentes significados? Por exemplo: Mercúrio (planeta) / Mercúrio (elemento químico) / Mercúrio (mitologia). 15.1 Se não, como é feita a diferenciação desses conceitos?	“Quando homógrafos são necessários como termos de tesouro, o significado de cada termo deve ser esclarecido, e a maneira tradicional de fazer isso é adicionando a ele um qualificador entre parênteses. O qualificador deve ser o mais breve possível, idealmente consistindo de uma única palavra”. (ISO 25964-1, 2011, p. 22, tradução nossa).
16. Há alguma orientação a respeito de como os termos devem ser gramaticalmente? Por exemplo: termos no plural e singular, podem ser utilizados verbos?	“Os termos do tesouro devem ser expressos da forma mais inequívoca possível. É especialmente importante formular o termo preferido de um conceito específico de maneira a transmitir o escopo pretendido”. (ISO 25964, 2011, p. 21, tradução nossa).
17. Como são empregados os termos compostos (como Gestão de tesauros)? Há momentos em que são divididos (como gestão + tesauros)?	“Conceitos complexos são frequentemente transmitidos por termos compostos, ou seja, termos que podem ser morfológicamente divididos em dois ou mais componentes. Alguns destes são termos de várias palavras, e outros consistem em apenas uma palavra” (ISO 25964-1, 2011, p. 37, tradução nossa).

Grupo gestor	
18. Há um grupo de pessoas responsáveis pela gestão e manutenção da L.I.?	Segundo a norma ISO 25964-1 (2011) durante a fase de planejamento de tesouro, um dos recursos-chave são os recursos humanos que devem incluir: a) um gestor principal responsável, de preferência com vínculo institucional/corporativo com o projeto da linguagem; b) Especialistas em diferentes áreas do conhecimento; c) Especialistas na elaboração de linguagens de indexação, para assumir as principais tarefas na construção e manutenção da ferramenta; d) Equipe de técnicos em informática para ajudar na implementação ou na “customização” do <i>software</i> de gestão da linguagem.
19. Você participa ou participou do grupo de gestão e manutenção da L.I.?	
20. Quando foi formado o grupo gestor?	
21. Quem foram/são os principais responsáveis?	
22. Houve algum treinamento ou capacitação dos membros do grupo?	Pergunta adicionada a partir da entrevista piloto.
Critérios e frequências para atualização	
23. Quais são os procedimentos adotados para a <u>inclusão</u> de termos novos na linguagem de indexação?	Segundo a norma ANSI/NISO Z39.19 (R2010), a garantia literária deve optar pelos termos que correspondam o mais próximo possível de termos utilizados na literatura. Já a garantia de uso está relacionada aos termos mais utilizados pelos usuários do sistema de informação. De acordo com Valléz, Pedraza-Jiménez, Codina, Blanco e Rovira (2015, p. 870) É necessário revisar as diversas abordagens utilizadas para identificar os descritores que compõem um vocabulário controlado, por exemplo, a unidade léxica utilizada para designar conceitos com assunto específico de domínio. Os diferentes métodos de extração de terminologia podem ser classificados em três abordagens: linguística, estatística e híbrida.”
24. Quais são os procedimentos adotados para a <u>exclusão</u> de termos?	
25. Quais os critérios que <u>você</u> utiliza para incluir um termo novo na L.I.?	
26. São realizadas correções dos termos da linguagem de indexação?	Pergunta adicionada a partir da entrevista piloto.
Software de gestão e manutenção	
27. É utilizado algum <i>software</i> para a gestão e manutenção da L.I.?	Segundo Leise (2003) é preciso avaliar qual <i>software</i> irá comportar a sua linguagem de

28. O <i>software</i> atende as necessidades da instituição?	indexação, os seus termos e a capacidade do programa de realizar constantes atualizações. Segundo Campos <i>et al.</i> (2006, p.80) “os <i>softwares</i> voltados para a construção de tesouros surgem para suprir a demanda de diminuir a complexidade da tarefa de criação, armazenamento, atualização e divulgação dos vocabulários concebidos”.
Disponibilização online da linguagem para consulta	
29. A linguagem de indexação está disponível no catálogo da biblioteca para o usuário realizar consulta?	Segundo Córdova e Lázaro (2012, p. 26) a linguagem de indexação possui “papel fundamental como intermediária entre as demandas dos usuários e a coleção. Nesse sentido, são linguagens de busca essenciais para a recuperação da informação”.
30. Você acredita que a L.I. é conhecida pelos usuários da biblioteca e que eles sabem que ela pode ser uma ferramenta aliada na busca e recuperação da informação?	

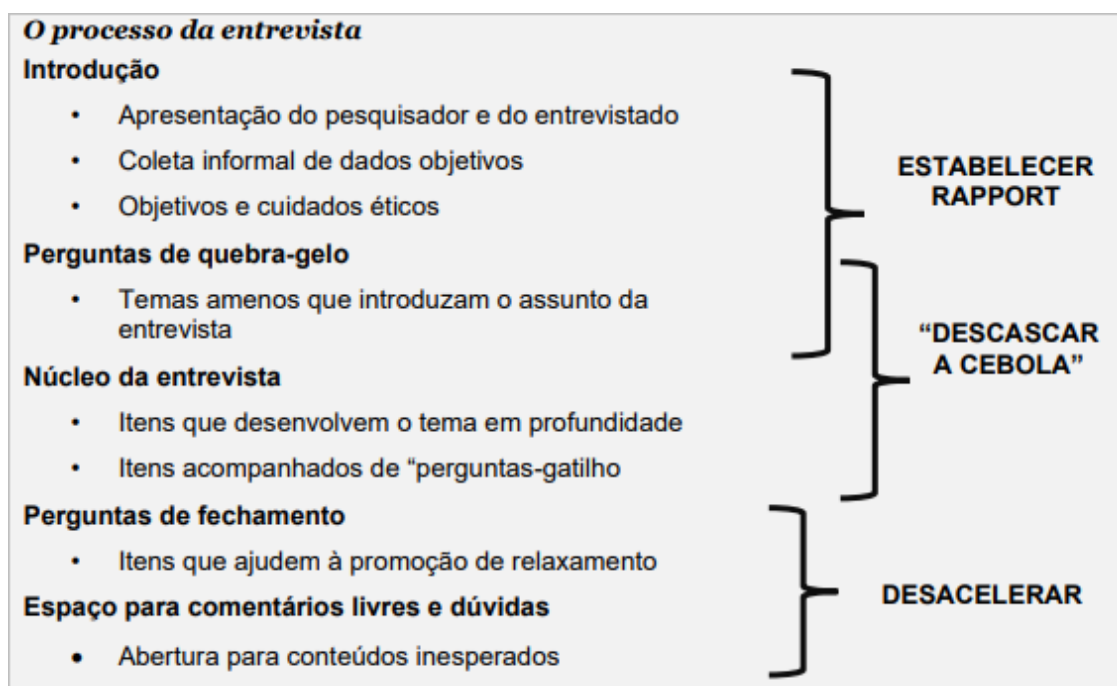
Fonte: elaborado pela autora.

- **A coleta de dados: a condução da entrevista**

Foi enviado e-mail para os bibliotecários, como relacionado no Anexo 1. A partir do contato com as bibliotecas, houve o acordo do dia e horário melhor para ambas as partes, priorizando sempre os horários disponíveis pelo participante.

Com isso, os procedimentos da entrevista seguiram o modelo de Leitão (2021):

Figura 5 - O processo da entrevista.



Fonte: Leitão (2021, p. 19).

Sendo assim, antes de iniciar as perguntas propriamente ditas, foram feitas apresentações pessoais, o objetivo da pesquisa foi exposto, a entrevistada informou o seu nome (que será preservado e apenas será divulgada a universidade a qual pertence), a biblioteca em que atua e a sua função dentro da biblioteca.

Ao final da entrevista, no momento de “perguntas de fechamento”, foi perguntado à entrevistada se houve alguma dúvida ou falta de clareza sobre as questões realizadas. Além disso, foi exposto que a partir da entrevista piloto com ela, foi possível incluir duas questões extras que poderiam contribuir com a entrevista como um instrumento de coleta de dados.

- **Análise: a transcrição e a análise dos depoimentos coletados**

Esta é a última fase da entrevista elencada por Leitão (2021, p. 19), na qual destaca que:

todo pesquisador que realiza entrevistas (ou, de modo mais amplo, que faz estudos qualitativos) deve estar consciente de que a análise dos dados resultantes das entrevistas não é a mera exposição reprodutiva dos pontos de vista dos participantes. Implicado no processo de investigação, o pesquisador lança seu olhar sobre esses pontos de vista para produzir novos significados a respeito do material coleta.

A transcrição da entrevista utilizou auxílio nas tecnologias disponíveis na *internet* gratuitamente. Inicialmente foi realizada a extração do áudio do vídeo utilizando o software de reprodução audiovisual chamado “VLC”. Após isso, a transcrição foi realizada a partir de uma ferramenta disponível pelo aplicativo Telegram, onde é feito um cadastro e inicia-se uma “conversa” com o usuário de nome *Transcribot* (ou @Transcriber_Bot), que é uma ferramenta de Inteligência Artificial (IA) do Telegram. Nessa página, é feito o *upload* do arquivo de áudio e a IA transcreve automaticamente o áudio como resposta.

A transcrição realizada pela IA não é de todo coerente, pois em alguns momentos apresenta erros ortográficos ou não traduz de forma correta uma palavra ou outra. Apesar disso, aproveita-se muito mais o tempo, mesmo que tenha sido necessária a checagem e correção da transcrição.

Com isso, a análise de coleta de dados será realizada a partir da relação explícita entre os elementos coletados, ou seja, dos trechos da entrevista e os significados a eles atribuídos, organizados por meio de categorias de análise previamente definidas. Segundo Leitão (2021) as categorias de análise previamente se chamam “análise *top-down*”.

Na Seção 6, serão apresentados os dados obtidos a partir da aplicação da entrevista e a discussão das respostas obtidas.

5 RESULTADOS E DISCUSSÃO DA ANÁLISE DESCRITIVA DO GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DAS LINGUAGENS DE INDEXAÇÃO DA USP E UNESP NA PERSPECTIVA DA LITERATURA

O uso de linguagens de indexação compreende uma tarefa a mais para o bibliotecário que pode utilizar conceitos de uma linguagem existente e disponível online, como a *Library of Congress Subject Headings* (LCSH), ou optar por uma linguagem construída pela instituição a qual pertence. Em ambos os casos, há esforços para traduzir os termos adequadamente para conceitos consistentes passíveis de uma recuperação com resultados relevantes. Nas considerações a respeito desse tema, Fujita (2021) evidencia aos pesquisadores e profissionais da área de Biblioteconomia e Ciência da informação as vantagens que a elaboração de linguagens de indexação tem para seu ambiente organizacional, dentre elas a garantia de uma recuperação com maior eficácia de informações por assunto nos catálogos online. Além disso, a autora incentiva a contribuição teórica e metodológica para a criação, implementação e avaliação dessas linguagens pelos gestores dos projetos, assim como enfatiza a necessidade de capacitar o pessoal responsável pelo tratamento da informação, com o objetivo de estabelecer concordância entre a política de indexação com suas ferramentas de controle de vocabulário.

No âmbito brasileiro há duas linguagens de indexação que foram desenvolvidas no universo universitário e que serão destacadas nesta discussão, tendo em vista sua relevância, a documentação da trajetória percorrida e dos processos embasados na literatura até o produto final com suas constantes atualizações. As linguagens em questão são resultados de pesquisas no âmbito de bibliotecas universitárias brasileiras: o Vocabulário Controlado da USP (VocaUSP), implementado no ano 2000/2001 (Lima *et al.* 2006) e o Tesouro Unesp da Universidade Estadual Paulista (UNESP), em 2013 (Fujita, 2021).

Ao tratar dessas duas linguagens, pretende-se verificar os procedimentos adotados para o projeto inicial do desenvolvimento da linguagem, ou seja, o que levou as instituições perceberem a necessidade de aprimoramento do tratamento da informação; o histórico das atividades realizadas; os participantes do projeto que compõem o grupo de gestão; como as instituições trabalham com a atualização da linguagem e quem é responsável por esse procedimento; além dos manuais e normas que amparam a tomada de decisões.

Para melhor compreensão, decidimos fazer a análise de cada linguagem de indexação a partir da análise da literatura selecionada com base em sete categorias formuladas conforme estudo normativo sobre construção e manutenção de linguagens de indexação (ISO 25964, 2011): *Background* das linguagens de indexação; Compatibilização de linguagens; Grupo

gestor; Critérios e frequência para atualização; Software de gestão e manutenção; Disponibilização online para consulta; Diretriz normativa adotada. Cada categoria apresenta uma ementa explicativa sobre o conceito e objetivo e na sequência a organização da análise da literatura para cada linguagem de indexação. Ao final apresenta-se as considerações sobre as duas linguagens.

CATEGORIA 1: *Background* das linguagens de indexação

Essa primeira categoria busca verificar o que levou a elaboração das linguagens de indexação e como foi o decorrer das atividades desenvolvidas, como projetos de pesquisa ou necessidades encontradas pelos bibliotecários. A partir das experiências passadas podemos comprovar a relevância das linguagens de indexação, especialmente no tratamento temático da informação e para o sucesso da recuperação da informação.

- **Tesouro Unesp**

As bibliotecas da UNESP fazem parte de uma rede automatizada e compartilham de normas, procedimentos em catalogação semelhantes e padronizadas. Contudo, constatou-se um problema de falta de padronização na indexação dos assuntos. Isso causava inconsistências na forma como os temas eram indexados na base utilizada, dificultando a recuperação da informação tanto dentro de cada biblioteca quanto na rede como um todo (Jesus, 2022).

Com isso, o Tesouro Unesp, anteriormente chamado de Linguagem Unesp, foi pensado a partir da política de indexação desenvolvida para a rede de bibliotecas da Universidade Estadual Paulista (UNESP). O projeto de implantação da política resultou no Manual de política de indexação de bibliotecas universitárias da Unesp (Fujita, 2014). A definição de uma linguagem de indexação dotada de um vocabulário em constante evolução foi parte estratégica da política de indexação das bibliotecas universitárias da UNESP (Fujita, 2014).

Os estudos no âmbito do projeto de política de indexação foram compostos por três fases:

“definição dos elementos da política de indexação que causam influência na recuperação dos assuntos no catálogo ATHENA, o processo de indexação para a catalogação de assuntos dos documentos das bibliotecas e a linguagem de indexação como instrumento de controle de vocabulário para mediar a representação temática na indexação e na estratégia de busca durante a recuperação por assuntos (Fujita, 2021, p. 21).

Anteriormente, o processo de indexação e representação de assuntos do catálogo online da rede de biblioteca da Unesp era realizado por meio da Linguagem da Rede BIBLIODATA (LCARB). Essa linguagem de indexação, além de desatualizada devido a sua inativação pela Fundação Getúlio Vargas, instituição que a alimentava, não era utilizada na estratégia de busca, pois não estava disponibilizada na web junto à ferramenta de busca ao catálogo. Consequentemente, não era conhecida pelos alunos, docentes e pesquisadores e não havia mediação da linguagem utilizada na representação dos documentos com a utilizada no momento da busca (Fujita, 2021).

Sendo assim, em 2013 foi criado o grupo de estudos para a elaboração da linguagem de indexação da Unesp, no qual atuou através da pesquisa-ação até a efetivação da ferramenta (Fujita, 2021).

● **VocaUSP**

As bibliotecas da USP são caracterizadas por seus acervos descentralizados e especializados, porém teve como objetivo de uma linguagem comum para a indexação de assuntos, considerando os recursos informacionais da graduação, especializados para pós-graduação e pesquisa. Para isso, foi criado um vocabulário unificado baseado nas linguagens usadas pelas bibliotecas, com o objetivo de que essa ferramenta fosse dinâmica e passível de atualizações, contando ainda com uma estrutura de relações lógico-semânticas entre áreas e subáreas, acompanhadas por regras de uso claras e compartilhadas (Vocabulário Controlado da USP, s.d.).

Para isso, no início dos anos 1991, o Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBi/USP) lançou como meta o aprimoramento do Banco de Dados Bibliográficos da USP (Dedalus), na qual foi observada a necessidade de atualização e expansão da lista de assunto utilizada pelas bibliotecas da rede.

O VocaUSP teve a sua trajetória em duas etapas fundamentais, sendo a primeira resultante de um projeto inicial do que participaram bibliotecários da Universidade, pesquisadores e docentes, da área de Biblioteconomia da USP e Unesp, que tiveram o papel de explicar a metodologia aplicada no projeto; além de docentes e especialistas de outras áreas do conhecimento que auxiliaram nas elucidações dos assuntos tratados na linguagem, bem como na garantia da consistência terminológica e conceitual, contribuindo para a dimensão teórico-metodológica exigente na elaboração de uma linguagem de indexação. Neste primeiro momento também houve a participação da equipe de tecnologia, responsável pela dimensão operacional e sistêmica do projeto. O resultado implementado em 2001 era uma base de dados

em CD-ROM ⁴ e foi integrado ao catálogo online das Bibliotecas do SIBi/USP, o Banco de Dados Bibliográficos da USP (DEDALUS), fato que resultou na ampliação da sua utilidade (Lima *et al.* 2006).

A segunda etapa é composta dos passos para a definição e implementação de um modelo padronizado de manutenção e gerenciamento do VocaUSP, tendo em vista a necessidade de continuidade das tarefas realizadas e buscando a garantia do crescimento do vocabulário de forma controlada, harmoniosa e consistente (Lima *et al.* 2006).

Após a elaboração do vocabulário, foi desenvolvido o “Manual de indexação de assuntos com o uso do vocabulário controlado da USP”, organizado por Kobashi, Lima e Leme (2006). Esse manual contempla elementos norteadores para a indexação dos recursos informacionais da USP, bem como explicita a estrutura do vocabulário e como deve ser feita a sua aplicação e uso.

CATEGORIA 2: Compatibilização de linguagens

De acordo com a norma ISO 25964-1 (2011) a compatibilização de linguagens de indexação pode ser uma fonte para a compilação de termos que estruturam uma linguagem de indexação que está sendo desenvolvida (ou sob manutenção). Por entender que essa é uma etapa fundamental da construção das linguagens de indexação, essa categoria tem como objetivo investigar se esse recurso foi ou é utilizado na gestão da construção e manutenção das ferramentas em questão.

- **Tesouro Unesp**

O plano de ação para elaboração da linguagem de indexação da Unesp foi com a compatibilização de três vocabulários semelhantes em sua raiz e estrutura, inicialmente a partir dos registros de autoridade da LCARB, do BIBLIODATA, com a *Library of Congress Subject Headings* (LCSH), a Terminologia da Biblioteca Nacional (TBN) e a *Medical Subject Headings* (MeSH). Em seguida, o objetivo era a criação de uma taxonomia das áreas do conhecimento separadas por categorias e subcategorias e seus respectivos termos de registros de autoridades. A partir disso, foi definido o *software* para comportar a macroestrutura, um dos requisitos para

⁴ Desde 2003, o SIBi/USP não mais editou o VocaUSP em suporte CD-ROM.

a escolha da ferramenta foi a sua capacidade de importar os registros de autoridade já compatibilizados.

- **VocaUSP**

O processo de elaboração e organização do vocabulário configurou-se em elaborar estruturas temáticas por área do conhecimento e compatibilizadas por subgrupos. Em seguida, foi realizada a hierarquização dos blocos de assuntos, alinhando-os com a lista de assuntos do sistema DEDALUS. Um software desenvolvido pela equipe de informática foi usado para auxiliar nessa tarefa. Os blocos de assuntos hierarquizados foram integrados na estrutura temática unificada de cada subgrupo, estabelecendo relações lógico-semânticas entre os termos e definindo termos ambíguos por meio de fichas terminológicas. As estruturas temáticas de cada subgrupo foram compatibilizadas com suas áreas complementares, fazendo com que a macroestrutura do VocaUSP fosse definida. Em seguida, cada descritor foi codificado de acordo com esse processo (Lima *et al.* 2013). A respeito desses procedimentos, a organização fica mais explícita quando observamos pelos seguintes pontos metodológicos, conforme exposto no site do vocabulário:

- a. Organização das bibliotecas da USP em nove sub-grupos
- b. Elaboração da estrutura temática de cada área e compatibilização das estruturas por sub-grupos
- c. Inclusão dos blocos de assuntos, gerados em ordem hierárquica, na estrutura temática unificada
- d. Estabelecimento de relações lógico-semânticas entre os termos
- e. Definição dos termos ambíguos (em ficha terminológica) e compatibilização das estruturas temáticas dos subgrupos com as áreas complementares (Vocabulário Controlado da USP, s.d.).

Ressalta-se que a compatibilização das estruturas temáticas ocorreu a partir dos descritores existentes nos catálogos das bibliotecas da Universidade e contou como fonte de referência outros “tesauros, sistemas de classificação, dicionários especializados, coleções básicas de cada área, estruturas curriculares, linhas de pesquisa das Unidades e especialistas da Universidade nas áreas do conhecimento consideradas” (Vocabulário Controlado da USP, s.d.).

Também é válido citar que no início dos trabalhos de desenvolvimento do vocabulário, foi ministrado aos 50 bibliotecários participantes, um curso sobre "Princípios de Compatibilização de Linguagens Documentárias" para que houvesse um entendimento em comum dentre os envolvidos em relação às metodologias e orientações compartilhados pelo grupo. Esse curso foi ministrado pelas docentes Anna Maria Marques Cintra, Maria de Fátima Gonçalves Moreira Tálamo, Marilda Lopes Ginez de Lara e

Nair Yumiko Kobashi, do CBD/ECA/USP, e Mariângela Lopes Fujita, da UNESP. Com isso, como resultado foi elaborado o "Projeto para Aprimoramento da Lista de Assuntos USP" (Vocabulário Controlado da USP, s.d.).

CATEGORIA 3: Grupo gestor

É imprescindível que haja pessoas responsáveis pela linguagem de indexação, independente do momento, seja na fase de projeto, de construção ou mesmo já finalizada e necessite de responsáveis pela sua manutenção. Isso faz parte dos recursos humanos que a ferramenta exige. Por isso, essa categoria tem como objetivo verificar os responsáveis pelas linguagens de indexação e a incumbência de cada parte.

- **Tesouro Unesp**

Em 2021, foi constituída oficialmente a Comissão Permanente do Tesouro Unesp, em substituição ao Grupo de Linguagem da Rede de Bibliotecas da Unesp, composto por catalogadores das bibliotecas da Unesp e pesquisadores da área de Organização do Conhecimento, com assessoria da Professora Doutora Mariângela Spotti Lopes Fujita (Fujita, 2021).

- **VocaUSP**

O grupo da USP que se dedica ao gerenciamento e atualização do VocaUSP é a Agência de Bibliotecas e Coleções Digitais (ABCD), ligada a Reitoria, e por bibliotecários da rede separada por áreas do conhecimento e representantes da Coordenadoria da ABCD. De acordo com Lima (2013), o grupo divide-se essencialmente em três instâncias principais:

- Equipe de coordenação: responsável pela elaboração e atualização de manuais, bem como analisar o uso dos descritores utilizados pelos indexadores; a realização de treinamentos para a capacitação dos bibliotecários da rede; validação dos termos propostos pela equipe de atualização e estabelecimento de procedimentos que possam aprimorar cada vez mais a estrutura do vocabulário.

- Equipe de atualização: responsável pela revisão permanente dos aspectos que envolvem o vocabulário, como sua atualização através da inserção de termos novos, alterações, exclusões, padronização de descritores e criação de novas áreas dentro da estrutura do vocabulário.

- Equipe de infraestrutura: destina-se a tomada de decisões referente ao

desenvolvimento e atualização de softwares que abrangem o vocabulário, manutenção periódica da base e viabilização da implementação das mudanças estabelecidas pela equipe de coordenação.

CATEGORIA 4: Critérios e frequências para atualização

A manutenção das linguagens de indexação se refere à sua atualização, o que implica na adição de termos novos, tendo em vista novos trabalhos desenvolvidos no campo científico, a correção ortográfica ou mesmo a exclusão de termos que estão obsoletos. Com isso em vista, essa categoria pretende relacionar os critérios e procedimentos que foram estabelecidos para manutenção de ambas as linguagens.

- **Tesouro Unesp**

A atualização da linguagem inicialmente se deu pela elaboração de um formulário de sugestão de inserção/alteração de termos por parte dos bibliotecários responsáveis pela indexação dos recursos informacionais de suas respectivas unidades. Esse formulário era enviado para o Grupo de Linguagem avaliar a possibilidade de integração na estrutura do vocabulário.

Com a alta demanda apresentada pelas bibliotecas, foi decidido que os próprios bibliotecários das unidades fariam a inserção ou modificação dos termos, seguindo o Padrão de Registros de Autoridades de Assuntos. Este documento apresenta todas as orientações para que os descritores sejam inseridos de forma correta pelos indexadores.

- **VocaUSP**

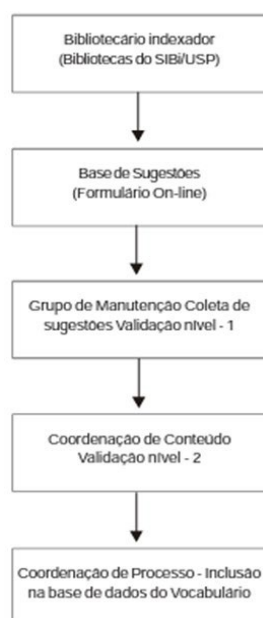
A atualização do VocaUSP é realizada diariamente pelos bibliotecários responsáveis pela indexação da ABCD, sendo que esse processo é revisto pelos coordenadores do grupo de gerenciamento da linguagem (Agência de Bibliotecas e Coleções Digitais da Universidade de São Paulo, s.d.).

Segundo Lima *et al.* (2013) a política para o gerenciamento e atualização da linguagem deve prever: o gerenciamento do processo e o controle terminológico. Segundo a autora, o gerenciamento do processo envolve estabelecer as diretrizes do processo, o modo como será a revisão e atualização do sistema, definição de prazos para resoluções, comunicação das tomadas de decisões aos integrantes do grupo e do sistema. O controle terminológico está relacionado a análise dos termos em si, ou seja, dos descritores e não-descritores, verificando a relação entre

eles conforme o desenvolvimento da linguagem sua consistência, além de consultar as fontes de informação e especialistas da área que fundamentam as mudanças no vocabulário.

O controle terminológico da linguagem também é regido pela garantia literária e garantia de uso estabelecidos nas diretrizes dos manuais de política de indexação. A equipe do VocaUSP, divide as tarefas de atualização em subgrupos, seguindo os aspectos da “dimensão social, que orienta as ações da Coordenação de Conteúdo; a dimensão cognitiva, principalmente desenvolvida pelo Grupo de Manutenção; e a dimensão operacional, que é fundamentalmente conduzida pela Coordenação do Processo” (Lima, et al. 2013, p. 20). A Figura 7, ilustra como é o processo de atualização, destacando os subgrupos e suas funções.

Figura 7 – Manutenção e gerenciamento do Vocabulário Controlado da USP: fluxo de trabalho.



Fonte: Lima (2013).

Segundo Santos *et al.* (2010) destaca sobre o sistema para gerenciamento do vocabulário, especificamente sobre o formulário online que através dele e de metadados padronizados para a inclusão, alteração e exclusão de descritores, foi desenvolvido no formato PHP, com o sistema de banco de dados de código aberto MySQL.

CATEGORIA 5: Software de gestão e manutenção

O software para a gestão da linguagem de indexação é algo importante para que seja possível ou facilitada a operacionalização das relações entre os termos, sua gestão e manutenção. Por isso, essa categoria tem como objetivo verificar se foi utilizado algum software ou desenvolvido algum programa internamente para abranger a linguagem, além de verificar quais foram os motivos para a escolha de tal recurso.

- **Tesouro Unesp**

Em Fujita, Santos, Cruz e Moreira (2017) foram analisadas as possibilidades do mercado em relação a dois softwares de gestão e manutenção de tesouros, o TemaTres e MultiTes. Esse estudo buscou verificar se ambos os softwares atendem à maioria aos requisitos necessários para o controle de autoridades em bibliotecas universitárias, especialmente no que diz respeito à recuperação de informações por assunto nos catálogos online. Os resultados indicaram que ambos os softwares atendem à maioria dos requisitos, mas que se sugere o teste prévio das duas ferramentas, para avaliar seu desempenho na construção do tesouro e na recuperação de informações por assunto, incluindo o uso com registros de autoridade.

Sendo assim, a partir dos testes realizados foi possível definir o *software* que abrangeria a macroestrutura da linguagem de indexação da Unesp, o TemaTres.

- **VocaUSP**

A Base de Dados do Vocabulário USP foi desenvolvida a partir das necessidades e demandas trazidas pelo grupo gestor à equipe de tecnologia e informática que desenvolveu um *software* específico para a linguagem, chamado de SIBIX.

Atualmente, o vocabulário passou a utilizar o *software* TemaTres para abranger a sua macroestrutura.

CATEGORIA 6: Disponibilização online para consulta

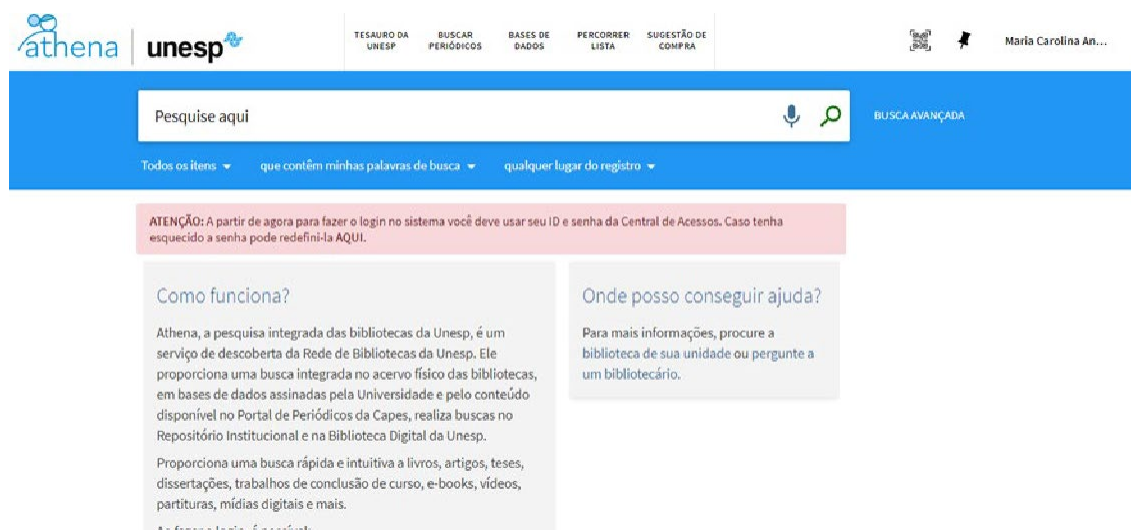
A partir da disponibilização online para consulta dos usuários ou bibliotecários, é possível que a estratégia de busca seja compatível com os termos utilizados para a representação da informação dos conteúdos informacionais. Por isso, buscou-se averiguar na literatura se as bibliotecas disponibilizam a linguagem, como ela é apresentada ao público ou se pretendem fazer isso em algum momento

- **Tesouro Unesp**

Segundo Fujita (2021) um aspecto importante considerado da elaboração do Tesouro Unesp foi a importância de divulgar essa ferramenta online para ser utilizado na interface de

buscas, ou seja, integrada à serviço da base de dados Athena para que alunos da graduação, da pós-graduação, docentes internos e externos, bibliotecários e demais usuários utilizem seus recursos.

Figura 8 – Interface do catálogo Athena



Fonte: <https://unesp.primo.exlibrisgroup.com/>

- **VocaUSP**

O VocaUSP está disponível online para sua consulta e esse é um dos resultados positivos elencados por Lima (2013) que destaca que a divulgação do vocabulário na *web* proporciona o uso da ferramenta por outras bibliotecas de outras instituições, sendo elas públicas ou privadas. Além disso, a disponibilização *online* proporciona a ampliação sobre o uso de ferramentas de controle de vocabulário e discussões a respeito da sua construção e manutenção.

Figura 9 – Banco de Dados Bibliográficos da USP (Dedalus)

Fonte: <https://dedalus.usp.br/>

CATEGORIA 7: Diretriz normativa adotada

As normas publicadas no âmbito das ferramentas de controle de vocabulário são destinadas a orientar a tomada de decisões e auxiliar no planejamento da linguagem pretendida. Por isso, considera-se fundamental que a literatura normativa seja considerada na gestão da construção e manutenção da linguagem. Com isso, essa categoria busca verificar se alguma diretriz normativa foi seguida e qual(is) dela(s) foi utilizada pelas instituições.

- **Tesouro Unesp**

Segundo Fujita (2021) os estudos a respeito do gerenciamento e manutenção da linguagem Unesp, especialmente em relação à elaboração de manuais e orientações a serem seguidas posteriormente, foram fundamentados a partir da ISO 25964-1, Parte I - Tesouro para recuperação da informação.

A partir dos estudos de Jesus (2022, p. 95) sobre análise do Tesouro Unesp, sob uma perspectiva da abordagem facetada, constatou-se que “é possível observar que esse tesouro utiliza a norma ISO 25964-1 e está em conformidade com a mesma”.

- **VocaUSP**

Apesar de verificar densidade e consistência teórica e metodológica nos processos de gerenciamento e manutenção do vocabulário controlado da USP, não foram encontradas informações a respeito do uso de diretrizes normativas.

5.1 Considerações a respeito do VocaUSP e Tesouro Unesp

Ambas as linguagens de indexação foram desenvolvidas por um grupo de pessoas empenhadas em construir uma macroestrutura capaz de representar os recursos informacionais de suas instituições. Assim como evidenciado pelas duas partes (Lima, *et al.* 2013; Fujita *et al.* 2021) o trabalho para desenvolver uma ferramenta consistente é uma tarefa que exige recursos humanos, financeiros e disponibilidade de tempo dos envolvidos.

O VocaUSP representa os esforços de uma iniciativa em biblioteca universitária não comum no Brasil. A Universidade impulsionada pelas Tecnologias de Informação e Comunicação (TICs), no início dos anos 2000, se mostrou a favor da evolução das ferramentas de controle de vocabulário e da importância de se estabelecer metodologias aplicáveis para o benefício do processo de indexação, com a mediação das linguagens utilizadas na representação de assuntos e busca nas bases de dados.

A Unesp desenvolveu o Tesouro Unesp a partir de uma necessidade observada através do projeto de pesquisa relacionada à elaboração de uma política de indexação da Universidade. Observou-se que seria preciso estabelecer uma linguagem capaz de atender as necessidades dos bibliotecários indexadores da rede e que essa linguagem fosse capaz de ser atualizada constantemente com os avanços científicos e tecnológicos da produção científica.

Pode-se notar que cada Universidade possui sua forma e critérios para a inserção de termos em suas linguagens. O Tesouro Unesp é mantido por um grupo que atende às dúvidas da rede sobre a ferramenta e decide sobre modificar ou inserir novos descritores na macroestrutura. Já o VocaUSP possui uma equipe que reúne as sugestões trazidas pelos indexadores e a partir de um fluxo de trabalho que acata ou não essas sugestões. Os softwares utilizados uma vez distintos no passado, agora é o mesmo, o TemaTres. Segundo o Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT (2023) esse software é amplamente utilizado, e o Instituto tem trabalhado para apoiar a iniciativa dessa ferramenta com o objetivo de integrá-lo a outros sistemas, como o DSpace, “para bibliotecas digitais e sistemas de bibliotecas, operando como base de autoridades”. Santos, Moreira e Fujita (2018) também consideram que o software TemaTres é uma ferramenta tecnológica que além de contemplar as macroestruturas, ela é responsável por possibilitar a gestão e manutenção da linguagem com a capacidade de importação de registros de autoridade de assuntos.

Observou-se também que ambas as linguagens de indexação prezam pela disponibilização online em seus catálogos. Sobre isso Santos, Moreira e Fujita (2018, p. 1132) enfatizam que o tesouro deve ser utilizado tanto pelo indexador como o pesquisador com o

objetivo de estabelecer um protocolo de comunicação ao utilizar os mesmos termos para se expressarem. Segundo os autores, para atingir esse objetivo, o tesauro deve expressar os termos controlados e suas relações conforme as ocorrências de seu domínio, pois isso possibilitará maior “precisão na identificação e descrição do campo nocional, na discriminação dos conceitos e de suas relações”.

A partir da consulta nos sites⁵ de cada linguagem, foi possível verificar que o Tesauro Unesp conta, atualmente, com 133.129 termos, sendo sua última alteração em novembro de 2023. O VocaUSP apresenta 47.187 termos em sua estrutura, com a última alteração datada em abril de 2024.

As duas Universidades, sendo de natureza pública, contribuem uma com a outra, especialmente no que diz respeito a estudos e pesquisas realizadas, para promover iniciativas capazes de promover avanços no campo da técnico-científico. As linguagens de indexação são ferramentas que auxiliam na mediação entre o conhecimento contido nos recursos informacionais e seus leitores. A biblioteca viabiliza esse “encontro” com o tratamento de todas as facetas que envolvem a informação.

⁵ <https://www.biblioteca.unesp.br/tesauro/vocab/sobre.php>

<https://vocabulario.abcd.usp.br/pt-br/sobre.php>

6 RESULTADOS E DISCUSSÃO DA ANÁLISE DESCRITIVA DO GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DAS LINGUAGENS DE INDEXAÇÃO DA USP E UNESP NA PERSPECTIVA DOS GESTORES E BIBLIOTECÁRIOS

Os resultados referem-se às informações obtidas na entrevista realizada com quatro bibliotecárias. A entrevista foi realizada por *Google Meet* e salva para posterior transcrição. Nesta aplicação das questões foi possível reunir dados da *práxis* da elaboração e gestão de uma biblioteca universitária brasileira, além de conhecer alguns dos percalços enfrentados pela equipe responsável e pelas bibliotecárias que utilizam a linguagem de indexação de suas respectivas instituições em seus cotidianos.

A discussão desses resultados será realizada a partir de categorias previamente definidas na própria estrutura da entrevista (Quadro 12), com base no estudo teórico e metodológico e na análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e da UNESP. Essas categorias são: “Informações sobre a biblioteca”, “Política de indexação”, “*Background* das linguagens de indexação”, “Aspectos técnicos da linguagem de indexação”, “Grupo gestor”, “Critérios e frequências para atualização”, “Software de gestão e manutenção”, “Disponibilização *online* da linguagem para consulta”. Cada categoria apresenta uma ementa explicativa sobre o conceito e objetivo e na sequência as análises das entrevistas acompanhadas de trechos selecionados das transcrições que caracterizam as categorias.

Conforme já mencionado, foram 4 bibliotecárias entrevistadas, duas atuantes na UNESP e as outras duas na USP. Foram selecionadas duas que participaram ativamente na construção de suas linguagens de indexação e outras duas que utilizam dessa ferramenta diariamente no tratamento temático da informação de suas respectivas unidades. Serão descritas da seguinte forma:

Bibliotecária Gestora do Tesouro UNESP - G1

Bibliotecária Gestora do VocaUSP - G2

Bibliotecária indexadora da UNESP - In1

Bibliotecária indexadora da USP - In2

Cada bibliotecária foi responsável por contribuir com essa pesquisa a partir de sua vivência profissional com o uso e percepção da linguagem utilizada. Destaca-se que inicialmente foi realizada uma pesquisa piloto com a G1, bibliotecária que participou ativamente da construção do Tesouro Unesp e contribuiu para a formação e consolidação dessa

entrevista. As respostas obtidas na pesquisa piloto foram consideradas nas discussões realizadas junto às demais.

❖ INFORMAÇÕES SOBRE A BIBLIOTECA

Essa categoria buscou uma familiarização com os bibliotecários, sua função na biblioteca/rede e um breve panorama profissional e as características da coleção da biblioteca em que atuam.

A G1 informou que é responsável pelo grupo técnico de informação e documentação e pelo gerenciamento da biblioteca em que ela trabalha, sendo que também faz parte da Comissão Permanente da manutenção e atualização do Tesouro Unesp. A G2 é uma bibliotecária da Escola Politécnica da USP, atuante no processamento técnico e no grupo de gestão do vocabulário e relata estar envolvida desde o início, até porque ingressou na Universidade em 1975, fazendo com que dessa forma acompanhasse a evolução dos sistemas de informação, juntamente com os meios que envolvem o tratamento da informação.

“Na Reitoria agora eu tô como responsável do grupo técnico de informação e documentação que é vinculado à Secretaria Geral, tá? Uma das atribuições do grupo é gerenciar a biblioteca da reitoria [...] E eu faço parte também da Comissão Permanente do Tesouro Unesp onde a gente entre várias atribuições, a gente trabalha com a manutenção e atualização do tesouro.” (G1)

“Eu adoro trabalhar no processamento. Eu adoro organizar essas coisas e eu fiz parte do vocabulário desde o começo até hoje eu estou envolvida com o grupo de gerenciamento do vocabulário. Na ABCD, né na agência de bibliotecas” (G2)

A In1 é uma bibliotecária da Rede Unesp, há pouco mais de 11 anos e é supervisora da Seção de Técnica de Aquisição e Tratamento da Informação (STATI). Dentre as atividades que realiza, a indexação de assuntos faz parte do seu cotidiano profissional. A biblioteca em que atua é voltada especialmente para a área de Ciências Biológicas, mas apresenta recursos informacionais também na área de Humanas. A In2 é bibliotecária na Escola de Comunicações e Artes da USP e também é supervisora de seção que comanda o tratamento da informação. O acervo tanto da USP como da Unesp faz parte de um catálogo coletivo online. A biblioteca em que atua a In2 é composta por recursos que abrangem especialmente duas áreas, a Comunicação e a Artes. Ela explica que rede de Bibliotecas da USP era chamado de Sistema Integrado de Bibliotecas da USP (SIBI-USP), mas recentemente passou a se chamar Agência de Bibliotecas

e Coleções Digitais (ABCD), sendo que essa Agência é responsável por orientar as demais bibliotecas da rede. Destaca ainda que a unidade em que trabalha está ligada à Divisão Geral de Tratamento da Informação da ABCD.

“Meu nome é In1, sou bibliotecária, trabalho há 11 anos e meio na Divisão Técnica de Biblioteca e sou supervisora técnica de seção.” (In1)

“Voltado para a área de ciências biológicas, mas também tem assuntos voltados para área de humanas, mas a sua maioria Biológicas.” (In1)

“Sim, a gente tá inserido no antigo SIBI da USP, né? Que agora é a Agência de Bibliotecas e Coleções Digitais (ABCD). Então tem um setor daqui que dá apoio para as bibliotecas, né? A gente tá ligada a Divisão Geral de Tratamento da Informação da ABCD” (In2).

A coleção da biblioteca em que a G1 é responsável é voltada principalmente para a área de Direito. Segundo a participante, a coleção atende à assessoria jurídica da Universidade, desde a sua criação. Mas, também ressalta outras áreas abrangidas como Administração, Administração Pública e pouco material de literatura. Ressalta que fora as áreas mencionadas, a compra de livros, como os da área de educação, é realizada por meio de solicitação dos docentes.

A G2 trabalha em uma biblioteca em que o acervo é constituído por recursos informacionais da área de Engenharia e cita que apesar da Escola Politécnica de São Paulo não ter o curso de Engenharia de Aeronaves, curso ofertado na unidade de São Carlos, eles apresentam materiais sobre esse tema também.

“[...] a gente tem bastante coisa também na área de educação, mas a gente só atende a solicitações dessas áreas se tiver uma solicitação, na verdade a gente só faz a compra se tiver uma solicitação. Então basicamente é Direito e Administração Pública”. (G1)

“É engenharia, todas as áreas da engenharia” (G2)

A biblioteca está inserida na rede de bibliotecas da Unesp e, assim como relatado pela In1, possui o seu catálogo coletivo, contudo, a G1 ressalta que apesar de estar ligada à Universidade, a biblioteca não recebe alunos por ser de natureza especializada com o foco em atender a parte jurídica e administrativa da rede. A G2 também aponta que o catálogo da sua instituição é coletivo e a biblioteca está inserida em rede. Além disso, a participante G2 fez um resgate histórico e contou que antes da base de dados atualmente utilizada pela USP, o Dedalus,

a Escola Politécnica já tinha um catálogo coletivo próprio dentre as bibliotecas que trabalhavam (e trabalham) no âmbito das engenharias.

“[...] a biblioteca está inserida em uma rede de bibliotecas sim agora a gente não tenha alunos né, que frequentem porque a gente é uma biblioteca administrativa mas a gente atende toda a rede da UNESP [...]” (G1)

“[...] O catálogo da USP é coletivo. A Poli enquanto não existia Dedalus ou até existir Dedalus, a Poli tinha um catálogo coletivo da Poli, porque nós éramos oito nove bibliotecas nesse momento [...] Para você ter uma ideia, somos 13 departamentos. A engenharia civil reúne, a biblioteca da engenharia civil que reúne os quatro departamentos na área de civil, a biblioteca da Elétrica reúne mais quatro departamentos e ao longo desses anos nós fomos reunindo acervos também”. (G2)

A partir do relato das participantes foi possível perceber que cada uma tem uma apresenta uma realidade diferente em relação ao seu ambiente de trabalho, especialmente por cada acervo ter as suas particularidades, uma biblioteca com o acervo voltado às áreas humanas, outra de biológicas, exatas e uma especializada em Direito, podemos considerar assim. As gestoras dos tesouros atuam em outras frentes que a biblioteca exige, como a administração da própria biblioteca, como é o caso da G1, e a responsável pela seção de tratamento da informação, como a G2.

Em suma, houve grande contribuição das bibliotecárias, tanto das responsáveis pela gestão das linguagens como das indexadoras, tendo em vista a bagagem de conhecimento dos procedimentos realizados durante a trajetória de desenvolvimento e o uso no cotidiano das linguagens em suas respectivas universidades. As indexadoras são experientes em suas áreas de atuação e estão há bastante tempo na seção de tratamento da informação utilizando das orientações que são passadas por grupos de trabalhos da rede e aplicando aos seus ambientes de trabalho, cada uma com a sua percepção do que é eficiente ou precisa de ajustes, como veremos a seguir.

❖ POLÍTICA DE INDEXAÇÃO

Essa categoria tem como objetivo verificar a presença de uma política de indexação e a sua ligação com a linguagem de indexação utilizada pela instituição ou pela biblioteca, caso a política seja restrita à uma unidade. A questão documental é fundamental para que os procedimentos adotados sejam realizados de forma uniforme e consistente, sendo uma política interna ou para todas as bibliotecas integradas da rede.

Ao perguntar se há uma política de indexação documentada, G1 disse que sim, há uma política de indexação oficial e que o documento é responsável por padronizar os procedimentos da indexação de toda a rede. A bibliotecária da Unesp, In1 também respondeu essas questões em consonância com o que foi dito pela G1. Anteriormente, a rede de bibliotecas não tinha uma linguagem específica que era padronizada na política de indexação, mas atualmente, a linguagem é definida no documento.

Sobre a presença de um documento que guia a gestão, uso e manutenção da linguagem de indexação foi dito que há documentos disponíveis para os bibliotecários, como manuais de procedimentos internos e a publicação de um livro sobre a política de indexação que contém essas informações.

“A política é pra todos da rede [...] É oficial para toda a rede de bibliotecas. Sim” (G1)

“[...] nós recebemos orientações de como proceder para alimentar, para corrigir, enfim de como fazer o trabalho com a linguagem, então os bibliotecários têm acesso a essas informações.” (In1)

“Sim. E se não me engano no início ainda a gente não tinha uma linguagem específica, tá? Então provavelmente nos primeiros documentos a gente ainda vai ter algo é, como uma possível, né, linguagem, já teve outro nome, já foi um pouco diferente.” (G1)

“Sim, é, principalmente no último documento que né? Que virou até um livro [...] mas a gente tem um manual, né? Da política de indexação e lá tem assim, a discussão de como os bibliotecários devem fazer.” (G1)

Com as bibliotecárias da USP, houve uma disparidade nas respostas da G2 e In2 sobre a política de indexação e a sua documentação. A G2 informou que há uma política de indexação disponibilizada pelo grupo do Vocabulário Controlado da USP para todos da rede, como uma política geral. A In2 relata que há uma política local e que não está documentada em documento formal, mas em um manual de procedimentos da seção que abrange decisões que foram tomadas ao longo dos anos em relação à indexação de assuntos e demais aspectos. A linguagem de indexação utilizada é o Vocabulário Controlado da USP, porém há uma base de dados com assuntos locais específicos da ECA. De acordo com In2 a biblioteca reuniu esses termos e enviou ao vocabulário. Por isso, compreendeu-se que a biblioteca utiliza o VocaUSP, mas também apresenta uma segunda “linguagem” que aplica quando necessário.

“A gente aqui tem uma política local que tem assim os assuntos pertinentes a biblioteca da ECA. A gente tem uma base de dados com assuntos locais com decisões que foram tomadas antes da biblioteca integrar o sistema, antes da constituição do vocabulário unificado da USP. Então a gente tem esse que é um parâmetro para a gente trabalhar localmente, para a gente usar a indexação

voltada para a nossa área e tem o vocabulário controlado da USP. Que aí se juntou todos os termos das bibliotecas da USP e a gente mandou termos para lá e a gente trabalha com esse vocabulário, mas a nossa política é local, tá?” (In2)

“[...] tem do manual da seção tem algumas orientações, mas eu não as classifico como uma política de indexação. É só assim decisões que foram tomadas ao longo dos anos e que a gente procura seguir.” (In2)

“Política de indexação nós seguimos a do grupo do vocabulário não tem nada que seja específico da Poli, não é uma política geral uma política geral seguimos a universidade. Ela é documentada. Ela está escrita em um documento para todos da rede.” (G2).

A questão seguinte abordou se os procedimentos de manutenção da linguagem estão presentes em algum documento que guie os bibliotecários ao fazer a inserção de termos novos, exclusão ou mesmo modificação. A G2, gestora do VocaUSP, afirmou que há um documento que apresenta essas orientações e também explicou que a tarefa de atualização da linguagem foi algo que não foi realizado com muita frequência, pois as bibliotecas passaram por dificuldades no período de pandemia da Covid-19. Relata que na Universidade houve programas de incentivo à demissão voluntária, aposentadorias e, infelizmente, óbitos de alguns funcionários. Tudo isso refletiu em serviços que eram realizados com frequência dentro da rede de bibliotecas e precisou ser interrompido.

Com o fim da pandemia, as rotinas voltaram ao normal, em especial, o grupo de gestão do vocabulário da USP. Outra situação relatada pela participante, foi no pós pandemia que apesar do retorno progressivo da normalidade, percebeu-se outras demandas importantes no cotidiano dos funcionários e alguns pediram para sair do grupo por não conseguirem assumir as exigências de serviço que acompanham a linguagem e seu grupo.

Isso corrobora com a premissa que há muitos desafios que permeiam uma organização, e que além dos quesitos que já são previstos na fase de planejamento da linguagem, como apontado pela ISO 25964 (2011), sobre os recursos humanos, financeiros e tecnológicos, junto ao estabelecimento de responsabilidades, nos deparamos com adversidades que muitas outras universidades enfrentam (com exceção da pandemia, pois fora algo fora do comum) como a falta de funcionários, grandes demandas de serviços e atividades, falta de tempo e etc. Mesmo assim, sobre isso, a G2 relata sua perspectiva pessoal em que reserva um tempo para se dedicar às tarefas de atualização da linguagem.

“É um trabalho grande? É um trabalho grande é para quem gosta muito? É para quem gosta muito. E eu costumo dizer que a gente tem muita coisa para fazer, mas às vezes você fala assim, olha hoje eu vou parar e vou mexer nisso, né?” (G2)

“Sim, e no dentro da ABCD tem uma página né que tem lá a política de indexação da USP, né? E

qual é o caminho para a gente inserir termos e sugerir troca de termos, né? Às vezes a gente precisa adequar.” (In2)

Na fala acima da In2, algo interessante é relatado e que em outro momento da entrevista com a G2 é falado também, sobre a adequação dos registros que já estavam presentes no catálogo e precisavam ser adequados com termos novos. Ambas costumam a fazer esse serviço no catálogo da rede, não apenas na linguagem de indexação.

Sobre esse bloco de perguntas, percebe-se a importância da documentação dos procedimentos para que sejam realizados de forma padronizada pelos bibliotecários que compõem a rede. Dessa forma, o trabalho dos indexadores é amparado por questões previamente estabelecidas facilitando e possibilitando a tomada de decisões.

❖ **BACKGROUND DAS LINGUAGENS DE INDEXAÇÃO**

Essa categoria apresenta questões de contextualização das linguagens de indexação, com o objetivo de verificar a trajetória de construção da ferramenta, considerando as necessidades que impulsionam o início do projeto, quais linguagens foram consideradas para compatibilização na etapa de desenvolvimento e aporte normativo.

A necessidade apontada pela G1 que contribuiu para a elaboração de uma linguagem própria da Unesp foi a recuperação da informação que apresentava disparidades nos resultados de busca. Foi uma demanda levantada pelos próprios bibliotecários. Com isso, na intenção de sanar os problemas que envolviam a indexação de assuntos; o grupo que antes era denominado “Grupo de terminologia”, buscou desenvolver uma política de indexação para a rede de bibliotecas da Unesp. Sobre isso, a In1, como uma funcionária mais recente na Universidade, não apontou o real motivo para o início desse projeto, mas declarou com razão que este seria um facilitador para o processo de indexação dos bibliotecários dentro da rede.

Além disso, assim como apontado pela In1 e pela G1, anteriormente, era utilizada a linguagem da Rede BIBLIODATA na indexação dos documentos pelas bibliotecas da UNESP. Uma concessão da Fundação Getúlio Vargas (FGV) que disponibilizava a base com os termos à Unesp. Essa linguagem não teria mais atualizações. Esse também foi um fator que contribuiu com que medidas fossem tomadas para a construção de uma linguagem própria. Com isso, já havia uma estrutura de termos do BIBLIODATA para a compatibilização, juntamente com o *Library of Congress Subject Headings* (LCSH) e por conta da área de assuntos da saúde também foi decidido utilizar na compatibilização o “*Medical Subject Headings*” (MeSH).

“Já havia sido feito uma pesquisa com os bibliotecários da rede, essa primeira pesquisa eu não participei e através né? Dos resultados identificou-se a necessidade de construção de uma política de indexação. Pessoal reclamava bastante, né? Da recuperação da informação” (G1)

“Eu não sei exatamente o que levou, me lembro, não me recordo disso, mas eu acredito que a busca é facilitar o trabalho dos bibliotecários dentro da rede, buscar trazer para rede o maior número de termos daquilo que a gente trabalha no nosso dia a dia” (In1).

“Antes de ter uma linguagem própria, a linguagem da Rede Bibliodata, né? E tratativas com ela na época da Coordenadoria Geral de Bibliotecas, a FGV, que na época em que era gerenciada já tinha linguagem né? Eles cederam pra UNESP a base de com os termos né?” (G1)

“Então a gente já já teve esse carga inicial dos registros, né?” (G1)

“[...] Mas basicamente a gente começou a trabalhar com LCSH e MeSH.” (G1)

Apesar de não se lembrar do número da norma, a participante G1 afirmou que a norma ISO foi utilizada para guiar a tomada de decisões na construção da linguagem. O percurso e os avanços do projeto foram registrados em vários documentos, a In1 mencionou a publicação do livro pela Prof. Mariângela Fujita e do Prof. Walter Moreira.

“Tem norma ISO que eu não lembro o número de cabeça agora, tá? Mas tem documentado lá no manual.” (G1)

“Não sei, mas acredito que sim. (In1)

“Sim, tem vários. Primeiro como... Depois da implementação da política, acho que a gente já teve uns dois documentos. Depois posso passar direitinho para você o histórico.” (G1)

“Tem o livro. Tem o livro publicado pela professora Mariângela.”

Sobre o resgate histórico do que impulsionou a elaboração do VocaUSP, a In2 aponta que na década de 80, com a informatização das bibliotecas da USP, todas foram criando suas próprias listas de cabeçalho de assuntos, de acordo com o acervo de cada instituição. Em 90, tendo em vista que seria necessária a elaboração de um vocabulário controlado para abranger todos esses assuntos, foi formado um denso grupo, com um representante de cada biblioteca e um docente do curso de Biblioteconomia, para realizar o trabalho de adequação dos termos, estruturação da linguagem e definição responsabilidades e de regras de utilização. A partir disso se deu a primeira versão do vocabulário. Em complemento com a fala de In2, a gestora G2 relata que a pioneira desse projeto, que se originou a partir de suas pesquisas de mestrado e doutorado, foi a Prof. Dra. Vânia Mara Alves Lima, orientada pela Prof. Dra. Nair Kobashi, da

ECA.

A G2 trouxe a informação de que antigamente, os assuntos utilizados para a indexação de assuntos, eram retirados de uma tabela do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) que não trazia a especificidade necessária para a descrição de assuntos e dificultava a recuperação. Isso também foi um fator crucial para impulsionar os estudos e as atividades que deram origem ao VocaUSP.

“Em 85 teve a informatização das bibliotecas aí todas as bibliotecas inseriram seus assuntos e foi criado uma lista de cabeçalho de assuntos, aí com o passar dos anos foi visto a necessidade de constituir um vocabulário controlado e diretrizes para fazer a indexação.” (In2)

“[...] foi um grupo de trabalho bem grande constituído por um bibliotecário de cada biblioteca e um docente do curso de Biblioteconomia aqui da ECA e a uma bibliotecária do Siro. Foi um trabalho longo que resultou na primeira versão do vocabulário controlado da USP.” (In2)

“[...] eu acho que o nome pioneiro desse trabalho é o da Vânia. Você deve conhecer a Vânia Mara Alves de Lima, professora da ECA. [...] trabalho dela de mestrado e doutorado foi em cima disso e ela era orientada e trabalhava com a Nair Kobashi, também da ECA.” (G2)

“[...] os assuntos que eram usados para a indexação no Dedalus eram era uma tabela do CNPQ. [P: Então era muito simplificado para a indexação] Nossa, eu tinha que por só Engenharia! Ou engenharia civil ou engenharia elétrica. Quer dizer uma coisa assim que não tinha cabimento. Como que faz para recuperar né?” (G2)

Em relação ao uso de outras linguagens de indexação e a compatibilização para a elaboração do vocabulário da USP, as bibliotecárias tiveram duas visões interpretações diferentes sobre o mesmo ponto. A In2 afirmou que foi utilizado o LCSH na indexação de assuntos da ECA, sendo que esses termos eram separados para a serem utilizados na descrição dos recursos informacionais de sua biblioteca e posteriormente, de forma indireta, foi utilizado no vocabulário da USP. Além de outras bibliotecas utilizavam outros vocabulários para atender às suas respectivas áreas do conhecimento.

A G2 foi para o mesmo sentido, mas disse que não foi utilizado nenhum vocabulário para a compatibilização, pois o objetivo era fazer um Vocabulário Controlado da USP, com a “originalidade” da Universidade. Mas que sua unidade, a Escola Politécnica da USP, utilizava-se a LCSH e outro ponto fundamental para a compilação e adequação dos termos foi o auxílio de docentes. Essa iniciativa de consulta com especialistas partiu da G2 e foi uma sugestão bem recebida pelo grupo de gestão do vocabulário.

“Foi. Foi sim, porque a como as outras bibliotecas, né do sistema tinha suas políticas e já tinha os seus vocabulários controlados, é algumas né? Não sei se todas. Por exemplo, aqui na ECA, a gente usava a lista da LC.” (In2)

“E aí foi feito uma análise dos termos que já eram usados e das suas linguagens que já eram adotadas ou por eles com a LC ou com a Biblioteca Nacional, o que a biblioteca originalmente adotava. E aí eles decidiram quais os termos que iam ser, mas aí já não se segue um vocabulário de assunto ou uma terminologia. Aí vai depender muito da área da biblioteca.” (In2)

“No início não. Nós queríamos fazer um vocabulário controlado da USP, em português, porque primeiro porque nós reunimos todas as especialidades. E segundo por causa das particularidades de cada área [...]” (G2)

“[...] subdivisões todas foram trabalhadas com os docentes aqui na Poli até hoje eu trabalho com docente porque eu sou bibliotecária, eu não sou engenheira.” (G2)

“O da engenharia foi trabalhado com os docentes, com cada especialista de cada área, sabe porque é o que você estava falando, uma coisa é a teoria e outra coisa é a prática. Então o assunto na prática é por isso que eu acho muito rico” (G2)

Sobre a utilização de normas para a elaboração do vocabulário houve uma certa incerteza, a In2 explicou que como não participou ativamente do grupo não tem essa informação, mas que acredita que sim, pois foi um trabalho realizado sob muitos estudos. A G2 esclareceu que a elaboração do vocabulário foi algo muito particular e que utilizou a experiência própria deles, então acredita que não foram utilizadas normas.

“A gente sempre esteve voltado a buscar o que nos atendesse mais, a nossa realidade [...] Eu acho que não, não foi uma coisa muito de experiência nossa” (G2)

“Eu acredito que sim, mas eu não como eu não participei do processo, eu não tenho esse conhecimento né, mas como foi um trabalho bem estruturado e tinha o professor da área de Biblioteconomia e as bibliotecárias que participaram elas eram bem comprometidas e engajadas com estudos e todos os processos. Eu acredito que tenha tido algo que norteou.” (In2)

Em um comentário feito pela In2, não necessariamente relacionada às perguntas realizadas nesse bloco, diz que sente falta de um controle de autoridades, pois ainda não possuem esse recurso e isso pode dificultar a recuperação da informação.

Toda a trajetória dos estudos realizados e resultados alcançados foram apresentados em eventos científicos da área de Biblioteconomia e Ciência da Informação, além de palestras oferecidas pelas gestoras do vocabulário para bibliotecários da rede e de outras instituições.

Em síntese, diante da trajetória das universidades, verificou-se o caminho percorrido para a elaboração das linguagens de indexação, enfatizando as dificuldades que antes havia na recuperação temática da informação, a falta de especificidade e consistência dos termos utilizados e a intenção de, além contribuir do aprimoramento da indexação e recuperação de assuntos, facilitar e amparar o trabalho do indexador. A partir dessas situações adversas, da

participação e percepção dos bibliotecários foi impulsionado a construção da política de indexação para posterior construção da linguagem de indexação, no caso da Unesp. Grandes etapas que foram trabalhadas de forma coerente, tendo em vista a dependência de uma etapa para concretização da outra. Essas etapas foram desenvolvidas por meio de projetos de pesquisa e deram frutos não somente com a política e linguagem de indexação da Unesp, como a publicação de livros, artigos, manuais de procedimentos e outros documentos que padronizam e retratam os procedimentos no tratamento da informação realizado pela rede. Com a USP, pioneira no Brasil na construção e disponibilização de um vocabulário controlado característico da universidade, o processo se iniciou pela própria construção da linguagem, baseada em estudos realizados no âmbito de pesquisa de docentes e no cotidiano dos bibliotecários. Além da pesquisa, utilizou-se da vivência e experiência dos bibliotecários integrantes da rede que contribuíram para esse projeto que até hoje é referência para outras bibliotecas e também são encontrados em diversas publicações do grupo precursor do vocabulário.

❖ ASPECTOS TÉCNICOS DA LINGUAGEM DE INDEXAÇÃO

Nos aspectos técnicos da linguagem de indexação busca-se compreender como são apresentadas as particularidades da forma das linguagens de indexação de cada biblioteca universitária, com os critérios e procedimentos adotados para a sua gestão e atualização.

A linguagem utilizada pela Unesp é o “Tesouro Unesp”, ou “Linguagem Unesp”, como chamado pela In1. Caracteriza-se por ser uma ferramenta multidisciplinar na abrangência das áreas do conhecimento. A bibliotecária G1 destacou a dificuldade de gerir a área de saúde, pois nessa área, representada pelos termos da linguagem do MeSH, os termos são muito específicos, dificultando a compatibilização, ocasionando alguns problemas de consistência. Relatou que há registros de autoridade, ou seja, são registros que contemplam vários outros com termos equivalentes e remissivas. G1 explica que a Linguagem Unesp não trabalha com termos isolados e únicos, pois devem estar relacionados com outros termos dentro da hierarquia, por isso, termos muito específicos podem ser difíceis de serem compatibilizados.

“[...] a linguagem é uma abrange todas as áreas do conhecimento.” (In1)

“[...] a gente não trabalha com termos isolados, a gente trabalha com registros de autoridade, e dentro de um registro você tem vários termos. Você não tem só o termo principal que é autorizado, você tem o equivalente, sinônimos, as remissivas. Assim você tem problemas quando é muito específico de uma área e você está compatibilizando com uma linguagem que não é tão específica né? Então a ideia é que o tesouro seja multidisciplinar”

O Tesauro Unesp conta com o recurso de siglas para orientar o uso dos termos pelos bibliotecários e/ou usuários. Apresenta-se em Língua Portuguesa, mas em vários casos os termos são inseridos em inglês no formato de remissivas, o que pode contribuir na hora da indexação de livros estrangeiros e da busca. Além disso, são utilizados qualificadores para diferenciar termos com escrita semelhantes. Segundo a G1, há um software chamado TemaTres que recebe os registros de autoridade que são elaborados na plataforma utilizada na rede de bibliotecas da Unesp, o Alma. Esse software possui uma interface em que é possível visualizar as tags que orientam o bibliotecário acerca do uso dos termos autorizados.

“O tesauro, ele vai para web em um software diferente do que a gente usa dentro da biblioteca para fazer o gerenciamento dos registros de autoridade, né? Que é o Alma. O Alma não tem uma versão pesada. Vamos dizer assim... A gente usa um Software livre chamado TemaTres [...] ele foi configurado para receber esses registros de autoridade que são trabalhados no formato Marc 21, ele foi configurado para receber esses registros e para que a gente visualize em formato de Tesauro e nesta visualização você tem as tags” (G1)

Sobre o Tesauro Unesp ser multilíngue, a In1 apresentou um pouco de dúvidas, pois como a biblioteca em que trabalha contempla em seu acervo recursos da área médica, ela utiliza muito em seu cotidiano o DeCs que é multilíngue e por ser uma linguagem mais específica e autorizada para ser consultada e posteriormente inserir o termo no Tesauro Unesp. A G1 explica que no Tesauro Unesp há termos em inglês, mas que geralmente são termos equivalentes, especialmente por serem importados de outras linguagens como da LCSH.

“Nossa, eu confundo com o DeCs que tem tudo isso. Eu ia até fazer uma comparação com DeCs que ele tem a nota de escopo, a nota de indexação que ele dá uma ideia melhor que eu não me lembro de ter visto uma linguagem no UNESP, né? Mas até nessa questão da nota de escopo, o “use este” e tal, o DeCs faz de uma forma com muito mais clareza...” (In1)

“Ele não tem a versão. Mas ele tem termos em quase todos os registros, né? Que eles têm um termo equivalente, equivalente em outra língua, tá geralmente é o inglês porque veio da LC então ele quando você vê o tesauro você vê termos em inglês também, mas não que ele seja...” (G1)

A partir do relato de In1, percebe-se como as notas de escopo são importantes para guiar as decisões dos indexadores ou mesmo de quem está consultando a linguagem, pois é de acordo com a definição apresentada nessas notas que será possível ter maior clareza sobre o significado

dos termos e se é plausível a sua utilização para descrever o tema desejado.

No tesauro são utilizados qualificadores que ficam entre parênteses para diferenciar um termo com a mesma grafia, mas com diferentes significados, como o exemplo utilizado na pergunta: Mercúrio (planeta) / Mercúrio (elemento químico) / Mercúrio (mitologia). Sobre a orientação de como os termos devem se apresentar gramaticalmente, por exemplo, termos no plural ou singular e o uso de verbos, a In1 aponta que essas informações estão definidas na política, a G1 reforça que essas orientações foram baseadas na norma ISO, a partir do momento em que se começou a estudar a norma para a elaboração da linguagem e que isso está definido no manual de construção e atualização do Tesauro Unesp. Também relatou que constantemente são realizadas correções de termos que foram importados de outras linguagens, como o BIBLIODATA e que não estão de acordo com as diretrizes técnicas estipuladas no manual, portanto são feitas essas adequações.

“Eu acho que na política deve ter alguma coisa a respeito disso e na própria linguagem a gente tem de ver o USE desta forma, eu acho que tem sim as siglas, né para falar qual que é o certo.” (In1).

“Sim, a partir do uso quando a gente começou a conhecer a norma ISO, tá?! É, nessa norma existe essa orientação, então o manual, no caso, o manual de construção e atualização do Tesauro que a gente tem hoje que é o documento mais recente, ele tem essas orientações para o bibliotecário tá?”

“E aí a gente vem corrigindo quando a gente detecta o erro, mas existe sim hoje uma orientação para isso, né? É basicamente é coisas contáveis no plural não contáveis do singular, algumas exceções para termos que usam os nomes científicos, né? É partes do corpo, doenças, algumas exceções.”

Em relação aos termos compostos, a entrevistada G1 os chama de termos com subdivisão e ressalta que alguns deles são de fato essencialmente compostos, mas busca-se por meio de recomendações evitar termos subdivididos e que são utilizados como termo principal na linguagem de indexação, diferente da combinação de termos realizada pelo catalogador no momento da indexação de assuntos. Relatou-se que uma dificuldade encontrada ao lidar com os termos compostos uma vez que são derivados de linguagens traduzidas, pois pode haver imprecisões na qualidade da tradução.

A In1 relata que há disparidades das informações contidas na política e na permissão do sistema (Alma) utilizado pelas bibliotecas em fazer isso, isto é, há possibilidade de utilizar termos compostos ou unir termos no momento da indexação, mas o sistema muitas vezes não permite essa ação.

“[...] a gente chama de subdivisão, né. Alguns termos são compostos já na raiz e outros são compostos porque foi usado uma subdivisão no termo principal, tá? A ideia é evitar esses termos com a subdivisão. A gente deixa isso para o catalogador fazer na hora da catalogação da obra dele, porque não há alguns desses termos que não são principais que são usados só como subdivisão, eles podem ser usados sobre vários termos né? Então você não precisaria você teria a orientação para como usá-lo separadamente e o catalogador faria essa combinação na hora de catalogar a obra, mas existem os termos que já na sua natureza são compostas.” (G1)

“É a gente encontra aí um pouco de dificuldade de trabalhar com esses termos, porque alguns eles são termos diretos por exemplo na LC, mas aqui já foi traduzido com quebrado, sabe já foi colocado hífen, então a gente fica um pouco em dúvida de como quem fez isso a instituição que fez isso primeiro porque que ela chegou a essa conclusão” (G1)

“Você pode juntar, você pode juntar. Então pode, mas nós estamos tendo problemas no Alma para isso em termos compostos e a informação que se tem é que isso vai ser revisto na política [...] Então a política ela prevê isso, mas o sistema hoje de trabalho ele não aceita da forma como está na política. Então tem que ser revisto ou um ou outro.” (In1)

A linguagem de indexação da USP, como já dito, é o VocaUSP que abrange todas as áreas do conhecimento que são desenvolvidas na Universidade, apresentam algumas *tags* e é uma linguagem em português. Os qualificadores, de acordo com a G2, são utilizados no vocabulário e não na indexação. Há orientações de como devem ser os termos gramaticalmente e estão disponíveis para os bibliotecários da rede e são consultadas essas diretrizes sempre que for sugerido ou termo ou analisar para compor o vocabulário.

Sobre termos compostos, ambas as bibliotecárias da USP falaram sobre o termo ser consolidado de forma composta que nesse caso não há como separá-los para não correr o risco de se perder o significado.

“Sim, utilizamos. Isso está na hierarquia, ele não é acrescentado na indexação, ele é montado desse jeito dentro da área dele porque, por exemplo, eu podia por... existir o termo uma vez só e cada vez que eu fosse indexar eu usasse o qualificador de área, não. Já está na árvore hierárquica o qualificador identificando.” (G2)

“Também existe isso está lá na orientação de como usar, como montar e como usar.” (G2)

“Mas no nosso vocabulário tá explicando quais são os critérios [...] Ela não está em nota pública, ela não tá pública” (In2)

“Existe algumas situações daquilo que está consolidado, como área de estudo ou de interesse é colocado “gestão de projetos”, “gestão de pessoas”, não é separado porque é um conceito só.” (G2)

“Eles não gostam que a gente use um termo composto. A não ser que seja bem sacramentado na área para não perder o significado, né?” (In2)

Com isso, verificou-se que tanto o Tesouro da Unesp como o VocaUSP, são apresentados na língua portuguesa e multidisciplinares. Na realidade da Unesp, constatou-se ainda desafios presentes na compatibilização da linguagem com outras mais específicas, como é o caso do MeSH. A equipe da Comissão que apesar de pequena, é muito empenhada em fazer a manutenção e atualização do Tesouro. Em ambos os casos, percebe-se a necessidade de recomendações aos bibliotecários de como contribuir para a gestão da linguagem de indexação, especialmente no que diz respeito a questões técnicas como visto nesta categoria.

GRUPO GESTOR

O grupo gestor possui importância e responsabilidade muito grande para elaborar o projeto e continuar com a manutenção constante da ferramenta. Por isso, essa categoria busca conhecer como foi montado o grupo de gestão das linguagens de indexação e seus principais contribuidores.

Segundo a G1, o grupo de gestão do Tesouro Unesp foi iniciado por volta de 2020. Iniciou-se a partir de um grupo que fazia os registros de autoridade, posteriormente esse grupo acabou e formou-se a Comissão Permanente do Tesouro. O grupo, também chamado de comissão, conta com seis bibliotecários, a maioria estando em diferentes unidades da Universidade. Atualmente, a bibliotecária entrevistada, G1 é a principal responsável que está à frente dos assuntos que concernem a linguagem e possui autonomia para tomar decisões em prol da gestão da ferramenta. A Prof^ª Dr^ª Mariângela S. L. Fujita que no início coordenou o projeto da Linguagem Unesp, hoje atua como consultora na comissão. Assim como relatado pela In1, a principal responsável é a G1. Além disso, também informou que o grupo ofereceu treinamentos para os bibliotecários da rede.

“Hoje existe uma comissão que são seis pessoas tá? São seis pessoas, são todos bibliotecários, é de unidades distintas da Unesp, né? Se bem que a gente tem três pessoas de uma mesma unidade, mas a gente tem. Hum, a ideia que tivesse uma pessoa para cada área do conhecimento, mas a gente não conseguiu.” (G1)

“Olha 2020?! Foi no começo da pandemia, porque uma vez implantada a linguagem a gente conseguiu fazer o Tesouro ficar disponível.” (G1)

“A gente tem hoje a prof. Mariângela como consultora na comissão e os seis bibliotecários que trabalham com a manutenção, né?” (G1)

“Assim faz tempo, mas eu lembro que teve sim.” (In1)

O grupo de gestão do VocaUSP possui um representante de cada área do conhecimento, apesar das baixas de funcionários que tiveram por causa da pandemia de Covid-19, como apontado pela G2. Não são todas as bibliotecas que elegeram um representante, mas sim se considerou as áreas do conhecimento. A In2 explica como funciona a dinâmica do grupo, em que são contemplados três níveis de serviço e cada um possui uma função para em relação à sugestão do termo realizado pelo bibliotecário indexador. Relata que sempre há uma devolutiva por parte do grupo, mesmo se o termo não for considerado para compor a linguagem. Essa resolução é interessante e divide a trabalhosa tarefa de manutenção do tesauro, garante a participação de outros profissionais e apresenta uma forma sistêmica de análise dos termos.

Apesar da In2 não ter participado do grupo de gestão e manutenção do VocaUSP, ela presenciou por muitos anos a participação de duas bibliotecárias de sua unidade, a ECA, e também falou que em alguns momentos essas colegas participaram de treinamentos e reuniões, além de orientações passadas do grupo aos bibliotecários da rede.

“Atualmente tem um representante na área de ciências humanas, um ciências biológicas e três das ciências exatas que são bibliotecários de bibliotecas da USP, além deles tem o pessoal do DGTI que é da Divisão Geral de Tratamento da Informação.” (In2)

“E se por um acaso o termo não for acatado a gente tem uma devolutiva. E aí eles falam o que a gente precisa fazer, se é sugerir um novo termo, se é mudar a hierarquia, se ele é se eu de repente sugerir na área de ciências humanas, mas ele é um termo também da área de exatas” (In2)

“[...] a gente teve uma reunião geral com todos os bibliotecários da USP que trabalhavam com a indexação participaram de um treinamento também, mas o grupo gestor ele tem esses treinamentos teve um acho que faz uns dois anos.” (In2)

A partir dos depoimentos das bibliotecárias, foi possível constatar a importância da formação de um grupo gestor da linguagem, com especialidades distintas e trabalho colaborativo. Todos com um objetivo em comum. Por serem linguagens multidisciplinares, percebe-se a relevância da participação de bibliotecários com a vivência em cada área do conhecimento, como visto na experiência relatada da USP, além do apoio de especialistas para melhor compreensão no estudo realizado no momento de avaliar um possível termo para compor a linguagem.

❖ CRITÉRIOS E FREQUÊNCIAS PARA ATUALIZAÇÃO

Essa categoria almeja identificar os critérios utilizados pelos bibliotecários para a inserção e exclusão de termos, visando a comparação entre aqueles que são previamente estipulados e a perspectiva pessoal do bibliotecário. Também busca-se verificar quais os

procedimentos necessários para a atualização e se são feitas adequações em termos já existentes na linguagem.

Em uma perspectiva mais ampla, G1 relata que a inclusão de termos novos geralmente ocorre por meio de pedidos de alunos e docentes. Lembrou-se que através do Repositório Institucional há constante entrada de produções científicas que possuem originalidade e termos novos ainda não estejam presentes no Tesauro. Contudo, quando perguntado quais eram os critérios que ela mesma utilizava para considerar a inclusão de um termo, ela se voltou à prática da catalogação de livros. Apesar de estar com outras atribuições fora a catalogação, ela se ateve a entrada de novos documentos. Mas destacou que para isso faz uma busca extensa nas bases de dados, bibliografias e até mesmo na CDD para conseguir respaldo para a sua decisão.

A In1 relata que há um formulário disponibilizado aos bibliotecários da Unesp que deve ser preenchido com o pedido e suas justificativas, seja para adicionar um termo, modificá-lo ou excluí-lo. Em sua perspectiva pessoal, In1 utiliza o recurso de solicitação de novos termos de acordo com a sua demanda de trabalho, em especial quando não encontra um termo que seja específico para a descrição de um trabalho. Mesmo assim, diz que costuma consultar bases de dados que são internacionais, como o DeCs e a LCSH, bases utilizadas no Tesauro Unesp, além de lembrar que já houve casos que consultou um docente para embasar o seu pedido.

“Quem detecta a necessidade de inclusão de um termo novo, geralmente é o bibliotecário ou por solicitação de um docente ou de um aluno. Às vezes na hora de fazer o autoarquivamento no repositório o aluno não consegue encontrar o termo que precisa porque existe uma orientação, né? Tem que ter pelo menos três termos do Tesauro e os outros podem ser livres [...]” (G1)

“Faz tempo que eu não catalogo, né? Mas é, às vezes é a falta mesmo você começa olhar ali falando, poxa, caramba, esse tema, então eu vejo todo dia fluindo e tão utilizado porque que não tem ainda. E aí você tem que buscar. Será que não tem mesmo?” (G1)

“De acordo com a demanda do nosso dia a dia de trabalho. Então eu tô fazendo a indexação de um trabalho e é um termo muito específico e que eu não localizo em nenhuma das bases inclusive nas bases internacionais. Então eu tento buscar essas informações normalmente eu já tive acho que duas situações dessa e eu pedi inclusive ao próprio docente” (In1)

Sendo assim, o procedimento para a inclusão ou mesmo exclusão de um termo ocorre via formulário. Esse formulário foi elaborado especialmente pelo o grupo, utilizando um método chamado “termográfico”, como explicado pela G1. Esse formulário deve ser preenchido pelos bibliotecários e enviado para a Comissão que irá analisar os pedidos. A partir disso, ele é incluído pela Plataforma Alma, para posteriormente ser lançado no *software*

específico do tesouro. A In1 apesar de não conhecer o nome do software utilizado para a disponibilização da linguagem, diz ser fácil de visualizar as informações, mas que o ideal seria apresentar mais notas explicativas sobre os assuntos.

Segundo G1, a frequência de atualização não é muita alta atualmente, tendo em vista que a linguagem já está bem estruturada. Por isso, a atualização se dá quatro vezes ao ano e considerou que o processo não é complexo, mas demorado.

“E esse formulário ele é... tem lá os campos que o pessoal tem que preencher é... dizendo o porquê é importante aquele termo de fato que é de consultado para chegar à conclusão de que aquele tempo que ainda não existe é necessário se tem equivalente em outra linguagem. Enfim, a gente tem esse formulário que a partir desse formulário a gente é uma comissão, né, mas o grupo de linguagem avalia a inclusão ou não do termo.” (G1)

“Acredito que as informações poderiam estar melhor disponibilizadas ali ou poderia ter, como eu citei há pouco, uma nota de indexação, sabe? Para alguns termos eu acho que faltam alguns elementos a mais nas informações colocadas ali” (In1)

As correções são realizadas sempre que necessárias, pois é possível gerar relatórios pelo sistema e verificar possíveis erros que precisam ser corrigidos, alguns podem ser feitos manualmente, porém outras situações que estão em maior quantidade são alteradas em lote com o auxílio do profissional de informática que dá assistência necessária para essa atividade. A G1 ressalta que é complexa a tarefa de excluir um termo, por exemplo, para um termo que é politicamente incorreto ou obsoleto, a orientação é colocá-lo como uma remissiva, um termo não preferido, na qual indica-se o termo preferido.

“E aquele termo que não é mais utilizado que tá obsoleto, ele vai constar como uma remissiva.” (G1)

“Algumas coisas a gente consegue pedir pra ele tirar em lote. Detectou isso nesse campo. Apaga. Agora tem coisas que você precisa corrigir manualmente. [...] o termo aparece às vezes em vários outros indícios, né? [Uhum.] Então você não pode arrumar por exemplo os temas que eu falei obsoleto você não pode arrumar o principal e esquecer dele. Tem outros”

De acordo com G2 e In2, a forma de solicitação é a mesma onde será necessário apresentar uma pesquisa prévia embasada sobre o termo e colocar o maior número de informações possíveis para que seja analisado pelo grupo técnico da ABCD. Esse formulário se encontra online e se chama “base de sugestões”. A G2 disse que por ser uma das gestoras do

vocabulário, ela tem permissão para realizar essas alterações da área de domínio de sua biblioteca, as Engenharias, sem esperar uma posição do grupo. Sobre uma perspectiva pessoal das bibliotecárias a respeito dos critérios utilizados para solicitação de inclusão de algum termo, a In2 explica que esse processo ocorre sob demanda de trabalhos, se há muitos registros que estão falando sobre um determinado tema, ela opta por inserir. A G2 diz que é necessário passar um tempo e observar se um termo se torna consolidado pela área. A devolutiva sobre a solicitação dos bibliotecários é realizada através de e-mail. Então sempre haverá uma resposta às solicitações, seja ela positiva, que necessita de maiores informações ou negada.

Como dito anteriormente e reforça-se aqui, uma prática de ambas as bibliotecárias da USP é realizar correções dentro dos registros que anteriormente estavam com termos considerados inadequados ou não suficientes para a descrição daquele tema. In2 exemplifica os registros com o tema indígena, em que antes utilizava-se “índios” como um termo autorizado. Contudo, com os recentes estudos e discussões a respeito disso, verificou-se que ideal seria utilizar “indígenas” na representação desse assunto. Por isso, a adequação retrospectiva da base de dados é necessária.

“O que eu uso é assim, se eu vejo que tem muito material que tem um termo que não existe no vocabulário da USP e ele tá sendo muito utilizado pelos pesquisadores, aí a gente para pra fazer essa análise dentro da literatura da área e pedir a inclusão...” (In2)

“[...] eu indico qual é a área dele, eu indico eu dou uma definição para este termo, eu indico a bibliografia de onde foi tirada essa definição, ele é um formulário que você preenche tem os campos prontos para você preencher é bem feito aquilo então aquilo vai ...” (G2)

“Eu acho muito importante quando a gente tem material para pedir um assunto, quando esse assunto é acrescentado na hierarquia, a gente tem que entrar na base e colocar esse assunto dos nossos registros, sabe? se não perde o sentido”

Nos quatro relatos, percebe-se que esses critérios são subjetivos, ainda mais por serem critérios pessoais. Aqui verificamos as questões que norteiam os conceitos de garantia de uso e garantia literária, vistos na seção de revisão de literatura desta tese: Quantos trabalhos são necessários para merecer estarem contemplados na linguagem de indexação? Quanto tempo é necessário passar para que um assunto se torne consolidado na área?

Em ambos os casos, as solicitações dos bibliotecários são realizadas por formulários, o que facilita a tarefa do grupo gestor. Na realidade das bibliotecas da USP, percebe-se que há uma equipe maior para trabalhar com a gestão e manutenção do vocabulário, o que pode contribuir para a realização das incumbências da linguagem.

❖ SOFTWARE DE GESTÃO E MANUTENÇÃO

O software de gestão e manutenção é uma ferramenta que viabiliza o funcionamento e a usabilidade da linguagem de indexação. Por isso, essa categoria busca verificar se os bibliotecários têm conhecimento desse recurso e saber se este atende as necessidades da instituição.

O software de gestão e manutenção da do Tesouro Unesp é o Tematres, um software livre que atende muito bem as necessidades e especificidades da linguagem multidisciplinar da Unesp, segundo a participante G1. Porém, por ser um *software* livre ele não possui suporte para eventuais dúvidas ou adequações.

“o software TemaTres é um software que atende as necessidades da instituição”

“Que deixa desejar em algum aspecto o TemaTres... Eu acho que isso eu sinto falta de ainda saber mexer um pouquinho mais no software.”

“[...] Ele é bem completo. A gente só não sabe mexer em todas as funções que ele tem, as funcionalidades.”

“Embora ele seja livre, mas no sistema ele é livre, né? Então não tem um suporte, por isso essa instituição, a gente não sabe não tem nem para quem perguntar sabe?”

O *software* utilizado pela rede de bibliotecas da USP, assim como o da Unesp, é o Tematres. Recentemente passou a ser utilizado, mas de acordo com In2 é um *software* que ela já tinha o conhecimento devido ao uso por outras bibliotecas e elogia a interface amigável do programa. G2 conta que muitos estudos pelo grupo são realizados a respeito de melhorias no vocabulário.

“Sim, a gente já tinha o conhecimento do Tematres em outros vocabulários, inclusive aqui na ECA. A gente tá a gente passou o nosso a nossa base de dados local que era em Isis pro TemaTres, né? E também utilizando no vocabulário da Biblioteca digital de imagens, ele é muito mais amigável do que a gente tinha antes, a interface.” (In2)

O software é algo que pode ser considerado como uma questão técnica e específica para aqueles que fazem a gestão das linguagens, não chegando ao conhecimento dos demais bibliotecários da rede. Mas é uma ferramenta imprescindível para facilitar a compreensão dos

usuários da linguagem, sejam eles bibliotecários, pesquisadores ou alunos. Por isso, é importante que seja verificado na fase de projetos da linguagem, a possibilidade de investimento em um software capaz de abranger a linguagem de indexação de acordo com as necessidades da instituição.

❖ **DISPONIBILIZAÇÃO ONLINE DA LINGUAGEM PARA CONSULTA**

A disponibilização da linguagem de indexação no catálogo online da biblioteca contribui para o usuário realizar as suas buscas de forma que seja possível elaborar uma estratégia de busca condizente com os termos utilizados para a representação dos documentos na indexação. Por isso, essa categoria tem como objetivo verificar se as bibliotecas dispõem dessa ferramenta de forma acessível e *online* e se os bibliotecários percebem que ela é uma ferramenta conhecida e utilizada pelos usuários.

O Tesouro Unesp está disponível online no catálogo de buscas, assim como relatado por G1 e In1. Na questão relacionada se os usuários conhecem esse recurso dentro do catálogo Athena, a In1 acredita que os pesquisadores conhecem a existência e a importância de outros vocabulários, mas não tem certeza sobre o uso do Tesouro Unesp. Segundo G1, houve a intenção de fazer um treinamento com os ingressantes à distância na época da pandemia, a respeito do uso do Tesouro, mas não foi possível por questões técnicas com a desatualização do *software*. Mas que os esforços de se gerir e manter uma linguagem é para que, também, o usuário conheça e se beneficie dessa ferramenta. Ressaltou-se que há tutoriais sobre o tesouro e seu uso tanto na página do Repositório, como no programa de solicitação de ficha catalográfica.

“A gente trabalha para isso, tá? É a gente fez treinamentos no início com os bibliotecários, para que eles orientassem a comunidade” (G1)

“Então não teve um treinamento é para a comunidade, mas a gente treina e orienta os catalogadores para aqui a comunidade saiba da existência” (G1)

“Eu acredito assim, os pesquisadores e para os alunos que fazem pesquisa, sim, eles têm esse conhecimento, né? [...] se usam é outra história, mas ele tem esse conhecimento, mas conhecimento do Tesouro Unesp... Não sei não, acredito que não chega até eles...” (In1)

Segundo G2, o VocaUSP está na internet e pode ser utilizado a qualquer hora, em qualquer lugar pelos usuários. Acredita que seus usuários conhecem essa ferramenta, pois há uma bibliotecária que realiza treinamentos com os estudantes de graduação e pós-graduação e que é apresentado a linguagem e suas formas de uso. Em contrapartida, a In2 apesar de também dizer que as bibliotecárias de referência apresentam esse recurso em seus treinamentos, acredita que eles desconhecem a função do vocabulário. Essa percepção ocorre quando chegam os trabalhos acadêmicos com as fichas catalográficas geradas de forma automática. Nos campos dos descritores, os alunos autores utilizam termos livres.

“Está na internet disponível para qualquer um qualquer lugar do mundo.” (G2)

“Aqui na Poli nós temos uma bibliotecária que faz esta apresentação é para alunos de graduação e de pós-graduação.” (G2)

“Não, eles não têm esse conhecimento. Eu acho que assim é divulgado, a gente divulga que na biblioteca, as bibliotecárias de referência falam quando orientam uma pesquisa, mas em geral o usuário não sabe que tem esse controle de vocabulário que ele poderia procurar de alguma outra forma ter um resultado melhor na pesquisa no catálogo”. (In2)

“Dá para ver também assim pelas fichas catalográficas que são geradas automaticamente pelo formulário da biblioteca da ECA que eles não utilizam os termos do vocabulário da USP eles usam os termos, nas palavras-chaves, termos livres.” (In2)

Nas duas realidades há a disponibilização online das linguagens de indexação e são realizados treinamentos com dos bibliotecários, especialmente na seção de referências, para ensinar os usuários da biblioteca, o que mostra comprometimento e conhecimento dos profissionais na função de mediação das ferramentas de controle de vocabulário.

Síntese dos resultados da entrevista

As considerações sobre as respostas obtidas pelas entrevistas concedidas, já foram feitas em cada categoria que organizou esta seção. Por isso, na síntese dos resultados, buscamos fazer uma compilação e comparação, se possível, dos principais pontos trazidos pelas gestoras da linguagem e indexadoras. Dessa maneira, será possível apoiar a elaboração da próxima seção da tese.

Quadro 13 - Pontos de destaques da entrevista

Categorias	Destaques	
	Indexadoras	Gestoras
Informações sobre a biblioteca	Indexadoras experientes em suas áreas de atuação na biblioteca.	Gestoras com experiência e conhecimento sobre a trajetória da L.I.
Política de indexação	Política de indexação local (In2) Linguagem da rede e linguagem local (In2)	Política de indexação geral (G2)
Background das linguagens de indexação	Auxiliar a tarefa do indexador (In1) Informatização das bibliotecas (In2) Inclusão dos bibliotecários da rede na gestão da construção da L.I. (In2) Compilação dos termos: - Uso de outras L.I. (In2) - Lista de cabeçalho de assuntos das próprias bibliotecas (In2) - BN e LCSH (In1) Amparo normativo (In1, In2) Publicação dos trabalhos realizados (divulgação) Uso de outras linguagens (MeSH/DeCs, LCSH) (In1) Falta de notas de escopo (In1)	Disparidades na recuperação da informação (G1) Desatualização da L.I. (anteriormente) utilizada (G1) Falta de especificidade e inconsistências dos termos (G2) Projetos, início do grupo e cronogramas de atividades (G1 e G2) Inclusão dos bibliotecários da rede na gestão da construção da L.I. (G2) Adversidades falta de funcionários e pandemia (G2) Compilação dos termos: - Adequação dos termos importados (G1) - Compatibilização (LCSH, MeSH) (G1) - Lista de cabeçalho de assuntos das próprias bibliotecas (G2) - Auxílio de especialistas (docentes) (G2) Amparo normativo (G1) Não utilizaram normas, mas a vivência (G2) Publicação dos trabalhos realizados (divulgação)
	Termos compostos: - Disparidades da política de indexação com as permissões do sistema “unir” termos na catalogação (In1) - Utiliza termos compostos quando consolidado na área (In2)	Especificidade de áreas do conhecimento (saúde) e dificuldade de compatibilização; (G1) Língua Portuguesa, mas apresenta termos em inglês nas remissivas e como termos equivalentes (G1);

Aspectos técnicos da linguagem de indexação		Uso de qualificadores a partir da norma ISO (G1) Qualificadores apenas no VC e não na indexação (G2) Correção de termos importados de outras bases (G1) Termos compostos: - Dificuldades com os “termos com subdivisões”, em especial na tradução de outras linguagens (G1) - Utiliza termos compostos quando consolidado na área (G2)
Grupo gestor	Houve treinamentos para os indexadores e todos os bibliotecários da rede (In1, In2) Bibliotecários representantes por áreas do conhecimento (In2) Divisão do fluxo de trabalho (In2)	(6) Bibliotecários em diferentes campus da Unesp (G1) Bibliotecários representantes por áreas do conhecimento (G2) Houve treinamentos (G1)
Critérios e frequências para atualização	Inclusão de termos novos: - Entrada de novos recursos informacionais (In1, In2) Correções, exclusões e inclusão são via formulário (In1, In2) Respaldo em outras linguagens e especialistas (In1, In2) Correção retrospectiva dos registros com termos modificados (In2) Falta de controle de autoridades (In2)	Inclusão de termos novos: - Pedidos de alunos e docentes (G1) - Entrada de novos recursos informacionais (G1) - Avaliação do termo por um período (G2) Correções, exclusões e inclusão são via formulário (as gestoras têm autonomia de pular esse processo) Respaldo em outras linguagens e especialistas (G2) Correções são periódicas, sendo manuais ou em lote (G1) Correção retrospectiva dos registros com termos modificados (In2 e G2)
Software de gestão e manutenção	- Software Tematres: - Atende bem (In1) - Fácil visualização (In1, In2)	- Software Tematres: - Atende bem (G1) - Falta de suporte técnico (G1) - Falta aprender todas as funcionalidades (G1)
Disponibilização online da linguagem para consulta	Está disponível online Usuários conhecem, mas não usam (In2)	Está disponível online Usuário tem conhecimento (G1, G2) - O conhecimento e uso pelos usuários é um objetivo (G1 e G2)

	Usuário conhece outras linguagens, mas não da instituição (In1)	- Disponível tutoriais e treinamentos (bibliotecária de referência) para seu uso (G1 e G2)
--	--	--

Fonte: Elaborado pela autora.

A partir do Quadro 13 com o compilado identificando os destaques das respostas obtidas com a entrevista, é possível considerar que tanto as gestoras como as indexadoras têm uma experiência singular sobre as linguagens de indexação de suas instituições, cada uma com seus conhecimentos adquiridos. Percebe-se que as gestoras possuem uma visão mais voltada para a própria ferramenta, suas qualidades, possibilidades, recursos, melhorias e etc. As indexadoras têm uma percepção maior dos resultados que a linguagem oferece dentro de suas unidades, como usabilidade, experiências, dificuldades e facilidades, por exemplo.

Sendo assim, essa rica contribuição que as bibliotecárias ofereceram através da entrevista possibilitou a elaboração das recomendações sobre o gerenciamento da construção e manutenção da linguagem de indexação, apresentadas na próxima seção, considerando não apenas a parte teórica e normativa vista nos capítulos de revisão de literatura, como também a vivência de duas importantes redes de bibliotecas brasileiras que documentaram todo o processo do percurso percorrido almejando melhorias no tratamento da informação.

7 RECOMENDAÇÕES SOBRE O GERENCIAMENTO DA CONSTRUÇÃO E MANUTENÇÃO DE LINGUAGENS DE INDEXAÇÃO

Nesta seção, buscou-se reunir os conhecimentos adquiridos nas discussões realizadas por meio da fundamentação teórica e dos resultados alcançados com as metodologias de pesquisa aqui empregadas, considerando o terceiro objetivo específico determinado em: “elaborar orientações sobre o gerenciamento da construção e manutenção de linguagens de indexação na política de indexação de bibliotecas universitárias, tendo como parâmetro a literatura técnico-científica e a *práxis* das bibliotecas universitárias.”

Com isso essas recomendações são divididas em quatro fases: política de indexação, planejamento, gestão da construção e gestão da manutenção.

→ 1ª fase: Política de indexação

Elaborar uma política de indexação como uma etapa anterior a linguagem de indexação pode ser vantajoso, tendo em vista que é através da política que serão definidos parâmetros que influenciam o processo de indexação e consequentemente a própria linguagem. Nesse sentido, a linguagem de indexação estará atuando como um espelho das diretrizes e decisões prescritas no documento por seus administradores. Por isso, considera-se importante seguir uma política única na instituição, mesmo que esta seja composta por diversas unidades, como visto na realidade da USP e da UNESP, para que isso garanta a uniformidade dos procedimentos empregados no processo de indexação por todos os bibliotecários da rede. Assim como o uso de uma única linguagem de indexação.

→ 2ª fase: Planejamento

O planejamento como etapa preliminar é relevante e atuará como um mapa com onde começar e para onde ir. Incluir bibliotecários que sejam experientes em determinadas áreas do conhecimento pode ser um fator engrandecedor para a linguagem e também poderá contribuir de forma mais efetiva para os bibliotecários da rede, no seu trabalho de indexação. Após a implantação da política de indexação algumas informações já estarão previamente definidas. Com isso, será possível partir para os seguintes pontos que foram baseados essencialmente na ISO (25964):

1- Objetivos

- Responder as perguntas:
 - Qual a necessidade da instituição em construir uma linguagem de indexação?
 - Por quem será utilizado?
 - Será possível a colaboração de outros bibliotecários da rede?
- Delimitar o objetivo do projeto;
- Área(s) temática(s) cobertas pela linguagem de indexação (são as mesmas abrangidas pela coleção da(s) biblioteca(s)).

2 - Características da linguagem de indexação

- Tipologia da linguagem: será uma lista de cabeçalho de assuntos, um tesouro, anéis de sinônimos?
- Escopo: o propósito da linguagem de indexação;
- Estilo de exibição: se no formato impresso ou no eletrônico (recomenda-se que seja uma ferramenta online para que seja passível de atualizações constantes e a participação de outros membros do grupo, como forma mais facilitadora);
- Periodicidade das atualizações: refere-se ao intervalo de tempo que se almeja fazer as atualizações e manutenções.

3 - Recursos

- Recursos humanos: fundamental não somente na fase de gestão da construção da linguagem de indexação, mas também em sua continuação, na gestão da manutenção com as atualizações e tomadas de decisões posteriores.
- Nos recursos humanos, considerar também a demanda de trabalho dos responsáveis, para que seja possível o comprometimento com a L.I.
- Recursos financeiros: previsão de gastos com a ferramenta ou mesmo com capacitações dos responsáveis.
- Aporte tecnológico: fazer um estudo na literatura e no mercado para avaliar quais tecnologias são necessárias para comportar esse projeto.

4 - Estabelecer as responsabilidades

- Coordenador(a) do grupo: responsável pela organização do grupo, gestão do projeto, fiscalização das atividades e possui autonomia para a tomada de decisões.

- Grupo de revisão: bibliotecários que estão no grupo de gestão da linguagem de indexação e possuem responsabilidades com a manutenção da ferramenta, com a validação dos termos.
- Consultoria: especialista, possivelmente um docente, para tirar dúvidas e auxiliar nas decisões a serem realizadas.
- Equipe: considera-se todos os bibliotecários que farão uso da linguagem de indexação e são responsáveis por sugerir adequações e termos novos.

5 - *Software* de gerenciamento do tesouro: operacionalização da linguagem

- Para abranger os recursos e estrutura almejada, o *software* de gestão para a construção e manutenção da linguagem deverá ser definido a partir das necessidades e objetivos previamente traçados, aspectos que podem contribuir para essa escolha são: interface amigável, fácil visualização, possibilidade de importação e exportação de dados através da interoperabilidade semântica, por exemplo. Considerar suporte técnico da instituição quando o *software* for livre ou pago, mas que possua suporte.

6 – Compilação dos termos

Fonte dos termos: definir o ponto de partida para compilação dos termos, considerar:

- Listas de cabeçalho de assuntos já utilizados pela unidade/biblioteca;
- Enriquecimento do vocabulário com outras linguagens de indexação: que possam ser utilizadas de acordo com sua similaridade da área temática, considerando também a relevância e credibilidade das linguagens utilizadas;
- Seleção automática: há *softwares* especializados em extrair de palavras e frases automaticamente de acordo com a frequência de incidência;
- Índice de pesquisa dos usuários;
- Referências tradicionais: livros e dicionários especializados.

7 - Construção

- Amparo normativo: qual norma atende melhor os objetivos em questão?
- Organização dos termos e agrupamentos dos termos: organização necessária para poder visualizar a estrutura e manuseá-la.
- Relacionamento entre os termos.
- Idiomas abrangidos e como aparecerão na estrutura da linguagem.
- Compatibilização de linguagens: recurso definido quando se utiliza mais de uma

linguagem de indexação.

8 – Introdução

- Apresentação da ferramenta ao público.
- Exibição das regras utilizadas para que outros possam compreender a dinâmica da linguagem.
- Formas de utilização: além da indexação, auxiliar na estratégia de busca dos pesquisadores nas bases de dados da instituição.
- Contato dos responsáveis

9 - Disseminação

- Divulgação: meios de divulgação dos resultados alcançados, ou seja, da elaboração da linguagem de indexação.
- Preservação: amparo nas tecnologias para preservar a ferramenta.

10 - Manutenção

- Equipe responsável
- Atualização da linguagem
- Frequência
- Critérios para as adequações

→ 3ª fase: Gestão da construção

1 - Definir os fluxos de trabalho

- Considera-se fundamental a comunicação entre gestores e indexadores, para que ambos estejam empenhados em solucionar problemas, como registros de autoridades e disparidades entre o sistema e a política de indexação ou mesmo na busca por melhorias, como incluir notas de escopo para os termos autorizados e termos em inglês como equivalentes, por exemplo. Aspectos que envolvem a linguagem de indexação e podem passar despercebidos para aqueles que gerenciam a linguagem e que fazem pouco uso no dia a dia. Algumas formas de alinhar o conhecimento dos gestores e indexadores são por meio de treinamentos periódicos, canais de comunicação interna, a capacitação de novos funcionários e a participação dos bibliotecários da rede no grupo. Por isso, deve haver um grupo gestor para concentrar os fluxos

de trabalho e tomar importantes decisões considerando não somente as necessidades da linguagem, mas também daqueles que a utilizam. Dessa forma o trabalho será mais colaborativo e mesmo que o número de bibliotecários esteja reduzido, os que não podem participar do grupo poderão participar e cooperar com suas análises, sugestões e contando ainda com a atualização da linguagem, pois quem irá alimentar a ferramenta são os indexadores da rede de acordo com as demandas apresentadas através de novos recursos informacionais, assuntos recém descobertos, como muitas vezes presentes em trabalhos de pós-graduação e solicitação de alunos e docentes.

2 - Construção

- Construção dos registros de autoridade: para garantir a correção persistente no catálogo de registros bibliográficos e possibilitar a interoperabilidade em *softwares* de construção e manutenção de linguagens de indexação.
- Estabelecer regras na seleção de conceitos dos termos anteriormente reunidos, considerar os princípios de especificidade desejado na linguagem de indexação;
- Agrupamento e organização dos termos de forma hierárquica (ou conforme o tipo de linguagem). A construção da linguagem pode ser mais eficiente quando dividida em áreas do conhecimento para cada responsável do grupo, para que sejam trabalhadas a partir dos termos mais abrangentes para os mais específicos na estrutura hierárquica;
- Correção dos termos importados: especialmente dos termos em inglês em que na tradução pode ter havido algum equívoco. Além da adequação desses termos de acordo com as orientações previamente definidas na política, como uso de plural ou singular, termos complexos, erros gramaticais e etc.
- Alimentação do software com essa estrutura, buscar auxílio da equipe técnica de informática ou responsáveis pelo produto.

3 - Divulgação

- Disponibilização e capacitação para os pesquisadores: bem como disponibilizar a linguagem no catálogo online da biblioteca para os pesquisadores utilizarem, observou-se que é preciso apresentar e ensinar o porquê da linguagem de indexação na busca pela informação e como utilizá-la. Percebe-se a importância da participação dos bibliotecários de referência nesse processo.

- É interessante que as pessoas responsáveis por esse trabalho façam publicações sobre esses projetos para que outras instituições possam se embasar e se relacionar com realidades semelhantes.

→ **4ª fase: Gestão da manutenção**

- Estabelecer as linguagens ou fontes de dados permitidas para consulta dos indexadores;
- Formulários disponíveis para os indexadores para que estes auxiliem na manutenção da linguagem de indexação através da solicitação via formulário, com todas as informações que são necessárias para facilitar a avaliação dos responsáveis;
- Avaliação do grupo de revisão: verificação pelos responsáveis desta etapa a adequação do termo para compor a linguagem de indexação ou não. Em todo caso, dar uma devolutiva para o bibliotecário que fez a solicitação;
- Correções de forma retroativa em registros que possuam termos modificados para que não seja afetada a recuperação da informação desses materiais.

8 CONSIDERAÇÕES FINAIS

Como visto na introdução desta tese e com a premissa de que as bibliotecas universitárias brasileiras estão buscando aprimoramentos no tratamento da informação, como a elaboração de linguagens de indexação, foi delimitado o problema de pesquisa em: como as bibliotecas universitárias da USP e da UNESP fazem o gerenciamento da construção e a manutenção das linguagens de indexação?

A hipótese levantada foi de que as bibliotecas universitárias brasileiras não tinham procedimentos sólidos para construir suas linguagens de indexação, apesar de constatar na problemática da pesquisa que havia intenções de melhoria nesse âmbito. A escolha da USP e da UNESP ocorreu para corroborar com a tese de que através da experiência dessas duas universidades na construção e gerenciamento das linguagens de indexação poderiam contribuir como subsídio para outras bibliotecas que desejam desenvolver uma ferramenta própria para o controle e representação de seus recursos informacionais.

Com base nisso, pode-se considerar que foi uma escolha assertiva das duas universidades, primeiramente pela documentação dos casos em artigos, eventos científicos, manuais e livros publicados, ou seja, uma gama de produções que auxiliaram as discussões aqui realizadas; e em particular pelo caráter universitário brasileiro de ambas. O segundo ponto que pode ser destacado é que se pôde analisar e verificar como as redes de bibliotecas da USP e da UNESP fazem o gerenciamento da construção e manutenção de suas linguagens de indexação no decorrer do trabalho desenvolvido. Por isso, destaca-se que elaborar uma linguagem de indexação configura-se em um projeto ambicioso e que requer equipe determinada e responsável por manter a linguagem de forma atual e dinâmica.

Para atender os objetivos específicos desta pesquisa utilizou-se como amparo às metodologias empregadas: a pesquisa bibliográfica e a pesquisa descritiva. Com a pesquisa bibliográfica foi possível reunir estudos diversos em esfera nacional e internacional sobre aspectos que envolvem as linguagens de indexação. Para a análise desses trabalhos optou-se por criar categorias seguindo esses aspectos verificados na literatura para que facilitassem a sua sistematização e organização das ideias

trabalhadas de forma eficiente, considerando a quantidade de textos recuperados e que precisavam ser examinados e discutidos. Em algumas categorias de análise se obteve um maior número de trabalhos nacionais, como por exemplo a segunda, intitulada “o uso de linguagens de indexação por bibliotecas universitárias”, pelo fato de que foram recuperados mais trabalhos com essa temática, contudo buscou-se incluir literatura estrangeira também. Sobre isso, verificou-se que o Brasil tem produzido vários

trabalhos que envolvem ferramentas de controle de vocabulário, principalmente no âmbito dos projetos de pesquisa de docentes universitários, como recentemente da UNESP.

A pesquisa bibliográfica contribuiu para a solução do primeiro objetivo específico em “realizar estudo teórico e metodológico a partir da revisão de literatura nacional e internacional sobre o gerenciamento da construção e manutenção de linguagens de indexação na perspectiva da política de indexação no âmbito da Ciência da Informação”. Em continuidade, como parte da fundamentação teórica e considerando ainda o primeiro objetivo específico, foi produzida a seção teórica “Gerenciamento da construção e manutenção de linguagens de indexação em bibliotecas universitárias”, nesta, além de apresentar os estudos de cunho teórico e metodológico a partir da revisão de literatura nacional e internacional, incluiu também diretrizes normativas da ISO 25964 (2011). Considera-se basilar a consulta das diretrizes normativas, como a trazida neste trabalho, para amparar às bibliotecas e aos sistemas de informação que desejam construir suas próprias linguagens de indexação ou mesmo buscar formas para fazer a gestão a longo prazo, em conjunto com a atualização e manutenção da ferramenta.

Mesmo assim, apesar das significativas contribuições obtidas pela literatura normativa, considera-se crucial a exposição da realidade enfrentada pelas bibliotecas frente a aspectos complexos da indexação, destacamos aqui o uso das linguagens de indexação, e como elas lidam com isso. A *práxis* das bibliotecas, especialmente das universitárias, tendo em vista que são ambientes muitas vezes multidisciplinares em que trabalham com diversos tipos assuntos em suas coleções, é algo que merece destaque devido às suas particularidades, além de que com o exemplo de outras bibliotecas é possível ajudar e incentivar outras a tomarem iniciativa para construírem ou mesmo utilizarem uma linguagem de indexação consistente e adequada à realidade da biblioteca, na representação e na busca pela informação.

A pesquisa descritiva foi empregada para atender ao segundo objetivo específico da pesquisa em: desenvolver estudo analítico descritivo sobre gerenciamento da construção e manutenção de linguagens de indexação pelas bibliotecas universitárias da USP e UNESP. Com isso, houve dois momentos centrais na pesquisa descritiva: a análise descritiva do gerenciamento da construção e manutenção das linguagens de indexação da USP e da UNESP na perspectiva da literatura (1) e nas perspectivas dos gestores e bibliotecários (2).

Com a análise descritiva da literatura recuperada foi possível se obter um panorama documentado dos procedimentos tomados em cada fase do desenvolvimento das linguagens de indexação das duas universidades. A maioria desses trabalhos foram realizados por pesquisadores que participaram ativamente do desenvolvimento das linguagens, desde sua fase de projetos até a etapa de sua divulgação. As publicações desses relatos de caso sintetizam as

decisões de profissionais bibliotecários e pesquisadores que se dedicaram a desenvolver uma ferramenta que unisse as contribuições da literatura e de suas respectivas realidades. De fato, propósitos esses que buscamos ter alcançado com essa tese. Com isso, na seção que envolve a análise na perspectiva da literatura, buscou-se condensar essas informações em categorias o que também facilitou o processo de análise e a elaboração de considerações a respeito de ambas as ferramentas de controle de vocabulário. O VocaUSP, como já mencionado anteriormente, foi fruto de um estudo pioneiro no Brasil, pela Universidade de São Paulo, em que foram utilizados os recursos da época, como a base de dados em CD-ROM, recorrendo e obtendo o apoio da equipe de informática que desenvolveu um software específico para para auxiliar na tarefa de organização dos termos em blocos de assuntos hierarquizados, para posteriormente passar a compor o catálogo online das Bibliotecas do SIBi/USP, e atualmente ser disponibilizado utilizando por um *software* específico para essa finalidade, o Tematres. A UNESP iniciou o projeto de construção de sua linguagem a partir de uma base sólida que foi de sua política de indexação, “Manual de política de indexação de bibliotecas universitárias da Unesp” (Fujita, 2014) com a adequação e compatibilização dos registros de autoridade. Foi um projeto com etapas pré-definidas e seguidas. Enfatiza-se que nos dois casos foi necessária a participação dos bibliotecários da rede e de pesquisadores, cada um responsável por uma determinada tarefa. Verificou-se que ambas empregaram diferentes métodos, em diferentes momentos, mas que mesmo assim produziram linguagens de indexação consistentes e suscetíveis a atualizações e aprimoramentos, além de estarem disponíveis para a consulta dos usuários.

Não é o caso das duas universidades, pois antes de sugerir termos, são consultadas fontes de referência e outras linguagens de indexação autorizadas, mas destaca-se que o uso de linguagem natural é um indicativo de que o sistema adota uma abordagem híbrida, visando a atualização contínua das linguagens de indexação durante a gestão e manutenção. Isso reflete a necessidade de adaptação às mudanças na terminologia, garantindo que o sistema permaneça consistente e eficaz. Portanto, é crucial que a manutenção e atualização das linguagens de indexação sejam incorporadas à política de indexação. Essa prática assegura que o sistema continue a atender às necessidades dos usuários e da instituição, melhorando a recuperação da informação e a eficiência na organização temática dos recursos informacionais.

Em sequência, para obter um resultado ainda mais preciso, foi decidido utilizar o método de entrevista para contactar com aqueles(as) que utilizam a linguagem de indexação diariamente. Por isso, a segunda etapa da pesquisa descritiva buscou a perspectiva das gestoras e indexadoras sobre o gerenciamento da construção e manutenção das linguagens de indexação. Observou-se que as indexadoras tiveram mais facilidade e até mesmo propriedade para falar da

atualização da linguagem, pois é a partir delas que são verificadas as necessidades de modificar, incluir ou até mesmo desconsiderar um termo contido na estrutura da linguagem. As gestoras têm uma visão voltada mais para o gerenciamento “operacional” da linguagem, tendo em vista que reúnem as informações trazidas pelos indexadores da rede, principalmente por meio dos formulários de solicitação e estudam caso a caso para fazer as alterações. Ressaltaram os desafios enfrentados como gestoras, como tempo para se dedicar às solicitações que chegam da rede, bibliotecários para compor o grupo de gestão e até mesmo questões técnicas como falta de registros de autoridades e inclusão de termos específicos, como os da área da saúde. Outro ponto de destaque é a importância de ambas as partes, indexadores e gestores da rede, estarem cientes de suas atribuições, para não haver desconhecimento do papel de cada e de seu próprio, considerando ainda as diretrizes estipuladas em seus manuais de política de indexação. Por isso, é considerável que as capacitações e treinamentos sejam periódicos ou marcados sempre que houver necessidade para alinhar todas essas partes, em especial quando ocorre mudanças nos responsáveis pela gestão ou na política e bibliotecários recém-contratados, por exemplo.

Portanto, considera-se que essa etapa da pesquisa foi crucial para ter acesso a vivência das bibliotecas com o uso das ferramentas que foram desenvolvidas. Contudo, ao ponderar sobre essa etapa da pesquisa, considera que seria interessante ter um número maior de participantes entrevistados, para reunir outros pontos de vista de indexadores e gestores. Inclusive daqueles que participaram e contribuíram no início do projeto, mas que não compõem mais o grupo gestor, de bibliotecários com mais tempo em suas funções, mais experientes e dos novatos na seção de tratamento da informação. Uma limitação encontrada foi o tempo para realizar mais entrevistas em tempo hábil para realizar todos os procedimentos necessários, como a transcrição, correção, tabulação e análise do material coletado.

Ainda sobre essa fase, acredita-se que em pesquisas futuras possam ser considerados também os usuários e pesquisadores para observar como é a percepção deles em relação à linguagem e a forma como a utilizam.

O terceiro e último objetivo específico definido no início desta pesquisa foi “Elaborar orientações sobre o gerenciamento da construção e manutenção de linguagens de indexação na política de indexação de bibliotecas universitárias, tendo como parâmetro a literatura técnico-científica e a *práxis* das bibliotecas universitárias”.

Diante de todas as informações reunidas, relatos e métodos de pesquisa empregados, foi produzida a seção que antecede a esta, onde foi priorizada a sistematização dos procedimentos em tópicos e com breves explicações sobre cada um, para que isso seja um facilitador para quem busca aprimorar o contexto em que se encontra o tratamento temático da informação em

sua instituição, seja em construir uma linguagem própria ou aperfeiçoar a existente. Essa etapa foi simples de executar devido à fundamentação previamente estabelecida, que incluiu uma ampla revisão de literatura, uma análise descritiva dos estudos específicos sobre as experiências da USP e da UNESP, e os depoimentos obtidos nas entrevistas. Esses elementos facilitaram a identificação e organização dos pontos essenciais em recomendações.

Espera-se que essas recomendações possam ser úteis para diversas bibliotecas ou sistemas de informação, independentemente de sua natureza. Assim, outras instituições, além das universitárias, também poderão se beneficiar das discussões e resultados apresentados.

REFERÊNCIAS

ACHILLES, D.; SOUSA, B. P.; SABBAG, D. M. A. Interação pelo catálogo on-line dos processos de seleção e representação temática: exploração da literatura lésbica enquanto assunto. *In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (ENANCIB)*, 20., 2019, Florianópolis. **Anais [...]**. Florianópolis: UFSC, 2019. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/944>. Acesso em: 18 ago. 2023.

ACOSTA BRAVO, A.; DÍAZ JIMÉNEZ, A.; LEIVA MEDEROS, A. A. Propuesta de procedimiento para la construcción semiautomática de tesauros en bibliotecas universitarias. **Anales de Investigacion**, Havana, v. 17, n. 2, p. 56-80, 2021. Disponível em: <http://revistas.bnjm.sld.cu/index.php/BAI/article/view/380/377>. Acesso em: 25 mar. 2024.

AITCHISON, J.; CLARKE, S. D. The thesaurus: a historical viewpoint, with a look to the future. **Cataloging & Classification Quarterly**, Philadelphia, v. 37, n. 3-4, p. 5-21, 2004. Disponível em: https://doi.org/10.1300/J104v37n03_02. Acesso em: 25 mar. 2024.

ALEXIEV, V.; ISAAC, A.; LINDENTHAL, J. On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). **International Journal on Digital Libraries**, New York, v. 17, n. 1, p. 39-48, 2016.

ANDRADE, J.; LARA, M. L. G. Interoperability and mapping between knowledge organization systems: metathesaurus-unified medical language system of the national library of medicine. **Knowledge Organization**, Wurzburg, v. 43, n. 2, p. 107-112, 2016.

BARITÉ, M. Sistemas de organización del conocimiento: una tipología actualizada. **Informação & Informação**, Londrina, v. 16, n. 3, p. 122-139, 2011. DOI: <http://dx.doi.org/10.5433/1981-8920.2011v16n2p122>. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/9952>. Acesso em: 25 mar. 2024.

BARITÉ, M.; FERNÁNDEZ-MOLINA, J. C.; GUIMARÃES, J. A. C.; BATISTA, J. B. E. Garantia literária: elementos para uma revisão crítica após um século. **Transinformação**, Campinas, v. 22, n. 2, p. 123-138, 2010. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/6201>. Acesso em: 21 abr. 2024.

BISCALCHIN, R.; MOREIRA, W. Construção de vocabulários multilíngues: perspectivas culturais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 25, n. 4, p. 47-67, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/151936>. Acesso em: 17 nov. 2023.

BOCCATO, V. R. C.; FUJITA, M. S. L.; GIL LEIVA, I. Avaliação comparada do uso de linguagens de indexação em catálogos de bibliotecas universitárias para recuperação por assunto. **Scire: representación y organización del conocimiento**, Zaragoza, v. 17, n. 1, p. 55-64, 2011.

BOCCATO, V. R. C.; TORQUETTI, M. C. Interoperabilidade entre linguagens de indexação como recurso de construção de instrumento de representação temática de clippings de coordenadorias de comunicação social em ambientes universitários: uma proposta metodológica. **Informação & Informação**, Londrina, v. 17, n. 3, p. 76-101, 2012. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/10800>. Acesso em: 30 jul. 2020.

BRASCHER, M.; GUIMARÃES, J. A. C. Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. **Liinc em Revista**, Rio de Janeiro, v. 14, n. 2, p. 241-258, 2018. Disponível em: <https://revista.ibict.br/liinc/article/view/4347/3949>. Acesso em: 13 fev. 2024.

CAMPBELL, H. M.; DIECKMAN, C. S.; TEAL, W.; WINTERMUTE, H. E. Improving subject headings for iowa indigenous peoples. **Library Resources and Technical Services**, Chicago, v. 66, n. 1, p. 48-59, 2022.

CAMPOS, M. L. A.; GOMES, H. E.; MARTINS, A. E.; CAMPOS, M. L. M.; CAMPOS, L. M.; SALES, L. F. Estudo comparativo de softwares de construção de tesouros. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 1, p. 68-81, 2006. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&%20pid=S1413-99362006000100006&lng=en&nrm=iso&tlng=pt. Acesso em: 30 jul. 2023.

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, 1985.

CENTRAL DE PESQUISA GOOGLE. **Guia detalhado sobre como a pesquisa Google funciona**. 2023. Disponível em: <https://developers.google.com/search/docs/fundamentals/how-search-works?hl=pt-br#:~:text=Indexa%C3%A7%C3%A3o%3A%20o%20Google%20analisa%20os,relevantes%20para%20a%20consulta%20dele>. Acesso em: 15 set. 2023.

CERVANTES, B. M. N. **A construção de tesouros e a integração de procedimentos terminográficos**. 2009. 209 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

CERVO, A. L.; BERVIAN, P. A.; SILVA, R. **Metodologia científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2006.

CHETI, A.; VITI, E. Functionality and merits of a faceted thesaurus: the case of the *Nuovo Soggettario*. **Cataloging & Classification Quarterly**, Philadelphia, v. 61, n. 5-6, p. 708-733, 2023.

CHO, J. H. Cataloging for a celebration: metadata for an institutional repository from the ground up. **Cataloging and Classification Quarterly**, Philadelphia, v. 60, n. 2, p. 141-163, 2022.

CHOI, I.; PARK, M. S. Specificity and exhaustivity of bibliographic classifications: a cross- cultural comparison with text analytic approach. *In: iCONFERENCE 2020*:

sustainable digital communities, 2020, Borås, Sweden. **Proceedings** [...]. Borås: University of Borås, Oslo Metropolitan University, 2020. Disponível em: <https://core.ac.uk/download/pdf/288433168.pdf>. Acesso em: 30 set. 2023.

CLARKE, S. G. D. The information retrieval thesaurus. **Knowledge Organization**, Wurzburg, v. 46, n. 6, p. 439-459, 2019.

CLARKE, S. G. D.; ZENG, M. L. From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modeling. **Information Standards Quarterly**, Baltimore, v. 24, n. 1, p. 20-26, 2012. Disponível em: <https://strathprints.strath.ac.uk/2623/>. Acesso em: 21 jun. 2018.

CÓRDOVA, R. D.; LÁZARO, C. J. R. El uso de los sistemas tradicionales de organización del conocimiento en las bibliotecas peruanas. **Biblios: revista de Bibliotecología y Ciencias de la Información**, Florianópolis, n. 46, p. 26-32, 2012. Disponível em: <http://biblios.pitt.edu/ojs/biblios/article/view/38>. Acesso em: 2 jun. 2024.

CORRÊA, R. F. ; FUJITA, M. S. L. Método para avaliação direta da indexação automática via julgamento por indexadores. **Encontros Bibli**, Florianópolis, v. 29, p. e96485, 2024. DOI: <https://doi.org/10.5007/1518-2924.2024.e96485>. Disponível em: <https://www.scielo.br/j/eb/a/zNVDGw4BVt5FbQtLbr96pQq/#>. Acesso em: 2 jun. 2024.

CRUZ, M. C. A. Política de indexação em bibliotecas universitárias: estudo diagnóstico na região Sul e Sudeste. In: ENCONTRO REGIONAL DOS ESTUDANTES DE BIBLIOTECONOMIA, DOCUMENTAÇÃO, CIÊNCIA E GESTÃO DA INFORMAÇÃO, 2., 2015, São Carlos. **Anais** [...]. São Carlos: UFSCar, 2015. p. 25-32. Disponível em: <http://www.2erebd.ufscar.br/index.php/erebd/erebd/paper/viewFile/50/9>. Acesso em 01 abr. 2019.

CRUZ, M. C. A. **Linguagem de indexação no contexto da política de indexação: estudo em bibliotecas universitárias**. 2017. 61 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2017.

CRUZ, M. C. A. **Linguagens de indexação em bibliotecas universitárias: estudo analítico em território nacional**. 2019. 82 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2019.

CRUZ, M. C. A.; FERNEDA, E.; FUJITA, M. S. L. A disponibilização de vocabulário controlado aos usuários para a recuperação da informação. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 15, n. 1, p. 266-282, 2022. DOI 10.26512/rici.v15.n1.2022.42464. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/42464>. Acesso em: 18 nov. 2023.

CRUZ, M. C. A.; FUJITA, M. S. L. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. **Informação & Informação**, Londrina, v. 26, n. 1, p. 574-600, 2021.

CRUZ, M. C. A.; FUJITA, M. S. L.; SANTOS, L. B. P. Linguagem de indexação no contexto da política de indexação: estudo em bibliotecas universitárias. *In*: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). **Memória, tecnologia e cultura na organização do conhecimento**. Recife: UFPE, 2017. p. 217- 224.

CRUZ, M. C. A.; TOLARE, J. B.; FUJITA, M. S. L. Analysis of indexing language in Covid-19 subject representation. *In*: BOBCATSSS, 32., 2024, Coimbra. 2024. No prelo.

CRUZ, M. C. A.; SANTOS, L. P.; FUJITA, M. S. L. Linguagens de indexação em bibliotecas universitárias brasileiras: diagnóstico preliminar das regiões sul e sudeste. *In*: ENCONTRO DA ASSOCIAÇÃO DE EDUCAÇÃO E PESQUISA EM CIÊNCIA DA INFORMAÇÃO DA IBEROAMÉRICA E CARIBE, 10., 2016, Belo Horizonte. **Anais [...]**. Belo Horizonte: Universidade Federal de Minas Gerais, 2016. p. 1821-1835.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.

DEMO, P. **Introdução à metodologia da ciência**. São Paulo: Atlas, 1985.

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

FERNANDÉZ, L. M. M.; ALONSO, M. I.; MONTÁVEZ, A. J. V. La consistencia como indicador de la fiabilidad (corrección) de la indización: los casos de LISA, PASCAL y VOCED. **Scire: representación y organización del conocimiento**, Zaragoza, v. 12, n. 1, p. 85-98, 2006.

FOSKETT, A. C. **A abordagem temática da informação**. Tradução: Antonio Agenor Briquet de Lemos. São Paulo: Polígono, 1973.

FOSTER, A. K. A controlled vocabulary for an electronic resources problem reporting system creation, implementation and assessment. **Library Resources & Technical Services**, Chicago, v. 65, n. 1, p. 23-32, 2021. Disponível em: <https://journals.ala.org/index.php/lrts/article/view/7382>. Acesso em: 14 jan. 2024.

FREDERICK, D. E. **Managing eBook metadata in academic libraries: taming the tiger**. Amsterdam: Elsevier, 2016.

FUJITA, M. S. L. (coord.). **Política de indexação para bibliotecas**. Marília: UNESP; Brasília: CNPq, 2010. Projeto de Pesquisa.

FUJITA, M. S. L. (coord.). **Linguagem de indexação para bibliotecas na perspectiva da política de indexação**. Marília: UNESP; Brasília: CNPq, 2015. Projeto de Pesquisa.

FUJITA, M. S. L. Linguagem de indexação para bibliotecas na perspectiva da política de indexação. *In*: FUJITA, M. S. L.; MOREIRA, W. (org.). **Manual do planejamento, construção e manutenção do Tesauro Unesp para bibliotecas: do conceitual a práxis**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2021. Disponível em: https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/view/234/2148/3632.

Acesso em: 19 abr. 2024.

FUJITA, M. S. L. Política de indexação para bibliotecas: funções e finalidades. *In*: FUJITA, M. S. L. (org.). **Política de indexação para bibliotecas: elaboração, avaliação e implantação**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016.

FUJITA, M. S. L.; BOCCATO, V. R. C.; RUBI, M. P. O contexto de indexação para catalogação de livros usando uma abordagem sociocognitiva. **Brazilian Journal of Information Science: research trends**, Marília, SP, v. 4, n. 2, p. 22–40, 2010. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/485>. Acesso em: 2 jun. 2024.

FUJITA, M. S. L.; CRUZ, M. C. A.; PATRÍCIO, B. O. M. A construção de tesouros na perspectiva dos manuais de indexação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais [...]**. Marília, 2017. v. 1, p. 1-26. Disponível em: <http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/schedConf/presentations>. Acesso em: 16 out. 2018.

FUJITA, M. S. L.; CRUZ, M. C. A.; PATRÍCIO, B. O. M.; RIO BRANCO, L. B. P. R. Linguagens de indexação em bibliotecas universitárias: estudo analítico. **Informação & Informação**, Londrina, v. 24, n. 1, p. 190-225, 2019. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/31771>. Acesso em: 6 abr. 2024.

FUJITA, M. S. L.; GIL LEIVA, I. Políticas de indexação na América Latina. **Ibersid: revista de sistemas de información y documentación**, Londrina, v. 3, p. 155-162, 2009. Disponível em: <https://www.iversid.eu/ojs/index.php/iversid/article/view/3735>. Acesso em: 2 jun. 2024.

FUJITA, M. S. L.; MOREIRA, W.; SANTOS, L. B. P.; CRUZ, M. C. A.; RIBAS, R. R. B. Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: an experience of the Sao Paulo state university. **Knowledge Organization**, Wurzburg, v. 45, n. 3, p. 220-231, 2018. DOI: <http://doi.org/10.5771/0943-7444-2018-3>. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2018-3-220.pdf>. Acesso em: 6 abr. 2024.

FUJITA, M. S. L.; MOREIRA, W.; CAMPOS, M. L. A. Aspectos conceituais do tesouro Unesp. *In*: FUJITA, M. S. L. ; MOREIRA, W. (org.). **Manual do planejamento, construção e manutenção do Tesouro Unesp para bibliotecas: do conceitual a práxis**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2021. Disponível em: https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/view/234/2148/3632. Acesso em: 19 abr. 2024.

FUJITA, M. S. L.; RUBI, M. P. O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário. **Perspectivas Ciência da Informação**, Belo Horizonte, v.11, n. 1, p. 48-66, 2006.

FUJITA, M. S. L.; SANTOS, L. B. P.; ALVES, R. V. Linguagem de indexação e

linguagem documentária são sistemas de organização do conhecimento? Uma análise bardiana da variação terminológica. **Scire: representación y organización del conocimiento**, Zaragoza, v. 24, n. 2, p. 23-33, 2018. Disponível em: <http://www.ibersid.eu/ojs/index.php/scire/article/view/4577>. Acesso em: 6 abr. 2019.

FUJITA, M. S. L.; SANTOS, L. B. P.; CRUZ, M. C. A.; MOREIRA, W. Avaliação das características do TemaTres e Multites para o controle de autoridades nas bibliotecas universitárias. **Scire: representación y organización del conocimiento**, Zaragoza, v. 23, n. 2, p. 63-73, 2017.

GARCÍA MARCO, F. J.; LÓPEZ DEL RAMO, J. Developing an interoperable taxonomy by mapping empiric taxonomies: a case study on the websites of the associations of friends of the Way of St James. **Scire: representación y organización del conocimiento**, Zaragoza, v. 28, n.1 , p. 55-73, 2022.

GIL LEIVA, I. **La automatización de La indización de documentos**. Gijón:Trea, 1999.

GIL LEIVA, I.; FUJITA, M. S. L.; ORTUÑO, P. D.; REIS, D. M. The abandonment of the assignment of subject headings and classification codes in university libraries due to the massive emergence of electronic books. **Knowledge Organization**, Wurzburg, v. 47, n. 8, p. p. 646-667, 2020.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GIVEN, L. M.; OLSON, H. A. Knowledge organization in research: a conceptual model for organizing data. **Library & information Science Research**, Norwood, v. 25, n. 2, p. 157-176, 2003. Disponível em: http://ac.els-cdn.com/S0740818803000057/1-s2.0-S0740818803000057-main.pdf?_tid=9f288cc8-5d98-11e6-95d5-00000aacb35f&acdnat=1470682256_d2529d056accb393654f95f369d7f32c. Acesso em: 8 ago. 2023.

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**, Brasília, v. 1, n. 1, p. 77-99, 2008.

GUIMARÃES, J. A. C.; EVANGELISTA, I. V. Exaustividade e especificidade na indexação: uma análise de conteúdo como perspectiva de investigação sobre o tema. In: CONGRESO ISKO ESPAÑA-PORTUGAL, 2., CONGRESO ISKO ESPAÑA, 12: organización del conocimiento: sistemas de información abiertos. **Actas [...]**. Murcia: Universidad de Murcia, 2015. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/134993/1/ISKO-aplicacao-teoria.pdf>. Acesso em: 16 ago. 2024.

HJØRLAND, B. Indexing: concepts and theory. **Knowledge Organization**, Frankfurt, v. 45, n. 7, p. 609-639, 2018. Disponível em: <http://www.isko.org/cyclo/indexing#refW>. Acesso em: 5 abr. 2024.

HO, J.; STOKES, C. Core metadata element recommendations for institutional repositories at Texas A&M University Libraries. **Journal of Library Metadata**,

Philadelphia, v. 19, n. 3-4, p. 187-214. DOI 10.1080/19386389.2019.1651499.
Disponível em: <https://www.tandfonline.com/doi/full/10.1080/19386389.2019.1651499>.
Acesso em: 5 abr. 2024.

HODGE, G. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington, DC: Council on Library and Information Resources, 2000. Disponível em: <http://www.clir.org/pubs/reports/pub91/contents.html>.
Acesso em: 25 jun. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1: information and documentation - Thesauri and interoperability with other vocabularies - part 1: Thesauri for information retrieval**. Genebra, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-2: information and documentation - Thesauri and interoperability with other vocabularies - Part 2: Interoperability with other vocabularies**. Genebra, 2013.

JESUS, S. A. **Análise do Tesauro UNESP sob a perspectiva da abordagem facetada**. 132 f. 2022. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2022.

KIM, G. Relationship between index term specificity and relevance judgment. **Information Processing Management**, New York, v. 42, p. 1218-1229, 2006.

KLOSE, A. C.; GOLDSTEIN, S.; LEVY, M. S. Numismatics & bibliographic description: how Rutgers University Libraries described coins with MODS. **Journal of Library Metadata**, Philadelphia, v. 22, n. 1-2, p. 75-104, 2022.

KOBASHI, N. Y.; LIMA, V. M. A.; LEME, M. A. T. **Manual de indexação de assuntos com uso do Vocabulário Controlado USP: versão preliminar**. São Paulo: Universidade de São Paulo, Sistema Integrado de Bibliotecas, 2006. 72 p. Disponível em:
https://edisciplinas.usp.br/pluginfile.php/2184590/mod_resource/content/1/Vocabulario.pdf. Acesso em: 30 mar. 2023.

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

LEISE, F. **Creating a controlled vocabulary**. 7 abr. 2003. Disponível em:
<https://boxesandarrows.com/creating-a-controlled-vocabulary/>. Acesso em: 16 ago. 2024.

LEITÃO, C. A entrevista como instrumento de pesquisa científica em Informática na Educação: planejamento, execução e análise. In: PIMENTEL, M.; SANTOS, E. (org.) **Metodologia de pesquisa científica em Informática na Educação: abordagem qualitativa**. Porto Alegre: SBC, 2021. (Série Metodologia de Pesquisa em Informática na Educação, v. 3). Disponível em: <https://metodologia.ceie-br.org/livro-3/>. Acesso em: 10 nov. 2023.

LIMA, V. M. A. **O Vocabulário Controlado USP: elaboração, implantação e**

gerenciamento: uma experiência profissional. São Paulo: USP, 2013.

LIMA, V. M. A.; KOBASHI, N. Y.; COUTTO, M. L. M.; SANTOS, C. A. C. M.; AMARAL, M. C.; TOKAREVICZ, S.; DELLA TORRE, S. R. S.; GUERRA, S. R. Y.; BOCCATO, V. R. C.; BARCELLOS, J. C. H. Estudos para implantação de ferramenta de apoio à gestão de linguagens documentárias: vocabulário controlado da USP. **TransInformação**, Campinas, v. 18, n. 1, p. 17-25, 2006. Disponível em: <http://eci.ufmg.br/wp-content/uploads/2020/03/jornadaDeRelatosEDebatesDaPraticaBibliotecaria-2013-VocabularioControladoUSP.pdf>. Acesso em: 10 nov. 2023.

MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. semantic web methodological framework to evaluate the support of integrity in thesaurus tools. **Journal of Information Science**, Amsterdam, v. 46, n. 3, p. 378- 391, 2020. DOI 10.1177/0165551519837195. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/0165551519837195>. Acesso em: 10 nov. 2023.

MARTÍNEZ-GONZÁLEZ, M. M.; ALVITE-DÍEZ, M. L. A. The support of constructs in thesaurus tools from a Semantic Web perspective: framework to assess standard conformance. **Computer Standards and Interfaces**, Amsterdam, v. 65, p. 79-91, 2019.

MARTINS, G. A.; THEÓPHILO, C. R. **Metodologia da investigação científica para ciências sociais aplicadas**. 3. ed. São Paulo: Atlas, 2016.

MOREIRA, W. Relações conceituais como elementos constitutivos essenciais dos sistemas de organização do conhecimento. **Informação & Informação**, Londrina, v. 24, n. 2, p. 1-30, 2019. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/37989>. Acesso em: 27 abr. 2024.

NISO - National Information Standards Organization. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Bethesda: NISO Press, 2010. 184 p. (ANSI/NISO Z39.19-2005 R2010).

OLSON, H. A.; BOLL, J. J. **Subject analysis in online catalogs**. 2. ed. Englewood: Libraries Unlimited, 2001.

PIOVEZAN, L. B.; FUJITA, M. S. L. Exaustividade, precisão e consistência em indexação: sistematização conceitual. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais [...]**. Belo Horizonte, 2014. p. 1120. Disponível em: <https://brapci.inf.br/index.php/res/download/186302>. Acesso em: 27 abr. 2024.

PPGCI - Programa de Pós-Graduação em Ciência da Informação. Marília: Faculdade de Filosofia e Ciências, 2022. Disponível em: <https://www.marilia.unesp.br/#!/pos-graduacao/mestrado-e-doutorado/ciencia-da-informacao/programa/linhas-de-pesquisa/>. Acesso em: 12 jan. 2024.

RIO-BRANCO, L. B. P. **Interoperabilidade semântica entre linguagens de**

indexação para bibliotecas universitárias. 2020. 142 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2020.

RUBI, M. P. **Política de indexação na perspectiva do conhecimento organizacional.** 2004. 135 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2004.

RUBI, M. P.; FUJITA, M. S. L.; BOCCATO, V. R. C. Elaboração do manual de política de indexação na formação continuada do catalogador. *In*: FUJITA, M. S. L.; GIL LEIVA, I. (ed.). **Política de indexação.** São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 217-227.

SALEH, A.; SULEIMAN, H. Roles of academic libraries towards enhancing knowledge management in higher institution of education. *In*: SHIVRAJ, K. S.; SULEIMAN, A. A.; GUPTA, P. (ed.). **Knowledge management in higher education institutions.** Índia: Manipal University Jaipur, 2022. v. 1, p. 63-67.

SANTOS, L. B. P.; FUJITA, M. S. L. A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. **Scire: representación y organización del conocimiento**, Zaragoza, v. 22, n. 2, p. 37-46, 2016.

SEVERINO, A. J. **Metodologia do trabalho científico.** 23. ed. rev. atual. São Paulo: Cortez, 2007.

SILVA, E. G. **O escritório de comunicação científica como perspectiva de atuação para bibliotecas universitárias brasileiras.** 2023. 199 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2023. Disponível em: <http://hdl.handle.net/11449/243319>. Acesso em: 12 jan. 2024.

SOERGEL, D. Indexing and retrieval performance: the logical evidence. **Journal of the American Society for Information Science**, New York, v. 45, n. 8, p. 589-599, 1994.

SOUZA, F. P.; HILLESHEIM, A. I. A. Tratamento da informação e o uso das tecnologias da informação e comunicação. **Biblionline**, João Pessoa, v. 10, n. 2, p. 81-96, 2014. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/biblio/article/view/16748>. Acesso em: 1 abr. 2019.

STACHOKAS, G. The problem for libraries in the twenty-first century: the need to accept a paradigm shift. *In*: STACHOKAS, G. **After the book: information services for the 21st Century.** Sawston: Chandos, 2014. cap. 3, p. 33-47.

STRADER, C. R. Author-Assigned keywords *versus* Library of Congress Subject. **Library Resources & Technical Services**, Chigago, v. 53, n. 4, p. 243-250, 2009. Disponível em: <https://journals.ala.org/index.php/lrts/article/view/5183>. Acesso em: 5 mar. 2024.

SUNNY, S. K.; ANGADI, M. Evaluating the effectiveness of thesauri in digital information retrieval systems. **The Electronic Library**, Oxford, v. 36, n. 1, p. 55-70, 2018.

TARTAROTTI, R. C. D.; FUJITA, M. S. L. Subject cataloguing & indexing in university libraries: a verbal-protocol comparative study. **Scire: representación y organización del conocimiento**, Zaragoza, v. 23, n. 1, p. 57-66, 2017.

THENAGLIA, G. Diseño de tesauros como estrategia didáctica para fortalecer su comprensión. **Palabra Clave**, La Plata, v. 12, n. 1, p. e174, 2022. Disponível em: <https://doi.org/10.24215/18539912e174>. Acesso em: 16 ago. 2024.

TOLARE, J. B.; CRUZ, M. C. A.; FUJITA, M. S. L. Evaluation of indexation consistency in publisher subject metadata. **Central European Journal of Educational Research**, Debrecen, v. 4, p. 28-34, 2022.

TOLARE, J. B.; FUJITA, M. S. L. A utilização de linguagens de indexação por bibliotecas universitárias: revisão bibliográfica. **Scire: representación y organización del conocimiento**, Zaragoza, v. 28, n. 1, p. 45-54, 2022.

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”. Coordenadoria Geral de Bibliotecas. In: FUJITA, M. S. L. (coord.). **Manual de política de indexação para as bibliotecas universitárias da Unesp**. São Paulo: Unesp, 2014. Disponível em: <http://www.biblioteca.unesp.br/portal/arquivos/manual-politica-indexacao>. Acesso em: 9 out. 2023.

VALLÉZ, M.; PEDRAZA-JIMÉNEZ, R.; CODINA, L.; BLANCO, S.; ROVIRA, C. Updating controlled vocabularies by analysing query logs. **Online Information Review**, Bradford, v. 39, n. 7, p. 870-884, 2015.

VAN SLYPE, G. **Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales**. Traducción del francés: Pedro Hípola, Félix de Moya. Madrid: Fundación Germán Sanchez Ruipérez, 1991.

WILL, L. The ISO 25964 data model for the structure of an information retrieval thesaurus. **Bulletin of the American Society for Information Science and Technology**, Silver Spring, v. 38, n. 4, p. 48-51, 2012. DOI 10.1002/bult.2012.1720380413. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/bult.2012.1720380413>. Acesso em: 16 ago. 2024.

ZENG, M. L. Compatibility of indexing languages in an online access environment: a review of the Approaches. In: ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP, 3., 1992, Pittsburgh. **Proceedings** [...]. Medford, N. J.: Learned Information, Inc., 1992. Disponível em: <https://journals.lib.washington.edu/index.php/acro/article/view/12604>. Acesso em: 16 ago. 2024.

ANEXO 1 - E-mail para as Bibliotecas

Bom dia, (*nome do bibliotecário*)!

Sou doutoranda em Ciência da Informação no Programa de Pós-Graduação em Ciência da Informação da Unesp, câmpus de Marília-SP. Sou orientada pela Profa. Dra. Mariângela Spotti Lopes Fujita e coorientada pela Profa. Dra. Franciele Marques Redigolo.

Em 2019, você contribuiu com a minha pesquisa de Mestrado respondendo a um questionário sobre “Linguagem de Indexação no contexto das bibliotecas universitárias brasileiras”. A partir das respostas obtidas pelo questionário reuni algumas bibliotecas para essa nova fase da pesquisa.

Atualmente, estou desenvolvendo uma pesquisa mais específica sobre essas ferramentas de controle de vocabulário. Por isso, peço, por favor, que você contribua mais uma vez participando de uma entrevista que estou fazendo com bibliotecários que possuem **projetos para a elaboração de linguagem de indexação (vocabulário controlado)** ou que já possuem linguagens de indexação própria, ou seja, vocabulário controlado construído pela própria instituição de vocês.

A entrevista não é longa, com aproximadamente 40 minutos de duração pelo Google Meet. A sua identidade será preservada, sendo analisadas apenas as respostas obtidas na entrevista. Por favor, responda esse e-mail para marcarmos o melhor horário para você.

Desde já agradeço imensamente,
Maria Carolina Andrade Cruz

ANEXO 2 - Transcrição da entrevista com G1

Pesquisadora: Bom, está gravando aqui. É... eu vou falar rapidinho sobre quais são os objetivos da pesquisa pra você ficar ciente. O objetivo geral é contribuir com as bibliotecas universitárias que buscam o desenvolvimento da sua própria linguagem de indexação e também com os estudos acerca da organização e representação da informação em CI. É o objetivo específico dessa pesquisa um deles né? Que eu que eu acho que é o mais importante é a elaboração de orientações sobre o gerenciamento da construção e manutenção de linguagem juntamente com a política de indexação de bibliotecas universitárias brasileiras. Tendo como parâmetro a literatura técnico-científica e práxis das bibliotecas. Então por cima eh basicamente é essa parte da pesquisa, né? Que eu considero e vamos começar com as perguntinhas. É, primeiramente queria que você se apresentasse. Nome, biblioteca e quais qual sua função? Na na rede.

Entrevistado: Tá. Então eu sou a B1, tá? Eu trabalho na Reitoria, eu sou bibliotecária já há bastante tempo, mas na reitoria agora eu tô como responsável do grupo técnico de informação e documentação que é vinculado à Secretaria Geral, tá? Uma das atribuições do grupo é gerenciar a biblioteca da reitoria, tá? É, eu já tô como responsável há uns cinco anos mais ou menos, tá? E eu faço parte também da Comissão Permanente do Tesouro Unesp, onde a gente entre várias atribuições, a gente trabalha com a manutenção e atualização do tesouro.

É quais são os assuntos que a coleção abrange.

Entrevistado: Tá. É basicamente direito. Tá? Direito. É porque ela é mais utilizada, utilizada pela assessoria jurídica, né? Desde a criação da Unesp, né? É, mas acho que seria a gente tem alguns outros assuntos, administração, administração pública, a gente também tem um pouco de literatura, né? De ficção romance, mas são o de doação ao longo dos dos anos, né? Que a gente até manteve, eh tinha a gestão que e aqui os funcionários eh tivessem eh pra lazer, né? Então a gente tem então mas a gente não não compra eh livros que não sejam da área de direito ou eh por solicitação de alguma área específica tá? É... de repente o pessoal de orçamento pede livre de orçamento público, é a gente tem bastante coisa também na área de educação, mas a gente só atende a solicitações dessas áreas eh eh se tiver uma solicitação, na verdade a gente só faz a compra se tiver uma solicitação. Então basicamente é direito e administração pública.

A biblioteca está inserida em uma rede?

Entrevistado: Tá, a biblioteca está inserida em uma rede de bibliotecas sim agora a gente não tenha alunos né, que frequentem porque a gente é uma biblioteca administrativa mas a gente atende toda a rede da UNESP então a gente é uma das bibliotecas da rede.

Pesquisadora: Uhum e o catálogo eh da biblioteca é coletivo?

Entrevistado: Sim. Né? Está inserido na rede também.

Pesquisadora: Certo. Agora eu vou fazer umas perguntinhas sobre política de indexação. A biblioteca tem um documento com a política de indexação, se sim a política de indexação é local ou é para todos da rede?

Entrevistado: A política é pra todos da rede, ela foi implantada é... a data exata eu não tenho mais posso conseguir pra você eh na rede de bibliotecas pra todo mundo utilizar, ela é tem um documento, né? É oficial para toda a rede de bibliotecas. Sim.

Pesquisadora: A linguagem de indexação é descrita nessa política? Ela é definida?

Entrevistado: Sim. E se não me engano no início ainda a gente não tinha uma linguagem específica, tá? Então provavelmente nos primeiros documentos a gente ainda vai ter algo é, como uma possível, né, linguagem, já teve outro nome, já foi um pouco diferente. Hoje, a gente chama de né, mas ela já teve um um histórico aí e talvez no documento mais atual ela já seja descrita assim.

Pesquisadora: A forma como fazer a gestão, o uso e a manutenção, atualização da linguagem é descrita em algum documento? Como os bibliotecários devem trabalhar com linguagem?

Entrevistado: Sim, é, principalmente no último documento que né? Que virou até um livro eh [...] mas a gente tem um manual, né? Da política de indexação e e lá tem assim, a discussão de como os bibliotecários devem fazer. Fazer a indexação.

Pesquisadora: Tá. Normalmente eu perguntaria aqui se a pessoa pode me enviar esse manual. Mas como já está disponível e eu tenho acesso daí eu não vou. O que levou a construção de uma linguagem de indexação da da universidade. Houve alguma necessidade assim que vocês viram?

Entrevistado: Sim quando entrei, né? No grupo de na época chamava grupo de terminologia, depois mudou pra grupo de linguagem. Já havia sido feito uma pesquisa com os bibliotecários da rede, essa primeira pesquisa eu não participei eh e através né? Dos resultados identificou-se a necessidade de construção de uma política de indexação. Pessoal reclamava bastante, né? Da recuperação da informação, né? Do de termos que eh sei lá, temos eh similares, é plural singular e isso atrapalhava, né? Por essa biblioteca que necessidade que foi levantada pelos próprios bibliotecários. Aí quando eu entrei nesse grupo já tinha essa eh já tinha essa finalidade. Já estava detectada então só tinha a finalidade de desenvolver uma política pra rede. Sabe?

Pesquisadora: Foram utilizadas linguagens, outros vocabulários no desenvolvimento do sim, a linguagem da Unesp, a UNESP usava, né?

Entrevistada: Antes de ter uma linguagem própria, a rede biblio, a linguagem da rede bibliodata, né? E tratativas com a na época com a né? Coordenadoria Geral de Bibliotecas, a FGV que na época em que é gerenciada já tinha linguagem né? Eles cederam pra UNESP a base de com os termos né? De assunto e de autor. Hm-huh. Acho que é de assunto de autor... Então a gente já já teve esse essa carga inicial dos eh registros, né? Porque a gente trabalha com registro de autoridade em formato MAC eh da FGV e aí a partir desse dessa primeira carga a gente começou a fazer o que a gente chama de compatibilização com linguagens. A né? A LC que a gente chama de LC porque a base da da linguagem do Bibliodata era a na verdade era traduziam os registros da LC. Tá? Então a gente começou a fazer uma compatibilização com os registros da LC com os dados da Biblioteca Nacional que também são advindos da LC e por uma necessidade que foi levantada na época pelo pessoal da área de de saúde e como a gente tem uns problemas hoje com o mas na época foi o MeSH foi apontado como é... uma linguagem que a gente poderia fazer uso, né? Compatibilizar os nossos registros. Então a gente hoje até dentro do Alma, né? Que é o sistema que a gente usa, cê consegue ver outras linguagens, tá? Mas basicamente a gente começou a trabalhar com LCSH e MeSH, beleza?

Pesquisadora: Foi utilizado alguma norma pra guiar as decisões na elaboração? Alguma norma internacional de linguagem?

Entrevistado: Isso. Ou da política você está falando da linguagem. É do tesouro. Tem norma ISO que eu não lembro o número de cabeça agora, tá? Mas tem documentado lá no manual. É e a gente também ... não sei se vai ficar bem na sua pergunta, a gente teve um treinamento em terminologia, né? A gente fez um um método terminográfico para que a gente também É... eu vou dizer assim, não é aceitar ou não um termo, né? Uma solicitação de um termo, mas é pra gente estudar mesmo. Analisar exatamente se aquele termo é candidato ou não para entrar na linguagem.

Pesquisadora: Então vou aproveitar para perguntar, mas aqui não tem essa pergunta, mas você diz que esse treinamento foi feito para o grupo de linguagem?

Entrevistado: Para o grupo de linguagem com a professora Mariângela Fujita. A gente desenvolveu um método termográfico, a gente desenvolveu um formulário, né? Que hoje ele é disponibilizado para os bibliotecários. São os bibliotecários da rede que fazem as solicitações, né? E esse formulário ele é... tem lá os campos que o pessoal tem que preencher é... dizendo o porquê é importante aquele termo de fato que é de consultado para chegar à conclusão de que aquele tempo que ainda não existe é necessário se tem equivalente em outra linguagem. Enfim, a gente tem esse formulário que a partir desse formulário a gente é uma comissão, né, mas o grupo de linguagem avalia a inclusão ou não do termo.

Pesquisadora: Existe algum documento publicado ou não que tem a trajetória de elaboração da linguagem da política?

Entrevistado: Sim, tem vários. Primeiro como... Depois da implementação da política, acho que a gente já teve uns dois documentos. Depois posso passar direitinho para você o histórico.

Pesquisadora: Qual a linguagem de indexação da sua instituição?

Entrevistado: É o Tesauro Unesp.

Pesquisadora: Quais são os campos de assuntos abordados na L.I.?

Entrevistado: É multidisciplinar. A gente tem até alguns problemas de consistência. A gente estava falando do MeSh né? É que como MeSh muito específico para área de Ciências da Saúde, às vezes a gente acaba tendo problemas para incluir os termos que vem do MeSh porque eles acabam vencendo muito específicos e a gente não trabalha com termos isolados, a gente trabalha com registros de autoridade, e dentro de um registro você tem vários termos. Você não tem só o termo principal que é autorizado, você tem o equivalente, sinônimos, as remissivas. Assim você tem problemas quando é muito específico de uma área e você está compatibilizando com uma linguagem que não é tão específica né? Então a ideia é que o tesauro seja multidisciplinar, mas a gente claro, né? Trabalha, se houver necessidade, com termos que os bibliotecários encontrem em vocabulários. Ou de repente sei lá glossários mais específicos, mas a gente tem que ter esse cuidado para fazer a inclusão no tesauro.

Pesquisadora: O tesouro utiliza as siglas para esquematizar e organizar seus elementos como “USE” o “UP” essas *tags*?

Entrevistado: É o tesauro, é ele vai para web em um software diferente do que a gente usa dentro da biblioteca para fazer o gerenciamento dos registros de autoridade, né? Que é o Alma. O Alma não tem uma versão pesada. Vamos dizer assim... A gente usa um Software livre chamada Tematres ou Tematrês, não sei exatamente a pronúncia... e ele foi configurado para receber esses registros de autoridade que são trabalhados no formato Marc 21, ele foi configurado para receber esses registros e para que a gente visualize em formato de Tesauro e nesta visualização você tem as *tags* o UP. Eu acho que tá até em inglês lá, depois eu posso abrir aqui para ver. “Usado para”, “use”, “TG - Termo geral”, “TE - Termo específico” né?

Pesquisadora: O Tesauro da Unesp, ele é multilíngue? Tem a versão em inglês, em espanhol?

Entrevistado: Ele não tem a versão. Mas ele tem termos em quase todos os registros, né? Que eles têm um termo equivalente equivalente em outra língua, tá geralmente é o inglês porque veio da LC

então ele quando você vê o tesouro você vê termos em inglês também, mas não que ele seja... É que tenha uma versão que traduza vai digamos o Tesouro em outra língua.

Pesquisadora: A linguagem utiliza qualificadores para diferenciar um mesmo termo com diferentes significados? Por exemplo: Mercúrio (planeta) / Mercúrio (elemento químico) / Mercúrio (mitologia).

Entrevistado: Usa, os qualificadores ficam entre parênteses.

Pesquisadora: Há alguma orientação a respeito de como os termos devem ser gramaticalmente? Por exemplo: termos no plural e singular, podem ser utilizados verbos?

Entrevistado: Sim, a partir da do uso que a gente é de quando a gente começou a conhecer a norma ISO, tá?! É essa norma existe essa orientação, então o manual, no caso, o manual de construção e atualização do Tesouro que a gente tem hoje que é o documento mais recente, ele tem essas orientações para o bibliotecário tá? A gente ainda trabalha com muita fazendo muitas correções, porque o que a gente já recebeu do BIBLIODATA e o que a gente já fez durante todo esse tempo. Eh, não tinha essa orientação. Então a gente tem muito tempo que tava no singular e correto é no plural. E aí a gente vem corrigindo quando a gente detecta, né o erro, mas existe sim hoje uma orientação para isso, né? É basicamente é coisas contáveis no plural não contáveis do singular, algumas exceções para termos que usam o nome científicos, né? É partes do corpo, doenças, algumas exceções. Nós temos essas orientações hoje.

Pesquisadora: Como são empregados os termos compostos (como Gestão de tesouros)? Há momentos em que são divididos (como gestão + tesouros)?

Entrevistado: A gente tem algumas coisas... Eu não sei se entra também na sua pergunta a questão dos termos que a gente chama de subdivisão, né alguns termos são compostos já na raiz e outros são compostos porque foi usado uma subdivisão no termo principal, tá? A ideia é evitar esses termos com a subdivisão. A gente deixa isso para o catalogador fazer na hora da catalogação da obra dele, porque não há alguns desses termos que não são principais que são usados só como subdivisão, eles podem ser usados sobre vários termos né? Então você não precisaria você teria a orientação para como usá-lo separadamente e o catalogador faria essa combinação na hora de catalogar a obra, mas existem os termos que já na sua natureza são compostas, como você falou.

E existem algumas orientações também no manual, mas ainda a gente tem um pouco de dificuldade. É a gente encontra aí um pouco de dificuldade de trabalhar com esses termos, porque alguns eles são termos diretos por exemplo na LC, mas aqui já foi traduzido com quebrado, sabe já foi colocado hífen, então a gente fica um pouco em dúvida de como quem fez isso a instituição que fez isso primeiro

porque que ela chegou a essa conclusão. [Qual que foi a leitura deles, né?] Exatamente, qual foi ... exatamente. Então aí a gente tem um pouco de dúvida em relação a isso, mas na teoria existe essas orientações a gente tem um pouco de dificuldade de colocar isso em prática porque a gente já traz é boa parte dos nossos registros de outra instituição, né?

Pesquisadora: É sobre o grupo gestor. Há um grupo de pessoas responsáveis pela gestão e manutenção da L.I?

Entrevistado: Sim. Hoje existe uma comissão que são seis pessoas tá? São seis pessoas, são todos bibliotecários, é de unidades distintas da Unesp, né? Se bem que a gente tem três pessoas de uma mesma unidade, mas a gente tem. Hum, a ideia que tivesse uma pessoa para cada área do conhecimento, mas a gente não conseguiu. Então a gente tem pessoas que trabalham desde o início tem pessoas que entraram depois, a gente é ... são poucas pessoas, né para bastante trabalho, mas a gente tem uma boa equipe assim.

Pesquisadora: Tá a outra pergunta e já foi falado a resposta, né? Se você participa do grupo de Gestão Manutenção de linguagem. Você participa desde o início que começou o grupo?

Entrevistado: O esse grupo que você fala é do tesauro? [sim, do tesauro] Ta, os grupos que deram origem a esse trabalho. Eh, eu já comecei a participar no início, e eu sou responsável pela Comissão hoje. Então se eu precisar... mas tem uma pessoa do grupo que é da Coordenadoria Geral da Biblioteca que faz parte da comissão, então tem essa ponte né com a Coordenadoria. Então a partir de comunicados e treinamentos essas coisas a gente tem é tem essa ponte com a Coordenadoria. Mas eu sou responsável pela Comissão. Digamos que eu tenho autonomia pra, sei lá, fazer alguma solicitação, passar uma orientação.

Pesquisadora Quando foi formado o grupo gestor?

Entrevistado: Olha 2020?! Foi no começo da pandemia, porque uma vez implantada a linguagem a gente conseguiu fazer o Tesauro ficar disponível. O grupo que trabalhava com registro de autoridades foi extinto e aí foi formada essa comissão. Esse grupo ele tinha docentes, ta? Que era grupos que foram formados por causa de projetos de pesquisa, ta? Desde o início. Então ele tinha docentes, orientandos desses docentes, outros pesquisadores convidados e tinha o pessoal da Rede de Bibliotecas, os bibliotecários da Rede. Então esse grupo foi extinto e quando o tesauro tornou-se público, essa comissão permanente foi criada para fins de atualização e manutenção do tesauro, ta? A gente tem hoje a prof. Mariângela como consultora na comissão e os seis bibliotecários que trabalham com a

manutenção, né? Mas não tem mais aquele grupo de pesquisa, que deu origem ao trabalho, então hoje temos bibliotecários e a prof. Mariângela como consultora.

Pesquisadora: Agora sobre os critérios de frequência para a atualização. Quais são os procedimentos adotados para a inclusão de termos novos na linguagem de indexação?

Entrevistado: Tá quem detecta a necessidade de inclusão de um termo nova, geralmente é o bibliotecário ou por solicitação de um docente ou de um aluno. Às vezes na hora de fazer o autoarquivamento no repositório o aluno não consegue encontrar o termo que precisa porque existe uma orientação, né? Tem que ter pelo menos três termos do Tesouro e os outros podem ser livres e tal. Então geralmente essas essas solicitações chegam via biblioteca, chegou na biblioteca aí o bibliotecário ele precisa fazer uma solicitação para a comissão do Tesouro, via um formulário. E aí na comissão, a gente vai fazer análise da solicitação tá? Ver se realmente não tem um equivalentes, né? É só um caso de incluir o termo como remissivo. Por exemplo: a gente aceita o termo aí o termo é incluído na base do Alma, na plataforma Alma. E para ir para o tesouro as bases não se conversam. [e como que passa do Alma para o Tesouro] Então, dentro do Alma quem mexe é o analista de sistemas da CGB. Ele tem um programinha que foi desenvolvido na verdade por um orientando, se eu não me engano do Professor Walter. Eu também não sei se já foi orientando da professora Mariângela também, mas acho que era do Professor Walter. Ele desenvolveu um programinha eu chamo de programinha porque eu não sou da área de TI. Não sei exatamente como né? Ele já fez algumas atualizações algumas melhorias, mas ele que pega os registros do Alma e transforma numa linguagem que, eu acho que é XML, se eu não me engano, e leva para o Tematres. Não, ele me manda por email e eu carrego no Tematres. Então ele manda essa carga e são quatro vezes por ano essa atualização, tá? Porque não é uma coisa rápida. Então a gente já está numa fase que não tem grandes cargas de atualizações. A gente mexeu muito com o registro, é tirar uma coisa em lote fazer correções em lote então a gente parou um pouco com isso já tá mais estabilizado, então a gente não vê necessidade de fazer uma atualização mensal. 27:56

Pesquisadora: Quais os critérios que você utiliza para incluir um termos novo na L.I.?

Entrevistado: Faz tempo que eu não catalogo, né? Mesmo, porque às vezes não dá tempo, né? E agora a gente vai ter uma bibliotecária nova lá. Mas é, às vezes é a falta mesmo você começa olhar ali falando, poxa, caramba, esse tema, então eu vejo todo dia fluindo tão utilizado porque que não tem ainda. E aí você tem que buscar. Será que não tem mesmo? Eh as vezes eu pego até registros da de outras bases pra ver esse mesmo livro em outras bibliotecas pra ver como que ele foi catalogado com fogo né? Quais termos? É ver se não existe mesmo. Às vezes eu vou até na obra porque quando é muito nosso que a gente não encontra lá, né? Aham. É a ver como que ele foi eh às vezes até classificado lá a gente da CDD pra tentar dar uma ajuda, né? Mas aqui se a pessoa falar da minha a versão do software

estava desatualizada e o analista ele não conseguia fazer as atualizações trabalhando remotamente e né? E aí acabou eu não tinha muito o que fazer né? Aí também eu acho que o grupo que estava trabalhando com isso né? A mudou de de foco né depois que voltou pro presencial pra comunidade, mas a gente treina o os de orienta os catalogadores pra que eh a comunidade saiba da existência. Inclusive, por exemplo, eh no repositório, tem um tutorial pra pra uso, você já áureo, né? Lá na página do naquela página que o pessoal pode pedir as fichas catalográficas no gerador automático de ficha também. Tutorial. Estão fazendo submissão de de extensão tem um sistema de extensão me engano. Hã? Está errado lá a informação mas existe a informação do eu estou aham. Inclusive tentando contato com o pessoal porque eles colocaram o e-mail quando a pessoa não encontra o termo mandar email pra gente que não é por aí, né? Tem que contratar primeiro a biblioteca. A biblioteca local. É, porque aí você vê a pessoa que está fazendo a submissão no último dia, não consegue achar. E aí fica mandando e-mail pra gente. Excluir as coisas e não é assim, né? É sempre no último dia. Então, acorda no pescoço pra divulgação. Sim. É verdade.

Pesquisadora: Então, são essas as perguntas que eu tinha pra fazer pra você do questionário que eu montei a partir da literatura, a partir das normas, né? E conversando com você eu acrescentei duas perguntas. Uhum. Né? Que você ajudou a formar na sua fala que são os cursos de capacitação que eu acho interessante perguntar se as bibliotecárias tem isso porque é uma coisa eh específica, né? É igual você fazer uma catalogação, você precisa ter um um conhecimento prévio ali no manual, alguma coisa pra nos orientar, né? E e se a correção dos termos pelo grupo né? Se é feita essa parte. Então você tinha falado que sim né?

Entrevistado: Sim é... a gente faz correções, a gente tira relatórios. Eu tenho como tirar alguns relatórios pelo próprio software e eu de alguns relatórios internos pro analista pelo Alma. Então, a gente tira alguns relatórios e acompanhar por por essa eu posso pedir assim relatórios de termos que não tenham ligação com outros termos. Existem os temas que a gente chama de termos órfãos de lixo que estão lá ligação com nada não está numa hierarquia assim tem uma lista de termos que a gente vai eh trabalhando a gente também recente eu puxei um relatório de termos que não tinham termos é genérico, mas genéricos eh indicados então ele tem um regi construído sem termos eh como é que fala? É equivalentes e têm as remissivas mas esses equivalentes por exemplo não existe a hierarquia indicada não sabe se enquadra então alguns é por falha mesmo de você indicar no no campo lá no no subcampo correto. Indicador indicador do campo e hoje é porque faltou mesmo. Às vezes você copiou um registro da BN e na BN não tinha. Aí quem copiou não presta atenção e ficou lá. E então a gente vai tirando esses relatórios nessas correções. Eh tirei o nosso relatório também recente. Ah esse foi mais mais técnico que eu pedi pro pro analista tinha um aparecendo uma sujeirinhas né? Digamos assim no uhum. No cenário aí a gente viu que tinha um dos indicadores e em alguns estavam gravados de forma errada. Aí a gente tem que arrumar trabalho simplesmente. O trabalho é constante. Constante.

Algumas coisas a gente consegue pedir pra ele tirar em lote. Detectou isso nesse campo. Apaga. Agora tem coisas que você precisa corrigir manualmente. Até porque [...] o termo aparece às vezes em vários outros indícios, né? Uh hm. Então você não pode arrumar por exemplo os temas que eu falei obsoleto você não pode arrumar o principal e esquecer dele. Tem outros. Não é. Não ajuda.

Pesquisadora: E agora terminou o questionário só queria saber sua opinião você acha que está clara as perguntas você acha que foi muito específico? É que você achou?

Entrevistado: Eu eu acho que está claro sim eu eu como eu tenho essa essa visão mais assim de de de trabalhar em de ter participado de repente da constituição, né? Das coisas, desde a política. Talvez eu tenha eu não tive nenhum problema assim pra entender. Tá? Eh mas eu acho que isso aqui está tranquilo, que está claro mesmo. Tá. [...]

Agradecimentos e finalização da entrevista.

ANEXO 3 – Transcrição da entrevista com G2

P: Pronto. Você aceita a gravação? [**B2:** com certeza, autorizadíssimo, autorizada] Essa parte é só para eu poder transcrever depois todas as respostas para poder analisar mais fácil, né? [...] Vamos lá então primeiro vou falar do objetivo dessa minha pesquisa que eu estou labutando em concluí-la que é sobre linguagem de indexação vocabulários controlados, né? Então o objetivo dessa pesquisa é ajudar que outras bibliotecas façam o seu gerenciamento elaboração de suas próprias linguagens, como exemplo da USP que tem tudo tão documentado e o da Unesp também [...] **P:** Então, vou iniciar aqui pedindo para você se apresentar a biblioteca que você trabalha e o cargo.

B2: (nome completo), meu CRB é bem baixinho. Eu trabalho na Universidade de São Paulo, na Escola Politécnica e eu sou supervisora técnica do processamento técnico, tratamento da informação como você queira chamar né? E muitos anos de Universidade. Você tem uma ideia eu estou na Universidade desde 1975. Legal, é eu gosto muito do que eu do que eu faço. Eu adoro trabalhar no processamento. Eu adoro organizar essas coisas e eu fiz parte do vocabulário desde o começo até hoje eu estou envolvida com o grupo de gerenciamento do vocabulário. Na ABCD, né na agência de bibliotecas

P1: Perfeito, vamos lá que tem outras perguntas que relacionado a isso vamos chegar lá eu só para contar para você que eu iniciei minha. Iniciei aqui na Unesp como bibliotecária o ano passado consegui passar no concurso. Então também estou trabalhando aí no processamento técnico, né? Tratamento da informação então estou aprendendo muito.

B2: Que bom que eu encontrei mais alguém que goste disso porque o pessoal hoje em dia só quer ser gestor, sabe e a gente precisa ter quem trabalha gerenciar aquilo que vai ser disponibilizado.

Bom é sobre a biblioteca tem umas categorias aqui e a primeira é sobre a biblioteca o catálogo da biblioteca da Politécnica é coletivo catálogo da USP é coletivo?

B2: O catálogo da USP é coletivo. A Poli enquanto não existiu Dedalus ou até existir Dedalus, a Polly tinha um catálogo coletivo da Poli, porque nós éramos oito nove bibliotecas nesse momento. Eu não sei éramos 8 ou 9 bibliotecas. Para você ter uma ideia somos 13 departamentos... A engenharia civil reúne, a biblioteca da engenharia civil reúne os quatro departamentos na área de civil, a biblioteca da Elétrica reúne mais quatro departamentos e ao longo desses anos nós fomos reunindo acervos também.

P: E quais são os principais assuntos que a coleção da sua biblioteca que ela abrange?

B2: É engenharia, todas as áreas da engenharia. Apesar da gente não ter aqui em São Paulo o curso de engenharia de aeronaves. A engenharia de aeronaves está lá em São Carlos, ta? Mas é uma área que eu tenho material também.

P: A biblioteca está inserida em uma rede de bibliotecas?

B2: ela está inserida em uma rede de bibliotecas da universidade.

Política de indexação

P: Perfeito, agora sobre a política de indexação. A biblioteca tem um documento com a política de indexação se sim a política é local ou para todos da rede?

B2: Política de indexação nós seguimos a do grupo do vocabulário não tem nada que seja específico da Polly, não é uma política geral uma política geral seguimos a universidade. Ela é documentada. Ela está escrita em um documento para todos da rede.

Vocês têm um documento?

Para a universidade sim, é um manual, eu posso ver com o pessoal da ABCD para passar para você, tá?

P: Beleza, a linguagem de indexação, o vocabulário, está definida nessa política?

Está! Uma pessoa se você tiver tempo hábil, uma bibliotecária que pode te ajudar muito também, eu eu repeti mencionei a ABCD duas ou três vezes a pouco, é a Roseli lá na ABCD. Ela também está conosco desde o começo do vocabulário. E ela manja assim muito. Ela é da agência da agência de bibliotecas, a agência de bibliotecas e coleções digitais. O antigo SIBI, né o sistema de biblioteca.

P: As coisas vão mudando, né? E a forma como faz a gestão o uso e a manutenção da linguagem é descrita em algum documento, como fazer essa atualização.

Olha ao longo do tempo o ano passado nós tentamos voltar ... como é que fala?! Reativar todo este controle, então formamos novamente um grupo com alguns representantes de cada área. Mas como todo este trabalho de gestão do vocabulário ficou um pouco de lado parado, até porque por causa disso a universidade teve dois programas de incentivo demissão voluntária, muitas aposentadorias, morte também de funcionários tivemos a pandemia. Então tudo isso ficou um pouco parado. A gente tem esse grupo e a Roseli tá nesse grupo que eu falei para você lá na ABCD é que cada vez que alguém precisa de um assunto novo é solicitado. Existe uma base de sugestões, tá? E aí é a unidade que precisa solicita já identificando, qual é a área desse descritor e aonde ele deve ser colocado. É essa a orientação que é dada e onde ele deve ser inserido na árvore, vamos dizer assim, né na hierarquia. Então isso vai sendo feito a medida que as pessoas estão pedindo nós estamos numa fase de muitas alterações nesse trabalho todo. É um trabalho grande? É um trabalho grande é para quem gosta muito? É para quem gosta muito. E eu costumo dizer que a gente tem muita coisa para fazer, mas às vezes você fala assim, olha hoje eu

vou parar e vou mexer nisso, né? Então nós estávamos num grupo mais ou menos grande de pessoas o ano passado trabalhando nisso e as pessoas foram pedindo para sair porque justificaram que não tinham condição de trabalhar ou até de fazer uma reunião assim online como eu tô fazendo com você agora né? Então o que aconteceu para a Poli, eu pedi e eu mantive isso, a Valéria de São Carlos também da escola engenharia de São Carlos também me ajudou a trabalhar um pouco nisso, mas nós fomos mexendo em alguns assuntos da área de engenharia que não estavam bem colocados, vamos dizer assim, ou que estavam sendo usados de forma inadequada, né? Então a gente vai mexendo e cada vez que eu pego é um assunto. Pra você ter uma ideia, o trabalho do vocabulário começou a um pouco mais de 30 anos. Eu me lembro bem disso esse trabalho, essa ideia de um vocabulário consulado na universidade começou há quase 30 anos mais ou menos [P: É a USP foi pioneira, no Brasil, né?] É, assim exatamente, de 25 a 30 anos deve ter isso mais ou menos então a gente vê não é já se falou muito. Vamos mudar. Precisa mexer com a plataforma precisa trocar precisamos fazer bilíngue pelo menos né? Se não for mais do que isso é exatamente pelo menos precisa, pelo menos no inglês.

Você quer ver uma outra coisa, que a gente mexe muito e porque como é de todas as áreas cada área tem um vocabulário internacional específico para sua área. Trabalha de formas variadas, na verdade. A gente tá indo, mas a Roseli também pode te dar mais informações sobre a política

P: Você estava falando da linguagem, né do vocabulário que começou há 30 anos o que que levou essa USP e os bibliotecários perceberem que precisava fazer esse projeto de construir o vocabulário.

B2: Olha, naquela época, a verdade é um grupo de bibliotecários, nós estávamos ... A Cibele também ajudou desde o começo, mas eu acho que o nome pioneiro desse trabalho é o da Vânia. Você deve conhecer a Vânia Mara Alves de Lima, professora da ECA verdade a pioneira disso aí ela, porque ela que estudando e trabalhando. Ela trabalhava na FAU nessa época, né? Todo o trabalho dela de mestrado e doutorado foi em cima disso e ela era orientada e trabalhava com a Nair Kobashi, também da ECA. Então uma pessoa tão querida e que tá aposentada já alguns anos, né da época então elas é que pensando nisso começaram a porque a gente usava para o Dedalus, um controle vamos dizer entre aspas ou os assuntos que eram usados para a indexação no Dedalus eram era uma tabela do CNPQ. [P: Então era muito simplificado para a indexação] Nossa, eu tinha que por só Engenharia! Ou engenharia civil ou engenharia elétrica. Quer dizer uma coisa assim que não tinha cabimento. Como que faz para recuperar né? E aí elas que começaram com isso aí fizeram o quê? Trouxeram meu nome daquela professora Ana Maria da PUC, de linguagem também você me desculpa falha de memória, mas eu a gente pode conversar com a Vânia também, se você quiser um pouco de história, elas vão lembrar disso dessas informações isso deve estar tudo muito bem guardado em trabalhos, porque nós escrevemos vários trabalhos que nós apresentamos em congresso, foram publicados... Em então o que que elas fizeram? Elas chamaram e reuniram esse pessoal é? Apresentaram inicialmente para os bibliotecários

que trabalhavam indexação. Nós fizemos cursos rápidos de na verdade muito mais orientações para tudo isso. [P: até essa é uma pergunta minha daqui para mais para frente] Então, vamos lá a sua pergunta porque senão eu vou falando e fica solto.

P: No desenvolvimento do vocabulário, né? É que aqui eu uso o termo linguagem, mas é a mesma coisa foram utilizadas outras linguagens para no desenvolvimento do vocabulário da USP?

B2: No início não. Nós queríamos fazer um vocabulário controlado da USP, em português, porque primeiro porque nós reunimos todas as especialidades. E segundo por causa das particularidades de cada área, então eu naquela época eu falei gente. É nós trabalhamos na Poli com o Subject Headings da LC, e ele tem um trabalho bem feito de hierarquização de termos relacionados maiores e menores. Sem problema. Uma grande atualização e eu estou falando da Library of Congress, né? Podia aproveitar para todo mundo, mas aí o pessoal da área médica tem o DECS, né? E que vai atender aquilo lá, olha mas a biblioteca do congresso, o Subject Headings da biblioteca do Congresso. Eu sou fã do Subject Headings, tudo que eu preciso que eu não sei aonde pôr que eu preciso pedir a mais, eu vou primeiro ao Subject Headings da LC para depois solicitar o acréscimo e às vezes eu acho que um assunto que tá no vocabulário da USP não tá bem colocado eu me guio pelo Subject Headings. Aliás o ano passado quiseram retornar com essa discussão de bases das áreas e eu falei: Olha, eu acho que ainda é o melhor mesmo ele não sendo específico para engenharia porque na engenharia a gente... [P: Mas ele foi usado como na seleção de termos assim para compor a linguagem da USP?] Não, não é que eu resolvi eu usava para a indexação aqui quando eu vim para a Poli, é que eu comecei a usar isso porque eu trabalhei no instituto de matemática estatística na organização daquela biblioteca e naquela altura com a diretora que eu tinha lá de vasta experiência nós usamos a classificação, o sistema de classificação da biblioteca do congresso, aí a grande hierarquia e a subdivisões todas foram trabalhadas com os docentes aqui na Poli até hoje eu trabalho com docente porque eu sou bibliotecária eu não sou engenheira. E ao contrário de outros lugares do mundo que o especialista da área faz uma pós-graduação em biblioteconomia. [P: E na composição da estruturação do vocabulário. Aonde então que foi buscar esses termos para compor o vocabulário?

B2: O da engenharia foi trabalhado com os docentes, com cada especialista de cada área, sabe porque é o que você estava falando, uma coisa é a teoria e outra coisa é a prática. Então o assunto na prática é por isso que eu acho muito rico a gente trabalhar ... e sabe e se colocar no mesmo nível de docente. Professor, é o seguinte especialista é você, eu sou bibliotecária. Eu preciso de hierarquizar isso. Então, a gente sempre trabalhou assim e aqueles estranharam muito que o pessoal não estava acostumado com isso, mas eu vinha de uma experiência de uma outra unidade que isso funcionou muito bem. [P: E no fim você conseguiu trazer isso?] Eu trouxe isso o grupo aceitou aderiu e vários docentes de todas as áreas passaram a ajudar todo mundo a montar. Por isso o vocabulário da USP é da USP, é diferente. Você sabe que com esse trabalho eu aprendi que o engenheiro de Formação mais completa é o

engenheiro naval. Engenheiro naval porque o engenheiro de estruturas trabalha com estruturas o engenheiro de telecomunicações trabalha com telecomunicações o engenheiro mecânico trabalha com motor. O engenheiro naval tem que entender de estrutura de eletricidade, de eletrônica, de telecomunicações. Não é a toa que a Marinha do Brasil tem um convênio de desenvolvimento de cursos, né com a Poli e o escritório da Marinha aqui dentro. [P: Eu não sabia, que legal.]

P: Bom vamos lá então Silvia, essa daqui é um pouco específica, mas foi usado alguma Norma para guiar a elaboração da linguagem da USP?

Eu acho que para algumas coisas. A gente sempre esteve voltado a buscar o que nos atendesse mais, a nossa realidade. Normas que você diz...? [P: no sentido de construção mesmo. Sabe Norma da ISSO, ANSI/NISO, ABNT, essas agências normalizadoras.] Eu acho que não, não foi uma coisa muito de experiência nossa, mas uma coisa é muito de linguagem natural também, sabe aquela coisa o termo no geral, vai no plural o termo específico vai no singular. [P: Essa pergunta também mais para frente. Vamos lá hahaha]

P: Essa trajetória toda de essa história do vocabulário, ela é descrita em algum artigo, algum livro?

B2: Nós temos várias publicações disso daí e eu vou te dizer a Vânia e eu demos uma palestra de dia inteiro, o pessoal ficou por aqui de nós especialmente de mim, porque eu falo para caramba, né? E como diz a Vânia. Fala você! Nós demos uma palestra na Fundação Escola de Sociologia e Política aqui em São Paulo. Que é quem tem o curso mais antigo tradicional da biblioteconomia antes até da USP ter a sociologia política, já tinha esse curso. Nós demos um nós passamos um dia lá falando sobre isso ensinando passando essa informação. Então isso é muito legal também, mas quem tem todo esse controle é Vânia, porque a Vânia quero a coordenadora do grupo. Então, todo esse material tudo isso.... Você conhece a Vânia? [P: não conheço pessoalmente só publicações] Ah, mas olha. Então ela é uma pessoa extremamente sensível, um amor de pessoa a semana passada, encontrei com ela na hora do almoço. Ela ainda falou você tá aqui ainda? Eu falei tô, nós estamos aqui ainda. Mas olha, você pede para Cibele fazer o seu contato com a Vânia que ela também pode te ajudar.

Aspectos técnicos da linguagem

P: Qual a linguagem de indexação da sua instituição?

B2: É o da Universidade, o VocaUSP ou o SIBIX que o pessoal fala.

P: Quais são os campos de assuntos abordados nesse vocabulário?

B2: Eu queria lembrar e a subdivisões dele, não sei se é nessa pergunta que te interessa a subdivisões, mas ele quer dizer ele te dá a possibilidade de indicar o assunto com divisão de forma, com divisão geográfica, O que mais que tem ali? [**P:** Mas então ele abrange vários assuntos, várias áreas do conhecimento]. As áreas do conhecimento todas que são desenvolvidas na universidade.

P: É utilizado no vocabulário siglas como USE e UP?

B2: Sim, nós usamos o USE, especialmente o USE, nós queremos que ainda não conseguimos porque é o que tá por trás a estrutura dele que ainda não deixou, mas a gente quer usar especialmente USE também ou VER também ver também.

P: Você tinha falado disso antes, a linguagem de indexação é multilíngue?

B2: Ela não é, neste momento ela tá só em português.

P: A linguagem utiliza qualificadores para diferenciar o mesmo termo?

B2: Sim, utilizamos. Isso está na hierarquia, ele não é acrescentado na indexação, ele é montado desse jeito dentro da área dele porque, por exemplo, eu podia por... existir o termo uma vez só e cada vez que eu fosse indexar eu usasse o qualificador de área, não. Já está na árvore hierárquica o qualificador identificando.

P: Entendi. Há alguma uma orientação de como os termos devem ser gramaticalmente, por exemplo os termos no plural, no singular?

B2: Tinha começado a falar, né? Também existe isso está lá na orientação de como de como usar como montar e como usar. [**P:** Tá depois eu vou pedir por e-mail. Se você puder me mandar essa parte.] A gente tem isso. Por isso que eu falei para você, isso eu peço para Roseli e peço para a gente tem tudo isso. Espero que tenha isso guardado.

P: Como são empregados os termos compostos? Há algum momento que eles são divididos, por exemplo: gestão de tesouros, pode ser gestão de tesauros ou gestão + tesauros.

B2: Existe algumas situações daquilo que está consolidado, como área de estudo ou de interesse é colocado “gestão de projetos”, “gestão de pessoas”, não é separado porque é um conceito só. Então ele já está ele já aparece desta forma alguns outros é que não existe a área de estudo, aí você conta. Eu posso ver sobre um monte de exemplo disso depois mandar. [**P:** Eu agradeço!]

P: Há um grupo de pessoas responsáveis pela Gestão e Manutenção do vocabulário?

B2: Sim. Sim foi o que eu falei para você. O pessoal quando viu que como nós tivemos depois da pandemia, pois a pandemia foi um caos, foi gente que morreu gente que saiu. Então esse nosso grupo que estava grande foi reduzindo até o final do ano que passou várias pessoas. “Olha eu tô à disposição, se você quiser tirar uma dúvida, mas eu não posso trabalhar me comprometer”, mas a gente tem sim lá na é um grupo de trabalho. Como é que fala que para gerir tudo isso é um grupo de trabalho de gente lá da ABCD. **[P1:** Você participa ou participou desse grupo?] Desde o começo.

P: E quando foi formado o grupo gestor? Pode ser aproximadamente.

B2: Isso é da década de 90, não sei te dizer. **[P1:** Nossa foi muito foi muito pioneira mesmo.] É isso é daquela época, ele é uma coisa que não existia, né?

P: E os principais responsáveis atualmente? Por comandar esse grupo e as decisões...

B2: Olha nós estamos divididas por área. Então na verdade não tem alguém. A Roseli e a Cristina que são bibliotecárias que trabalham na ABCD que é o grupo gestor das bibliotecas da universidade, é que vamos dizer, mais ou menos são as comandantes do grupo. E aí eu a Valéria, na área de engenharia depois na área médica, eu não sei te dizer um nome agora, mas eu tenho isso, elas têm por escrito.

P: Houve algum treinamento ou capacitação dos membros do grupo?

B2: Na verdade são as pessoas que estão acostumadas a trabalhar com a indexação, né, então o treinamento e a experiência é de quem já trabalha com isso. Não é nada específico para gerir, entendeu?

P: você precisa incluir um novo termo que não tenha lá, qual que é o procedimento que você usa?

B2: É assim e serve para qualquer um para qualquer um de nós, né? **[P:** Na verdade já juntando com a outra pergunta para inclusão e exclusão de termos]. É assim é para inclusão como o desenvolvimento da ciência nós aqui então na engenharia é trabalhar com tecnologia de ponta tem coisa nova todo momento todo momento. Então o que eu tenho que fazer eu preciso a gente tem o vocabulário tem online uma coisa que a gente chama de base de sugestões, tá? Então eu entro lá. Eu, B2, com a sigla da minha unidade, com a minha senha, tá? E vou pedir um novo termo. Então eu tenho que identificar se é um descritor, se é um qualificador. Qual é a tabela que ele é para ser colocado porque às vezes eu preciso colocar uma área geográfica dependendo do assunto é importante que você tenha uma área geográfica. Então, o que que eu faço? Se é um descritor normal, eu indico que ele é um descritor, eu indico qual é a área dele, eu indico eu dou uma definição para este termo, eu indico a bibliografia de onde foi tirada essa definição, ele é um formulário que você preenche tem os campos prontos para você preencher é bem feito aquilo então aquilo vai ... agora, por exemplo, estou pedindo um termo da área de

engenharia que é a minha especialidade. Então eu mesma tenho autorização para pedir. Eu tenho autorização para dar o OK, meu caso. Eu não sou muito exemplo por causa disso, sabe? Como eu estou lá desde o começo, eu já fui coordenadora do grupo e tal. Então qualquer outra pessoa de uma outra área pede para o indexador pede um termo da área médica, só que não é um bibliotecário da Medicina. Mas é da área médica biomédica, ele pede indica a definição do termo diz aonde tem que ser colocado e fica inclusive o código numérico de onde ela acha que aquilo deveria ser inserido. Aí é alguém lá da Medicina vai dizer está certo? Pode ser assim. O que é tudo isso? Cada vez que você dá esse comando passa de nível e vai para o outro responsável daquele nível maior. Então é essa pessoa da Medicina disse tá certo? O que a área da Psicologia pediu? Quando ele diz o Tá certo ele passa por um nível que vai direto para o pessoal lá da agência que é quem mexe e acrescenta naquela plataforma, naquela base aquele assunto com todas aquelas indicações. Quando sou eu B2 que peço um ... por exemplo, outro dia uma das bibliotecárias que trabalha aqui comigo disse “eu preciso... está na moda veículos elétricos. E a gente não tem no vocabulário os elétricos”. Não tem problema, o que aconteceu? Eu fui conversar com o professor da área dos veículos elétricos. Porque agora não bastasse ser elétrico tem os veículos movidos a hidrogênio Verde. Então o que acontece tem uma empresa chinesa desses veículos que entrou no Brasil e que colocou uma montadora moderníssima lá na Bahia. Já tá funcionando é o tal do BYD que tá na moda moderno bonito, Luciano Huck tá distribuindo BYD porque eles estão pagando e fazendo uma projeção nacional ou então é uma beleza a Globo tá mandando ver com eles. Aí eu fui conversar com o professor como eu ainda não tinha uma definição em português nisso o professor que está se especializando nisso, me deu aí eu pus como fonte bibliográfica o professor. Ele falou para mim, você vai colocar? Eu falei professor a minha hierarquia de assunto é essa, onde eu ponho Veículos Elétricos? Você coloca aqui Veículos Elétricos. Muito bem. Então eu acertei o código tá? De subordinação. Mande na base de sugestões, eu mesmo autorizei, então eu passei no nível de solicitação para o nível de autorização. E aí passou pro nível 3 que é o que vai direto lá para turma da gestão que vai intercalar e vai acrescentar aqui, muito bem. Só aqueles lá a Roseli, por isso que eu falei. A Roseli é uma pessoa fantástica para você conversar. Ela falou para mim assim: B2, você tá pensando em veículos elétricos como automóveis? Eu falei estou. E ela falou e a bicicleta elétrica? E o trem elétrico? E o ônibus elétrico? Sabe, e a Rose e também tem 20 anos ou mais trabalhando com isso, aí eu falei “Roseli, é seguinte minha filha, eu vou entrar em férias, ver a minha turma lá fora do país e você vem com essas perguntas com isso na véspera de eu sair de férias, você põe o Veículos Elétricos ali que eu preciso atender a minha produção intelectual aqui da Poli e quando eu voltar a gente rever tudo isso. E aí o que que aconteceu? Ela pois Veículos Elétricos lá eu pedi de manhã, ela botou de tarde. No dia seguinte, eu entrei no Dedalus, pesquisei toda a produção tudo que tinha de acerto da Poli, que eu podia mexer no acervo dos outros, quando eu não tô ligando para o meu atrevimento eu faço isso, mas era fim de ano eu falei eu vou sair de férias, eu vou mexer no que é da Poli. Aí eu achei lá é 80 e poucos registros 70 e poucos registros de material que tem na Poli sobre Veículos Elétricos. Sabe o que eu fiz? Entrei registro por registro, acrescentei o assunto Veículos Elétricos e fechei. Porque sabe o que acontece Carolina? Eu

acho muito importante quando a gente tem material para pedir um assunto, quando esse assunto é acrescentado na hierarquia, a gente tem que entrar na base e colocar esse assunto dos nossos registros, sabe? se não perde o sentido, tá na moda hidrogênio Verde. A Toyota botou um carro lá na entrada do prédio da engenharia mecânica para test drive movido a hidrogênio Verde. Então esse como eu cheguei de férias da semana passada, tá lá em cima da minha mesa a anotação. **[P: E para exclusão de termo? Um que você vê não esta politicamente correto ou é defasado...]** Aí a gente diz assim exclusão a função do mesmo jeito, eu vou entrar lá e vou pedir a exclusão de um termo, eu tenho que justificar. Às vezes eu tô excluindo aquele mas ele na verdade não é escusa. Ele é uma alteração, ele mudou exatamente pelo desenvolvimento da ciência. Então não é excluído. Ele é modificado. Ele é modificado. É às vezes a gente pede exclusão porque o mesmo assunto. Isso é mais comum acontecer era né? Isso já faz tempo que eu não vejo acontecer, é aquela história do plural e do singular eu tenho o mesmo assunto no plural e no singular e não faz sentido. O certo é só plural ou é só singular. Aí a gente pede para excluir passa a valer só daquela forma correta da linguagem natural.

P: E assim uma perspectiva mais pessoal qual que é o critério que você usa para fazer a inclusão do termo? É frequência que vai aparecendo de trabalho?

B2: Por exemplo, se olha você quer ver eu tenho um exemplo muito muito claro disso aí. Há alguns anos apareceu um assunto que fez uma revolução na engenharia de produção e na administração, que foi a “Reengenharia”. Verdade pessoal que chamar de reengenharia, eles quiseram mexer numa vamos dizer numa racionalização de serviço de pessoal de otimização aí a descobrir o que não precisava de 10 pessoas para fazer um serviço bastava sete. Então, eles mandavam três embora reduziam o custo. Só que em vez disso ser feito com critério isso não foi só no Brasil, foi no mundo inteiro, tá? Era uma coisa para ser feita com critério que juntam os setores, não ser demitir pessoa, mas você junta áreas que trabalham assim, sabe? Como é que? É não precisa ter dois chefes, basta um mas você não precisa mandar embora o outro, ele pode continuar trabalhando. E isso foi feito sem critério. O que aconteceu? Demitiram um monte de gente foi um tempo de experiência do que eles acharam que ia ser uma nova uma nova um novo modelo de administração. Chegar à conclusão que não. Mas foram publicados inúmeros, trabalhos e livros sobre isso. Mas isso perdeu o sentido, mas eu não posso pedir a exclusão do termo reengenharia porque tem todo aquele acervo lá. Agora “hidrogênio verde” é uma coisa muito nova e eu ainda não tenho 20 ou 30 trabalhos sobre hidrogênio verde. Eu tenho que esperar mais um pouco, mas eu já estou estudando, ao contrário de veículos elétricos. **[P: Então você não vai introduzir ainda por ser não ter muitos trabalhos sobre?!]** Não, quando você faz uma pesquisa no Dedalus, você recupera por todos os termos que tiverem dentro do registro, não necessariamente precisa estar no assunto. Agora se que com a preocupação o aquecimento global, com a sustentabilidade isso começasse desenvolver muito e realmente mostrarem que se existe, muito sério, é como o “Ecoponto”, ecoponto é aquele lugar que você é aquelas estações de reciclagem, né? Mas isso demorou muito para parecer como

Ecoponto. Então eu tenho que esperar ter. É nesse sentido aí... eu falei demais, né? **[P: Não tá ótimo assim tem bastante material para nós.]**

P: Faltam só mais quatro perguntinhas tá? É utilizado algum software para gestão do vocabulário?

B2: É, mas isso eu não sei responder para você isso tem que ser o pessoal lá da ABCD.

P: Você acha assim questão pessoal de bibliotecário de quem usa que esse software atende as necessidades da instituição?

B2: Eu não sei te dizer exatamente. Não por isso é que tanto se estuda se tão estamos trabalhando para tentar mexer porque o Dedalus é um pouco restrito, né? Então a Sara que é a bibliotecária vamos dizer acima da Cristine e da Roseli, estava trabalhando muito nisso, e ela estava pensando até no Tematres para controlar o vocabulário. Mas isso não está definido que a gente depende de dinheiro, você sabe como é que é isso.

P: O vocabulário ele tá disponível no catálogo online da biblioteca para os usuários fazerem consultas?

B2: Esta internet disponível para qualquer um qualquer lugar do mundo.

P: E o vocabulário, ele é conhecido pelos usuários da biblioteca? Eles sabem que ela pode ser uma ferramenta aliada na busca e na recuperação da informação?

B2: Aqui na Poli nós temos uma bibliotecária que faz esta apresentação é para alunos de graduação e de pós-graduação. A Cidinha faz umas palestras bem legais para isso e eles sabem inclusive na apresentação dos trabalhos acadêmicos deles existe a indicação de procure a bibliotecária que atende a biblioteca do seu departamento para definição dos seus assuntos de acordo com o vocabulário controlado da USP. A gente indica isso então, eles vêm atrás buscar então, eles conhecem.

P: B2, é isso! Essa turma de perguntas, foi ótimo...

Agradecimentos e finalização da entrevista.

ANEXO 4 - Transcrição da entrevista com In1

P: Pronto tentar mais próximo de você ir para pegar o assunto. Questões de praxe, você permite gravar nossa entrevista para depois eu transcrever, não vai ser divulgado os dados do bibliotecário.

B3: Sim, permito.

P: Primeiro falar um pouquinho sobre o que que é a pesquisa tá né? O objetivo principal assim o produto que eu quero tirar da tese é como se fossem orientações para as bibliotecas que querem construir suas linguagens, que querem fazer suas próprias, como o tesauro UNESP e da USP. Inicialmente eu ia pegar várias bibliotecas assim, só que POR CAUSA de tempo esse tipo de coisa eu vou focar só USP E NA UNESP. Então o objetivo principal é elaborar orientações sobre o gerenciamento da construção e manutenção de linguagem de INDEXAÇÃO para bibliotecas universitárias usando a literatura e usando a prática das bibliotecas. E eu queria começar falando para você se apresentar, no nome da BIBLIOTECA que você trabalha, o seu cargo...

B3: Meu nome é B3, sou bibliotecária, trabalho há 11 anos e meio na Divisão Técnica de Biblioteca e sou supervisora técnica de seção.

P: As perguntas, elas estão divididas em categorias que nem agora é informações sobre a biblioteca. O catálogo da biblioteca que ele é coletivo?

B3: É coletivo.

P: Quais são os principais assuntos que a coleção abrange aqui também da biblioteca?

B3: Voltado para a área de ciências biológicas, mas também tem assuntos voltados para área de humanas, mas a sua na sua maioria biológicas.

P: Certo, a biblioteca está inserida em uma rede biblioteca?

B3: Sim, rede de bibliotecas UNESP.

P: Sobre a política de indexação. A biblioteca tem um documento com a política, se sim a política de indexação é local ou para todos da rede?

B3: Tem o documento, mas é documento, de uma política elaborada para rede de biblioteca

P: A linguagem de indexação é descrita na política, qual linguagem deve ser utilizada está escrita na política de indenização?

B3: é escrita, quanto tempo que eu não leio a política, sim a linguagem está sim. A linguagem tá definida tá definida sim. Sabe quando você precisa recuperar o documento para ter certeza que você tá falando? Mas tem sim.

P: A forma como fazer o uso da linguagem a manutenção atualização que eu quero dizer é é descrita em algum documento?

B3: Ah sim, eu não me lembro agora se isso está na própria política. Tá, mas sim, nós recebemos orientações de como proceder para alimentar, para corrigir, enfim de como fazer o trabalho com a linguagem, então os bibliotecários tem acesso tem acesso essas informações.

P: Você sabe o que levou a construção do Tesouro da Unesp, porque o que impulsionou a necessidade assim que eles começaram a pensar em fazer isso?

B3: Eu não sei exatamente o que levou, me lembro, não me recordo disso, mas eu acredito que a busca é facilitar o trabalho dos bibliotecários dentro da rede, buscar trazer para rede o maior número de termos daquilo que a gente trabalha no nosso dia a dia, ou seja, enfatiza o que eu falei antes que é facilitar realmente o processo de indexação, acredito que a ideia foi realmente isso essa.

P: E foram utilizadas outras linguagens no desenvolvimento dos seus dados, né? Se você sabe que por exemplo?

B3: Para ter um parâmetro!? Sim, foram a Biblioteca Nacional na época ainda funcionava bastante o BIBLIODATA, o catálogo da USP, eu acho que teve alguma coisa também, mas não sei dizer assim muito claro ali de uso com a USP. É acho que a Biblioteca Nacional e o BIBLIODATA foram os mais utilizados.

P: Você sabe se foi usado alguma Norma para guiar as decisões na elaboração da linguagem?

B3: Não sei, mas acredito que sim, não sei, tá?

P: Existe algum documento publicado ou um livro ou projeto que tem a trajetória da elaboração da linguagem de indexação?

B3: Tem o livro. Tem o livro publicado pela professora Mariângela.

P: Agora uns aspectos mais técnicos da linguagem. Bom, qual é a linguagem de indexação da sua instituição?

B3: Linguagem UNESP.

P: Quais são os assuntos os campos de assuntos abordados na linguagem há uma linguagem específica única assim só daquele tema?

B3: É uma geral, você tá querendo dizer isso? É aquilo que eu acabei de falar, os assuntos agora de todas as áreas do conhecimento a linguagem é uma abrange todas as áreas do conhecimento.

P: A Linguagem utiliza-se como “use” para esquematizar e organizar elementos ou notas?

B3: Sim

P: Ela é multilíngue?

B3: Nossa, eu confundo com o DeCs que tem tudo isso. Eu ia até fazer uma comparação com DeCs que ele tem a nota de escopo, a nota de indexação que ele dá uma ideia melhor que eu não me lembro de ter visto uma linguagem no UNESP, né? Mas até nessa questão da nota de escopo, o “use este” e tal, o DeCs faz uma forma com muito mais clareza do que ele tá sendo colocado do que a linguagem Unesp, mas eu não consigo lembrar se mantermos em inglês. Porque o que eu mais uso, realmente é o DeCs.

P: A linguagem utiliza qualificadores para diferenciar ao mesmo termo com diferente significados, por exemplo um mercúrio pode ser um planeta, pode ser um Deus Grego, o mercúrio pode ser um elemento químico... Ele tem esses qualificadores?

B3: Carol que dúvida cruel. Porque eu sempre vou me confundir com o DECS, porque o DECS tem os qualificadores. Na Unesp não me lembro de ter visto isso. Talvez de uma forma diferente do que a gente visualiza no DeCs, uma forma de observação diferente. Deve ser mais claro, eu acho que não até a forma de colocação ali poderia ser mais assertiva não sei pelo menos eu tenho essa sensação.

P: Mas é porque a gente tá tão acostumada a mexer, né?

B3: E às vezes os nossos termos a gente pode adotar, a gente acha muito mais no DeCs do que na Linguagem Unesp.

P: Ela tá crescendo ainda, né? Há, alguma uma orientação de como os termos devem ser, por exemplo, singular ou plural, podem ser utilizados verbos?

B3: Eu acho que há. Eu acho que na política deve ter alguma coisa a respeito disso e na própria linguagem a gente tem de ver o USE desta forma, eu acho que tem sim as siglas, né para falar qual que é o certo.

P: Como são empregados os termos compostos como por exemplo gestão de tesouros há momentos que eles são divididos como gestão + tesouros, quando não encontram o termo único?

B3: Você pode juntar, você pode juntar. Então pode, mas nós estamos tendo problemas no Alma para isso em termos compostos e a informação que se tem é que isso vai ser revisto na política. Então eu não sei exatamente como isso vai ficar. Então hoje se não acha o termo composto, né o conjunto, né da ali completo. A gente consegue fazer, mas ele já não fica de forma correta voce ve que o Alma ele nem dá o linkzinho de que aquele termo tá correto. Então ele precisa ser revisto na política. Então a política ela prevê isso, mas o sistema hoje de trabalho ele não aceita da forma como está na política. Então tem que ser revisto ou um ou outro.

P: até porque o ALMA é novo

B3: É novo no Brasil, é novo na Unesp, mas não é novo na catalogação mundial no processo, então não sei como isso é feito lá fora, como que é feito essa questão de termos compostos.

P: Sobre o grupo gestor. Há um grupo responsável pela Gestão Manutenção da Linguagem Unesp?

B3: Havia um grupo, hoje há... Porque o grupo da linguagem mesmo é a Rosane, eu não sei se o grupo da Mariângela ainda está ativo.

P: Eu acho que passou a responsabilidade para Rosane.

B3: Só a Rosane que responde para o grupo. Então eu não sei se o grupo ainda está ativo tá? Mas existia sim um grupo, mas hoje quem responde pela linguagem de indexação é única exclusivamente a Rosana

P: E você participa ou já participou do desse grupo?

B3: Não.

P: Não participou. Você sabe quando foi formado esse grupo?

B3: Olha, ele é anterior de 2010 ou até anterior é isso quando eu cheguei em 2012, o grupo já tava bem adiantado nesse processo, então ou era 2010 acho que é 2010. Tá eu vou trabalhar com essa data de 2010. 2012 eu cheguei ele já estavam em andamento

P: Então é quem foram os principais responsáveis que você é você falou a Rosane, né?

B3: Então Mariângela tá na frente disso a Sulamita, a Rosane, a Telma de Marília o Alan e Rosana. De Araçatuba... o Cláudio e eu acho que a Marcinha de Rio Claro, deve ter mais. Não lembro... a Helen, mas não tenho certeza da participação da Helen.

P: Houve algum treinamento, capacitação dos membros do grupo assim para os outros bibliotecários?

B3: Deles para a rede? [**P:** Sim.] Assim faz tempo, mas eu lembro que teve sim.

P: Quais são os procedimentos adotados para a inclusão de novos termos da linguagem?

B3: Então, assuntos tópicos, eles devem ser encaminhados para o grupo para um e-mail, né direcionado a esse grupo aonde você deve dar o maior número de informações sobre aquele termo né? A referência de onde veio, né? O significado. Enfim o maior número de informações sobre o termo e enviar solicitação por um formulário.

P: Então tem um formulário que deve ser preenchido?

B3: Tem, se ele ainda estiver disponibilizado. Mas sempre teve um formulário e a gente mandado por ali.

P: E quais são os procedimentos adotados para exclusão de algum termo?

B3: Correções. A mesma coisa enviar, né? Quais são os termos, o porquê, justificar o pedido e encaminhar para o grupo, o grupo vai analisar e se julgar procedente eles vão fazer exclusão, as condições.

P: E agora, quais são os critérios que você utiliza para pedir um termo? Assim você vê que está precisando?

B3: De acordo com a demanda do nosso dia a dia de trabalho. Então eu tô fazendo a indexação de um trabalho e é um termo muito específico e que eu não localizo em nenhuma das bases inclusive nas bases internacionais. Então eu tento buscar essas informações normalmente eu já tive acho que duas situações dessa e eu pedi inclusive ao próprio docente quando ele fez questão de ter e eu expliquei como que funcionava o processo e eu disse eu vou ter que fazer isso, então se o senhor tiver essa as informações, né? essas referências onde considerou? onde pegou, né? a definição do termo. e aí eu então eu junto esses elementos, então se eu tiver a solicitação de um docente eu já peço para ele, então me colocar essas informações. Mas eu também busco as minhas e enfim. E se for meu mesmo é enviar de acordo, mas eu assim não todas as vezes todos esse tempo de trabalho, eu não me lembro de ter batido com o termo que eu sozinha sem a questão do docente eu não encontrasse não, encontrar na Linguagem

UNESP, mas eu consigo encontrar mesmo colocando em inglês e buscando na LC. Aí eu consigo já consegui encontrar.

P: Está acabando. São realizadas as correções dos termos da linguagem? Você sabe se o grupo faz essa correção?

B3: Eu acredito que assim pelo grupo eles fazendo esse levantamento... não sei, não acredito que aconteça, mas sim, a partir das indicações dos bibliotecários da rede.

P: Sobre o *software* do tesouro. É utilizado algum software de que abrange o tesouro?

B3: Eu não conheço. [**P:** É aquele programinha que a gente entra na interface, ele se chama Tematres]. Não conhecia, desconheço. Você pode até colocar que eu desconheço.

P: Você acha que esse software atende as necessidades da instituição? Por exemplo, é fácil de você visualizar o que aparece para você?

B3: Quanto a isso sim, é fácil. Acredito que as informações poderiam estar mais melhor disponibilizadas ali ou poderia ter, como eu citei há pouco, uma nota de indexação, sabe? Para alguns termos eu acho que faltam alguns elementos a mais nas informações colocadas ali, mas eu acho que ele é bacana sim.

P: Agora disponibilização online. Ela tá disponível no catálogo da biblioteca para o usuário para poder fazer uma consulta?

B3: No catálogo não, mas na página da página da biblioteca ele encontra lá, pode ter acesso por lá.

P: E você acredita que a linguagem é conhecida pelos usuários?

B3: Só me corrigindo, no catálogo tem sim. Pelo menos eu buscando internamente eu busco pela autoridade, eu nunca fiz uma busca como usuário, mas eu acho que ele tem pelo site é eu acho que são se eu tenho a opção autoridade ali dentro.

P: E você acredita que essa ferramenta, a linguagem é conhecida pelos usuários? Eles sabem que pode ser uma ferramenta aliada na busca de informações, na busca da recuperação da informação?

B3: Eu acredito assim, os pesquisadores e para os alunos que fazem pesquisa, sim, eles têm esse conhecimento, né? Ah do Tesouro eu não estou falando do Tesouro Unesp, dá dessa questão de tempo vocabulário controlado da importância de se usar termos dentro desse vocabulário controlado, eu acho que boa parte dos pesquisadores eles têm esse conhecimento sim, se usam é outra história, mas ele tem esse conhecimento, mas conhecimento do tesouro Unesp... Não sei não, acredito que não chega até eles. Eu vou ser muito sincera. Eu mesma estou falando que eu acho muito mais termos no DeCs do que na Linguagem UNESP, então eu não vou fazer uma indexação sem abrir os dois. Eu pesquiso nos dois e eu utilizo o que está na linguagem porque às vezes dá uma divergência do que tá no DeCe com o que tá na Linguagem UNESP assim. Aí eu vou dentro do que é nosso, do que é recomendado.

Agradecimentos e finalização da entrevista.

ANEXO 5 – Transcrição da entrevista com In2

P: Então vamos lá eu vou falar primeiramente dos objetivos da minha pesquisa para você ficar ciente. Eu não sei se você lembra do questionário que você respondeu quando eu fiz do mestrado. Agradeço também por ter participado dessa parte da pesquisa, mas foi a partir dela que eu selecionei as bibliotecas que tem um vocabulário controlado que tem uma linguagem de indexação e esse é o objeto de estudo as linguagens e como as bibliotecas universitárias constroem as suas linguagens essa investigação então o objetivo geral dessa pesquisa é contribuir com as biblioteca universitárias que buscam o desenvolvimento de sua própria linguagem de indexação. E também acerca dos estudos sobre esse tema, ciência da informação. O objetivo específico principal, né? Que se Deus quiser vai ser um fruto da tese é elaborar orientações sobre o gerenciamento da construção e manutenção de linguagens juntamente com a política de indexação, tendo como parâmetro a parte da literatura e a parte da experiência das bibliotecárias e bibliotecários. E e para começar, eu queria que você se apresentasse o seu nome a biblioteca que eles trabalham Universidade. Qual que é a sua função dentro da biblioteca?

B4 - Eu sou B4, eu sou bibliotecária na biblioteca da Escola de Comunicações e Artes da USP e eu sou a chefe de seção do serviço de tratamento da informação, então é onde a gente trabalha com a catalogação classificação e a indexação na biblioteca.

P: Aí também chama STATI?

B4: Não aqui é esse STTI, a seção de tratamento sessão de tratamento da informação, eu acabei falando o serviço porque antes a gente era um serviço, mas mudou organograma, passamos a seção de tratamento da informação.

P: Ah, tá, beleza aqui na Unesp STATI. Bom, agora o primeiro bloco de perguntas é sobre informações da biblioteca. O catálogo da biblioteca de vocês é coletivo?

B4 Sim, ele é coletivo.

P: Quais são as principais os principais assuntos que a coleção da sua biblioteca abrange?

B4: Ela abrange a área de Comunicações e a área de Artes. Então são áreas um pouco distintas, né? Então são áreas de Relações Públicas, Propaganda, Turismo, Educomunicação, Biblioteconomia, Cinema, Teatro, Artes Plásticas e Música. E com essas duas vertentes a de Comunicação e Artes.

P: Perfeito. A biblioteca está inserida em uma rede de bibliotecas?

B4: Sim, a gente tá inserido no antigo Sibi da USP, né? Que agora é a Agência de Bibliotecas e Coleções Digitais (ABCD). Então tem um setor daqui que dá apoio para para as bibliotecas, né? A gente tá ligada a Divisão Geral de Tratamento da Informação da ABCD.

P: Então deixou de ser sibi?

B4: passou para a Águia. E o ano passado passou para ser ABCD.

P: Interessante, não sabia disso. Bom, agora sobre política de indexação. A biblioteca tem um documento com a política de indexação? Se sim, a política é local, é para sua biblioteca ou é para todos da rede?

B4: A gente aqui tem uma política local que tem assim os assuntos pertinentes a biblioteca da ECA. A gente tem uma base de dados com assuntos locais com decisões que foram tomadas antes da biblioteca integral sistema antes da constituição do vocabulário unificado da USP. Então a gente tem esse que é um parâmetro para a gente trabalhar localmente, para a gente fazer a usar a indexação voltada para a nossa área e tem o vocabulário controlado da USP. Que aí se juntou todos os termos das bibliotecas da USP e a gente mandou termos para lá e a gente trabalha com esse vocabulário, mas a nossa política é local, tá?

P: Ela está documentada?

B4: Não, não está documentada, em um documento formal de decisão, ele é um catálogo em Isis. E aí é aonde tem quais termos eu posso qualificar. Quais termos eu não qualifico. Quais são as remissivas e tem do manual da seção tem algumas orientações, mas eu não não as classificam como uma política de indexação. É só assim decisões que foram tomadas ao longo dos anos e que a gente procura seguir.

P: Tá entendi. A linguagem de indexação é descrita ness política, nesse manual de vocês?

B4: Sim, a gente utiliza o vocabulário da USP, né?

P: Bom, a forma como faz a gestão, o uso e a inserção de termos no vocabulário é descrito em algum documento? Como vocês devem fazer esse procedimento?

B4: Sim, e no dentro da ABCD tem uma página né que tem lá a política de indexação da USP, né? E qual é o caminho para a gente inserir termos e sugerir troca de termos, né? Às vezes a gente precisa adequar. E como que se faz a indexação a partir do vocabulário controlado da USP. Porque a gente aqui tem a biblioteca local com a sua política, mas a gente também tá dentro de um catálogo global, né? E aí cada biblioteca, ela tem uma política, mas ela tem também que a adotar a política da USP mesmo, mas pode acontecer de você ter uma um mesmo livro com edições diferentes e tá uma indexação diferente, porque o sistema não tem uma indexação igual para toda a obra cada biblioteca, vai inserir os termos que está relacionado com a sua biblioteca.

B4: Ah, então é que aqui estamos falando do vocabulário, como que a gente insere. A gente tem um grupo que gere o vocabulário da USP.

P: Tem uma pergunta mais para frente sobre isso. Bom, eu já eu já faço a perguntinha do grupo. Para saber o um pouco da parte da história das linguagens da linguagem de indexação de vocês. O que levou à construção de uma linguagem própria? Teve alguma necessidade que desencadeou o vocabulário? Você sabe o que levou?

B4: Em 85 teve a informatização das bibliotecas aí todas as bibliotecas inseriram seus assuntos e foi criado uma lista de cabeçalho de assuntos aí com o passar dos anos foi visto a necessidade de constituir um vocabulário controlado e diretrizes para fazer a indexação. Então em 93 foi criado um grupo com 40 bibliotecas, cada biblioteca tinha um representante. E aí eles foram analisando os termos e inserindo eles dentro de hierarquias, né e tomando as decisões de termo adotar de que termo não adotar questão de plural não usar plural e E então foi um grupo de trabalho bem grande constituído por um bibliotecário de cada biblioteca e um docente do curso de Biblioteconomia aqui da ECA e a uma bibliotecária do Círio. Foi um trabalho longo que resultou na primeira versão do vocabulário controlado da USP.

P: Constituição do vocabulário foram utilizadas outras linguagens, outros vocabulários no desenvolvimento do vocabulário da USP como a BN, LC...?

B4: Foi. Foi sim, porque a como as outras bibliotecas, né do sistema tinha suas políticas e já tinha os seus vocabulários controlados, é algumas né? Não sei se todas. Por exemplo, aqui na ECA, a gente usava a lista da LC. Ah sim, então tudo isso foi pensado com os vocabulários das bibliotecas porque tinha biblioteca que da área de medicina que usava um outro vocabulário, elas foram contemplando conforme as suas Áreas, porque o grupo foi de subdividido em Ciências Humanas, Biológicas e Arte e Arquitetura. Então cada área do conhecimento foi contemplada dentro desse grupo. E aí foi feito uma análise dos termos que já eram usados e das suas linguagens que já eram adotadas ou pela eles com a LC ou com a Biblioteca Nacional, o que a biblioteca originalmente adotava. E aí eles decidiram quais os termos que iam ser, mas aí já não se segue um vocabulário de assunto ou uma terminologia. Aí vai depender muito da área da biblioteca. A gente tem muita liberdade de pesquisar o nosso termo em outros tesouros, né seguindo as políticas para você pedir a inclusão desse termo, mas ele não precisa estar necessariamente, por exemplo, aqui na ECA a gente já nem usa mais o termo da da LC. A gente pesquisa, mas se ele não existir na LC e ele for necessário para nossa indexação, a gente procura em outros vocabulários que nem a gente incomoda a arte a gente usa muito GET.

P: Entendi, então anteriormente foi utilizada a lista da LC pela sua biblioteca na época.

B4: Foi, inclusive eu tenho ainda os volumes impressos. Histórico!

P: Foi utilizado alguma norma, você tem conhecimento se foi utilizado alguma norma para guiar as decisões de elaboração da linguagem?

B4: Eu acredito que sim, mas eu não como eu não participei do processo, eu não tenho esse conhecimento né, mas como foi um trabalho bem estruturado e tinha o professor da área de Biblioteconomia e as bibliotecárias que participaram elas eram bem comprometidas e engajadas com estudos e todos os processos. Eu acredito que tenha tido algo que norteou.

P: Existe alguma publicação ou não que tem a trajetória da da do vocabulário da USP como livro artigo, não precisa falar o título, mas assim você sabe que você tem?

B4: tem foi publicado em em congressos, né no CBB, não é SNBU, se eu não me engano e tem no portal de livros abertos da USP. Eu acho que tenha algum Manual ou algum livro que do próprio SIBI que conta essa trajetória.

P: Agora o outro bloquinho de perguntas que a gente vai sobre os aspectos mais técnicos do Tesouro. Aliás, qual é a linguagem de indexação da sua instituição? É um tesouro, é uma lista alfabética de assuntos, como que classifica o seu vocabulário?

B4: não, ele não é um tesouro. Ele é uma lista, né? Ele tem hierarquia, mas eu não vejo ele como um tesouro.

P: Quais são os campos de assuntos abordados no vocabulário?

B: São todos os campos do conhecimento, por conta do seu usado pelo sistema, né? Então é Ciências Humanas, Exatas e Biológicas.

P: No vocabulário, utiliza siglas como “use” “up” para esquematizar e organizar os elementos?

B4: Tem algumas remissivas, né? Tem o “use” que a gente que usar o ver, né? Eu acho que é o ver, em alguns termos não são todos os termos que tem remissiva e deveriam ter né? Isso ele peca um pouco, tá? Nesse sentido entendi

P: O vocabulário é multilíngue?

B4: Só em português.

P: Há termos em inglês, você tem conhecimento?

B4: Quando a área adulta adota um termo em inglês, ele tá no inserido no vocabulário, mas a preferência é para o termo em português.

P: Ok. A linguagem utiliza qualificadores para diferenciar um mesmo termo com diferentes significados? Por exemplo “mercúrio”, pode ser planeta mercúrio, pode ser um elemento químico e pode ser um personagem da mitologia.

B4: Sim, usa qualificadores.

P: Há alguma orientação a respeito de como os termos devem ser gramaticalmente que nem você tava falando do plural e do singular, né? É se pode como que deve ser essa parte se pode ser utilizados verbos?

B4: Sim. Tem escrito o que você pode sugerir no termo, se você já vai sugerir geralmente a gente sugere no plural, né? Se eu não me engano é aqui eu vejo em alguns termos o plural, em outros não. Então tem assim, é na verdade é o singular, mas eu tenho visto muito termo que a gente tem pedido no plural, tô até abrindo aqui para não falar besteira, né? Mas eu a orientação eu tenho aqui inclusive. Hoje eu tô sem óculos, então não tô enxergando nada aqui no computador. Mas no nosso vocabulário tá explicando quais são os critérios.

P: Tá, se eu entrar isso eu consigo ver essas essas orientações ou é só para os bibliotecários.

B4: Ela não está em nota pública, ela não tá pública. O nosso vocabulário é público, né? Inclusive agora foi lançado uma nova plataforma que é o Tematres. [**P:** Aê? Nossa, eu não sou eu também não sabia disso.]

B4: Foi semana passada, acabou de sair sim aqui nem me familiarizei ainda com essa nova interface porque a gente usa um outro programa para no trabalho, né? Não o que o público, mas sim tem depende da área tem o plural e tem singular

P: e como são empregados sermos compostos, como gestão de tesouros. Há momentos são divididos termos gestão + tesouros?

B4: Alguns termos estão compostos, outros a gente usa qualificador. Então posso eu sentir a necessidade de um termo composto, eu vou sugerir um termo composto, mas no trabalho de indexação eu posso usar tanto com qualificador como usar termos diferentes que na busca vão ser permutados, né? Ou então sugerir que se cria um termo composto. Eles não gostam que a gente use um termo composto. A não ser que seja bem sacramentado na área para não perder o significado, né? E também de repente o

termo passa a não ser mais utilizado, mas se você olhar o vocabulário da USP você vai encontrar tanto termo composto como separado, tá?

P: Perfeito. bom sobre o grupo agora. Há um grupo de pessoas responsáveis pela gestão e manutenção da linguagem?

B4: Sim, tem um grupo que é responsável pela gestão e manutenção. então tem representantes de algumas bibliotecas não de todas. Então, é por área de conhecimento. Atualmente tem um representante na área de ciências humanas, um ciências biológicas e três das ciências exatas. QUE SÃO BIBLIOTECARIOS de bibliotecas da USP, além deles tem o pessoal do DGTI que é da Divisão Geral de Tratamento da Informação que aí tenha as três bibliotecárias que trabalham lá e o vocabulário é por nível de as pessoas estão em níveis então, vocês sugere o termo, a bibliotecária do nível 1 ela analisa esse termo, ela vai ver se tá conforme as políticas do vocabulário, se o termo é pertinente, se não tem um outro termo novo vocabulário que substitua esse termo e se a definição se ele tá na hierarquia correta, se a definição do termo que está correta, ela vai conferir a bibliografia, vai conferir a bibliografia da área, ela vai consultar outros vocabulários para ver se se a sugestão é atende ou não a inserção desse termo na USP. Depois dela aprovando ele vai para o nível 2 que também vai fazer essa mesma pesquisa do nível 2, ele vai para o nível três que vai conferir o trabalho dos dois níveis. E aí ela acatando esse termo que vai para divisão de tratamento da informação que vai inserir o termo. E se por um acaso o termo não for acatado a gente tem uma devolutiva. E aí eles falam o que a gente precisa fazer, se é sugerir um novo termo, se é mudar a hierarquia, se ele é se eu de repente sugerir na área de ciências humanas, mas ele é um termo também da área de exatas. Então eu poderia colocar ali na hierarquia da área de exatas. Então esse é todo o processo. O que a gente perdeu do começo do vocabulário da USP para agora é que não tem mais um docente do curso de Biblioteconomia fazendo parte do grupo gestor.

P: é tem uma uma pessoa assim que seja principal responsável ou é realmente o esse grupo que você fala?

B4: Ah, a gente tem uma pessoa do grupo que é a coordenadora, né?

P: Quem é a coordenadora?

B4: Agora eu vou ver aqui porque eles mudaram a equipe recentemente, a pessoa da coordenação agora, ela tá com o pessoal da ABCD. Então, eu acredito que seja a Sara que é a coordenadora da DGTI. Antes tinha um bibliotecário coordenador que era a bibliotecária da Geociências, mas eu acho que com a saída dela não encontraram ninguém para assumir esse papel então a Sara acabou incorporando.

P: E você participa do grupo ou já participou desse grupo de gestão?

B4: Não, eu nunca participei aqui na biblioteca da ECA. A gente sempre teve uma bibliotecária participando né? A Sara que agora é da DGTI e ela era bibliotecária aqui da ECA. Então ela por muitos anos quase 10 anos, ela fez parte como nível 2, e aí quando ela saiu a Lilian que trabalha aqui comigo, ela assumiu o lugar da Sara. Só que por conta de trabalhos e falta de tempo ela pediu faz uns dois meses, ela pediu para sair e aí eles ainda não convidaram outra biblioteca cara da da área de humanas para assumir. Daí eu talvez assim eu não conseguiria também assumir no momento porque demanda há muito tempo. Porque além das suas atividades da biblioteca, você tem as atividades do vocabulário, do grupo né? O grupo também toma muito tempo, esse foi um dos motivos que a Lilian saiu, então se ela não tem tempo eu como a chefe da seção eu tenho menos tempo ainda.

P: você sabe se houve treinamento para essas pessoas ou alguma capacitação para as pessoas que participaram do grupo?

B4: Sim, outros treinamentos houve treinamento, mas não recente para as que entraram agora na nova gestão não teve um treinamento, mas eu lembro durante a pandemia de ter tido a Lilian ter colocado algumas reuniões que ela participou de realmente de orientação, né? E antes de acho que foi em 2019, a gente teve uma reunião geral com todos os bibliotecários da USP que trabalhavam com a indexação participaram de um treinamento também, mas o grupo gestor ele tem esses treinamentos teve um acho que faz uns dois anos.

P: Quais são os procedimentos adotados se você percebe que precisa de da inclusão de um termo novo?

B4: Ele pode ser feito a qualquer momento, né? Assim que a gente sente a necessidade de um termo novo tem um formulário na área técnica da ABCD que a gente preenche aí a biblioteca que tá sugerindo o termo faz toda a pesquisa de área, definição, hierarquia e a fonte de informação utilizada. E aí ela vai passar por todo aquele processo que eu falei nível 1, nível 2 e nível 3 e é assim, ele é inserido até que rapidamente. A não ser que seja um termo que gere dúvidas. E além disso a é comum a gente fazer com outras bibliotecas afins uma revisão de áreas. Mas isso é mais lento, né você pegar uma área que precisa que os termos sejam revisados. Eles estão agora com trabalho que eu que eu tive conhecimento de termos que a gente usava não controlado no sistema que é um campo local aqui do Marc que a gente usa que é o 952 que são as palavras-chaves do autor sim e eles vão analisar as a pertinência de algum desses termos por conta das de termos muitas ocorrências deles serem incluídos ou não no vocabulário

P: Entendi do campo 952... Partindo disso também que você tá falando e por exemplo se você se vocês percebem um termo que já não faz sentido para linguagem, a exclusão também é por formulário. Como funciona?

B4: Também é por formulário, é o mesmo formulário que a gente tem para inclusão para alteração. E para isso exclusão então, eu também posso sugerir alteração, por exemplo, faz uns quatro meses que a biblioteca da ECA sugeriu alteração do termo “índios” pelo termo “indígenas”. Porque a gente estava com muita dificuldade de usar o termo índios que é considerado pejorativo atualmente que na época da inclusão do termo no vocabulário não era, mas atualmente é inclusive com outros termos que poderiam substituir, como povos originários ou povos indígenas. Então a gente sugeriu que trocasse e realmente foi trocado. A única coisa que não ocorreu foi a troca disso no banco.

P: Como assim no banco?

B4: Eu tenho os termos indexados como índios e tenho registros indexados como indígenas. Os anteriores não foi feito uma troca retroativa e isso é geral no sistema. Toda vez que eu peço a inclusão de um novo termo, eu não tenho essa revisão de termos que por acaso foram utilizados no lugar desse novo termo. Aqui na biblioteca da ECA, quando a gente sugere um termo e a gente usou algum outro termo de indexação para quebrar um galho assim diga-se de passagem. A gente emite uma lista dos registros que tem o termo que precisa ser adequado e faz a correção manualmente, registro por registro.

P: Então é que tem uma perguntinha dessa aqui, se são realizadas as correções do termo dos termos então sim, se vocês sentem necessidade?

B4: São feitas as correções, mas quando é um termo que é usado muito, em muitos registros, a gente não consegue fazer, por exemplo, a gente tem o termo “Artes” e tem o termo “arte” que aqui na biblio “Artes” engloba as artes em geral, a cênicas, a da música e as artes plásticas e o termo “arte” no vocabulário da USP, ele é usado no lugar do termo a Artes Plásticas. Mas, como não existia o termo “arte” quando a ECA indexava, indexava por “Artes”, então é registros de 8, 9 anos atrás, você ainda vai encontrar Artes, mas ele o assunto do livro seria a “Arte - Artes plásticas” e a gente não consegue fazer essa correção porque são muitos termos, muitos registros.

P: E as correções na linguagem são feitas que nem você falou pro formulário e mandado para o grupo?

B4: Isso manda para o grupo analisa. Aí eles fazem a correção e eles mandam um e-mail para as bibliotecas comunicando a inclusão e alteração desse termo. Era prática eles mandarem mensalmente uma lista em Excel com todos os termos que foram incluídos, excluídos e alterados, mas isso passou e parou faz uns 5 anos. Eles pararam de fazer isso, né? E a gente sente falta porque de repente você não procura um termo no vocabulário porque você sabe que ele não existe e de repente ele tá lá, ele só a biblioteca que sugeriu que tem o conhecimento.

P: Quais os critérios que você pessoalmente utiliza para incluir termos na linguagem?

B4: O que eu uso é assim, se eu vejo que tem muita muito material que tem um termo que não existe no vocabulário da USP e ele tá sendo muito utilizado pelos pesquisadores, aí a gente para pra fazer essa análise dentro da literatura da área e pedir a inclusão. Isso ocorre muito quando a gente está catalogando tese e a produção acadêmica que são documentos que são mais dinâmicos que estão mais atualizados do que um livro, né? E outro tipo de documento. E também alguns termos que a gente poderia ter sugerido e a gente não sugeriu porque não precisou deles, né? Isso eu falo na questão da indexação de filmes, né? Inclusão de personagem, a inclusão de algo que seja mais voltado para a área de imagem, né? Porque o vocabulário da USP é mais pro bibliográfico. Inclusive, ele tá lá quando você entra no vocabulário da USP e ele na página do sobre o vocabulário ele fala mais que é para atender mais os materiais bibliográficos, mas outras bibliotecas têm outros materiais e a biblioteca da ECA é um deles, né? Então a gente sugere conforme a gente precisa, mas o caminho que eu que eu que eu sigo é: tem muita tese abordando um assunto e ele não existe no vocabulário? E eu sinto a necessidade de que ele faça parte porque é o termo que as pessoas vão procurar no banco, né? E ele não vai existir e às vezes eu não tenho nem como substituir com um termo e um qualificador ou dois termos... essa indexação ficaria inadequada.

P: Sobre o *software* até você contou uma novidade. É utilizado algum software para gestão e manutenção do vocabulário?

B4: Está sendo usado o Tematres.

P: Tá. Eu sei que é recente, mas você acha que ele atende as necessidades da instituição?

B4: Sim, a gente já tinha o conhecimento do Tematres em outros vocabulários, inclusive aqui na ECA. A gente tá a gente passou o nosso o nossa base de dados local que era em Isis pro Tematres, né? E também utilizando no vocabulário da Biblioteca digital de imagens, ele é muito mais amigável do que a gente tinha antes, a interface.

P: Última rodada de perguntinhas, são só duas é sobre a disponibilização online do vocabulário. O vocabulário está disponível no catálogo da biblioteca para usuário fazer uma consulta?

B4: Não, não está no catálogo no catálogo, ele tem uma lista de cabeçalho de assuntos e não é necessariamente os mesmos termos que estão no vocabulário, porque como a gente usa a importação de registro, durante muito tempo, o que acontecia era as bibliotecas não tinha autorização para retirar assuntos que vieram da biblioteca que fez a catalogação de origem, então às vezes acontece de ter termos que não fazem parte do vocabulário da USP dentro da lista de cabeçalho de assuntos no Aleph que é um sistema que a gente usa para CATALOGAR o material e também o que é uma pena a gente não tem um catálogo de autoridades, então todo esse trabalho que eu que tem de selecionar os termos de ter qualificador, de termo autorizado, o termo não autorizado, para o usuário não tá disponível. Então se ele usar um termo não autorizado no vocabulário, ele não vai recuperar o que ele precisa, porque ele não tem esse controle de autoridade, então se ele usar um termo que não faz parte do vocabulário da USP, ele não vai recuperar nenhum documento. A não ser que ele use o campo de, para pesquisar em todos os campos que aí vai pesquisar no campo de resumo, no campo de título e aí ele pode ter algum resultado, mas abrangente.

P: Bom, tem a ver com isso que você já falou, você acredita que o vocabulário é conhecido pelos usuários da biblioteca? E se eles sabem que ele pode ser utilizado como uma ferramenta aliada na busca de pela informação?

B4: Não, eles não têm esse conhecimento. Eu acho que assim é divulgado, a gente divulga que na biblioteca, as bibliotecárias de referência falam quando orientam uma pesquisa, mas em geral o usuário não sabe que tem esse controle, né de vocabulário que ele poderia procurar de alguma outra forma ter um resultado melhor na pesquisa no catálogo.

P: Os bibliotecários de referência eles apresentam assim né?

B4: Eles apresentam nas dicas de pesquisa, como fazer uma pesquisa. Se eu perceber que o usuário procurar eu não tô localizando o documento, eles falam “Olha a gente tem esse vocabulário” e explica como funciona, né? Dá para ver também assim pelas fichas catalográficas que são geradas automaticamente pelo formulário da biblioteca da ECA que eles não utilizam os termos do vocabulário da USP eles usam os termos, nas palavras-chaves, termos livres. São termos mais Livres, então não tem esse para o usuário ele não tá claro que tem todo esse trabalho por trás, né?

P: E como tem né?

B4: E como tem como tem! Mas o que eu acredito que atrapalhe muito é essa falta desse controle de autoridade. Tem que ter um banco de autoridade, porque se ele te existisse seria indiferente para o usuário porque a gente poderia ter esse índice melhor no catálogo, né? Teria como ter um índice do vocabulário USP inserido no catálogo o vocabulário USP, ele não tá no catálogo. Eu uso o software a parte. Então na hora de indexar, eu não tenho dentro do sistema como inserir um assunto. Eu tenho que sair do sistema para poder colocar o assunto no registro... Mas se ele (usuário) pesquisar no índice obviamente que os termos que a gente utiliza no local do vocabulário e os que são os termos que têm mais ocorrência então pelo índice de assunto, você consegue visualizar. Quais são os assuntos do documento então tem os índices, mas ele entrar no vocabulário USP, fora pelo catálogo, não tem como.

P: Tá, entendi. Então B4 foram essas as perguntas que eu que eu tinha separado. Fiz um teste preliminar, uma aplicação com uma bibliotecária aqui da Unesp que ela é responsável pela pelo

Tesouro da Unesp. E você é a segunda pessoa com o que eu estou fazendo e queria saber a sua opinião, foram claras as perguntas? Você achou que teve um pouco de dificuldade alguma coisa e que eu posso melhorar?

B4: Não, eu acho que as perguntas foram claras. Achei as perguntas pertinentes, eu não sei se as minhas respostas foram adequadas... [**P:** Foram sim!]

Agradecimentos e finalização da entrevista.