



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Marília



**CULTURA
ACADÊMICA**
Editora

Linguagem de indexação para bibliotecas na perspectiva da política de indexação

Mariângela Spotti Lopes Fujita

Como citar: FUJITA, M. S. L. Linguagem de indexação para bibliotecas na perspectiva da política de indexação. *In:* FUJITA, M. S. L.; MOREIRA, W. **Manual do planejamento, construção e manutenção do Tesauro Unesp para bibliotecas: do conceitual a práxis.** Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2021. p. 17-50.
DOI: <https://doi.org/10.36311/2021.978-65-5954-069-3.p17-50>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-Non Commercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.

1. LINGUAGEM DE INDEXAÇÃO PARA BIBLIOTECAS NA PERSPECTIVA DA POLÍTICA DE INDEXAÇÃO¹

Mariângela Spotti Lopes Fujita

1.1 Introdução

1.2 Linguagem de indexação em Biblioteconomia e Ciência da Informação

1.3 A construção do tesouro Unesp: plano de ação e resultados

1.3.1 Plano de ação para a construção do tesouro Unesp

1.3.1.1 Compatibilização de registros de autoridade

1.3.1.2 Construção da Taxonomia das áreas de conhecimento na Unesp: a macroestrutura

1.3.1.3 Software para gestão e manutenção do Tesouro Unesp

1.3.1.4 Disponibilização do Tesouro Unesp para consulta e uso na indexação e recuperação da informação

1.3.1.5 Manual do Tesouro Unesp

1.4 Considerações finais

Referências

Apêndice A – Manual do Tesouro Unesp – Estrutura Preliminar

Apêndice B - Lista das publicações que demonstram o desenvolvimento das etapas do plano de ação

¹ Projeto de pesquisa desenvolvido no período de novembro de 2014 a nov. 2017 com apoio da FAPESP e do CNPq

1.1 INTRODUÇÃO

Este capítulo traz o relato do desenvolvimento do projeto de pesquisa sobre linguagem de indexação para bibliotecas que se originou da evolução da política de indexação na rede de bibliotecas da Unesp. Um dos pontos principais da política é que o processo de indexação necessita de uma linguagem dotada de vocabulário continuamente atualizado para realizar a adequada representação dos conteúdos científicos cuja característica é justamente a evolução constante. Bibliotecas sempre utilizaram listas de cabeçalhos de assuntos, cuja pré-coordenação de termos, não acompanhava a rapidez de atualização do vocabulário científico. Esta situação mudou com a transição de importantes listas de cabeçalhos de assunto para tesouros tal como a Library Congress Subject Headings, considerada linguagem de referência mundial por bibliotecas. A decisão de ter uma linguagem de indexação que pudesse ser continuamente atualizada a partir de suas fontes originais é uma proposta que se tornou realidade na política de indexação adotada pelas bibliotecas da Unesp. O objetivo final alcançado proporcionou a redação deste manual de elaboração, implantação, uso e avaliação de linguagem de indexação para bibliotecas universitárias.

A investigação sobre o tema linguagem de indexação tem como objeto de análise a elaboração, manutenção e uso de linguagem de indexação em catálogos de bibliotecas universitárias que não dispõem de instrumental teórico e metodológico para essas finalidades.

A linguagem de indexação, como ferramenta de conversão, deve ser utilizada, tanto pelo indexador após a atribuição de conceitos extraídos do documento, quanto pelo usuário antes da atribuição de conceitos extraídos da necessidade de informação que orientará a estratégia de busca no catálogo.

O uso de uma linguagem de indexação para escolha do descritor ou cabeçalho de assunto reduz a diversidade e a ambiguidade do vocabulário e estabelece uma uniformidade de representação dos termos selecionados pelo indexador para descrever o assunto dos documentos, já que vários autores podem utilizar diferentes palavras para expressar uma mesma ideia, assim como os usuários podem apresentar diversidade de vocabulário quando da expressão de uma estratégia de busca.

A linguagem de indexação faz-se necessária uma vez que a qualidade dos serviços de disseminação da informação depende da linguagem utilizada pelo sistema para a indexação e recuperação dos documentos. Isso se dá quando a linguagem do sistema permite que se traduza a linguagem natural do autor sem que se perca a ideia principal e quando permite que se traduza a linguagem do usuário de modo que satisfaça suas necessidades de informação. Esse é o foco da representação da informação dentro do processo de indexação. Para isso, depende de vários componentes para completar o conjunto de procedimentos e atingir seus objetivos de recuperação, quais sejam, o indexador, a linguagem de indexação, elementos e variáveis da política de indexação da biblioteca ou serviço de informação.

A indexação, manual, semiautomática ou automática, tem por objetivo o armazenamento das representações conceituais dos documentos para atender as necessidades de informação (GIL LEIVA, 2008, p. 70), compreendendo, fundamentalmente, duas etapas: 1) Análise de assunto; 2) Representação de conceitos.

As bibliotecas, tendo em vista o modo como realizam a catalogação cooperativa no sistema de bibliotecas universitárias, tem a necessidade de sistematizar processos e condutas de indexação e para isso dependem de uma política elaborada em consenso por catalogadores para o planejamento, elaboração e implantação de normas, procedimentos, técnicas e manual da política de indexação com orientações gerais e específicas.

Considerando-se que a investigação sobre política de indexação em bibliotecas iniciou-se pelos aspectos da indexação, a ordenação do desenvolvimento da pesquisa foi realizada tendo em vista Lancaster (2004) e Gil Leiva (2008), examinando-se:

- prática de indexação: processo de indexação, tematicidade, normas sobre indexação, manuais de indexação, indexação automática;
- qualidades da indexação: exaustividade, especificidade, correção e consistência
- ferramentas para a indexação: listas de cabeçalhos de assunto e tesauros
- avaliação da indexação

A proposta de política de indexação desenvolvida e avaliada tornou-se o “Manual de política de indexação de bibliotecas universitárias da Unesp” (FUJITA, 2014), um conjunto de condutas, procedimentos e instrumentos imersos em um contexto de gestão e envolvidos pela cultura organizacional da biblioteca.

Pelas características de desenvolvimento da pesquisa e da articulação de ambos os grupos, de pesquisadores e de catalogadores, significou que a política de indexação emanou de uma negociação entre as pessoas que realizaram um consenso, pautados por fundamentação teórica e metodológica e respaldados pelo conhecimento prático, para promoverem um serviço estratégico com resultados visíveis para a comunidade. Por isso, a proposta foi subdividida em três partes principais: os elementos da política de indexação, o processo de indexação e o instrumento da indexação.

Neste sentido, é importante observar que a pesquisa desenvolveu um produto altamente qualificado de política de indexação cujos resultados são, principalmente, a modificação das condutas de catalogação de assuntos, pois o desenvolvimento do Manual de indexação pautou-se por sua implantação e avaliação, simultaneamente. Isso significa que a política de indexação foi aceita, implantada, avaliada e determinante para a mudança de realidade de atuação profissional dos catalogadores e do catálogo online ATHENA. Dentre os itens realizados, iniciou-se, em separado, o trabalho de sistematização da Linguagem Unesp.

A linguagem de indexação da rede BIBLIODATA, anteriormente adotada para a representação de assuntos no catálogo online da rede de bibliotecas da Unesp, por não estar disponibilizada na web junto à ferramenta de busca, nunca foi conhecida ou utilizada por usuários durante a estratégia de busca. Portanto, nunca existiu a mediação da linguagem no catálogo ATHENA e, em realidade, não precisaria ser utilizada pelos catalogadores, pois a linguagem natural era a mais utilizada pelos usuários, como demonstraram os resultados obtidos do estudo de avaliação realizado por Boccato, Fujita, Gil Leiva (2011). O estudo de avaliação verificou o desempenho pelos índices de precisão e revocação de recuperação da informação de áreas científicas especializadas em 3 áreas de conhecimento no catálogo Athena (GIL LEIVA, 2008; LANCASTER, 2004).

Em consequência desse problema, foi observado que, os catalogadores na tentativa de resolver o problema de recuperação para os

seus usuários passaram a utilizar o campo 690, além do campo 650 do MARC autorizado para os cabeçalhos de assuntos da Linguagem da Rede BIBLIODATA (LCARB)². Nesse campo 690, cada biblioteca resolveu colocar descritores de outras linguagens mais especializadas e conhecidas por seus usuários e até termos de linguagem livre por falta de procedimentos sistematizados de análise documental de conteúdos e, em consequência, a linguagem não era utilizada em toda a sua capacidade de vocabulário e de estrutura sintático-semântica. Por esse motivo, o “Manual da Política de Indexação das bibliotecas universitárias da Unesp” (FUJITA, 2014) se compõe de três partes: a 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.

Na parte de instrumento de indexação, o texto refere-se ao uso de somente uma linguagem de indexação que o Grupo decidiu por denominar, inicialmente, de “Linguagem Unesp” considerando resultados de avaliação da Linguagem da Rede BIBLIODATA (LCARB) em comparação com a Linguagem da Biblioteca Nacional. A decisão de usar uma linguagem construída pela rede de bibliotecas da Unesp foi tomada em função de dois motivos: a LCARB estava desatualizada e não teria mais manutenção tendo em vista a desativação do sistema BIBLIODATA pela Fundação Getúlio Vargas³; foi necessário realizar a compatibilização de vocabulário de uma mesma linguagem a ser utilizada na indexação e na estratégia de busca para obter resultados de recuperação mais precisa, correta e específica.

Nessa ocasião, em que foi constatado que a LCARB não atendia às necessidades das bibliotecas universitárias da Unesp, foi necessário realizar o processo de substituição dos termos de uma linguagem pela outra à medida do processo de indexação. Em consulta ao sistema BIBLIODATA, foi assegurada a importação de registros bibliográficos sem a obrigatoriedade de adoção de sua Lista de Cabeçalhos de Assuntos para realizar, durante a indexação, a substituição da LCARB pela linguagem

² A Linguagem da Rede BIBLIODATA (LCARB) foi utilizada até 2013 pelo sistema de bibliotecas da Unesp.

³ O sistema BIBLIODATA e sua linguagem de indexação está atualmente sob responsabilidade do IBICT.

da Biblioteca Nacional, Library of Congress Subject Headings (LCSH) e Medical Subject Headings (MeSH).

As linguagens citadas para compatibilização são atualizadas e utilizadas por várias bibliotecas no Brasil e no mundo, cujos catálogos são acessíveis pela web. Além disso, é expressivo o número de descritores que possuem. A Terminologia de assuntos da Biblioteca Nacional, no Brasil, adota a LCSH como fonte de vocabulário.

A compatibilidade de vocabulários entre linguagens de mesma raiz e estrutura formaram a estrutura léxica e lógica da Linguagem Unesp: Terminologia de assuntos da Biblioteca Nacional⁴, Library of Congress Subject Headings (LCSH)⁵ e Medical Subject Headings (MeSH)⁶. Dessa forma, a linguagem UNESP foi disponibilizada com o nome de “Tesauro Unesp”⁷ ao usuário para consulta com a garantia de compatibilização de vocabulário utilizado na indexação.

Quando um registro bibliográfico é criado, o catalogador faz a descrição bibliográfica do item documentário nos campos devidos e quando se depara com os campos de assunto, recorre à Base de Autoridades de Assuntos de Registros de autoridade que formam, atualmente, o Tesauro Unesp em formato MARC21.

Os registros de autoridade destinam-se a conter informação autorizada para padronizar nomes e assuntos a serem usados como pontos de acesso em registros bibliográficos, bem como permitir os relacionamentos entre eles. Compõem uma base de dados que é consultada, tanto pelo catalogador ao atribuir os termos que representam assuntos, quanto automaticamente pelo próprio sistema para correção de termos e inclusão de remissivas dentro dos registros bibliográficos. Isso significa que os 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. Abaixo se demonstra exemplo de registro de autoridade de termo tópico da Biblioteca Nacional na versão de ficha completa para o usuário e no formato MARC21:

⁴ http://acervo.bn.gov.br/sophia_web/busca/autoridades

⁵ <http://id.loc.gov/authorities/subjects.html>

⁶ <https://www.nlm.nih.gov/mesh/meshhome.html>

⁷ <https://www.biblioteca.Unesp.br/tesauro/vocab/index.php>

Figura 1- Ficha completa do termo “Contratos de afretamento”

Descrição	Contratos de afretamento (subdividido geograficamente)
Library of Congress	 Charter-parties
Remissiva Ver (US/UF)	Affreightment Afretamento - Contratos Chartering of ships Contratos de frete marítimo Fretes marítimos - ContratosShip chartering
Remissiva Ver Também (TR)	TG: Direito marítimo TG: Fretes
Fonte positiva dos dados (R)	LCSH

Fonte: Catálogo de Autoridades da Biblioteca Nacional

Figura 2 - Ficha no formato MARC21 do termo “Contratos de afretamento”

000 00563cz a2200217o 4500
001 000085744
003 Br
005 20070227114454.6
008 070208d anznbabn a ana d
035 __ a 2007020811064084med
040 __ a Br c Br f Br
150 __ a Contratos de afretamento
450 1_ a Affreightment
450 __ a Afretamento x Contratos
450 __ a Chartering of ships
450 __ a Contratos de frete marítimo
450 __ a Fretes marítimos x Contratos
450 __ a Ship chartering
550 1_ w g a Direito marítimo
550 __ w g a Fretes
670 1_ a LCSH
750 _0 a Charter-parties

Fonte: Catálogo de Autoridades da Biblioteca Nacional

Obedecendo ao formato e à Linguagem de indexação adotada, o catalogador pode formatar o cabeçalho, caso não esteja pronto, extraindo os termos autorizados da Base de Autoridades de Assuntos e os distribuindo nos campos e subcampos destinados ao assunto do registro bibliográfico.

Outro aspecto a ser colocado é referente à atualização do vocabulário controlado do Tesouro Unesp com termos novos. Considerando-se a constante evolução científica registrada pelos recursos de informação presentes nas coleções impressas e digitais das bibliotecas universitárias, surgiram termos que não são encontrados em nenhuma das linguagens de compatibilização. Foi necessário investigar procedimentos da criação de termos novos. Observamos pelos procedimentos que termos novos precisam estar vinculados à uma ou mais áreas de conhecimento para identificação, principalmente, de termos genéricos e específicos e isso depende de um conjunto de estruturas hierárquicas de áreas do conhecimento que oriente a inclusão e classificação dos termos. Nos Quadros 1 e 2 é possível verificar a estrutura hierárquica dos termos superordenados e subordinados que compõe a “Macroestrutura do Vocabulário USP”⁸:

Quadro 1- Macroestrutura com as relações Lógico-Semânticas com as áreas e terminologia em seus diferentes níveis

CA100 CIÊNCIAS AGRÁRIAS
AGRONOMIA
CA120 ENGENHARIA DE PESCA
CB200 BIOCÊNCIAS
CB210 BIOLOGIA
CB220 BOTÂNICA
CB230 IMUNOLOGIA
CB240 MICROBIOLOGIA CB250 ZOOLOGIA

Fonte: Vocabulário Controlado da USP: <http://vocabusp.sibi.usp.br/Vocab/Sibix652.dll/Mac>.

⁸ Verificar que este relato se remete à nossa experiência de participação inicial do desenvolvimento pioneiro do Vocabulário USP pela equipe de pesquisadores e profissionais da USP em “Histórico do Vocabulário Controlado do SIBI/USP”. <http://143.107.154.62/Vocab/imagens/Historico.htm>.

Quadro 2 - Vocabulário Controlado do SIBi/USP - Ordem Hierárquica dos Assuntos

CÓDIGO DA MACROESTRUTURA	DESCRITORES
CH712	ECONOMIA
CH712.1	CENSOS
CH712.1.1	CENSO AGRÍCOLA
CH712.1.2	CENSO AGROPECUÁRIO

Fonte: Vocabulário Controlado da USP: <http://vocabulary.sibi.usp.br/Vocab/Sibix652.dll/ARV?Hier=CH712>

Essa macroestrutura é uma ferramenta de classificação de organização intelectual e ordem hierárquica dos assuntos do Vocabulário SIBi/USP que pode ser denominada de taxonomia. A Linguagem de indexação Unesp não possuía uma taxonomia que precisava ser criada para orientar a alteração de termos importados da Linguagem da Rede BIBLIODATA (LCARB) e a inclusão de termos novos pela organização dos registros de autoridade. Aparentemente, a Terminologia de Assuntos da Biblioteca Nacional (BN) e a Library of Congress Subject Headings (LCSH) não possuem uma taxonomia que organiza a ordem hierárquica dos assuntos dentro de seus registros de autoridade, porém, todos os termos possuem uma organização hierárquica evidente nas fichas dos termos da Biblioteca Nacional que segue a Classification Web da Library of Congress⁹ como se observa nas Figuras 3 e 4:

Figura 3 - Ficha da autoridade – Termo tópico “Engenharia de Transportes” na Linguagem da Biblioteca Nacional

Descrição	Engenharia de transportes (subdividido geograficamente)
Library of Congress	 Transportation engineering
Remissiva Ver Também (TR)	TE: Engenharia de tráfego TE: Engenharia ferroviária TE: Engenharia rodoviária TG: Engenharia TG: Engenharia civil
Fonte positiva dos dados (R)	LCSH

Fonte: Terminologia de assuntos da Biblioteca Nacional .

⁹ <https://classweb.org/>

Figura 3 - Ficha da autoridade – Termo Tópico “Banks and banking, Cooperative” da LCSH

Banks and banking, Cooperative

From Library of Congress Subject Headings URI(s)
<http://id.loc.gov/authorities/subjects/sh85011651> info:lc/authorities/sh85011651 <http://id.loc.gov/authorities/sh85011651#concept>

Instance Of

MADS/RDF Topic

MADS/RDF Authority

SKOS Concept [↗](#)

Scheme Membership(s)

Library of Congress Subject Headings

Collection Membership(s)

LCSH Collection - Authorized Headings

LCSH Collection - General Collection

LCSH Collection - May Subdivide Geographically

Variants

-  Cooperative banks
-  Credit cooperatives
-  People's banks

Broader Terms

-  Banks and banking
-  Cooperative societies

Narrower Terms

-  Agricultural cooperative credit associations
-  Credit unions
-  Rotating credit associations

Closely Matching Concepts from Other Schemes

-  Banche cooperative [↗](#)
-  Banques coopératives [↗](#)
- 

Genossenschaftlicher Bankensektor

Fonte: Library of Congress Subject Headings.

O processo de substituição de uma linguagem pelas outras foi iniciado em junho de 2012 e está em andamento a partir da construção em formato MARC21 de registros de autoridade de assuntos tópicos e geográficos para descritores da linguagem Unesp. Esses procedimentos, descritos em Ribas e Parra (2016), são etapas da construção da Linguagem Unesp que altera termos importados da Linguagem LCARB ou inclui termos novos conforme procedimentos da terminografia.

Considerando-se a necessidade de construção da linguagem de indexação bem como a observação do problema de estruturação hierárquica de áreas de conhecimento da Unesp, foi realizada a investigação sobre o processo de construção do Tesauro Unesp a partir da compatibilização de registros de autoridade de outras linguagens de indexação de mesma fonte e elaboração da Taxonomia contendo as estruturas hierárquicas de áreas de conhecimento da Unesp na perspectiva de sua política de indexação.

Para isso, foi realizada sistematização teórica e metodológica de elaboração de linguagem de indexação com e para bibliotecas com os objetivos de: elaboração de linguagens para bibliotecas obterem boa recuperação por assunto de seus catálogos online; fornecer subsídios teóricos e metodológicos para elaboração, implantação e avaliação de linguagens de indexação em ambientes de bibliotecas; e avaliar estratégias de intervenção metodológica mediante formação de equipes de elaboração, implantação e avaliação de linguagens de indexação em bibliotecas.

1.2 LINGUAGEM DE INDEXAÇÃO EM BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO

A linguagem de indexação é o instrumento de conversão da linguagem natural do documento durante o processo de indexação e, por isso, exerce importante influencia na recuperação da informação cujos efeitos são controlados e avaliados pela política de indexação adotada pela biblioteca ou serviço de informação (FUJITA, 2012). No contexto de sistemas de recuperação da informação dotados de ferramentas e padrões de representação, as linguagens de indexação são fundamentais e aceitas no âmbito teórico e metodológico tanto da Biblioteconomia e Ciência da Informação quanto da Organização do Conhecimento (MAZZOCHI, 2018)

O uso de uma Linguagem de Indexação¹⁰, pelo indexador e pelo usuário do sistema de recuperação da informação, destina-se, de forma pragmática, à escolha do termo correto para descrever a fim de reduzir a diversidade e a ambiguidade de vocabulário. Desse modo, realiza uma mediação entre a linguagem do texto e a linguagem do usuário e estabelece uma representação precisa por meio de termos que correspondem aos termos utilizados pelo autor para expressar o assunto do texto.

A linguagem de indexação é formada de vocabulário e sintaxe. O vocabulário se refere à relação dos descritores usados para a identificação do conteúdo de um documento e a sintaxe se refere as regras utilizadas para a combinação dos descritores usados para a identificação do conteúdo de um documento.

Analizados seus aspectos de estrutura e funcionalidade, a linguagem de indexação como ferramenta de controle de vocabulário apresenta-se por meio de tipologias diferenciadas pela forma e objetivos diversos.

Nos sistemas pré-coordenados os termos são previamente combinados de acordo com regras sintáticas estabelecidas pela linguagem. Por outro lado, nos pós-coordenados os termos serão combinados de acordo com a sintaxe da lógica de busca. Gil Urdiciain (2004, p.24) conclui que “[...] nas linguagens pré-coordenadas a relação entre os termos é gramatical e que na linguagem pós-coordenada a relação é lógica.” As linguagens de estrutura combinatória, como os tesouros, permitem a combinação de seus termos a partir da necessidade de indexação.

Os tesouros constituem-se de descritores que podem ser combinados ou não entre si na indexação, relacionados entre si semanticamente e genericamente. Representam primeiramente os conceitos para depois especificar suas relações e são desenvolvidos para sistemas pós-coordenados. Constituem-se como Linguagem de indexação alfabética, controlada, pós-coordenada e combinatória.

Listas de cabeçalhos de Assunto são linguagens de indexação concebidas, anteriormente, a partir de cabeçalhos de assunto cuja composição prévia de dois ou mais termos realizava uma pré-coordenação.

¹⁰ Linguagem de indexação é denominação utilizada por Lancaster (2004) e Gil Leiva (2008), autores representativos da linha “Indexing” e Linguagem de indexação ou Documental é utilizada por autores como Guimarães (2008), Gil Urdiciain (2004), autores representativos da linha de “Análise Documentária”.

Apresentam sinônimos, quase sinônimos e homógrafos. Preveem regras específicas tais como forma de entrada do descritor, uso de singular/plural, uso de abreviaturas e de termos traduzidos. Geralmente são listas gerais que arrolam termos de todas as áreas do conhecimento, tendo em vista o contexto de bibliotecas que incluem várias áreas do conhecimento.

A linguagem de indexação criada por bibliotecas é a Lista de Cabeçalhos de Assuntos. Existem, portanto, muitas listas de cabeçalhos de assunto, algumas desenvolvidas por bibliotecas individuais que não tem acesso às redes de cooperação. As principais listas de cabeçalhos de assunto, atualmente existentes e atualizadas, são:

Quadro 1 - Principais Listas de cabeçalhos de assunto

Lista de cabeçalho de assunto	Breve descrição
Library Congress of Subject Headings (LCSH) (http://id.loc.gov/authorities/subjects.html)	“Os cabeçalhos de assuntos da Library of Congress (LCSH) foram desenvolvidos e são mantidos pela Biblioteca do Congresso dos EUA, inicialmente para as coleções da Biblioteca. LCSH, um vocabulário controlado usado para indexar, catalogar e procurar registros bibliográficos em catálogos de bibliotecas e bancos de dados eletrônicos, tornou-se um padrão de fato para bibliotecas que usam registros de catalogação divulgados pela Library of Congress ou que catalogam de acordo com os padrões estabelecidos colaborativamente pela Library of Congress e pela comunidade de bibliotecas. LCSH é atualizada diariamente na Classification Web” (AMERICAN LIBRARY ASSOCIATION, 2019).
Medical Subject Headings (MeSH) https://www.nlm.nih.gov/mesh/meshhome.html	“O tesauro Medical Subject Headings (MeSH) é um vocabulário controlado e hierarquicamente organizado, produzido pela Biblioteca Nacional de Medicina dos EUA. É usado para indexação, catalogação e pesquisa de informações biomédicas e relacionadas à saúde. O MeSH inclui os títulos dos assuntos que aparecem no MEDLINE / PubMed, no Catálogo NLM e em outros bancos de dados NLM.” (NATIONAL LIBRARY OF MEDICINE, 2019)

<p>Sears List of Subject Headings https://searslistofsubjectheadings.com/page/frontmatter</p>	<p>A lista de Sears, desde 1923 até a atualidade, foi desenvolvida para interpretar as necessidades das bibliotecas de médio e pequeno porte mantendo, ao mesmo tempo, as possibilidades de expansão. Com edições sucessivas, cada vez mais ampliadas, é destinada, também, para grandes bibliotecas. Uma adaptação da lista de Sears para a língua portuguesa, de autoria de Wanda Ferraz, publicada pelo Instituto Nacional do Livro em 1944 e atualizada em 1972 e 1977, foi uma contribuição às pequenas bibliotecas brasileiras. Possui versão impressa em formato livro, disponível em livrarias, e em formato eletrônico mediante subscrição em base de dados online.</p>
<p>Terminologia de assuntos da Biblioteca Nacional do Brasil http://acervo.bn.gov.br/sophia_web/busca/autoridades</p>	<p>A Biblioteca Nacional do Brasil (www.bn.br) disponibiliza online a Terminologia de Assuntos como linguagem de indexação atualizada pela equipe de bibliotecários em português e inglês com base na Library of Congress Subject Headings e coordena um consórcio formado por bibliotecas públicas, universitárias e governamentais.</p>
<p>Lista de cabeçalhos de Assuntos da rede BIBLIODATA (LCARB)</p>	<p>A uniformização e padronização das entradas de assunto são garantidas pela FGV pela manutenção de uma Lista de Cabeçalhos de Assunto para uso de bibliotecas brasileiras participantes da Rede. Esta</p>
	<p>Lista, representada como a Bases de Autoridades Assuntos, é fundamentada na Library of Congress Subject Headings (LCSH), mantendo a estrutura de linguagem pré-coordenada, respeitando as características da língua portuguesa (FUNDAÇÃO GETÚLIO VARGAS, 1995, p. 11-12). Atualmente, indisponível está sob responsabilidade do IBICT.¹¹</p>
<p>Descritores em Ciências da Saúde (DeCS) http://decs.bvs.br/</p>	<p>Utilizado para a indexação e recuperação da informação em fontes de informação disponíveis na Biblioteca Virtual em Saúde (BVS), foi desenvolvido a partir do Medical Subject Headings MeSH.</p>
<p>Vocabulário Controlado USP (VocaUSP) http://vocabusp.sibi.usp.br/vocab/</p>	<p>“[...] é uma linguagem documentária construída, a partir de procedimentos terminológicos e documentários, pelos bibliotecários do SIBi/USP, com a participação de especialistas de todas as áreas do conhecimento abrangidas pelos seus descritores.” (LIMA, BOCCATO, 2009). Surgiu com base no vocabulário de listas de assuntos compiladas pelas bibliotecas mais especializadas na elaboração dos catálogos de assuntos que tem uma mistura de princípios de cabeçalhos de assuntos e de tesouros.</p>

Fonte: Sistematização com base em informações coletadas na web e texto de Agustín-Lacruz, Fujita e Terra (2014).

¹¹ Conforme esclarecimentos de Cecília Leite Oliveira, Diretora do IBICT, em 01 de julho de 2020, o BIBLIODATA está passando por processo de revisão e atualização tecnológica e, por essa razão, o acesso ao serviço tem sofrido instabilidade.

A sintaxe dos cabeçalhos de assuntos se manifesta no universo dos cabeçalhos compostos de que são exemplos típicos os cabeçalhos-frase ou as expressões compostas. Os cabeçalhos de assuntos servem para expressar assuntos ou o conteúdo e/ou forma da informação registrada e entende-se, como sintaxe de cabeçalhos de assunto, a combinação dos elementos ou palavras que formam os cabeçalhos compostos.

Atualmente, em sistemas de busca online a pré-coordenação perde o efeito de sequência e de prioridade porque o programa de busca não se importa em como as palavras são combinadas durante a indexação, mas sim na combinação de termos durante a estratégia de busca avançada que os sistemas de descoberta aprimoraram. Esta evolução tecnológica dos catálogos online proporcionou às listas de cabeçalhos de assunto aproveitarem seus recursos de controle de vocabulário acumulados por muitos anos e adotarem padrões de organização do conhecimento desenvolvidas por tesouros. Nessa transição, foi necessária a inclusão de estruturas hierárquicas para se obter as ligações semânticas entre termos de uma mesma área de assunto.

A lista de cabeçalhos de assuntos é a forma mais antiga de linguagem de indexação baseado no sistema alfabético de assunto. Sua prática, desenvolvida na Library of Congress (LC) a partir da elaboração dos catálogos, sedimentou o uso do cabeçalho de assunto. Entretanto, sua atualização para uso nos catálogos online transformou sua configuração e visualização com uma aparência de linguagem de indexação póscoordenada tal como a dos tesouros que possuem estrutura lógica e hierarquizada entre os termos.

Ressalte-se que a Library Congress of Subject Headings – LCSH, a mais importante lista de cabeçalhos de assunto e a linguagem de indexação alfabética mais completa e específica da atualidade, evoluiu em suas últimas edições com a introdução e uso da nomenclatura de relações semânticas adotadas por tesouros, incluindo uma estrutura lógica de relações hierárquicas anteriormente atrofiada na linguagem.

Partindo da ideia de catalogação cooperativa e uso de recursos tecnológicos a Library of Congress (LC) criou a maior Central de Catalogação Cooperativa. A introdução da tecnologia computacional na década de 1960 e o desenvolvimento do formato MARC para padronização de registros bibliográficos e formatos de intercâmbio reflete-se até os dias

atuais nas bibliotecas e centros de informação. Atualmente os registros da LC são disponibilizados em outros formatos de intercâmbio e possibilitam acesso a outros vocabulários controlados que utilizam descritores idênticos e suas variações terminológicas.

O desenvolvimento de tecnologias de software e hardware juntamente com a Internet, proporcionaram às centrais de catalogação cooperativa disponibilizarem seus catálogos em meio digital, acelerando a troca de informações.

Para propiciar o intercâmbio de informações, foram criados, junto ao formato MARC, os protocolos Z39.50 e ISO-2709 que orientam sobre a estrutura de um registro e a transferência de dados bibliográficos, o que “[...] permite que um registro criado para uma máquina por um determinado software seja lido por qualquer outro software em qualquer tipo de máquina” (MARASCO; MATTES, 1995, p. 42). Essa evolução da LCSH, os formatos de intercâmbio de registros bibliográficos e os catálogos online demonstram que são importantes para a finalidade de recuperação da informação.

Na perspectiva da gestão e planejamento, a compatibilização dos registros de autoridade da LCARB feita com a adoção da LCSH, da Terminologia de Assuntos da Biblioteca Nacional e do MeSH, linguagens de indexação de mesma fonte, reduziu problemas estruturais e proporcionou resultados quantitativos e qualitativos para a criação do Tesouro Unesp, sobretudo no que se refere à atualização contínua do vocabulário especializado de modo confiável.

1.3 A CONSTRUÇÃO DO TESAURO UNESP: PLANO DE AÇÃO E RESULTADOS

O planejamento e a gestão de tesouros devem ser realizados por equipes multidisciplinares em que diferentes profissionais e pesquisadores especializados precisam realizar uma negociação para obter consenso de suas propostas e decisões. A aplicação da metodologia de pesquisa-ação permitiu que as ações fossem negociadas durante o desenvolvimento da pesquisa para elaboração de um manual da linguagem de indexação com vistas à sua constante atualização e manutenção.

Considerando que o desenvolvimento desta pesquisa contou com o apoio da Coordenadoria Geral de Bibliotecas e dos catalogadores das bibliotecas da Unesp em função dos resultados obtidos na pesquisa “Política de indexação para bibliotecas”¹² que os levou a decidir sobre a elaboração, manutenção e uso da Linguagem de indexação para suas bibliotecas, foi possível aproveitar a oportunidade única de envolvimento cooperativo entre pesquisadora, grupo de pesquisa e participantes em torno da situação de elaboração, implantação, manutenção e uso da linguagem de indexação.

Nesse sentido, esclarecemos que o Grupo de Política de indexação da Unesp que elaborou, implantou e avaliou a “Política de Indexação para as bibliotecas universitárias da Unesp” (FUJITA et al., 2016) discutiu e decidiu, no bojo da pesquisa-ação, pela elaboração da Linguagem de indexação da Unesp e para isso foi criado, em 2013, o Grupo de estudos da elaboração da Linguagem de Indexação que atuou na colaboração efetiva desta investigação.

A pesquisa-ação propiciou a comunicação de diferentes perspectivas do contexto de indexação pelos participantes, usuários, catalogadores, dirigentes de bibliotecas e bibliotecários, assim como revelaram o que pensavam sobre os problemas e restrições desse contexto, como seria possível resolvê-los e quais os instrumentos e métodos necessários, o que ajudou no aprimoramento da política de indexação. Mais do que isso, deu aos catalogadores a possibilidade de compreensão e, conseqüente domínio, de como desenvolver a organização e representação do conhecimento no contexto de bibliotecas universitárias.

A intervenção da pesquisa-ação foi possível mediante reuniões quinzenais com equipe formada por catalogadores dedicados à elaboração de registros de autoridades e pesquisadores da área de Ciência da Informação de modo a promover a reflexão crítica com a finalidade de identificar problemas ou avaliar mudanças. À cada dúvida, problema ou necessidade de implantar novos procedimentos foram realizadas duas modalidades de seminários: de vivências profissionais e acadêmicas ou compartilhamento de experiências com pesquisadores convidados para transferência de conhecimento quando o desenvolvimento dependia de metodologias ou ferramentas a serem implantadas.

¹² Projeto de pesquisa desenvolvido no período de 2010 a 2018 (Processo CNPq: 305648/2009-8).

Os pesquisadores do Grupo de Linguagem da Rede de Bibliotecas da Unesp apresentaram a sistematização teórica e metodológica sobre linguagens de indexação e experiências de elaboração de linguagens de indexação com uso de softwares a partir das considerações analíticas e comparativas do desenvolvimento da pesquisa.

Nos seminários de vivências haviam, após as apresentações dos pesquisadores, discussões sobre o contexto profissional de modo a levantar as dúvidas, problemas e restrições para os quais eram apresentadas propostas que, dependendo da situação eram destacadas para possível ação dentro do plano. Em alguns casos, outros pesquisadores eram convidados para transferência de conhecimentos durante as reuniões quinzenais ou mediante cursos de capacitação que nos ajudassem, de modo viável, a ter outros pontos de vista para solucionar as questões a serem ajustadas para o bom desenvolvimento de uma determinada ação. Os pesquisadores convidados e respectivas ações foram:

Ações	Pesquisador convidado
Escolha do software	José Carlos Francisco dos Santos (PPGCI/UNESP); Cibele Araújo C. M. dos Santos (USP-São Paulo); Tiago Rodrigo Marçal Murakami (SIBI/USP)
Taxonomia de áreas especializadas	Maria Luiza de Almeida Campos (UFF) Deise Maria Antonio Sabbag (USP-Ribeirão Preto)

Os cursos de capacitação para o Grupo de Linguagem da Rede de Bibliotecas da Unesp ministrados pelos pesquisadores do grupo e convidados foram:

“A terminologia na elaboração de linguagens de indexação” - 28 de abril de 2014 -

Profa. Dra. Mariângela Spotti Lopes Fujita

“Procedimentos para autorização de termos novos considerando a metodologia terminográfica” - 28 de abril de 2014

Profa. Dra. Mariângela Spotti Lopes Fujita

“Construção e manutenção de tesauro” 31 ago. 2015

Prof. Dr. Walter Moreira

Profª. Dra. Mariângela Spotti Lopes Fujita

“Elaboração da macroestrutura da linguagem Unesp: análise de modelos de categorias” -

23 de outubro de 2015

Dr. Walter Moreira

“Vocabulário Controlado: principais aspectos para o tratamento e a recuperação de informações” - 12 de Abril de 2018

Dra. Maria Luiza de Almeida Campos

O resultado das discussões e intervenções possibilitou a construção de um plano de ação para elaboração, manutenção e uso da linguagem de indexação. Em resumo, o uso da linguagem viabilizou-se a partir da escolha de um software implementado em 2017 e avaliado pelas bibliotecas durante o ano de 2018. Com o software TemaTres foi possível a importação dos registros de autoridade de assunto em formato MARC21 cuja compatibilização havia sido iniciada em 2012. Essa importação marcou a criação do Tesauro Unesp para consulta de todos os usuários profissionais da rede de bibliotecas e os usuários da comunidade externa. O vocabulário, antes invisível para todos por estar em internalizado em linguagem de máquina, tornou-se aparente, acessível e disponível. Tudo isso foi possível mediante a elaboração e desenvolvimento do plano de ação no contexto da pesquisa-ação.

1.3.1 PLANO DE AÇÃO PARA A CONSTRUÇÃO DO TESAURO UNESP

O plano de ação, elaborado nas reuniões do Grupo de Linguagem da Rede de Bibliotecas da Unesp, definiu quatro metas principais:

- a) Compatibilização dos registros de autoridade da linguagem BIBLIODATA com as linguagens de mesmo vocabulário e estrutura: Library of Congress Subject Headings (LCSH), Terminologia da Biblioteca Nacional (TBN) e Medical Subject Headings;

- b) Criação e desenvolvimento de uma Taxonomia de áreas do conhecimento da Unesp formada por categorias e subcategorias para classificação dos termos contidos nos registros de autoridade compatibilizados;
- c) Escolha de um software para possibilitar a gestão e manutenção do Tesouro Unesp com a importação dos registros de autoridade compatibilizados;
- d) Disponibilização do Tesouro Unesp junto à interface de busca integrada que dá acesso ao catálogo.

1.3.1.1 COMPATIBILIZAÇÃO DE REGISTROS DE AUTORIDADE

O contexto de contínua geração de novos conhecimentos e de inovação da universidade foi determinante para a construção da Linguagem Unesp. Entretanto, foi necessário desenvolver uma metodologia de compatibilização para uso colaborativo que consiste na combinação dos registros de autoridade de assuntos em formato MARC21 de outras linguagens documentais para linguagem Unesp. O estudo de Fujita; Santos; Cruz; Moreira e Ribas (2018) desenvolveu processo inovador de compatibilização de vocabulários de três linguagens de indexação de mesma matriz terminológica utilizadas por bibliotecas: Terminologia de assuntos da Biblioteca Nacional, Library of Congress Subject Headings (LCSH) e Medical Subject Headings (MeSH).

Entretanto, a compatibilização demonstrou que é necessário incluir termos novos a partir de nova metodologia, criada especificamente para o Tesouro Unesp. Nesse sentido, foi realizado em abril de 2014, o curso de formação em serviço “A terminologia na elaboração de linguagens de indexação” para o grupo de catalogadores, participantes do Grupo de Política de indexação da Unesp com a finalidade de discutir a elaboração da Linguagem de indexação da Unesp. Este curso formou o grupo inicial de estudos da elaboração da Linguagem de Indexação Unesp, capacitando-os para a inclusão de termos novos e alteração de termos conforme Método Terminográfico.

A proposta teórica e metodológica para elaboração, manutenção e uso de linguagem de indexação em bibliotecas consiste na combinação

dos registros de autoridade de assuntos em formato MARC21 de outras linguagens documentais para linguagem Unesp descrita em Ribas e Parra (2016) e em Fujita (2017).

Esta meta, implementada pelos catalogadores da Unesp, possibilitou a compatibilização atual de 89.581 termos¹³ e está em continuidade.

1.3.1.2 CONSTRUÇÃO DA TAXONOMIA DAS ÁREAS DE CONHECIMENTO NA UNESP: A MACROESTRUTURA

A segunda meta, desenvolvida e descrita no Capítulo 5 deste livro, foi a elaboração da Taxonomia de áreas do conhecimento da Unesp utilizando os procedimentos para a construção da estrutura hierárquica do Tesauro Unesp com uso da ferramenta Classification Web da Library of Congress Subject Headings (LCSH), macroestrutura do Vocabulário USP e estrutura hierárquica da Terminologia de Assuntos da Biblioteca Nacional.

As principais considerações referem-se à proposta de uma metodologia de compatibilização de uso colaborativo para elaboração, manutenção e uso de linguagem de indexação para catálogos online de bibliotecas que consiste na construção de macroestrutura a partir da análise das estruturas hierárquicas das linguagens LCSH, Terminologia de Assuntos da Biblioteca Nacional (TBN), LCARB e Vocabulário USP.

A metodologia proposta utilizou o método de comitê ascendente da Diretriz ANSI/NISO2005 com uso do Modelo metodológico integrado, proposto por Cervantes (2009). A aplicação do método comparou as hierarquias da LCSH, TBN, LCARB e Vocabulário USP nas áreas de Física e Matemática, demonstrando ser apropriado para a construção de macroestrutura de linguagens de indexação que permita a classificação dos termos e seus relacionamentos.

Verificou-se que a escolha das linguagens LCSH, TBN, BIBLIODATA e Vocabulário USP foi adequada para a análise das estruturas de categorias e das hierarquias lógicas. A linguagem LCARB não possui completa hierarquização de termos e sua utilização é inviável.

¹³ <https://www.biblioteca.Unesp.br/tesauro/vocab/sobre.php>.

As linguagens de indexação destinadas à análise de macroestruturas devem estar constantemente atualizadas e direcionadas a contextos de uso semelhantes. A LCSH destaca-se por sua exaustividade e diversidade terminológica. A TBN, embora seja uma versão da LCSH, não apresenta diversidade terminológica. É necessária a combinação das duas linguagens. O Vocabulário USP apresenta estrutura diferenciada com características das áreas de domínio da universidade. As três linguagens demonstram ser importantes para a construção de estruturas hierárquicas com profundidade adequada ao âmbito das bibliotecas universitárias.

Com o avanço das discussões sobre a elaboração da estrutura hierárquica contamos com a colaboração da Profa. Dra. Maria Luiza de Almeida Campos da Universidade Federal Fluminense que passou a integrar institucionalmente o Grupo de Linguagem da Rede de Bibliotecas da Unesp. A participação da Profa. Maria Luiza tem sido muito importante pela sua experiência na construção de vocabulários controlado, sobretudo na modelização de conceitos. Muitas informações, conhecimentos e decisões das reuniões do Grupo foram registradas e forneceram os principais contributos do trabalho de pesquisa sobre construção e avaliação da estrutura lógica e hierárquica de linguagens de indexação para catálogos online.

A proposta teórica e metodológica para elaboração, manutenção e uso de linguagem de indexação em bibliotecas consiste na combinação dos registros de autoridade de assuntos em formato MARC21 de outras linguagens documentais para linguagem Unesp e a elaboração de estruturas lógico-hierárquicas das áreas de conhecimento existentes na universidade que permitam a classificação dos termos que representam assuntos e seus relacionamentos em registro de autoridade de assuntos. A elaboração de estruturas será, portanto, a de incluir as semelhanças e as diferenças entre as três estruturas para que a exaustividade e a especificidade sejam contempladas de modo a garantir profundidade no permanente desenvolvimento da linguagem de indexação que se pretende para bibliotecas universitárias.

1.3.1.3 SOFTWARE PARA GESTÃO E MANUTENÇÃO DO TESAURO UNESP

O estudo de softwares, desenvolvido e descrito no capítulo 6, foi realizado, inicialmente, por meio de revisão de literatura com enfoque nos

softwares TemaTres e MultiTes (FUJITA et al. 2017a). O enfoque dado aos dois softwares foi definido em função das características e funcionalidades que ambos apresentam para uso na interface de busca do catálogo online ATHENA e a existência de literatura específica sobre a avaliação dos dois softwares. O Grupo de Linguagem da Rede de Bibliotecas da Unesp escolheu o TemaTres com base em duas variáveis influentes: a) ser um software livre, gratuito e com atualizações constantes fundamentadas em normas internacionais que preconizam a interoperabilidade semântica e tecnológica; b) ter sido implementado pelo Sistema de Bibliotecas da USP para gestão e manutenção do Vocabulário USP e servir como interface de busca por assuntos pela linguagem de indexação.

Portanto, para realizar a implementação do software TemaTres realizamos reuniões com a participação da Profa. Dra. Cibele Araújo Camargo Marques dos Santos da ECA/USP, do servidor técnico e pesquisador do SIBI/USP, Tiago Murakami, do aluno de doutorado do PPGCI, José Carlos Francisco dos Santos, especialista em Informática, e do servidor técnico Oberdan Luiz May da Coordenadoria Geral de Bibliotecas, responsável pelo catálogo online ATHENA e especialista da área de informática. Atualmente, estamos realizando testes com a importação dos registros de autoridade compatibilizados e, também, está em desenvolvimento o aplicativo para fazer a interface do TemaTres com o Aleph.

1.3.1.4 DISPONIBILIZAÇÃO DO TESAURO UNESP¹⁴ PARA CONSULTA E USO NA INDEXAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

Tesauro é uma linguagem artificial utilizada em bibliotecas e em outros sistemas de informação para organizar e representar o conteúdo de diferentes publicações científicas de modo que seja possível, a qualquer usuário, recuperar a informação necessária acerca do conhecimento científico armazenado em suas bases de dados tais como catálogos, repositórios e bibliotecas digitais.

O tesauro é uma linguagem com a função específica de representação do conhecimento durante a armazenagem e a busca em bases de dados. Para isso, é composto de palavras-chave (descritores)

¹⁴ <https://www.biblioteca.Unesp.br/tesauro/vocab/index.php>.

relacionadas semântica e genericamente com as áreas de conhecimento especializado. Dessa forma, a escolha de uma palavra-chave do tesauro garantirá, ao usuário que realiza uma determinada busca, a coincidência com a palavra-chave utilizada para a representação conteúdo de um ou mais documentos que serão recuperados.

O Tesauro da Unesp tem vocabulário com termos especializados das áreas de conhecimento de suas atividades de ensino, pesquisa e extensão. Cada termo compatibilizado e operacionalizado por um registro de autoridade em formato de intercâmbio MARC21 (vide Figuras 5 e 6 do item 6.2 da Seção 6, p.). A compatibilização do termo mediante registro de autoridade é realizada, conforme resumido no item 1.3.1.1 e detalhadamente esclarecido na Seção 5 para que se possa ter atualização do vocabulário do Tesauro Unesp. O registro de autoridade compatibilizado, possibilita tanto o controle e correção internos do tesauro Unesp nas bases de dados em que atua, quanto sua exportação para qualquer outro sistema externo como é o caso do software TemaTres onde assume o formato customizado de um tesauro pronto para ser compreendido e usado por qualquer usuário que necessite empregar seus termos.

É utilizado pelos bibliotecários catalogadores para representar o conteúdo da informação mais significativa de livros, dissertações, teses, monografias, trabalhos de conclusão de curso, artigos de periódicos, documentos, legislação e etc.

O mais importante é que foi desenvolvido para ser utilizado nas interfaces de busca integrada das bases de dados das bibliotecas durante a consulta de assunto por alunos de graduação e de pós-graduação, bem como por pesquisadores para recuperar informações sobre livros, artigos, teses, dissertações, monografias, ou qualquer outro documento que tenha sido indexado em qualquer base de dados da Unesp. Para isso, basta acessar o link do tesauro Unesp, abaixo da interface de busca, digitar uma palavra, parte da palavra ou frase na caixa de pesquisa e escolher a palavra-chave mais representativa da necessidade de informação. Ao lado direito da palavra-chave tem o link de acesso para a escolha dos registros bibliográficos.

A construção e manutenção do Tesauro da Unesp é realizada desde 2013 pelo Grupo de Linguagem da Rede de Bibliotecas da Unesp composto por catalogadores das bibliotecas e pesquisadores da área de Organização

do Conhecimento, supervisionados pela Profa. Dra. Mariângela Spotti Lopes Fujita (Programa de Pós-Graduação em Ciência da Informação da Unesp – Campus de Marília).

A disponibilização do Tesouro Unesp para consulta de suas palavras-chave é parte fundamental da Política de Indexação das bibliotecas Universitárias da Unesp (<https://www.biblioteca.Unesp.br/portal/arquivos/manual-politica-indexacao-2017.pdf>) com finalidade de aprimoramento da recuperação da informação mais precisa por seus usuários durante a busca por assuntos.

O processo de construção e manutenção sempre dará continuidade à correção das palavras-chave e, principalmente da consistência dos relacionamentos hierárquicos entre os termos.

1.3.1.5 MANUAL DO TESAURO UNESP

A investigação sobre os manuais fundamentou-se na Norma ISO 25964-1, Parte I- tesouro para recuperação da informação com enfoque na seção 13 sobre gestão de tesouros que propiciou discussão de proposta de sistematização de elaboração, implantação, uso e avaliação de linguagem de indexação para bibliotecas no contexto da política de indexação a partir das orientações definidas no item 3, “O instrumento de indexação – linguagem Unesp”, do Manual de Política de Indexação da Unesp (FUJITA, 2014). Essa discussão resultou na recomendação de uma estrutura preliminar para elaboração do “Manual do Tesouro Unesp” (Apêndice A) desenvolvida pelo Grupo de Linguagem da Rede de Bibliotecas da Unesp cuja estrutura, finalizada, compõe este livro.

1.4 CONSIDERAÇÕES FINAIS

A indexação na catalogação é a finalidade desta pesquisa de cunho aplicado para que os catálogos possibilitem a recuperação por assuntos mais precisa e específica. A principal contribuição para a área de Ciência da Informação, em especial da Biblioteconomia, é este manual de elaboração, implantação e avaliação de linguagem de indexação que faça parte do Manual de política de indexação para assegurar a qualidade da recuperação por assuntos na atual conjuntura de catalogação cooperativa.

Recomenda-se aos demais pesquisadores e profissionais a necessidade de: - Elaboração de linguagens de indexação para bibliotecas obterem boa recuperação por assunto de seus catálogos on-line;

- Contribuir com mais subsídios teóricos e metodológicos para elaboração, implantação e avaliação de linguagens de indexação em ambientes de bibliotecas;
- Aplicar e avaliar estratégias de intervenção metodológica mediante formação de equipes de elaboração, implantação e avaliação de linguagens de indexação em bibliotecas;
- Propiciar à formação do indexador o conhecimento e prática suficiente de uma metodologia consistente elaboração, avaliação e controle de linguagens de indexação capacitando-o em metodologias da Ciência da Informação e das áreas interdisciplinares; - Contribuir para o desenvolvimento curricular de conteúdos programáticos das disciplinas da área de Organização da Informação de Cursos de Graduação em Biblioteconomia e de cursos de capacitação em indexação para formação em serviço de profissionais da informação;
- Disseminar a importância e a influência que a linguagem de indexação exerce sobre o desempenho da recuperação da informação por assunto em biblioteca;
- Conscientizar e capacitar os responsáveis pela política de indexação: dirigentes e catalogadores em bibliotecas e sistemas de bibliotecas.

Considera-se concluído o plano de trabalho do projeto de pesquisa em apreço com diversas contribuições científicas em forma de publicações, porém, as principais contribuições são as experiências de compartilhamento de conhecimentos entre os membros do Grupo de Linguagem da Rede de Bibliotecas da Unesp, sobretudo para os alunos de pós-graduação e de graduação envolvidos na pesquisa que tiveram suas pesquisas fortalecidas com a riqueza de conhecimentos teóricos, metodológicos e tecnológicos. Destaca-se, também, as soluções importantes para a construção de uma Linguagem Unesp com o catálogo em funcionamento para uso de muitos

usuários internos e externo. Esse projeto de pesquisa teve essa importante repercussão de conscientizar os catalogadores para a importância de um controle de vocabulário sempre atualizado e disponível para busca e recuperação da informação. Com certeza o Grupo de Linguagem da Rede de Bibliotecas da Unesp continuará atuante por muito tempo pois a linguagem de indexação necessita de atualização constante e de respaldo científico e tecnológico.

Conforme previsto no plano de trabalho, enfatiza-se a continuidade das reuniões do grupo de catalogadores agora na composição da Comissão Permanente do Tesauro Unesp cuja atuação segue os procedimentos de pesquisa participante com metodologia de pesquisa-ação. Essas reuniões propiciaram várias orientações, levantamento de dúvidas, decisões e relatos registrados em atas e anotações que conduziram vários dos trabalhos publicados, em manuais, artigos, livros e capítulos de livros com a colaboração dos integrantes do Grupo de Linguagem da Rede de Bibliotecas da Unesp em um trabalho coletivo de muita participação de todos os envolvidos. A produção científica realizada pelo Grupo (Apêndice B) atesta de forma inequívoca o acerto da escolha da pesquisa ação como indutora das decisões inovadoras que proporcionaram os relevantes resultados à Unesp e à toda comunidade científica e profissional.

REFERÊNCIAS

AGUSTÍN LACRUZ, M. del C.; FUJITA, M.S.L.; TERRA, A.L.S. Linguagens documentais para bibliotecas escolares: o caso da Espanha, Portugal e Brasil. *Informação & Sociedade*, João Pessoa, v.24, n.3, p.83-97, set./dez. 2014.

AMERICAN LIBRARY ASSOCIATION. *Cataloging tools and resources: subjectheadings*. 2019. Disponível em: <https://libguides.ala.org/catalogingtools>. Acesso em: 01 dez. 2020.

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*, Zaragoza, v.17, p.55 – 64, 2011.

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.

FUJITA, M. S. L. A política de indexação para representação e recuperação da informação. In: GIL LEIVA, I.; FUJITA, M.S.L. (ed.). *Política de indexação*. São Paulo:

- Cultura Acadêmica, 2012. p.17-28. Disponível em: http://www.culturaacademica.com.br/_img/arquivos/Livro%20Politica-deindexacao_ebook.pdf. Acesso em: 01 dez. 2020.
- FUJITA, M. S. L. *Manual da política de indexação para as bibliotecas universitárias da Unesp*. Marília: Faculdade de Filosofia e Ciências, Unesp; São Paulo: Coordenadoria Geral de Bibliotecas, Reitoria da Unesp, 2014.
- FUJITA, M.S.L (org.). *Política de indexação para bibliotecas: elaboração, avaliação e implantação*. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016. Disponível em: http://www.marilia.Unesp.br/Home/Publicacoes/politicas-de-indexacaopara-bibliotecas_ebook.pdf. Acesso em: 01 dez. 2020.
- FUJITA, M.S.L *et al.* Avaliação das características do TemaTres e Multites para o controle de autoridades nas bibliotecas universitárias. *Scire*, Zaragoza, v.23, p.63-73, 2017a. Disponível em: <http://www.ibersid.eu/ojs/index.php/scire>. Acesso em: 01 dez. 2020.
- FUJITA, M.S.L.; SANTOS, L.B.P; CRUZ, M.C.A.; MOREIRA, W.; RIBAS, R.R. de B. Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: an experience of the São Paulo State University (UNESP). *Knowledge Organization*, Baden-Baden, v.45, p. 220 - 231, 2018.
- FUNDAÇÃO GETÚLIO VARGAS. *Manual de cabeçalhos de assunto: normas e procedimentos*. Rio de Janeiro: FGV, 1995.
- GIL LEIVA, I. *Manual de indización: teoría y práctica*. Gijón: Trea, 2008.
- GIL URDICIAIN, B. *Manual de lenguajes documentales*. 2. ed. rev. y ampl. Madrid: Noesis, 2004.
- LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Tradução de Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos, 2004.
- LIMA, V. M. A.; BOCCATO, V. R. C. O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBi/USP nos processos de indexação manual, automática e semi-automática. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 14, n. 1, Apr. 2009. <http://dx.doi.org/10.1590/S1413-99362009000100010>. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362009000100010&lng=en&nrm=iso. Acesso em: 14 jun. 2014.
- MARASCO, L.C.; MATTES, R.N. Avaliação e seleção de software para automação de centros de documentação e bibliotecas. *Cadernos da FFC*, Marília, v. 4, n. 1, p. 40-49, 1995.
- MAZZOCCHI, F. Knowledge Organization Systems (KOS). *Knowledge Organization*, Baden-Baden, v.45, p.54-78, 2018. doi:10.5771/0943-7444-2018-1-54 .
- NATIONAL LIBRARY OF MEDICINE. *Medical Subject Headings*. 2019. Disponível em: <https://www.nlm.nih.gov/mesh/meshhome.html>. Acesso em: 01 dez. 2020.

RIBAS, R.R.deB.; PARRA, L.S. A construção e atualização da linguagem Unesp. *In*: FUJITA, M.S.L (org.). *Política de indexação para bibliotecas: elaboração, avaliação e implantação*. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016. p.6787. Disponível em: http://www.marilia.Unesp.br/Home/Publicacoes/politicas-deindexacao-para-bibliotecas_ebook.pdf. Acesso em: 01 dez. 2020.

APÊNDICE A

MANUAL DO TESAURO UNESP¹⁵

ESTRUTURA ERELIMINAR

Introdução [Capítulos, 1, 2 e 3]

Parte 1: Fundamentos conceituais e normativos da elaboração do tesauro

1. Visão geral e objetivos do tesauro [Capítulo 4]
 - 1.1 Símbolos, termos abreviados e outras convenções
 - 1.2 Aspectos conceituais do tesauro (diferenciar o que é conceito e o que é termo, explicitar o significado da terminologia utilizada em tesauros, tais como termo específico, tesauro multilíngüe e etc)
2. Abrangência temática do tesauro [Capítulo 5]
3. Termos do tesauro [Capítulo 4]
 - 1.1 Forma dos termos (termos simples e termos compostos)
 - 1.2 Forma gramatical de termos (substantivos, adjetivos, advérbios, gênero, número, maiúsculas e minúsculas, abreviaturas e siglas, hífen, parênteses, elementos numéricos)
 - 3.3 Esclarecimento e desambiguação dos termos do dicionário de sinônimos ou controle terminológico para seleção do termo preferido (Homônímia, antonímia, sinonímia, termos estrangeiros, gírias ou jargões, nomes comuns e comerciais, nomes populares e nomes científicos e etc)
 - 3.4 Relação de equivalência
 - 3.4.1 A relação de equivalência, num contexto monolíngüe (Sinônimos, Quasesinônimos, Termos específicos incluídos num conceito mais amplo)
 - 3.4.2 Equivalência entre idiomas
 - 3.5 Relações entre conceitos e termos
 - 3.5.1 A relação hierárquica
 - 3.5.2 A relação associativa
 - 3.6 Notas explicativas
 - 3.7 Definição
 - 3.8 Análise de faceta

Parte 2: Planejamento da estrutura e operacionalização do tesauro [Capítulo 6]

5. Estrutura do banco de dados
 6. Protocolos do banco de dados usados com o tesauro
 7. Apresentação e layout
 7. A escolha do software de gerenciamento do tesauro
 8. Formato de troca: registros de autoridade em formato MARC
 9. Interoperabilidade com outras linguagens de indexação
 10. Integração com aplicações de indexação e pesquisa
- Parte 3: Construção do tesauro [Capítulo 5]
10. A estrutura lógico-hierárquica
 11. Atualização do vocabulário
 - 11.1 Correção de termos existentes no banco de dados
 - 11.2 Compilação dedutiva de termos novos

¹⁵ Capítulos correspondentes da organização atual neste livro.

APÊNDICE B

LISTA DAS PUBLICAÇÕES QUE DEMONSTRAM O DESENVOLVIMENTO DAS ETAPAS DO PLANO DE AÇÃO

ARTIGOS EM REVISTAS CIENTÍFICAS INDEXADAS

FUJITA, M.S.L.; SANTOS, L. B. P. Política de indexação em bibliotecas universitárias: estudo diagnóstico e analítico com pesquisa participante. *Transinformação*, Campinas, v. 28, p. 5976, 2016. Disponível em: www.scielo.br/pdf/tinf/v28n1/0103-3786-tinf-28-01-00059.pdf. Acesso em: 21 nov. 2020.

FUJITA, M. S. L.; SANTOS, L. B. P. A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. *Scire*, Zaragoza, v.22, n.2, p.37-46, jul./dic. 2016. Disponível em: <http://www.iberid.eu/ojs/index.php/scire/article/view/4336/3866>. Acesso em: 21 nov. 2020.

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*, Zaragoza, v.23, p.63-73, 2017. Disponível em: <http://www.iberid.eu/ojs/index.php/scire>. Acesso em: 21 nov. 2020.

FUJITA, M.S.L.; SANTOS, L.B.P; CRUZ, M.C.A.; MOREIRA, W.; RIBAS, R.R. de B. Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: an experience of the São Paulo State University (UNESP). *Knowledge Organization*, Baden-Baden, v.45, p.220 - 231, 2018.

FUJITA, M.S.L.; CRUZ, M.C.A.; PATRÍCIO, B.O.M. Linguagens de indexação em bibliotecas universitárias: estudo analítico. *Informação & Informação*, Londrina, v.24, p.190 -225, 2019. <http://dx.doi.org/10.5433/1981-8920.2019v24n1p190>.

ARTIGOS EM REVISTAS CIENTÍFICAS NÃO INDEXADAS

FUJITA, M. S. L. A linguagem documentária na negociação de uma política de indexação para bibliotecas universitárias: procedimentos e estratégias da pesquisa-ação integral. *Revista Conhecimento em Ação*, Rio de Janeiro, v.1, n.1, p.2-17, jan./jun.2016. Disponível em: <https://revistas.ufjf.br/index.php/rca/article/view/3555/2782>

TRABALHOS APRESENTADOS EM CONFERÊNCIAS NACIONAIS

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 - EDICIC*, 10., 2016, Belo Horizonte. *Anais do Encontro da Associação de Educação e Pesquisa Em Ciência da Informação da Ibero-América e Caribe*. Belo Horizonte: ECI/UFMG, 2016. v. 1. p. 1821-1835. Disponível em: <https://pt.scribd.com/document/372647015/Anais-Edicic-2016>. Acesso em: 10 jul. 2020.

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: PINTO, F.A.; GUIMARÃES, J.A.C. (org.) Memória, tecnologia e cultura na organização do conhecimento*. Recife: Ed.UFPE, 2017. v.4. p. 217-224. (Trabalho apresentado no IV Congresso Brasileiro em Organização e Representação do Conhecimento – IV ISKO Brasil, realizado em setembro de 2017 em Recife). Disponível em: <http://isko-brasil.org.br/>. Acesso em: 21 nov. 2020.

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., Marília, 2017. Disponível em: <http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB>. Acesso em: 21 nov. 2020.

SANTOS, J.C.F. dos; CERVANTES, B.M.N.; FUJITA, M.S.L. Tesouro eletrônico: importação no TemaTres e disponibilização na web. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 19., 2018, Londrina. *Anais do XIX Encontro Nacional de Pesquisa em Ciência da Informação*. Londrina: UEL, 2018. v. 1. p. 1-19.

SANTOS, J.C.F. dos; FUJITA, M.S.L. Tesouro Unesp: Integração do registro de autoridade para o TemaTres. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 19., 2018, Londrina. *Anais do XIX Encontro Nacional de Pesquisa em Ciência da Informação*. Londrina: UEL, 2018. v. 1. p. 1-20.

FUJITA, M.S.L. Construção de linguagem de indexação para bibliotecas universitárias: uma experiência da Unesp com a pesquisa-ação integral. *In: SEMINÁRIO DO GRUPO DE PESQUISA MHTX*, 3., 2018, Belo Horizonte. *Anais do III Seminário do Grupo de Pesquisa MHTX: perspectivas em representação e organização do conhecimento: atualidades e tendências na relação universidade-empresa*. Belo Horizonte: ECI/UFMG, 2018. v. 3. p. 36-40.

FUJITA, M. S. L. Construção e manutenção de tesouro em bibliotecas universitárias: uma experiência de interoperabilidade semântica. *In: CAMPOS, M.L. de A. et al. (org.). Produção, tratamento, disseminação e uso de recursos informacionais heterogêneos*. Niterói: IACS/UFF, 2018, v. 5, p. 110-116. Disponível em: <https://app.uff.br/riuff/handle/1/9195>. Acesso em: 10 jul. 2020.

TRABALHOS APRESENTADOS EM CONFERÊNCIAS INTERNACIONAIS

FUJITA, M.S.L ; RUBI, M.P. Compatibilidade entre linguagens documentais para construção, atualização e adequação de vocabulário de bibliotecas universitárias. *In: RODRÍGUEZ MUÑOZ, J.V.; GIL-LEIVA, I; DÍAZ ORTUÑO, P; MARTÍNEZ MÉNDEZ, F.J. (ed.). Organización del conocimiento: sistemas de información abiertos. Actas del XII Congreso ISKO España y II Congreso ISKO España y Portugal, 19 y 20 de noviembre.* Murcia, 2015. p. 345-356. Disponível em: <http://www.iskoiberico.org/congresos/murcia-2015/>. Acesso em: 21 nov. 2020.

FUJITA, M.S.L.; SANTOS, L.B.P.; CRUZ, M.C.A.; MOREIRA, W. La gestión del language de indización por software: evaluación para uso en control de los registros de autoridad de catálogos en línea de bibliotecas. *In: ENCUENTROS INTERNACIONALES SOBRE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN (IBERSID 2017)*, 22., Zaragoza, 2 a 4 out. 2017. Disponível em: <http://www.ibersid.es/wpcontent/uploads/2017/06/Programa.pdf>. Acesso em: 21 nov. 2020.

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*, Zaragoza, v.23, p.63-73, 2017. Disponível em: <http://www.ibersid.eu/ojs/index.php/scire/article/view/4430/3901>. Acesso em: 21 nov. 2020.

FUJITA, M.S.L.; SANTOS, L.B.P.; CRUZ, M.C.A.; MOREIRA, W.; RIBAS, R.R. de B. Construction and evaluation of logical hierarchical structures of indexing languages for online catalogs of libraries *In: ISKO UK Conference 2017: Knowledge Organization: what's the story?*, 2017, Londres. [Trabalho submetido como artigo ao periódico Knowledge Organization]. Disponível em: <http://www.iskouk.org/content/iskomedia>. Acesso em: 21 nov. 2020.

FUJITA, M.S.L.; SANTOS, L.B.P. dos; SANTOS, N. de S. dos. A função do registro de autoridade de assunto na construção e uso de linguagens de indexação para catálogos online. *In: CONGRESO ISKO ESPANA Y III CONGRESO ISKO ESPANA-PORTUGAL*, 13., 2017, Coimbra. Tendências Atuais e Perspetivas Futuras em Organização do Conhecimento: atas do III Congresso ISKO. Coimbra: Universidade de Coimbra. Centro de Estudos Interdisciplinares do Século XX - CEIS20, 2017. v. 1. p. 577-586. Disponível em: <https://purl.org/sci/atas/isko2017>. Acesso em: 21 nov. 2020.

LIVRO PUBLICADO COM MEMBROS DA EQUIPE COMO ORGANIZADOR

FUJITA, M.S.L (org.) *Política de indexação para bibliotecas: elaboração, avaliação e implantação.* Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016. Disponível em: http://www.marilia.Unesp.br/Home/Publicacoes/politicas-de-indexacao-parabibliotecas_ebook.pdf. Acesso em: 21 nov. 2020.

CAPÍTULO DE LIVRO

SHINTAKU, M.; FUJITA, M.S.L; SCHIELSSL, M. Conceitos gerais. *In*: SHINTAKU, M. (org.). *Guia do usuário do TemaTres*. Brasília: IBICT, 2019. v. 1. p. 21-50. Doi: <http://DOI: 10.18225/9788570131553>.