

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA**

SARA DE FREITAS MARCELINO

**Avaliação comparativa de softwares livres para gerenciamento de tesouros:
Thesa e TemaTres**

Porto Alegre

2021

SARA DE FREITAS MARCELINO

**Avaliação comparativa de softwares livres para gerenciamento de tesouros:
Thesa e TemaTres**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de Bacharela em Biblioteconomia pela Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

Orientadora: Prof^a. Dr^a. Rita do Carmo Ferreira Laipelt

Porto Alegre

2021

SARA DE FREITAS MARCELINO

Avaliação comparativa de softwares livres para gerenciamento de tesouros:

Thesa e TemaTres

Trabalho de Conclusão de Curso
apresentado como requisito parcial para a
obtenção do título de Bacharela em
Biblioteconomia pela Faculdade de
Biblioteconomia pela Faculdade de
Biblioteconomia e Comunicação da
Universidade Federal do Rio Grande do Sul.

Aprovado em ___ de novembro de 2021.

BANCA EXAMINADORA

Prof^ª. Dr^ª. Rita do Carmo Ferreiro Laipelt (Orientadora)
Universidade Federal do Rio Grande do Sul

Prof. Dr. Rene Faustino Gabriel Júnior
Universidade Federal do Rio Grande do Sul

Prof^ª. Dr^ª. Sônia Elisa Caregnato
Universidade Federal do Rio Grande do Sul

Dedico este trabalho primeiramente a Deus que me deu sabedoria para concluí-lo e também a meus pais por nunca terem deixado de acreditar em mim.

AGRADECIMENTOS

Primeiramente a Deus por ter me dado forças e sabedoria para seguir em frente. Agradeço aos meus pais por todo o apoio e carinho dados durante essa longa jornada. Também agradeço a Universidade Federal do Rio Grande do Sul que me acolheu e me permitiu a realização da graduação.

Agradeço a todo o corpo docente do departamento de Ciências da Informação que compartilharam seus conhecimentos com o intuito de formar bibliotecários cada vez melhores. Em especial a professora Rita, pela paciência e por não ter desistido de mim e ao professor Rene, pois sem ele não teria sido possível o uso de um dos softwares analisados nesta pesquisa.

Gostaria também de agradecer ao professor João Henrique Ferreira Flores, por ter me orientado no período em que fui bolsista de monitoria acadêmica em estatística básica I e a professora Ana Eliza Pereira Bocorny, por ter me orientado durante a bolsa de monitoria acadêmica em inglês instrumental I.

Também merecedoras de agradecimentos são as bibliotecárias Thayná dos Santos Martinez, que me supervisionou durante a primeira parte do estágio curricular obrigatório e Francieli Ariane Lehnem Muck, que me supervisionou e orientou na segunda, e última, parte do estágio curricular obrigatório. Agradeço também a equipe do Colégio Batista de Porto Alegre e a equipe da Biblioteca Malvina Vianna Rosa que me acolheram na realização do meu estágio curricular obrigatório.

RESUMO

Este trabalho tem como objetivo investigar se softwares livres de gerenciamento de tesouros disponíveis em língua portuguesa, no Brasil, possibilitam a representação de um conjunto mais rico de relações semânticas do que as normalmente usadas em tesouros. Para a realização do trabalho foi realizado um levantamento bibliográfico e realizada a avaliação dos recursos dos softwares. Tendo como enfoque realizar uma avaliação comparativa, esta pesquisa foi uma pesquisa aplicada, do ponto de vista de sua natureza, com abordagem quantitativa. Sendo uma pesquisa descritiva, quanto aos objetivos e os dados foram coletados a partir de observação participativa. Para a realização do estudo foram comparados e avaliados o TemaTres e o Thesa. A análise dos dados foi realizada de forma simultânea e após são postas algumas discussões acerca dos resultados. Conclui que ambos os softwares atendem à algumas recomendações da norma ISO 25964-1, mas faz-se necessárias melhorias para atender as recomendações de forma plena.

Palavras-chave: TemaTres. Tesouro. Thesa. Softwares livres de gerenciamento de tesouros. ISO 25964-1.

ABSTRACT

This work aims to investigate whether free software for thesaurus management available in Portuguese, in Brazil, allows the representation of a richer set of semantic relations than those normally used in thesaurus. To carry out the work, a bibliographical survey was carried out and an evaluation of the software resources was carried out. Focusing on carrying out a comparative evaluation, this research was applied research, from the point of view of its nature, with a quantitative approach. Being a descriptive research, regarding the objectives and data were collected from participative observation. To carry out the study, TemaTres and Thesa were compared and evaluated. Data analysis was performed simultaneously and after some discussions about the results are put forward. It concludes that both software meet some ISO 25964-1 standard recommendations, but improvements are needed to fully meet the recommendations.

Keywords: TemaTres. Thesaurus. Thesa. Software for thesaurus management. ISO 25964-1.

LISTA DE ILUSTRAÇÕES

Figura 1 -	Meta-terminos.....	24
Figura 2 -	Apresentação sistemática, organizada hierarquicamente e ordenada de forma alfabética.....	39
Figura 3 -	Sugestão de Pesquisa.....	40
Figura 4 -	Pesquisa avançada.....	40
Figura 5 -	Mais opções de pesquisa avançada.....	41
Figura 6 -	Interface inicial do Thesa.....	41
Figura 7 -	Adição de colaboradores.....	42
Figura 8 -	Erro no relatório do TemaTres.....	43
Figura 9 -	Relatório do Thesa.....	44
Figura 10 -	Exemplo de uso de relações no Thesa.....	46
Figura 11 -	Fórum de discussões sobre o TemaTres.....	47
Figura 12 -	Contato do desenvolvedor do Thesa.....	48

LISTA DE SIGLAS

BRAPCI	Base de Dados em Ciência da Informação
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
E-Lis	Eprints in Library and Information Science
Enap	Escola Nacional de Administração Pública
IBICT	Instituto Brasileiro de Formação em Ciência e Tecnologia
ISO	International Organization for Standardization
LISA	Library and Information Science Abstracts
NISO	National Information Standards Organization
TE	Termo Específico
TG	Termo Geram
TR	Termo Relacionado
UFRGS	Universidade Federal do Rio Grande do Sul
UP	Use Por

SUMÁRIO

1 INTRODUÇÃO	111
1.1 IDENTIFICAÇÃO DO PROBLEMA	121
1.2 OBJETIVOS	12
1.2.1 Objetivo geral	12
1.2.2 Objetivos específicos.....	12
1.3 JUSTIFICATIVA	133
2 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO	14
2.1 TESAURO	16
2.2 NORMA DE ELABORAÇÃO DE TESAUROS.....	18
3 SOFTWARES PARA GERENCIAMENTO DE TESAUROS.....	28
4 METODOLOGIA	300
5 RESULTADOS.....	344
6 CONSIDERAÇÕES FINAIS	52
REFERÊNCIAS	54

1 INTRODUÇÃO

O uso de ferramentas que facilitem a criação e gerenciamento de tesouros se faz necessário, pois “[...] o uso de tesouros têm se destacado como ponto de apoio para a organização e acesso multifacetado da informação, bem como para a recuperação de conceitos relacionados.” (CAMPOS et al, 2006, p. 69) e os softwares para gerenciamento de tesouros facilitam o armazenamento, manipulação e apresentação dos termos definidos e suas relações.

Os tesouros têm a função de representar os assuntos dos documentos e das solicitações de busca (MOREIRA; MOURA, 2006), é um tipo de vocabulário controlado, ou seja, não é composto de palavras da língua natural, mas de uma lista de termos de um dado domínio. Tesouros são voltados para a organização e recuperação da informação, é expresso em linguagem documentária que formaliza os termos simples ou compostos e as relações entre eles e permitem a organização lógica e semântica de um domínio. Ao profissional, estabelece e define os termos e seus relacionamentos e ao pesquisador, auxilia na busca da informação (SHINTAKU, 2019).

Para haver uma melhor construção de tesouros, a norma ISO 25964-1 apresenta diretrizes a serem seguidas para a elaboração desses instrumentos. “A parte 1 da ISO cobre o desenvolvimento e manutenção de tesouros, tanto monolíngues como multilíngues. A parte 2 da ISO cobre a interoperabilidade entre diferentes tesouros e com outros tipos de vocabulários estruturados.” (JESUS, 2018, p. 29).

Os softwares de gerenciamento de tesouros facilitam a criação e manutenção de tesouros, bem como sua visualização. São importantes ferramentas para padronizar os vocabulários criados, bem como tornar esse processo mais rápido, já que não é necessário fazer tudo manualmente. Os softwares também tornam a recuperação mais rápida, pois esses programas computacionais armazenam mais de um vocabulário controlado. Os softwares de acesso livre em língua portuguesa no Brasil são o Thema, TemaTres e TECER.

O Thesa (Tesouro Semântico Aplicado) foi desenvolvido visando ser um instrumento elaboração de tesouros para os estudantes de graduação de Biblioteconomia da UFRGS, mais especificamente para ser usado na disciplina de

Sistema de Classificação II. Opera em ambiente Web e pode ser utilizado de forma gratuita tanto para fins didáticos, para pesquisa e para uso profissional (GABRIEL JUNIOR; LAIPELT, 2019).

O TemaTres é um software que permite o gerenciamento, a publicação, compartilhamento e a reutilização de ontologias, taxonomia, tesouros e lista de valores e é de uso livre e possui interface Web. Foi desenvolvido pelo bibliotecário Diego Ferreyra, atualmente docente adjunto regular na Faculdade de Filosofia e Letras da Universidade de Buenos Aires (SABBAG; GALDINO, 2018).

O TECER foi desenvolvido pelo Instituto Brasileiro de Informação e Tecnologia (Ibict), é um software livre para criar e gerenciar tesouros. Funciona em ambiente Web e permite o acesso de vários usuários ao mesmo tempo (Wiki).

1.1 IDENTIFICAÇÃO DO PROBLEMA

Quais as características dos softwares livres de gerenciamento de tesouros são compatíveis com a norma ISO 25964-1?

1.2 OBJETIVOS

Nos subitens abaixo, serão apresentados o objetivo geral e os objetivos específicos deste trabalho.

1.2.1 Objetivo geral

O objetivo geral desta pesquisa é investigar se os softwares livres de gerenciamento de tesouros disponíveis em língua portuguesa, no Brasil, possibilitam o atendimento das recomendações da norma ISO 25964-1.

1.2.2 Objetivos específicos

A partir do objetivo geral, foram desenvolvidos três objetivos específicos, sendo eles:

- a) descrever as características dos softwares livres de gerenciamento de tesouros encontrados;
- b) avaliar se os softwares livres de gerenciamento de tesouros possibilitam a representação de um conjunto mais rico de relações semânticas do que as normalmente usadas em tesouros tradicionais.

1.3 JUSTIFICATIVA

Os softwares são importantes para auxiliar na representação e recuperação da informação, considerando que a produção de informação eletrônica cresce bastante. Aumenta o número de bases de dados e repositórios com textos completos e a recuperação da informação se dá pelo assunto. Sendo assim, a existência de um vocabulário controlado se torna extremamente importante para que as buscas não gerem resultados tão abrangentes e que não sejam do interesse do usuário. Logo, quanto mais completo for o software, mais completo o tesouro será e melhor serão os resultados das buscas.

Também tendo em vista a presença de uma cadeira de elaboração de tesouros, no curso de Biblioteconomia da UFRGS, é dada a opção de utilizar softwares para construir o tesouro que é requisito da disciplina. Logo, é interessante fazer um estudo para saber qual software é o que possibilitará a construção de tesouros mais completos.

Após realizar buscas nas seguintes bases de dados: Banco de Teses e Dissertações da CAPES, Brapci, Scopus, E-Lis, Banco de Teses e Dissertações da CAPES, SABI+ e Google Acadêmico, os termos buscados foram: TemaTres, Thesa, TECER, avaliação de softwares e *software evaluation*. Foi diagnosticada uma escassez de trabalhos sobre avaliação de softwares de gerenciamento de tesouros. Sendo assim, por ser uma lacuna na área da ciência da informação, a presente pesquisa irá subsidiar novos estudos para o crescimento e desenvolvimento dessa temática no mundo.

2 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

A Organização do Conhecimento se fundamenta em análises de caráter semântico, com abordagens que resultaram de relações interdisciplinares. A relação da Organização do Conhecimento se deu devido a cooperação integrada de dados e controle da linguagem de especialidade, do vocabulário e da terminologia. Envolve as abordagens de caráter epistemológico, como prático sobre a produção, o tratamento, a preservação, a representação, a difusão e a recuperação da informação (NHACUONGUE; DUTRA, 2016).

Segundo Miranda (2006, p. 77) “a organização do conhecimento se constitui em disciplina, inter e transdisciplinar que pressupõe análise, reflexão e aplicação de fundamentos científicos na investigação das técnicas de planejamento, tratamento e recuperação da informação.” Ainda de acordo com Miranda (2006), é graças a interdisciplinaridade da organização do conhecimento é que se faz possível o equilíbrio com os estudos acerca de problemas ligados à representação do conhecimento buscando soluções para os mesmos.

O primeiro pesquisador a utilizar o termo “organização do conhecimento” foi Henry Evelyn Bliss. Após um longo período, o termo foi utilizado por Dagobert Soergel, em sua tese de doutorado. E dois anos mais tarde, o termo é utilizado por Ingetraut Dahlberg, como título de seu doutorado. Para Dahlberg, em obra de 2006, organização do conhecimento é a construção sistemática dos conceitos segundo seus elementos de conhecimento intrínsecos e a aplicação desses conceitos e classes de conceitos ordenados a objetos/assuntos (LIMA; ALVARES, 2012).

Ao contrário de Dahlberg e com uma aplicação mais específica, Hjørland (2008, p. 86) explica que:

No sentido restrito, Organização do Conhecimento (OC) se refere a atividades como descrição, indexação e classificação de documentos realizadas em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de “instituições de memória” por bibliotecários, arquivistas, especialistas em informação, especialistas no assunto, bem como por algoritmos de computador e leigos. OC como um campo de estudo está preocupado com a natureza e qualidade de tais processos de organização do conhecimento (POC), bem como os sistemas de organização do conhecimento (SOC) usados para organizar documentos, representações de documentos, obras e conceitos. (tradução nossa).

Por meio de conceitos e suas relações, que podem ser ordenadas em concordância com os mais diversos critérios, é que é fornecida a metodologia da organização do conhecimento (LIMA; ALVARES, 2012).

O primeiro teórico da classificação a reconhecer o papel relativo ao desenvolvimento e à estrutura de estudos de assunto na representação do conhecimento, foi Ranganathan. Ele entendia que, os usuários em um sistema de recuperação de informação buscam informações sobre um assunto ou conceitos e ideias. Além disso, recomendava o uso dos métodos, técnicas e instrumentos de recuperação da informação, baseando-os nos atributos dos assuntos, fossem utilizados de forma eficiente a demanda de informação do usuário. (BINWAL, 1992, p. 197 *apud* MIRANDA, 2006).

A representação do conhecimento é uma abordagem multidisciplinar, por isso existe uma dificuldade de encontrar uma terminologia unificada para representá-la. O termo tem, para cada campo científico, sido definido de forma distinta. Tem sido estudado por diferentes campos com a intenção de perceber como o processo do conhecimento da mente humana ocorre e como materializá-lo. A representação do conhecimento estabelece relações estruturais e sistemáticas de associação e distinção, envolvendo conceitos e estruturas. Sendo assim, tem como seus produtos os instrumentos terminológicos. Para a Biblioteconomia e Ciência da Informação, a Teoria da Análise Facetada e a Teoria do Conceito são os principais fundamentos teórico-metodológicos que os classificacionistas utilizam para modelagem e representação de um domínio do conhecimento (LIMA, 2020).

A indexação está diretamente ligada com a recuperação do conhecimento, tendo em vista que é a análise intelectual do assunto de um documento para reconhecer os conceitos representados e alocação dos termos de índice correspondentes para permitir que as informações sejam recuperadas. A indexação pode ser realizada tanto por usuários humanos quanto por agentes automatizados (ISO 25964-1, 2011). É o processo de organização do conhecimento mais utilizado para a descrição temática. Seu surgimento foi determinado pela representação por conceitos e teve seu desenvolvimento teórico e metodológico fora do contexto das bibliotecas e mais dentro das bases de dados bibliográficas com o propósito de acesso e recuperação da informação científica. Ocorre através de duas etapas, a etapa primeira é a análise de assunto para que possa ser realizada a identificação e seleção

de conceitos que serão representados por termos relevantes do assunto principal. Na segunda etapa, os termos serão representados por meio da tradução de um vocabulário controlado. A indexação facilita a recuperação por assunto através de uma padronização. Por meio da representação do seu assunto em sistemas da informação, garante a recuperação de qualquer documento ou informação (SHINTAKU, 2019).

Tesouro é um tipo de sistema de organização e representação do conhecimento “[...] que garante aos profissionais da informação e pesquisadores a padronização da linguagem utilizada num sistema de recuperação da informação.” (MIRANDA, 2006, p. 142). É uma ferramenta fundamental para o controle de vocabulário na indexação e na recuperação da informação.

Na seção 2.1, a seguir, as características dos tesouros são apresentadas de forma mais aprofundada.

2.1 TESAURO

O termo “*thesaurus*” tem origem no grego e latim e significa “tesouro” e durante muitos séculos foi utilizado para designar léxico, ou tesouro das palavras. Sua popularização ocorreu a partir da publicação do *Thesaurus of English Words and Phrases*, de Peter Mark Roget, 1852, que ao contrário dos dicionários tradicionais, parte do significado, de uma ideia, para chegar às palavras que melhor a representam (CAMPOS, 2001).

Tesouro é um tipo de linguagem documentária e é um instrumento utilizado para representar o conhecimento de uma dada área do saber e que possibilita ao usuário o acesso à informação (CAMPOS, 2001), servindo também como auxílio aos bibliotecários na realização da indexação (GABRIEL JUNIOR; LAIPELT, 2017). É um vocabulário controlado, ou seja, é composto de uma lista de termos de um dado domínio (SHINTAKU, 2019). Podem ser apresentados de formas: alfabética, sistemática (CAMPOS, 2001) ou hierárquica. Reúne termos escolhidos, de acordo com a preferência de quem o estiver desenvolvendo, “[...] destinados à indexação e recuperação de documentos e dados num determinado campo do saber, através da representação de seu conteúdo.” (MIRANDA, 2006, p. 142).

Para Roget (1925, p. XIX *apud* CAMPOS, 2001), a parte sistemática é composta por seis categorias: Relações Abstratas, Espaço, Matéria, Intelecto, Vontade e Afeições. Onde,

1. A primeira destas classes compreende idéias derivadas de Relações Abstratas e mais gerais entre as coisas, tais como: Existência, Semelhança, Quantidade, Ordem, Número, Tempo, Poder.
2. A segunda classe se refere a Espaço e suas várias relações incluindo Movimento, ou mudança de lugar.
3. A terceira classe inclui todas as idéias relacionadas ao Mundo Material, a saber: Propriedades da Matéria, tais como Solidez, Fluidez, Calor, Som, Luz e os fenômenos que eles apresentam, bem como as simples percepções a que elas dão origem.
4. A quarta classe abarca todas as idéias de fenômenos relacionados ao Intelecto e suas operações compreendendo a Aquisição, a Retenção e a Comunicação de Idéias.
5. A quinta classe inclui as idéias do exercício da Vontade, Forças Voluntárias e Ativas tais como: Escolha, Intenção, Utilidade, Ação, Antagonismo, Propriedade etc.
6. A sexta e última classe compreende todas as idéias derivadas das operações de nossas Forças Moral e do Sentido, incluindo nossos Sentimentos, Emoções, Paixões, Sentimentos Morais e Religiosos (ROGET, 1925, p. XIX *apud* CAMPOS, 2001, p. 88-89).

Para Currás (1995, p. 88), “Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos linguísticos que o compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.

Além de servir para a indexação de assuntos, o tesouro pode oferecer recursos variados como a exploração das relações entre os termos, através de notas de escopo, e até a origem do termo (CAMPOS. et al, 2006).

Para a Ciência da Informação, o tesouro é um vocabulário controlado que se qualifica pela individualidade e pela complexidade inerente do relacionamento entre os termos de determinado domínio. Serve como auxílio tanto para o usuário na realização de consultas quanto para o profissional que irá efetuar a indexação durante a classificação. Sua elaboração e manutenção segue as etapas de planejamento, levantamento do vocabulário, organização dos conceitos, apresentação do tesouro e critérios de avaliação. De acordo com a norma ISO 25964-1, as etapas de gestão e manutenção do tesouro correspondem ao planejamento, estágios anteriores de compilação, construção, introdução ao tesouro, disseminação e atualização. Logo, a seleção do software de gerenciamento que será utilizado ocorre durante o

planejamento, pois uma vez selecionado, a construção e atualização terá que seguir os métodos residentes (SHINTAKU, 2019).

A norma ISO 25964-1 será ponderada no subcapítulo seguinte, assim como as diretrizes para softwares de gerenciamento de tesouros.

2.2 NORMA DE ELABORAÇÃO DE TESAUROS

Geralmente, para realizar a construção de um tesouro, são obedecidas duas regulamentações: as Normas ANSI/NISO Z39.19 e ISO 25964 (GABRIEL JUNIOR; LAIPELT, 2017). Sendo esta última a que será levada em conta para a realização deste estudo.

A ANSI/NISO Z39.19 foi a primeira norma elaborada para a construção de tesouros impressos e eletrônicos. Nela observa-se diretrizes acerca gerenciamento, manutenção e construção de vocabulários controlados monolíngues (JESUS, 2018).

A norma ISO 25964-1 define tesouro como “vocabulário controlado e estruturado em que os conceitos são representados por termos, organizados de modo a tornar explícitas as relações entre os conceitos, e os termos preferidos são acompanhados de entradas iniciais para sinônimos e quase-sinônimos.” (ISO, 2011, p. 11, tradução nossa).

A norma é composta por duas partes:

[...] a primeira publicada em 2011 e descreve os padrões para construção de tesouros e a interoperabilidade com outros vocabulários, fornecendo recomendações para o desenvolvimento e manutenção de tesouros destinados a aplicações em Recuperação de Informações (RI). A segunda parte, publicada em 2013, apresenta as recomendações para o estabelecimento e manutenção de mapeamentos entre vários tesouros, ou entre tesouros e outros tipos de vocabulários. (GABRIEL JUNIOR; LAIPELT, 2017, p. 125).

Ao se construir um tesouro, o planejamento é fundamental. A norma ISO 25964 elenca os seguintes pontos que devem ser levados em consideração no planejamento de um tesouro (FUJITA; CRUZ; PATRÍCIO, 2017): “a) O tesouro deve ser usado para quê, e por quem; b) Ele será limitado pelas restrições do software existente com o qual deverá ser usado; c) Quão conhecedores serão os usuários da área de assunto do tesouro e de uso do tesouro.” (ISO 25964, 2011, tradução nossa).

A norma também diz que as pessoas envolvidas na elaboração do tesouro serão guiadas pelos objetivos, que são reconhecidos e definidos por meio de estudo

prévio, que é a tarefa melhor definida a todos, para que o trabalho seja realizado da melhor forma possível. As orientações são elaboradas para serem seguidas linearmente. Sendo assim, o passo seguinte é a determinação das características do tesouro. (FUJITA; CRUZ; PATRÍCIO, 2017). As características a serem delimitadas, segundo a ISO 25964-1 (2011, tradução nossa) são:

- a. Se o tesouro será usado em formato impresso ou eletrônico, ou ambos;
- b. O modelo de apresentação requerido;
- c. Se os formatos especiais são necessários para integração com outros sistemas, tais como sistemas de indexação ou de busca;
- d. Que formatos são previstos para atualizações e quão frequentemente são requeridas.

Para que o tesouro e os objetivos iniciais propostos se ajustem, essas definições são essenciais, sendo incluídas decisões sobre,

[...] as características opcionais, como armazenamento de informações sobre definições de termos, relações customizadas, presença dos rótulos de nó. Outras características a serem mais detalhadas, por exemplo, seriam o formato dos termos utilizados no plural ou singular, grafia, maiusculização, extensão máxima do termo, caracteres especiais, se necessários, e se haverá número limite de níveis da hierarquia (ISO 25964-1, 2011 apud FUJITA; CRUZ; PATRÍCIO, 2017, n.p).

Quanto aos recursos, “[...] os recursos chave, tais como pessoas, o financiamento, ferramentas de software e recursos de vocabulário devem ser determinados.” (ISO 25964-1, 2011, p. 89). Segundo Fujita; Cruz; Patrício (2017), as finanças devem ser voltadas aos recursos humanos e de vocabulário. O responsável pelo projeto é o que deve delegar as responsabilidades do grupo que está participando do projeto, onde cada pessoa deve ter atividades delimitadas a serem realizadas. Logo,

A escolha de software de gestão de tesouro é uma decisão importante, pois é através do suporte oferecido pelo software que a eficiência do serviço se manifestará. Por isso, é necessária uma avaliação dos softwares presentes no mercado, a fim de verificar qual se enquadra melhor nos objetivos e orçamento do projeto (FUJITA; CRUZ; PATRÍCIO, 2017, p. n.p)

A norma prevê diretrizes para softwares de gerenciamento de tesouros. O software deve ser capaz de gerenciar, importar e exportar os dados conforme descrito na ISO 25964-1, artigo 15 e “[...] deve ser escolhido com base nos requisitos do projeto em questão. Embora as características e funções descritas sejam geralmente necessárias, podem ocorrer exceções. Além disso, muitas vezes são necessários recursos adicionais” (ISO, 2011, p. 97). Abaixo serão discutidos os pontos, julgados pela autora, mais importantes presentes no artigo 15.

Em um tesauro, cada conceito é representado por um termo preferido por idioma e por seja qual for número de termos não preferidos. Os relacionamentos de termos mais amplos/restritos (no TemaTres, subordinados e no Thesa, específicos/relacionados), a notação e a nota de escopo são aplicadas ao conceito como um todo. O conceito pode ser identificado unicamente por seu termo preferido ou pelo identificador de seu termo preferido, em alguns sistemas. Entretanto, caso ocorra mudança na grafia de um termo, a continuidade durante as atualizações se dá ao se ter um identificador de termo. O formato de data recomendado é Ano/Mês/Dia (ISO, 2011).

A norma contextualiza que o controle de versão está relacionado principalmente à atualização sucessiva de um tesauro original. Assim:

- a) Atualizações feitas ou adotadas pelo proprietário / criador original, ou seja, novos conceitos, termos, relações, notas, alterações em partes existentes do tesauro e edição especial para finalidades diferentes, devem normalmente ser tratadas como novas versões de um determinado tesauro.
- b) Os verdadeiros subconjuntos, embora possam ter um subtítulo como "Versão resumida", são mais convenientemente tratados como grupos de conceitos, uma vez que se baseiam no conteúdo de uma versão existente.
- c) Da mesma forma, versões em diferentes idiomas, desde que correspondam à obra original e compartilhem a mesma estrutura conceitual (mesmo que possam ser referidas como "Versão em Inglês", "Versão em Francês", etc.), devem ser tratadas como pertencentes à mesma versão do tesauro. Para fins de troca de dados, eles devem ser diferenciados usando o atributo de idioma de tesauro.
- d) Qualquer trabalho que incorpore alterações de termos, conceitos ou relações, novas traduções, simplificação/redução de relações e semelhantes, se emitida por terceiros que não os proprietários/criadores originais, deve ser identificada como um tesauro separado. Alterações nos atributos do tesauro criador, criado, editor, direitos ou título podem indicar que se trata de um trabalho separado. Nesses casos, deve ser dado um identificador distinto e vinculado à obra original por dados em atributos como relação, descrição e fonte (ISO, 2011, p. 106, tradução nossa).

Ao aplicar uma notação expressiva ao meta-termo permitirá que sejam mostrados corretamente nas exibições hierárquicas. Caso não haja notação expressiva, outro meio deve ser encontrado para que os meta-termos sejam exibidos corretamente. Identificador, data, fonte, idioma, contribuidor, cobertura, criador, descrição, formato, editora, relação, direito, assunto, título e tipo são atributos pertencentes a classe tesauro destinados a atribuir metadados Dublin Core a um arquivo usando o formato ISO 25964. Com exceção de idioma e identificador, todos são opcionais (ISO, 2011).

Cada tesauro aplica a análise de faceta de forma distinta. Porém, se as facetas forem incluídas como termos principais, as mesmas deverão passar a ser tratadas como termos preferidos aplicando-se aos conceitos de nível superior. As facetas podem também ser nomes de grupos conceituais e/ou aparecer em meta-termos. Quanto aos grupos conceituais,

[...] usando uma estrutura de classificação que existe em paralelo com as hierarquias de conceitos do tesauro com base nas relações TG/TE. Os grupos criados pela classificação são geralmente baseados em disciplinas, áreas temáticas ou áreas de atividade empresarial. Às vezes são chamados de "categorias de assunto", "temas", "domínios", "grupos", "subconjuntos" ou "microtesauros". [...] Em geral, não existe uma relação TG/TE entre um grupo conceitual e os conceitos que ele contém. Os conceitos podem ser reunidos em grupos conceituais de muitas facetas ou hierarquias diferentes do tesauro, e a notação usada para a classificação em grupos pode ser bem distinta de qualquer notação que possa ser usada para os próprios conceitos. Os grupos podem ter subgrupos, sendo aninhados em qualquer nível. Cada grupo deve receber um rótulo verbal por idioma.

Se for importante excluir termos não preferenciais selecionados de um determinado subconjunto de conceitos, pode ser necessário marcar os termos selecionados, o que pode ser feito usando um atributo personalizado (ver 15.2.16) (ISO, 2011, p. 108, tradução nossa).

As relações comuns USE/USE POR (UP), Termo Geral (TG)/Termo Específico (TE) e Termo Relacionado (TR)/Termo Relacionado (TR) podem ser transmitidas através das classes de equivalência, relação hierárquica e relação associativa. Logo,

Cada ocorrência de um relacionamento é expressa em apenas uma direção, entre instâncias de uma classe de "origem" e uma classe de "destino".

[...]

Para tipos de relacionamento personalizado, o texto fornecido no atributo "papal" deve ser composto por (a) o nome do tipo de relacionamento principal, seguido por (b) o símbolo de barra "/" e, finalmente (c) o nome de o tipo de relacionamento personalizado. Se necessário, os tipos de relacionamento personalizados podem ser subdivididos ainda mais da mesma maneira (ISO, 2011, p. 108).

O software de gerenciamento de tesauro não deve estabelecer as limitações descritas abaixo:

- a. Não deve haver limitações quanto ao número de termos no vocabulário que o impeça de se expandir para o tamanho necessário. O mesmo se aplica a outros elementos, como meta-termos, notas de escopo, etc.
- b. De preferência, não deve haver limitações quanto ao comprimento dos termos ou meta-termos, notas, etc. Embora relativamente poucos termos excedam 40 caracteres, alguns precisam de 100 caracteres ou mais.
- c. Não deve haver limitação do número de níveis hierárquicos admissíveis, nem do número de relações hierárquicas, associativas ou de equivalência possuídas por qualquer termo preferido.
- d. O software deve ser capaz de lidar com todos os caracteres do Conjunto de Caracteres Universal, conforme definido na ISO / IEC 10646,

quando usado em qualquer um dos elementos do texto (incluindo termos, meta-termos e notas).

e. O editor deve ser capaz de escolher letras maiúsculas ou minúsculas, conforme apropriado.

f. Para tesouros multilíngues, não deve haver limite no número de idiomas que podem ser incluídos. Deve ser possível aplicar todos os idiomas desejados a algum ou todos os elementos de texto, como termos, meta-termos e notas (ISO, 2011, p. 98-99, tradução nossa).

Quanto às relações entre termos e entre conceitos, o software deve estar em conformidade com os requisitos especificados abaixo:

a. Não deve permitir a coexistência de termos duplicados para o mesmo idioma. Quaisquer duplicatas inseridas devem ser rejeitadas na entrada ou pelo menos submetidas ao editor para correção, fusão, adição de um qualificador ou outra ação corretiva. O algoritmo de correspondência para detecção de duplicatas deve ser personalizável de forma que, por exemplo, diferenças tipográficas como itálico ou maiúsculas possam ser ignoradas para fins de detecção de duplicatas.

b. Deve apoiar as relações básicas TG/TE, TR/TR, USE/UP conforme definido nas Cláusulas 8 e 10.

c. Deve apoiar os requisitos de reciprocidade descritos nas Cláusulas 8 e 10, ou seja, se o Conceito A tem relacionamento TG com o Conceito B, então o Conceito B deve ter um relacionamento TE com o Conceito A e vice-versa; se o Conceito C tem um relacionamento TR com o Conceito D, então o Conceito D deve ter um relacionamento TR com o Conceito C; se o Termo E tiver um relacionamento USE com o Termo F, então o Termo F deverá ter um relacionamento UP com o Termo E vice-versa. De preferência, o software deve disponibilizar as relações recíprocas automaticamente em resposta à inserção do editor em uma direção, mas, no mínimo, o software deve emitir um aviso se qualquer relação não recíproca estiver presente.

d. Quando um termo ou conceito é alterado ou excluído, a mudança deve ser propagada automaticamente em todos lugares em que esse termo ou conceito aparece relacionado a outro (seja como TG, TE, TR, USE ou UP). No caso de exclusão de termo ou conceito, todos os relacionamentos de e para esse tcm ou conceito devem ser excluídos. No entanto, se a exclusão de um termo ou conceito ou relacionamento deixaria qualquer conceito sem pelo menos um termo / conceito mais amplo, ou qualquer termo não preferido sem um termo preferido, o editor deve ser avisado.

e. Não deve haver limitação no número de relacionamentos que um determinado conceito ou termo preferido pode ter com outros. Assim, um conceito pode ter qualquer número de TGs TEs e TRs, e um termo preferencial pode ter qualquer número de termos não preferenciais [...] (ISO, 2011, p. 99, tradução nossa).

Quanto ao ponto acima, percebe-se um equívoco na norma, pois um termo pode ter apenas um Termo Geral (TG). Um termo pode ter vários TEs e TRs, porém, nunca mais de um TG. Ainda acerca das relações entre termos e entre conceitos:

[...]

f. Deve ser possível estabelecer relacionamentos recíprocos definidos pelo usuário, por exemplo, para distinguir entre diferentes tipos de TG / TE ou diferentes tipos de relacionamento associativo.

- g. As verificações de validação devem evitar a entrada de combinações de relacionamento inadmissíveis, como segue:
 - Se dois termos ou conceitos já possuem uma das relações básicas, nenhuma outra relação básica entre os mesmos termos ou conceitos é admissível.
 - Se o conceito A tiver o conceito B como TG nenhum dos conceitos na hierarquia TG acima do conceito B deve ser admissível como TG, TE ou TR do conceito A.
 - Os termos não preferenciais (ou seja, qualquer termo com uma relação de USE ou USE POR com outro termo) não podem ter relações com TG, TE, TR ou UP.
 - Caso ocorra a relação USE POR, o software deve verificar se a relação é pelo menos ternária.
- h. Nenhuma relação de um termo ou conceito com ele mesmo é admissível, ou seja, todas as relações básicas são irreflexivas.
- i. Apenas um termo preferido deve ser admissível para cada conceito, em cada idioma do tesouro (ISO, 2011, p. 99, tradução nossa).

Por exemplo, Universo Cinematográfico da Marvel é TG do termo Heróis que por sua vez é TG de Máquina de Combate. Logo, não existe relação entre Máquina de Combate e Universo Cinematográfico da Marvel. Patriota de Ferro possui uma relação de USE POR com Máquina de Combate. Ou seja, não possui nenhum outro tipo de relação com qualquer outro termo.

Acerca das notas que se aplicam a termos ou conceitos, o software deve suportar as notas a seguir:

- a) Deve apoiar a entrada de uma nota de escopo, associada a qualquer conceito.
- b) Se uma nota (de qualquer tipo) fizer referência a outro termo ou conceito no vocabulário, o software deve, preferencialmente, suportar a adição de um marcador ou hiperlink ao registro desse termo ou conceito. O software deve verificar a validade do destino do link.
- c) Deve apoiar a criação de notas definidas pelo usuário associadas a qualquer termo ou conceito, por exemplo, notas de história, notas editoriais, etc. (ISO, 2011, p. 100)

Quanto aos códigos e notação, o software deve suportar categorias de assunto, códigos e notações da seguinte forma:

- a) Deve ser possível associar pelo menos um código, número ou outro tipo de notação a qualquer termo ou conceito, grupo de conceitos ou matriz (ver 12.2.2 e 15.2.18). De preferência, mais de um tipo de codificação / numeração deve ser suportado.
- b) Deve ser possível associar uma ou mais categorias de assunto a qualquer termo ou conceito, grupo de conceitos ou matriz (ver 12.2.2 e 15.2.18).
- c) Deve ser possível atribuir um identificador único a cada termo e a cada conceito. De preferência, a atribuição do (s) identificador (es) deve ser automática sempre que um novo termo e / ou conceito for inserido, de forma a nunca duplicar nenhum dos identificadores existentes ou identificadores de termos ou conceitos anteriormente excluídos. O identificador não deve ser alterado quando atributos de termos ou conceitos ou relacionamentos são modificados de alguma forma.

d) Deve ser possível produzir o vocabulário usando a sequência de cada tipo de notação, codificação ou identificador (ISO, 2011, p. 100, tradução nossa).

Meta-termos são empregados para definir as áreas temáticas e categorias dos tesouros, são exibidos hierarquicamente e mostram como os termos foram organizados. Podem conter um dos dois tipos de informação: o nome de uma faceta à qual pertencem os seguintes termos ou a característica ou atributo de divisão pelo qual uma série de conceitos irmãos foi classificada ou agrupada (ISO, 2011). Cada meta-termo possui descritores relacionados. No TemaTres, por exemplo, os meta-termos possuem uma seta preta ao lado que ao ser clicada torna os descritores visíveis, em ordem alfabética, os descritores que apresentam uma seta ao lado são os que possuem termos subordinados (figura 1). O software deve suportar meta-termos da seguinte forma:

a) Os meta-termos não são considerados termos ou conceitos do tesouro (consulte a Cláusula 11, especialmente os exemplos na Figura 4) e, portanto, não estão sujeitos aos requisitos de relacionamento em 14.3. Além disso, eles não precisam ser exclusivos e, portanto, não devem estar sujeitos ao controle duplicado descrito em 14.3 a).

b) O software deve ter meios para localizar um meta-termo em monitores na posição correta em relação a qualquer termo principal e aos termos de nível mais alto que vêm dentro da faceta ou matriz correspondente (ISO, 2011, p. 100, tradução nossa).

Figura 1 - Meta-termos



Fonte: Elaborado pela autora, 2021.

Segundo a norma ISO (2011), em relação ao status das línguas, o software deve permitir que o idioma de um tesouro multilíngue tenha igual status, evitando o predomínio de um dos idiomas sobre os demais. Por exemplo:

- a) O número de termos não preferidos que se aplicam a qualquer conceito em cada idioma não deve ser determinado pelo número presente em outro idioma, e os termos não preferidos em idiomas diferentes não devem ser obrigados a corresponder uns aos outros.
- b) A existência de uma nota de escopo para um determinado conceito em um idioma não deve exigir uma nota de escopo correspondente em qualquer um dos outros idiomas (ISO, 2011, p. 100, tradução nossa).

A norma ISO salienta a importância de saber exportar e importar dados para que o conteúdo do tesouro não se restrinja ao software ou aplicativo inicialmente utilizado para desenvolvê-lo. Os seguintes mecanismos devem ser fornecidos

- a) Importação em massa de conjuntos de dados de vocabulários existentes, compreendendo termos, notas de escopo, meta-termos, relações padrão entre os termos e conceitos e todos os outros atributos dos termos, conceitos e meta-termos.
 - Todos os recursos mencionados devem ser mantidos após a importação, bem como todos os caracteres do Conjunto de caracteres universais, sempre que usados.
- b) Produzir relatórios ou exportar o vocabulário, incluindo todos os termos, notas de escopo, notação e relações padrão entre os termos e conceitos.
 - Também deve ser possível exportar subconjuntos definidos pelo editor, como termos não preferenciais apenas, ou termos preferenciais com suas notas de escopo e apenas TEs, ou termos preferenciais de conceitos que não têm relações hierárquicas ou associativas, etc.
- c) Edição em lote/exclusão em lote.
 - Deve ser possível excluir ou editar lotes de registros da mesma forma, de preferência usando um recurso nativo do software. Se isso não for possível, uma opção pode ser usar funções do sistema de gerenciamento de banco de dados subjacente ou exportar os registros selecionados, editá-los externamente e importá-los novamente. Essa técnica pode ser usada, por exemplo, para adicionar a mesma nota de histórico ou relacionamento a um lote de termos, ou talvez um mapeamento para termos em outro vocabulário. Se este método for usado, as validações habituais devem ser aplicadas na importação. [...] (ISO, 2011, p. 101, tradução nossa).

Em outras palavras, com a edição em lote, é possível substituir um texto por outro de forma que todos os termos que possuam este texto sejam modificados. Por exemplo, ao localizar o texto **Homem** e substituir por **homem**, os termos Homem de Ferro, Homem-Aranha e Homem-Formiga passarão a ser escritos com “h” ao invés de “H”. Excluir os dados em lote é, excluir todos os termos do tesouro de forma irreversível de uma única vez.

Outros mecanismos que devem ser fornecidos pelos softwares são:

- [...]
- d) Exportar todos os termos que foram alterados após uma determinada data, com ou sem seus detalhes completos.
 - Deve haver a opção de selecionar apenas certos tipos de alteração, por exemplo, apenas novos termos, ou apenas aqueles em um determinado idioma, ou de incluir todos os termos que sofreram qualquer alteração de atributos ou relações. Também deve ser possível relatar todos os termos que foram excluídos a partir de uma determinada data.

- e) Gerar exibições de tesouros, tanto em papel quanto na tela.
- Deve ser possível escolher entre uma variedade de sequências e layouts para a exibição, das quais a exibição em ordem alfabética é essencial e uma exibição hierárquica ou classificada é altamente recomendada. Consulte a Cláusula 12 para obter uma descrição dos vários tipos de exibição.
- f) Quando os termos irmãos em exibição alfabética, hierárquica ou classificada são apresentados em ordem alfabética, a convenção palavra por palavra ou letra por letra pode ser usada. A mesma convenção deve ser aplicada em todo o dicionário de sinônimos. (Uma explicação das convenções pode ser encontrada na ISO 999; consulte também o exemplo abaixo.) No entanto, deve ser possível forçar uma sequência diferente quando apropriado (consulte a Cláusula 11) (ISO, 2011, p. 101, tradução nossa).

No que concerne a navegação editorial e suporte, o software deve suportar as seguintes funções editoriais:

- a) O editor do tesouro deve ser capaz de localizar os detalhes completos de qualquer termo (e o conceito correspondente) pela entrada direta deste termo, de preferência apenas precisando inserir uma parte do termo, com o software auxiliando na seleção, identificando todos os termos que contêm a parte.
- b) O editor deve ser capaz de localizar os detalhes completos para qualquer termo (e o conceito correspondente) navegando através das relações termo/conceito. Os recursos de navegação devem, preferencialmente, permitir que ele comece a partir de uma lista de termos principais, mas na ausência disso ele deve, pelo menos, ser capaz de navegar de qualquer termo para o registro de qualquer um de seus TEs, s, TRs, USEs ou UPs.
- c) A interface editorial deve permitir a visualização de todo o contexto hierárquico de um termo ou conceito, ao mesmo tempo que o próprio termo, com todas as suas notas, códigos e relações apresentadas para edição.
- d) Deve ser possível mover um conceito (juntamente com todos os seus TEs, TEs desses TEs, etc.) facilmente de uma posição hierárquica para outra, preferencialmente usando arrastar e soltar. Isso requer quebrar um conjunto de relacionamentos TG/TE e fazer outro. Quaisquer relações TR assim invalidadas devem ser levadas ao conhecimento do editor.
- e) Os recursos de edição devem incluir recursos de processamento de texto padrão, como a capacidade de adicionar, modificar e excluir caracteres sem redigitar um campo inteiro. Uma função de verificação ortográfica também pode ser útil.
- f) Quando um editor executa etapas para excluir um termo ou conceito, o software deve primeiro buscar a confirmação de que a exclusão desse termo ou conceito é intencional, antes de concluir a ação.
- g) Quando se pretende estabelecer uma relação entre dois termos ou conceitos já presentes no vocabulário, deve ser possível fazê-lo por navegação e seleção, em vez de ter que digitar na íntegra um termo já conhecido do sistema. Isso é importante para evitar erros, bem como para apoiar a eficiência.
- h) O editor deve ser capaz de alternar facilmente de uma tela em um idioma para a tela correspondente em qualquer um dos outros idiomas do tesouro (ISO, 2011, p. 102, tradução nossa).

No que se refere a salvaguardas editoriais, o software deve incorporar as seguintes salvaguardas:

- a) Mudanças editoriais devem ser feitas em primeira instância em um banco de dados mestre, do qual as saídas são derivadas periodicamente para processos downstream, como indexação ou aplicativos de descoberta de recursos.
- b) Se mais de uma pessoa estiver editando a fase de dados mestre simultaneamente, um mecanismo embutido deve impedir o acesso de gravação simultâneo aos mesmos registros.
- c) Os controles de segurança/senha devem evitar alterações editoriais de pessoas não autorizadas.
- d) De preferência, o software deve permitir diferentes níveis de acesso para que possam ser feitas alterações provisórias, que não são admitidas definitivamente até serem aprovadas.
- e) Para cada editor, uma função de reversão (desfazer) deve permitir a reversão progressiva das alterações editoriais mais recentes que ele inseriu.
- h) Um log deve ser mantido para que o banco de dados possa ser restaurado de uma versão anterior (ISO, 2011, p. 102, tradução nossa).

No tocante às ferramentas de edição, de acordo com ISO (2011, p. 103, tradução nossa) “Deve ser possível obter relatórios da quantidade de termos com características particulares, particularmente do número total de termos preferidos e não preferidos.”

O próximo capítulo irá discorrer sobre os softwares para gerenciamento de tesouro, dando maior atenção aos softwares livres que serviram como objeto de estudo da pesquisa que será realizada.

3 SOFTWARES PARA GERENCIAMENTO DE TESAUROS

Os softwares de gerenciamento de tesauros são ferramentas que auxiliam a construção desses vocabulários, sendo necessários, pois facilitam o armazenamento, manipulação e apresentação dos termos definidos bem como suas relações (CAMPOS; *et al*, 2006). Para Campos e outros (2006, p. 70),

No processo de elaboração de software de tesauros há que se levar em conta os aspectos metodológicos, os quais variam de acordo com a linha adotada para a sua construção, e também outros aspectos operacionais independentes da metodologia adotada, como por exemplo, mecanismos para a criação e controle de atualização de termos, possibilidade de reutilização de outros tesauros e elaboração de relatórios.

Logo, a escolha do software que será utilizado deve ser feita levando em consideração se o mesmo atenderá as necessidades de quem irá construir o tesouro e também de quem irá utilizá-lo. A ferramenta escolhida deve apoiar, além das questões operacionais, a metodologia usada para sua elaboração (CAMPOS; *et al*, 2006).

O **Thesa** foi desenvolvido pelos professores Rene Faustino Gabriel Junior e Rita do Carmo Ferreira Laipelt ambos docentes do Departamento de Ciências da Informação da UFRGS. Foi criado com o intuito de reduzir o trabalho operacional e focar no desenvolvimento cognitivo e conceitual referente a modelagem do domínio. Seu desenvolvimento foi baseado nas normas ISO e NISO vigentes. Funciona em ambiente Web e pode ser usado gratuitamente, podendo ser utilizado para fins didáticos ou uso profissional. Pode ser traduzido para qualquer idioma, pois foi desenvolvido com o princípio multi-idioma, as traduções dependerão de convênios estabelecidos com instituições nativas de outros idiomas que demonstrem interesse pelo uso do software (GABRIEL JUNIOR; LAIPELT, 2017).

Segundo Gabriel Junior e Laipelt (2017, p. 134-135):

O Thesa utiliza uma concepção de múltiplos tesauros, ou múltiplos esquemas, ou seja, o usuário pode criar um número ilimitado de tesauros em diferentes áreas do conhecimento, os usuários/elaboradores desses tesauros, podem deixá-los para uso público ou privado, possibilitando o acesso de outros usuários. No Thesa partiu-se da concepção de URI, empregada pelo SKOS e sistemas baseados na Web Semântica, ou seja, cada conceito é associado a um endereço permanente na Internet e a um identificador único do conceito, e esse representado por termos por meio de propriedades.[...]

A estrutura do Thesa é baseada na concepção das relações entre os conceitos, partindo do pressuposto que um conceito pode ser representado

por um termo, uma imagem, um som, um link ou qualquer outra forma que possa ser explicitada. Nessa abordagem, o conceito é perene, enquanto a sua representação pode variar conforme o contexto histórico ou social, sendo definida uma forma preferencial, e inúmeras formas alternativas e ocultas.

O **TemaTres**, criado pelo bibliotecário Diego Ferreyra, é um software de uso livre. Gonzales-Aguilar, Ramíres-Posada e Ferreyra (2012, p. 319 *apud* GALDINO; SABBAG, 2017, p. 2569) dizem que “[...] o uso do TemaTres garante a coerência e a normalização da informação em um sistema distribuído, eliminando a ambiguidade, a polissemia e a sinonímia do vocabulário especializado”.

Devido a sua capacidade ilimitada para armazenar os termos, o TemaTres pode ser utilizado desde a construção de um mini tesouro até a elaboração de um no ambiente de trabalho do bibliotecário (GALDINO; SABBAG, 2017).

De acordo com a página wiki do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), o **TECER** foi desenvolvido pelo IBICT e é um software livre que possibilita a criação e gerenciamento de tesouros, funciona em ambiente Web e permite o acesso de vários usuários simultaneamente via programa de navegadores como, internet explorer ou mozilla. Foi desenvolvido com a linguagem JAVA e possui compatibilidade com os sistemas Windows e Linux. Ele é customizável e permite a alteração da apresentação do sistema, alterando o layout das páginas, facilidade de exportação de dados, gerenciamento de usuários e a seleção de idiomas. O TECER possui ferramentas de busca, gerenciamento de relatório e permite visualizar o tesouro por vários critérios, por exemplo: alfabéticos, hierárquicos e etc. Com relação à gestão, além de consultar pode-se adicionar termos, alterar, controlar duplicidade de e excluir vocábulos. Por ser uma ferramenta útil na gestão de tesouros, apresenta uma boa solução às instituições que procuram por um software livre para gestão de tesouros.

Os tesouros são importantes ferramentas para a recuperação da informação, por isso precisam ser bem estruturados e os mais completos possíveis. Logo, faz-se necessária uma análise para decidir qual software de gerenciamento utilizar, pois quanto mais recursos e compatibilidade com a norma tiver, mais completo será o tesouro construído e melhor será realizada a indexação e recuperação do conhecimento.

No próximo capítulo será descrita a metodologia utilizada nesta pesquisa.

4 METODOLOGIA

A pesquisa foi dividida em duas partes. Primeiramente foi realizado um levantamento bibliográfico para identificar trabalhos sobre a temática e sobre os softwares instrumentos desta pesquisa. Os termos utilizados no levantamento foram: TemaTres, Thesa, TECER, avaliação de softwares e *software evaluation*, as fontes utilizadas nesta sondagem foram Brapci, Lisa, E-Lis, Banco de Teses e Dissertações da CAPES. A segunda etapa consistiu em avaliar os recursos dos softwares para comparar se estavam de acordo com a norma ISO 25964. Um dos critérios que foi levado em conta na escolha dos softwares foi que estivessem livres para downloads.

Porém, o TECER apresenta problema em sua instalação. No processo para instalá-lo é necessário abrir o arquivo bd.sql no pgAdmin III, mas esse arquivo não existe. Logo, impossibilita a finalização da instalação do TECER, ou seja, não é possível abri-lo.

Para atingir os objetivos da pesquisa realizou-se uma avaliação comparativa entre os softwares livres de gerenciamento de tesouros identificados no levantamento utilizando a norma ISO 25964 como parâmetro para verificar se os softwares possuem os requisitos que a mesma apresenta.

Do ponto de vista da sua natureza foi uma pesquisa aplicada. Este tipo de pesquisa “[...]objetiva gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos.” (MORESI, 2003, p. 8). Sendo assim, foi realizado o download e a comparação entre os softwares.

A abordagem utilizada foi a qualitativa, pois não requereu o uso de métodos e técnicas estatísticas. Essa abordagem tem como focus o processo e seu significado. (MORESI, 2003). Quanto aos objetivos, a pesquisa foi descritiva. Descritiva, pois descreve quais as características dos softwares livres de gerenciamento de tesouros.

Os dados foram coletados a partir de observação participativa, pois foi aplicado o roteiro desenvolvido, que será apresentado adiante, nos softwares que foram selecionados para tal. Além disso, foi utilizado o tesouro: Universo Cinematográfico da Marvel: Os Vingadores - Vingadores: Guerra Infinita. Esse tesouro foi elaborado pela autora juntamente com Aléssia Silva dos Santos e Rodrigo Martin Branco para a cadeira de Linguagem Documentária III, agora Sistemas de Classificação III, e foi criado no Thesa. O mesmo foi escolhido devido a familiaridade da autora desta

pesquisa com os termos utilizados na criação do tesouro. Logo, foram utilizados os termos e suas relações e “criado” o mesmo tesouro no Thesa no TemasTres para possibilitar a avaliação de cada um em todos os aspectos abrangidos pela pesquisa.

Os termos escolhidos para a elaboração dos tesouros foram:

- Universo cinematográfico da Marvel;
- Filmes;
- Heróis;
- Vilões;
- Capitão América: Guerra Civil (2016);
- Feiticeira Escarlata;
- Soldado Invernal;
- Viúva Negra;
- Capitão América;
- Falcão;
- Gavião Arqueiro;
- Homem Aranha;
- Homem de Ferro;
- Homem Formiga;
- Máquina de Combate;
- Pantera Negra;
- Patriota de Ferro;
- Visão;
- Brock Rumlow “Ossos Cruzados”
- Helmut Zemo.

Sendo assim, os dados foram analisados a partir de uma comparação entre os dois softwares disponíveis para download e uso. Verificou se ambos atenderam aos requisitos da norma ISO 25964 e listaram-se sugestões do que poderia ser melhorado para facilitar o uso tanto por parte de quem criará o tesouro como de quem o utilizará.

Os critérios utilizados para avaliar os softwares e que serviram de roteiro para esta pesquisa foram retirados do artigo “Estudo comparativo de softwares de construção de tesouros” de Maria Luiza de Almeida Campos et al. Os critérios utilizados foram:

a) Características gerais

Em características gerais temos a avaliação de requisitos genéricos que são importantes para a construção dos termos de tesouro. Nesta categoria se enquadram as limitações genéricas (restrição de tamanho, número, de termos e possibilidade de uso de termos compostos, a existência da informação da fonte ou origem do termo, a possibilidade de criar classes de assuntos, e se os termos podem ser expressos em mais de um idioma.). (CAMPOS *et al.*, 2006, p. 74).

b) Tratamento de relações

Em tratamento de relações, estão os requisitos que indicam de que forma o software permite conceber as relações entre os termos, ou seja, se o software permite relações definidas pelo usuário, relações padrão e relações adotadas no padrão ISO 5964 (1985). (CAMPOS *et al.*, 2006, p. 75).

c) Tratamento de dados

Em tratamento de dados, estão os requisitos que indicam de que forma se dá o armazenamento e a aquisição dos dados, ou seja, se o software permite guardar o tesouro em banco de dados, se permite importação de termos de e para outras fontes, se existe crítica em relação à consistência de dados, se é feito algum controle de restrição de acesso para atualização do tesouro e se permite o gerenciamento da situação de termos, como por exemplo, termos candidatos ou termos obsoletos. (CAMPOS *et al.*, 2006, p. 75).

d) Interface/Manipulação dos dados

Em interface / manipulação dos dados, os requisitos estão voltados para avaliar a maneira que o software permite interagir com os seus usuários, tanto no projeto do tesouro quanto na pesquisa, ou seja, se permite busca flexível por termos, se permite reorganizar a hierarquia de termos, se possui interface Web, e se a criação do tesouro é um processo intuitivo. (CAMPOS *et al.*, 2006, p. 75).

e) Relatórios

Em relatórios, estão os requisitos voltados para avaliar a apresentação da estrutura do tesouro ao usuário final, ou seja, que tipos de relatório possuem, dentre os seguintes: alfabético, sistemático, KWIC, hierárquico e definido pelo usuário. (CAMPOS *et al.*, 2006, p. 75).

f) Características de implementação

Em características de implementação, os requisitos têm como objetivo avaliar os pré-requisitos que o software possui para ser utilizado, ou seja, se depende de software não livre, se roda em Unix, se roda em Windows e se é gratuito. (CAMPOS *et al.*, 2006, p. 75).

g) Apoio metodológico

Em apoio metodológico, os requisitos buscam avaliar se o software foi projetado com recursos que forneçam apoio ao uso de alguma metodologia de construção de tesouros [...]. Esses requisitos são: representação gráfica do escopo ajuda estabelecer o relacionamento entre termos, categorização com base no conceito e uso de características de divisão - que permitem agrupar termos de uma mesma hierarquia de acordo com algum atributo comum (GOMES, CAMPOS, MOTTA, 2004). Cabe ressaltar que este último requisito, diz respeito à etapa de categorização de termos do tesouro. (CAMPOS *et al.*, 2006, p. 75).

h) Apoio ao uso

Em apoio ao uso, os requisitos estão voltados para avaliar o tipo de apoio que o fabricante do software disponibiliza aos seus clientes na utilização do produto, ou seja, se o software possui: suporte, grupo de usuários ativo, ajuda online e documentação. (CAMPOS *et al.*, 2006, p. 76).

Também foram observadas as recomendações diretas da norma ISO 25964-1 para softwares de gerenciamento de tesouros. Estas estão listadas no subcapítulo 2.2 no referencial teórico.

O capítulo subsequente irá discorrer acerca dos resultados desta pesquisa e o capítulo seguinte irá tratar sobre as considerações finais da mesma.

5 RESULTADOS

Neste capítulo vamos apresentar os resultados da pesquisa, começando pelo levantamento dos softwares utilizados para a realização desta pesquisa e seguindo pela análise comparativa dos softwares utilizados para a realização deste trabalho.

O Thesa foi desenvolvido por dois docentes da UFRGS, os professores Rene Faustino Gabriel Junior e Rita do Carmo Ferreira Laipelt. Foi elaborado com base nas normas ISO e NISO vigentes. Funciona em ambiente Web e seu uso é gratuito tanto para fins didáticos quanto para uso profissional.

O TemaTres é um software espanhol multilíngue e de uso livre. Não possui restrição quanto a seu uso, ou seja, pode ser utilizado tanto para a construção de um mini tesouro quanto para a construção de um tesouro no ambiente de trabalho do bibliotecário.

Por sua vez, o TECER foi desenvolvido pelo IBICT e assegura ser uma ótima solução às instituições que procuram por um software livre para gerenciamento de tesouros já que foi criado com a linguagem JAVA, possui compatibilidade com os sistemas Windows e Linux e funciona em ambiente Web e permite o acesso de diferentes usuários de forma simultânea via navegadores como explorer e mozilla. Todavia, não é assim na prática, pois não é possível concluir a instalação do TECER, logo, não é possível seu uso.

Por isso, os resultados apresentados abaixo são acerca do Thesa e TemaTres.

a) Características gerais

Em relação a essa categoria, o TemaTres é bastante completo. Multilíngue e possibilita a inclusão de notas - podendo ser de escopo, do catalogador, histórica, bibliográfica, privada, de definição e de exemplo.

Assim como o TemaTres, o Thesa permite a inclusão de notas - podendo ser de escopo/nota de aplicação, definição do conceito, de exemplo, de alteração, garantia literária e de citação.

O Thesa possui a opção de habilitar ou desabilitar idiomas, esses serão utilizados quando um novo termo for adicionado - pois ao adicionar um termo um ou mais idiomas precisam ser selecionados -. Os idiomas disponíveis até a conclusão desta pesquisa, são: português (Brasil), inglês, espanhol, francês e alemão. Porém, é importante ressaltar que o uso de mais de um idioma ao termo está condicionado à

inclusão desses termos em outro idioma também, tendo em vista que o Thesa não faz a tradução automática dos termos inseridos na língua original em que o tesouro foi escrito. Tanto no Thesa quanto no TemaTres, a publicação ou não dos tesouros é uma opção de seus autores. No segundo também é possível escolher se o vocabulário será hierárquico.

b) Tratamento de relações

Em relação ao tratamento de relações, o TemaTres não permite a criação de relações definidas pelo usuário. Sendo assim, os termos no TemaTres podem ser gerais, subordinados (específicos), relacionados ou alternativos.

A norma ISO 25964-1 (2011) define Termo Geral como termo preferido que representa um conceito que é mais amplo do que aquele em questão, Termo Subordinado como termo preferido que representa um conceito que é mais restrito do que aquele em questão. Termo Relacionado como termo preferido que representa um conceito que tem uma relação associativa com aquele em questão e Termo Alternativo como o que não é atribuído a documentos, mas é fornecido como um ponto de entrada em um tesouro ou índice.

Quando se faz necessária a adição de um termo que seja subordinado hierarquicamente, este é um termo subordinado. Já um termo relacionado, é aquele que é relacionado a outro, porém, não de forma hierárquica e sim possui uma relação associativa.

O termo alternativo (ou termo não preferido) é utilizado para adicionar termos através da relação USE/USADO POR. No TemaTres existe a opção de especificar o tipo de relação que os dois termos possuem. As especificações são: *Spelling Variant* (é variação de), por exemplo, Patriota de Ferro é variação de Máquina de Combate, *MisSpelling* (erro ortográfico), tal como, mecher é erro ortográfico de mexer, *Abbreviation* (abreviação), tendo como CAPES é abreviação de Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, *Full form of the term* (é o extenso de), tendo como exemplo, cinquenta é o extenso de 50 e *Hidden* (é termo oculto), por exemplo, Ranganatam é o termo oculto de Ranganathan.

Tal qual o TemaTres, o Thesa não permite a criação de relações definidas pelo usuário. Um termo pode ser geral, específico, relacionado, equivalente (usado por) ou equivalente (oculto). O Thesa apresenta uma maior variação de termos relacionados

e termos equivalentes. A seguir é apresentada uma tabela de como o(s) termo(s) relacionado(s) pode(m) se relacionar, assim como exemplos:

TABELA 1 – Relações dos termos e exemplos

Coordenação	Biblioteca Pública e Biblioteca Municipal – são tipos de biblioteca
Ação/Produto	catalogação (ação) produz catálogos (produto)
Causa/Efeito	excesso de sal (causa) pressão baixa (efeito)
Oposição/Antonímia	processo aberto é antonímia de processo fechado
Afinidade	enfermeira tem afinidade com médico
Processo/Parte	acusação (processo) julgamento (parte de)
Características do produto	mesa é de madeira (característica do produto)
Ação/Paciente ou Objetivo	vacinação (ação) é realizada para que haja imunização (objetivo)
Ação/Propriedade	Censo demográfico (ação) gera o conhecimento das condições de vida da população (propriedades)
Campo de estudo/Objeto ou Fenômenos estudados	Ciências exatas (campo de estudo) estudam a matemática, química e física (fenômenos estudados)
Ação ou Objeto/Contra Agente	deterioração por mofo (ação) é combatida por ventilação (contra agente)
Conceito ou Objeto/Propriedade	revista (conceito) tem no mínimo 20 páginas (propriedade)

Objeto/Suas Partes	escada (objeto) é composta por degraus (suas partes)
Matéria Prima/Produto	madeira é matéria prima da mesa

Fonte: Elaborado pela autora, 2021.

A tabela abaixo apresenta as possíveis variações do termo equivalente (usado por), assim como exemplo:

TABELA 2 – Variações do termo equivalente

Plural de	vilões é plural de vilão
Extenso de	vinte é o extenso de 20
Abreviatura	fã é abreviatura de fanático
Notação (código)	CDU 33 é notação de educação
Sigla	IBICT é a sigla de Instituto Brasileiro de Formação em Ciência e Tecnologia
Variação de	vilões é variação de vilões
É traduzido por	Universal Decimal Classification é traduzido por Classificação Decimal Universal
É gerúndio de	estudando é gerúndio de estudar
É feminino de	vilã é feminino de vilão
É flexão verbal de	acharam é flexão verbal de achar

Fonte: Elaborado pela autora, 2021.

Percebe-se que quanto às relações o Thesa é mais completo do que o TemaTres, pois este possui poucas relações disponíveis, tornando o tesauro que for criado muito abrangente já que não é possível especificar as relações.

c) Tratamento de dados

Tanto no TemaTres quanto no Thesa, os termos ficam armazenados em servidor web e quanto a aquisição dos dados, podem ser importados ou adicionados manualmente.

O TemaTres permite importar tesouro de um arquivo txt nos formatos: texto tabulado, texto etiquetado, Skos-core e MARC 21 XML Schema (MarcXML).

Quanto a exportação, essa pode ser realizada nos seguintes formatos:

- Lista hierárquica (txt);
- Lista alfabética (txt);
- Sistemático (PDF);
- Alfabético (PDF);
- Moodle;
- Zthes;
- Skos-core;
- TopicMap;
- BS8723;
- Metadata Authority Description Schema (MADS);
- MARC 21 XML Schema (MarcXML);
- IMS Vocabulary Definition Exchange (VDEX);
- WXR (Wordpress XML);
- SiteMap;
- SQL (Backup).

O Thesa possibilita a importação nos formatos: Turtle RDF (.ttl), XML OWL (.owl), Term List (.txt), Skos-Core (.rdf) e Lista Alfabética (Tematres) (.txt). Já a exportação é possível nos formatos: (.xml), (.csv), (.txt), (.rdf), (.json), (.skos) e (.pdf).

Observa-se uma maior variação de formatos disponíveis para exportação no TemaTres, quinze diferentes formatos, enquanto o Thesa possibilita sete diferentes formatos.

Quanto a importação, o TemaTres só permite que ocorra em arquivo (.txt) e em quatro diferentes formatos, sendo esses: texto tabulado, texto etiquetado, Skos-Core e MARC 21 XML Schema (MarcXML). Já o Thesa, possibilita que a importação ocorra em arquivo (.txt), (.ttl), XML OWL (.owl), Term List (.txt), Skos-Core (.rdf) e Lista Alfabética (Tematres) (.txt).

A norma ISO 25964-1 evidencia que para que o conteúdo do tesauro não se limite ao software ou aplicativo inicialmente usado para desenvolvê-lo é importante saber exportar e importar dados. Um dos pontos frisados pela norma é que além de exportar o vocabulário incluindo todos os termos, notas de escopo, notação e relações padrões entre os termos e conceitos, deve ser possível exportar subconjuntos definidos pelo autor. Vê-se que apenas o TemaTres possibilita a exportação de subconjuntos definidos pelo autor, enquanto o Thesa exporta o vocabulário completo.

Sendo assim, caso o usuário queira exportar apenas uma parte do tesauro, por qualquer motivo que seja, isso não é possível no Thesa, apenas no TemaTres. Logo, este é um ponto que o Thesa precisa melhorar.

d) Interface/Manipulação dos dados.

O tesauro, no TemaTres, é apresentado de forma sistemática, que se organiza hierarquicamente e se ordena de forma alfabética. A navegação no TemaTres pode ser realizada através dos meta-termos que aparecem na página inicial - ao clicar na seta ao lado do meta-termo, os descritores ficam visíveis, em ordem alfabética - do alfabeto localizado acima e abaixo dos meta-termos ou de pesquisas (figura 2).

Figura 2 - Apresentação sistemática, organizada hierarquicamente e ordenada de forma alfabética

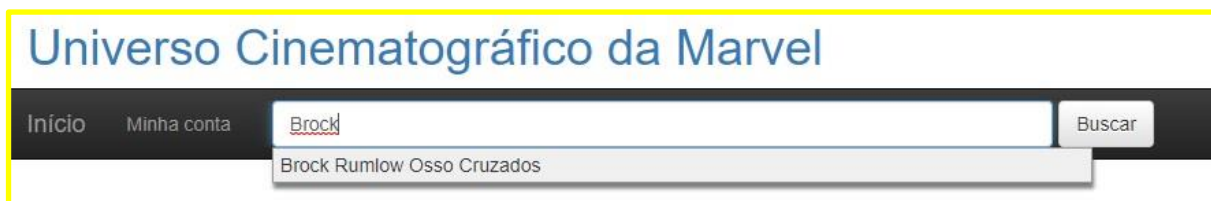


Fonte: Elaborado pela autora, 2021.

Ao digitar uma palavra, logo abaixo, aparecem sugestões de todos os termos que possuem aquela palavra (figura 3). A busca também pode ser realizada na aba de “pesquisa avançada” que apresenta a opção de filtrar a busca somente por termo, meta-termo, termo não preferido, podendo selecionar uma caixa de pesquisa “frase

exata” (figura 4). Pode-se restringir mais a busca ao selecionar “opções” onde se pode escolher por: “tem este termo superior”, “tipo de nota”, “criado em ou após” e “está localizado no nível mais profundo” (figura 5).

Figura 3 - Sugestão de pesquisa



Fonte: Elaborado pela autora, 2021.

Figura 4 - Pesquisa avançada

A imagem mostra a interface de "Pesquisa avançada". No topo, há o título "Pesquisa avançada". Abaixo dele, há a pergunta "O que buscar?". Há um menu suspenso com "Termo" selecionado e uma seta para baixo. O menu suspenso está aberto, mostrando as opções: "Termo", "Meta-termo", "Termo não preferido" e "Nota". Abaixo do menu, há um botão "Frase exata" com uma caixa de seleção vazia. No rodapé, há dois botões: "Enviar" (em azul) e "Opções" (em cinza).

Fonte: Elaborado pela autora, 2021.

Figura 5 - Mais opções de pesquisa avançada

Pesquisa avançada

O que buscar?
Termo

Pesquisar termo

Frase exata

Tem este termo superior
Todos

Tipo de nota
Todos

Criado em ou após
Todos

Está localizado no nível mais profundo
Todos

Fonte: Elaborado pela autora, 2021.

É permitido reorganizar a hierarquia do termo, porém para isto é necessário fazer de forma individual, ou seja, termo por termo. No Thesa não é possível alterar o TG de um termo a não ser que entre no Termo Geral e troque os Termos Específicos do mesmo.

O TemaTres possui uma interface Web, podendo ser utilizado remotamente por pessoas autorizadas a manipular o tesauro, sendo necessário saber o usuário e senha utilizados para aquele tesauro.

Diferentemente do TemaTres que assim que acessado já apresenta em sua página inicial sua organização sistemática, o Thesa é organizado apenas de forma alfabética em sua interface inicial - a apresentação sistemática ocorre apenas ao se fazer a exportação do tesauro em formato PDF. A navegação no Thesa se dá através do alfabeto localizado mais ou menos no meio da página inicial ou da aba de pesquisas, localizada no canto direito da página logo acima do alfabeto. Não possui, porém, a opção de “pesquisa avançada” (figura 6).

Figura 6 - Interface inicial do Thesa

Home Configurações Ferramentas adicionar termo buscar termos pesquisar

todas B C F G H M P S U V

Fonte: Elaborado pela autora, 2021.

O TemaTres para ser utilizado remotamente por diferentes pessoas, o usuário e senha é o mesmo para todos. Diferentemente do Thesa que há a possibilidade de adicionar mais de um colaborador no tesouro, cada um usando seu usuário e senha. Logo, o criador do tesouro precisa adicionar os outros colaboradores, podendo optar pela função que o usuário irá desenvolver, podendo ser: autor, orientador ou colaborador (figura 7).

Figura 7 - Adição de colaboradores

A imagem mostra uma janela de diálogo intitulada "Adicionar colaboradores". No topo, há um ícone de fechamento (x) e o título "Adicionar colaboradores". Abaixo, o texto "Informe o e-mail do colaborador" é exibido em azul. Segue um campo de entrada rotulado "email". Abaixo dele, a seção "function" contém três opções de função com botões de opção: "Autor", "Orientador" e "Colaborador". Na base da janela, há dois botões: "cancelar" (botão cinza) e "adicionar termo" (botão azul).

Fonte: Elaborado pela autora, 2021.

A ISO 25964-1 (2011, p. 102, tradução nossa) diz que “O editor do tesouro deve ser capaz de localizar os detalhes completos de qualquer termo [...] pela entrada direta deste termo, de preferência apenas precisando inserir uma parte do termo. com o software auxiliando na seleção, identificando todos os termos que contêm a parte”. Pensando nisso, ambos os softwares atendem este requisito, porém o TemaTres, ao digitar uma palavra, aparecem sugestões de todos os termos que possuem aquela palavra. Enquanto o Thesa, ao digitar uma palavra é necessário clicar no botão pesquisar para aparecer todos os termos que possuem aquela palavra. Ou seja, a recuperação de termos no TemaTres ocorre melhor do que no Thesa, pois segue com mais afinco a norma.

A ISO 25964-1 diz também que “O editor deve ser capaz de alternar facilmente de uma tela em um idioma para a tela correspondente em qualquer um dos outros idiomas do tesouro” (ISO, 2011, p. 102, tradução nossa). No TemaTres, uma vez selecionado o idioma que o tesouro será construído, não tem mais como mudar, porém existem dezesseis idiomas dos quais a página inicial pode ser apresentada. O Thesa só altera o idioma da página inicial e não do tesouro em si. Logo, faz-se

necessárias melhorias em ambos os softwares em relação aos idiomas de apresentação para ser possível a criação de tesouros multilíngues ou apenas da visualização por pessoas que não entendem os idiomas que o mesmo foi construído, mas que entenderiam se os próprios softwares fizessem essa conversão idiomática.

e) Relatórios

Os relatórios gerados pelo TemaTres são apresentados em planilhas no formato .csv. Porém, nesse formato alguns recursos podem ser perdidos. Para preservar estes recursos, a recomendação quando se abre o arquivo no formato .csv é que o arquivo seja salvo em um formato de arquivo Excel (figura 8).

Figura 8 - Erro no relatório do TemaTres

term_id	term_id	relType	relSubTyp	user_id	created
24	14	UP/USE	SP	1	#####
2	23	TG/TE		1	#####
1	22	TR		1	#####
22	1	TR		1	#####
1	2	TR		1	#####

Fonte: Elaborado pela autora, 2021.

Ao se criar um relatório em português faz-se necessário marcar a opção “Codificação latin1” para que os acentos e cedilhas apareçam no relatório.

As opções de relatórios são:

- Termos livres;
- Termos repetidos;
- Mais de um termo geral;
- Termos segundo quantidade de termos específicos;
- Termos por quantidades de palavras;
- Meta-termos;
- Termos preferidos;
- Termos relacionados;
- Termos não preferidos;
- Fontes de referência;
- Termo candidato;

- Termo não aceito;
- Todos os termos;
- Todas as relações entre termos;
- Todas as notas;
- Fonte de referência dos termos.

Quanto aos relatórios avançados, os filtros que podem ser aplicados são:

- ✓ Tem este termo superior;
- ✓ Tem nota do tipo;
- ✓ Criado em ou após;
- ✓ Inclui palavras - o termo pode ser buscado por um termo que “inclui palavras iguais a”, “inclui palavras que começam com” e “inclui palavras que terminam em”.

Também é possível gerar relatórios de todos os termos que não possuem um tipo ou nenhum de nota. Porém, ao contrário dos filtros de relatórios avançados, só se pode selecionar uma opção.

O único relatório disponibilizado pelo Thesa é um gráfico que mostra o crescimento da base. Não é permitido o download deste relatório (figura 9).

Figura 9 - Relatório do Thesa



Fonte: Elaborado pela autora, 2021.

A ISO 25964-1 (2011, p. 103, tradução nossa) diz que “Deve ser possível obter relatórios da quantidade de termos com características particulares, particularmente do número total de termos preferidos e não preferidos.” A partir dessa citação, constata-se que o Thesa não vai de encontro com a norma visto que o único relatório

que fornece é sobre o crescimento da base. Já o TemaTres está totalmente de acordo com a norma neste requisito.

f) Características de implementação

Podendo ser baixado em sistema Windows, Mac ou Linux, para ser possível a instalação do TemaTres, é necessário um servidor web que suporte PHP. O PHP instalado deve ser versão 4.3.0 ou superior sendo suportado o PHP 5. Além disso, é preciso um Sistema Gerenciador de Banco de Dados, o padrão utilizado pelo TemaTres é o MySQL, porém pode-se utilizar qualquer outro desde que o código-fonte seja ajustado (SHINTAKU, 2019).

Para a autora, também se fez necessária a ajuda do professor Rene F. Gabriel Junior para concluir a instalação do TemaTres, pois mesmo existindo um Guia de Usuário do Tematres, o mesmo ensina a fazer a instalação no sistema Linux apenas e para a realização desta pesquisa foi utilizado o sistema Windows. Mesmo com o auxílio de um vídeo encontrado no YouTube que mostrava a instalação do TemaTres no Windows e um profissional da Tecnologia da Informação, ao tentar abrir para concluir a configuração estava dando erro. Somente após pedir auxílio técnico ao professor Rene, que possui um conhecimento avançado de programação e sobre este tipo de base, que foi possível concluir a instalação e usar o TemaTres. Logo, entende-se que se faz necessário um conhecimento específico avançado para a instalação do TemaTres.

O Thesa se encontra 100% em servidor web, não se fazendo necessário o seu download ou de um Sistema Gerenciador de Banco de Dados. Para acessá-lo basta entrar em seu endereço web (<https://www.ufrgs.br/tesauros/index.php/thesa>) e realizar o cadastro. Com o cadastro realizado, o Thesa já está pronto para ser utilizado.

Devido à dificuldade encontrada na instalação do TemaTres, identifica-se que um manual de instalação em sistema Windows é necessário, visto que possui manual para sistema Linux, além de um melhor e mais fácil suporte por parte do desenvolvedor.

Quanto a este requisito, o Thesa está em vantagem visto que, não é necessário download nem instalação, apenas acessar o site que está disponível e criar seu usuário e senha.

g) Apoio metodológico

Quanto a essa categoria, tanto o TemaTres quanto o Thesa não possuem apoio abrangente em suas respectivas páginas. Porém o primeiro possui o Guia do Usuário do TemaTres, onde é realizada a conceituação geral de vocabulário controlado, termo, controle de autoridade, tesouro, taxonomia e ontologia. Também ajudará na instalação, configuração e manutenção do software, em sistema Linux, bem como administrá-lo. Este guia irá ensinar a gerir um tesouro no TemaTres e irá exemplificar a experiência do TemaTres na Enap.

Porém existe um artigo de 2019, dos desenvolvedores do Thesa que auxilia ao explicar quais relações existem no Thesa e apresenta exemplos de uso (figura 10). O nome do artigo é “Descrição das relações semânticas para aplicação de KOS: uso do tesouro semântico (Thesa)” e está disponível através do seguinte endereço Web: <http://doi.org/10.21721/p2p.2019v6n1.p117-135>.

Figura 10 - Exemplos de uso de relações no Thesa

Quadro 2 - Apresentação das Metacategorias Semânticas Hierárquicas e de Equivalências nos tesouros tradicionais e no Thesa

<i>Tesouro</i>	<i>Tesouro Semântico</i>	<i>Exemplos de aplicações</i>
TG	TG (é tipo de ...)	"Biblioteca Universitária" é um tipo de "Biblioteca"
	TG (é parte de ...)	"Porto Alegre" é parte do "Rio Grande do Sul"*
USE	UP (abreviatura)	"Pag." é abreviatura de "Página"
	UP (sigla)*	"AACR2" é a sigla de "Código de Catalogação Anglo Americano, segunda edição"
	UP (acrônimo)	"UFRGS" é a acrônimo de "Universidade Federal do Rio Grande do Sul"
	UP (é variação de)	"Catalogação" é variação de "Representação Descritiva" (Sinônimo)
	UP (é um gerúndio de)	"Catalogando" é um gerúndio de "Catalogar"
	UP (é plural de)*	"Livros" é plural de "Livro"
	UP (é singular de)*	"Livro" é singular de "Livros"
	UP (é masculino de)*	"Bibliotecário" é o masculino de "Bibliotecária"
	UP (é o feminino de)*	"Bibliotecária" é o feminino de "Bibliotecário"
	UP (é o extenso de)* (é o termo completo)	"Cinco" é o extenso de "5"
	UP (é a notação de)*	CDU02 é a notação de "Biblioteconomia"
	UP (é termo oculto)*	"Ranganatam" é o termo oculto de "Ranganathan"
	UP (é flexão verbal de)*	"Pesquisaram" é flexão verbal de "Pesquisar"

* Proposição dos autores

Fonte: Autores (2019).

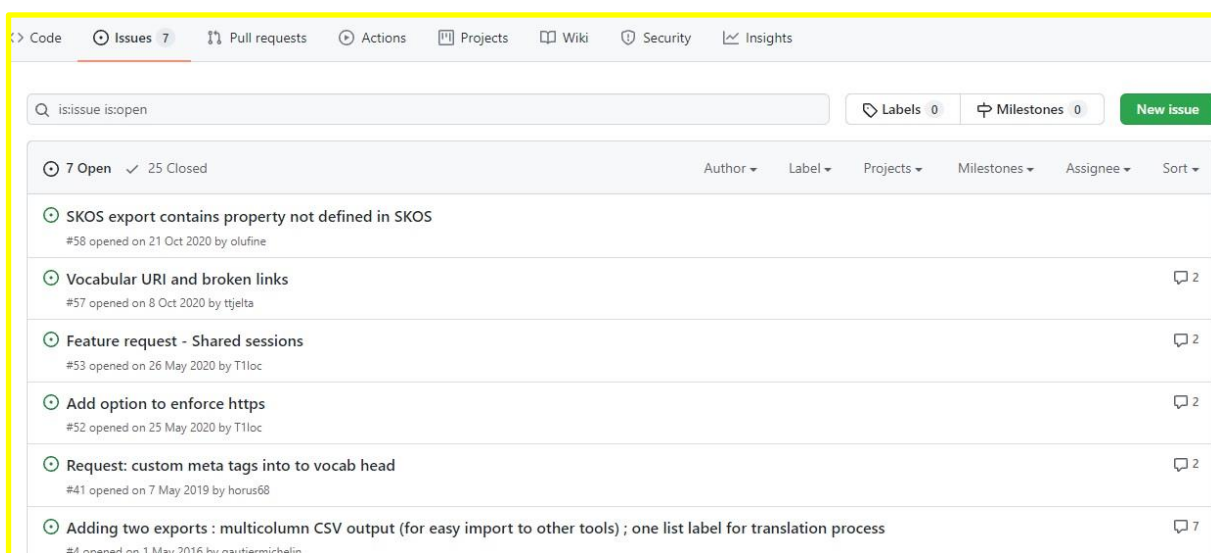
Fonte: GABRIEL JUNIOR, LAIPELT, 2019, p. 129.

Tendo em vista o que foi apresentado neste tópico, constata-se a necessidade de ferramentas de apoio metodológicos nas páginas dos dois softwares. Pois, mesmo o Thesa possuindo um artigo que auxilia quanto às relações, o mesmo não está disponível em sua página, então os usuários precisarão realizar uma pesquisa mais a fundo para encontrar esta ferramenta de apoio. Logo, faz-se necessário a criação de um guia de usuário do Thesa.

h) Apoio ao uso

O desenvolvedor do TemaTres diz que a melhor maneira de obter ajuda caso haja algum problema é acessar o fórum de discussões (figura 11). Caso não queira participar do fórum, o usuário pode se inscrever na lista de ajuda e esperar o retorno.

Figura 11 - Fórum de discussões sobre o TemaTres



Fonte: Elaborado pela autora, 2021.

Caso seja preciso algum apoio durante o uso do Thesa, na parte de cima da página existe a aba “contato” onde se encontra o e-mail do desenvolvedor Thesa (figura 12). Esta é a única maneira de procurar ajuda caso seja necessário.

Figura 12 - Contato do desenvolvedor do Thesa



Fonte: Elaborado pela autora, 2021.

Nenhum dos dois softwares possuem grandes ferramentas de apoio ao uso. Caso o usuário não queira esperar a resposta do desenvolvedor do TemaTres, existe um Guia de Usuário que serve como manual, porém se o problema for com a instalação e o sistema que o usuário estiver usando for Windows, o manual não irá ajudar e será necessário aguardar o contato do desenvolvedor. Em contrapartida, o Thesa não possui guia ou manual de uso, logo, a única forma de conseguir ajuda é entrando em contato com o desenvolvedor, que geralmente responde consideravelmente rápido.

Entretanto, é primordial que ambos os softwares possuam manual ou guia para todos os sistemas em que eles podem ser utilizados, pois caso o desenvolvedor de um deles ficar indisponível, a dúvida do usuário não será sanada.

Com base nos resultados desta pesquisa, a seguinte tabela foi proposta para comparar os softwares:

TABELA 3 - Comparativo entre TemaTres e Thesa

Crítérios	TemaTres	Thesa
a) Características gerais		
Multilíngue	sim	sim
Inclusão de notas	sim	sim
b) Tratamento de relação		
Relações definidas pelo usuário	não	não
Relação padrão do software	sim	sim
c) Tratamento de dados		

Permanência em banco de dados	sim	sim
Importação de dados	sim	sim
Exportação de dados	sim	sim
d) Interface e manipulação de dados		
Busca Flexível	sim	não
Recurso para reorganizar hierarquias	não	não
Facilidade de uso	sim	sim
Interface Web	sim	sim
Adição de colaboradores	não	sim
Pesquisa avançada	sim	não
Apresentação sistemática	sim	não
Apresentação alfabética	sim	sim
e) Emissão de relatórios		
Relatório alfabético	sim	não
Relatório sistemático	não	não
Relatório hierárquico	não	não
Relatório na Web	não	não
Relatório definido pelo usuário	não	não
Download de relatório	sim	não
f) Características de implementação		
Software livre e gratuito	sim	sim
Roda Linux	sim	sim
Roda Windows	sim	sim
Necessita conhecimento específico de informática	sim	não
Necessita download	sim	não
g) Apoio metodológico		
Ajuda a relacionar os termos	não	não

Ajuda a definir conceitos	não	não
h) Apoio ao uso		
Possui suporte	sim	sim
Possui ajuda online	sim	sim
Possui rede de usuários	sim	não

Fonte: Elaborado pela autora, 2021.

Foi possível perceber, através deste estudo, que não tem como definir qual o melhor software para ser utilizado, pois depende do objetivo que se almeja atingir. Se desejar relações definidas pelos usuários, nenhum dos dois softwares aqui estudados irão atender o desejo. Se for necessário fazer a exportação dos dados, de forma específica, o recomendado é o TemaTres, pois possui maior variação de formatos disponíveis para exportar.

Caso, um requisito para a criação do tesauro seja a apresentação sistemática dos termos, o Thesa não é indicado, considerando que ele apresenta os termos de forma alfabética. Caso possua muitos termos, ou apenas seja opção do criador que os termos possam ser buscados através de pesquisa avançada, o mais adequado é o TemaTres.

Se for necessária a emissão de relatórios, o Thesa não deve ser levado em consideração, visto que o único relatório que emite é o gráfico do crescimento da base. Enquanto isso, no TemaTres são emitidos relatórios especificando se são sobre termos livres, termos repetidos, mais de um termo geral, termos segundo quantidade de termos específicos, termos por quantidade de palavras, meta-termos, termos preferidos, termos relacionados, termos não-preferidos, fontes de referência, termo candidato, termo não aceito, todos os termos, todas relações entre termos, todas as notas, fontes de referência dos termos e termos sem notas. Nesta última, é possível definir qual tipo de notas deverão ser excluídas do relatório.

É importante ressaltar que antes de decidir qual software de gerenciamento de tesauro decidir utilizar, o usuário deve assegurar que o software irá rodar em seu computador/notebook. Se porventura, escolher o TemaTres é importante fazer a leitura do guia de usuário e caso não consiga instalá-lo é primordial pedir ajuda a quem possui familiaridade com informática.

Quanto ao apoio metodológico, o TemaTres sai um pouco a frente do Thesa, pois possui o guia de usuários. Enquanto, o segundo, tem como apoio um artigo escrito pelos desenvolvedores do mesmo que explica as relações existentes no software e exemplifica. Já sobre o apoio de uso, o TemaTres dá duas opções: participação no fórum ou entrada na lista de ajuda, onde aguardará retorno. O Thesa oferece apenas um meio de apoio ao uso, o e-mail do desenvolvedor do mesmo. Diante disso, a resposta à dúvida do usuário pode demorar a vir, tanto no TemaTres quanto no Thesa.

Após todas as discussões e apontamentos sobre avaliação comparativa de softwares livres para gerenciamento de tesouros no decorrer desta pesquisa. O capítulo seguinte trará as considerações finais deste trabalho.

6 CONSIDERAÇÕES FINAIS

Este trabalho teve como seus objetivos geral e específico, respectivamente, investigar se os softwares livres de gerenciamento de tesouros disponíveis em língua portuguesa, no Brasil, que possibilitam o atendimento das recomendações da norma ISO 25964-1; descrever as características dos softwares livres de gerenciamento de tesouros encontrados; avaliar se os softwares livres de gerenciamento de tesouros possibilitam a representação de um conjunto mais rico de relações semânticas do que as normalmente usadas em tesouros tradicionais.

Dos três softwares livres para gerenciamento de tesouros em língua portuguesa no Brasil, vale ressaltar que o TECER apresenta problema em sua instalação o que impossibilitou o seu uso. No processo para instalá-lo é necessário abrir o arquivo `bd.sql` no pgAdmin III, arquivo que é inexistente. Sendo assim, os desenvolvedores desse software precisam resolver o problema que impossibilita seu uso, pois existem apenas duas opções de softwares livres de gerenciamento de tesouros em língua portuguesa, no Brasil. Logo, a avaliação comparativa foi realizada entre o TemaTres e o Thesa.

Após a realização de todas as etapas propostas neste trabalho, percebeu-se que tanto o TemaTres quanto o Thesa estão indo ao encontro ao estabelecido na norma ISO 25964-1. Porém, ainda são necessárias atualizações nos softwares para que possam ficar em concordância completa com a norma. Portanto, o problema de pesquisa: “Quais as características dos softwares livres de gerenciamento de tesouros são compatíveis com a norma ISO 25964-1?” pôde ser respondida pela análise dos dois softwares analisados.

Ambos os softwares permitem a adição de notas, o Thesa permite o uso de até seis tipos de notas e o Tematres cinco. Entretanto, não são os mesmos tipos de notas.

Na configuração do TemaTres é escolhido o idioma no qual o tesouro será construído, não se exclui a possibilidade de adicionar um termo em outro idioma que não o escolhido. Já o Thesa, só é trilingue no que se refere a apresentação na página inicial. Caso queira adicionar termos em outra língua, que não o português, deve ser feito de forma manual tendo em vista que, mesmo habilitando outros idiomas, nada acontece.

Diante disso, mesmo ambos os softwares apresentando um tesouro relativamente bem construído, é necessário analisar qual dos dois melhores se encaixa nos propósitos de quem está desenvolvendo um tesouro. Pois, como foi apresentado, cada software tem seu ponto forte e suas limitações e diferentemente de um tesouro tradicional criado no word, no TemaTres e no Thesa as relações são padronizadas pelos sistemas. Porém, a representação de um conjunto mais rico de relações semânticas do que as normalmente usadas em tesouros tradicionais é possível apenas no Thesa, tendo em vista que o TemaTres não possibilita a explicitação das relações entre os termos conforme sugere a norma, apresentando apenas as relações tradicionais. Logo, ambos necessitam de melhorias para estarem inteiramente de acordo com a norma ISO 25964-1.

Vale ressaltar a importância de que mais estudos sobre a temática sejam realizados, não só utilizando os softwares livres, mas utilizando os softwares pagos também para que assim sejam possíveis avaliações mais ricas.

REFERÊNCIAS

CAMPOS, Maria Luiza de Almeida. **Linguagem Documentária: teorias que fundamentam sua elaboração**. Niterói, RJ: EdUFF, 2001.

CAMPOS, Maria Luiza de Almeida *et al.* Estudo comparativo de softwares de construção de tesouros. **Perspect. ciênc. inf.**, Belo Horizonte, v. 11, n. 1, p. 68-81, abr. 2006. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000100006&lng=en&nrm=iso. Acesso em 28 set. 2019.

CURRÁS, Emília. **Tesouros, Linguagens Terminológicas**. Brasília: IBICT, 1995.

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://hdl.handle.net/20.500.11959/brapci/104617>. Acesso em: 30 out. 2019.

GABRIEL JUNIOR, R. F.; LAIPELT, R. C. Descrição das relações semânticas para aplicação em KOS: uso do tesouro semântico aplicado (Thesa). **Revista P2P & Inovação**, Rio de Janeiro, v. 6, n. 1, p.117-135, set. 2019/fev. 2020. Disponível em: <http://doi.org/10.21721/p2p.2019v6n1.p117-135>. Acesso em: 03 nov. 2021.

GABRIEL JUNIOR, R. F.; LAIPELT, R. C. Thesa: ferramenta para construção de tesouro semântico aplicado interoperável. **Revista P2P & Inovação**, Rio de Janeiro, v. 3, n. 2, p.124-145, mar./set. 2017. Disponível em: <https://doi.org/10.21721/p2p.2017v3n2.p124-145>. Acesso em: 10 out. 2019.

GALDINO, Rejane; SABBAG, Deise Maria Antonio. TemaTres e a construção de tesouros: aspectos de aplicabilidade para o profissional da informação. **RBBB. Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 13, p. 2566-2579, jan. 2018. ISSN 1980-6949. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1040>. Acesso em: 28 out. 2019.

HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge organization**, Germany, v. 35, n.2/n.3, p. 86-1001, 2008. ISSN 0943-7444. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2008-2-3-86/what-is-knowledge-organization-ko-volume-35-2008-issue-2-3>. Acesso 5 abr. 2021.

INTERNATIONAL STANDARD. **Iso 25964-1: Information and documentary - Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval**. 2011.

JESUS, Rafaela Cristina de. **Construção e manutenção de tesouros brasileiros: análise na perspectiva das normas internacionais**. 2018. 84 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista / Faculdade de Filosofia e Ciências, Marília, 2018. Disponível em: <http://hdl.handle.net/11449/155854>. Acesso em 22 nov. 2021.

LIMA, Gercina Ângela de. Organização do conhecimento e da informação na web: teorias e técnicas. **Perspectiva em Ciência da Informação**, v. 25, p. 57-97, fev/2020. Número especial. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22283>. Acesso em: 15 abr. 2021.

LIMA, José Leonardo Oliveira; ALVARES, Lilian. Organização e representação da informação e do conhecimento. *In*: ALVARES, Lilian (org.). **Organização da Informação e do Conhecimento: conceitos, subsídios interdisciplinares e aplicações**. B4 Editores, 2012. p. (21-48). ISBN: 9788565358439.

MIRANDA, M. L. C. de. **Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais**. 2005. 353 f. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro / Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2005. Disponível em: <https://ridi.ibict.br/handle/123456789/690>. Acesso em 05 abr. 2021.

MOREIRA, M. P., MOURA, M. P. Construindo tesouros a partir de tesouros existentes: a experiência do TCI Tesouro em Ciência da Informação. **DataGramZero**, v.7, n.4, ago. 2006.

MORESI, E. (org.). **Metodologia da Pesquisa**. 2003. 108 f. Trabalho Científico (Especialização em Gestão do Conhecimento e Tecnologia da Informação) Universidade Católica de Brasília, 2003.

NHACUONGUE, J. A.; DUTRA, M. L. De paul otlet à web semântica: aportes teóricos sobre a organização do conhecimento. **Informação & Tecnologia**, v. 3, n. 1, p. 138-158, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41780>. Acesso em: 1 out. 2021.

SHINTAKU, Milton (org.). **Guia do usuário do TemaTres**. Brasil: Escola Nacional de Administração Pública (Enap), 2019. Disponível em: <http://repositorio.enap.gov.br/handle/1/4199>. Acesso em: 05 out. 2021.

Wiki. Instituto Brasileiro de Informação em Ciência e Tecnologia: TECER. Disponível em: <http://wiki.ibict.br/index.php/Tecer>. Acesso em: 28 out. 2019.