

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO & ORGANIZAÇÃO DO
CONHECIMENTO**

MAYARA OLIVEIRA SILVA GONÇALVES

**DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL
DE LIVRO POR MEIO DA INDEXAÇÃO SEMIAUTOMÁTICA**

Belo Horizonte
2020

MAYARA OLIVEIRA SILVA GONÇALVES

**DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL
DE LIVRO POR MEIO DA INDEXAÇÃO SEMIAUTOMÁTICA**

Dissertação apresentada ao Programa de Pós-graduação em Gestão & Organização do Conhecimento (PPG-GOC) da Escola de Ciência da Informação da Universidade Federal de Minas Gerais, para obtenção do grau de mestre em Gestão e Organização do Conhecimento.

Linha de Pesquisa: Arquitetura e Organização do Conhecimento (AOC).

Orientador: Profa. Dra. Gercina Ângela de Lima.

Belo Horizonte
2020

G635

Gonçalves, Mayara Oliveira Silva

Diretrizes para elaboração e estruturação de Índice de Final de Livro por meio de indexação semiautomática [recurso eletrônico]. / Mayara Oliveira Silva Gonçalves. - 2020.

1 recurso eletrônico (169f.: il., color): pdf.

Orientadora: Gercina Ângela de Lima

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 156-162.

Apêndices: f. 163-169.

Exigências do sistema: Adobe Acrobat Reader.

1. Ciência da informação - Teses. 2. Indexação automática - Teses. 3. Sistemas de recuperação da informação - Teses. I. Título. II. Lima, Gercina Ângela de. III. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU:025.4



UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO
CONHECIMENTO



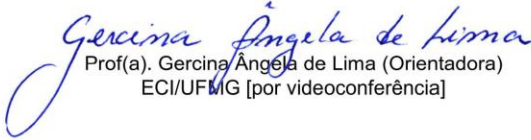
FOLHA DE APROVAÇÃO

DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL
DE LIVRO POR MEIO DA INDEXAÇÃO SEMIAUTOMÁTICA

MAYARA OLIVEIRA SILVA GONÇALVES


Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, como requisito para obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, área de concentração CIÊNCIA DA INFORMAÇÃO, linha de pesquisa Arquitetura e Organização do Conhecimento.

Aprovada em 08 de maio de 2020, pela banca constituída pelos membros:


Prof(a). Gercina Ângela de Lima (Orientadora)
ECI/UFMG [por videoconferência]


Prof(a). Benildes Coura Moreira dos Santos Maculan
ECI/UFMG [por videoconferência]


Prof(a). Cintia de Azevedo Lourenço
ECI/UFMG [por videoconferência]


Prof(a). Elisângela Cristina Aganette
ECI/UFMG [por videoconferência]

Belo Horizonte, 8 de maio de 2020.



ATA DA DEFESA DA DISSERTAÇÃO DA ALUNA MAYARA OLIVEIRA SILVA GONÇALVES

Realizou-se, no dia 08 de maio de 2020, às 14:00 horas, Videoconferência, da Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL DE LIVRO POR MEIO DA INDEXAÇÃO SEMIAUTOMÁTICA*, apresentada por MAYARA OLIVEIRA SILVA GONÇALVES [por videoconferência], número de registro 2016712451, graduada no curso de BIBLIOTECONOMIA, como requisito parcial para a obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Gercina Ângela de Lima - ECI/UFMG [por videoconferência] (Orientadora), Prof(a). Benildes Coura Moreira dos Santos Maculan - ECI/UFMG [por videoconferência], Prof(a). Cíntia de Azevedo Lourenço - ECI/UFMG [por videoconferência], Prof(a). Elisângela Cristina Aganette - ECI/UFMG [por videoconferência].

A Comissão considerou a dissertação:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.
Belo Horizonte, 08 de maio de 2020.


Prof(a). Gercina Ângela de Lima


Prof(a). Benildes Coura Moreira dos Santos Maculan


Prof(a). Cíntia de Azevedo Lourenço


Prof(a). Elisângela Cristina Aganette

Dedico ao meu filho Davi e aos meus pais Silvio e Lucimar.

AGRADECIMENTOS

Primeiramente agradeço a **Deus** por se fazer presente em minha vida e me conceder saúde, força e perseverança para alcançar meus objetivos.

À querida **Profa. Gercina Lima** que me acompanha desde a graduação. Pela paciência, atenção, dedicação e por todos os ensinamentos ao longo desses anos. Uma pesquisadora admirável, e um ser humano incrível. Agradeço à vida pelo privilégio de tê-la conhecido, e por ter ampliado meus horizontes para a Biblioteconomia!

Aos meus **pais**, Silvio Gonçalves e Lucimar Silva, por sempre me apoiarem com amor e dedicação. Gratidão por tê-los como pais, admiração pelas pessoas que são e por me presentear com ótimas lembranças. O amor de vocês pelo conhecimento e estudo me fizeram lutar e acreditar que os sonhos podem ser alcançados.

Ao meu amor **Enoque Oliveira** pelo companheirismo de longa data, por ser meu porto seguro e meu melhor amigo. Obrigada por entender e me apoiar em diversos momentos de ansiedade e ausência durante essa trajetória.

Ao Davi meu **filho**, minha motivação para ser melhor todos os dias, você é minha pedra preciosa o presente mais lindo que a vida e Deus poderiam ter me dado, tudo foi intensificado com sua chegada, principalmente meus sorrisos.

À minha **irmã** Yara Silva que desde crianças nos apoiávamos e isso não mudou depois da vida adulta. Fico orgulhosa em ver a mulher que se tornou, sonhadora e determinada, será uma engenheira brilhante!

Aos **amigos de longa data**, que se fazem presentes em minha vida e estão sempre na torcida pelas minhas conquistas: Amanda do Carmo, Amanda Reis, Denison Leone, Gabrielle Vianna, Gustavo Senna, Izabella Teixeira, Lorrane Teixeira, Tatiane Melo, Vinicius Alves e Williana Garcia. Obrigada pela amizade e pelos momentos de descontração.

Aos meus familiares, meus avós paternos, Francisco e Francisca, e avós maternos Maria e José, que me ensinaram o valor da família.

Aos **professores** da Escola Cidade dos Meninos, por encorajarem seus alunos a avançarem nos estudos.

Aos meus amigos do **CAED**, que me encorajaram nesse processo.

À Graça Moreira, por ser sempre disponível em me ajudar. E por suas valiosas contribuições no meu trabalho.

À Viviane Barrozo, por sua positividade e simpatia em momentos difíceis, pelas palavras de encorajamento e companheirismo.

Ao Thiago Belchior, por vivenciarmos juntos o processo de mestrado e “cantarolar” a vida comigo.

Ao Leonardo Nunes, pelas longas conversas e amizade, e, principalmente, pela preocupação e apoio constantes.

Aos membros do grupo de pesquisa **MHTX**, Celsiane Araujo, Junio Lopes, Rainer Finelli, Tereza Cristina e Webert Araujo, pelas discussões sempre tão produtivas e por reforçarem

que a pesquisa não precisa ser uma tarefa solitária. Em especial à Patrícia Lopes, pela leitura e auxílio na redação do trabalho.

A **Graciane Bruzinga**, amiga que me estendeu a mão em todos os momentos, me presenteando com palavras de esperança e passou comigo durante todo o percurso desde o egresso ao mestrado até a finalização, sua amizade é um presente.

A **Letícia Miranda**, que me auxiliou nesta dissertação com seu trabalho de excelência em diversos momentos.

Aos **bibliotecários** que se dispuseram a me ensinar durante minha graduação, não sabem o quanto fazem diferença em “alunos sonhadores”, em especial Irénquer Cruz, Madalena Rocha, Célio Diniz.

Aos meus **queridos amigos que a “biblio”** me apresentou Juliana Rodrigues, Marcus Moreira e Sara Isaqueline.

Aos **membros da banca** de qualificação e defesa, Profa. Benildes Maculan, Profa. Celia Dias, Profa. Cíntia de Azevedo e Profa. Elisangela Aganette, que tão gentilmente aceitaram participar e colaborar com esta dissertação.

Ao **Programa de Pós graduação** em Gestão e Organização do Conhecimento (PPGGOC) pela oportunidade de realização deste trabalho.

Aos **colegas da Pós**, Grazielle Ferreira, Fernanda Matos e Rosy Oliveira pelos trabalhos e disciplinas realizados em conjunto.

Aos **professores** da Escola de Ciência da Informação, pela dedicação, apoio e todo conhecimento compartilhado.

Aos **funcionários da ECI**, em especial Gildenara, Gisele, Lucimary, Vinícius e Claudia pela competência e disponibilidade.

Ao **Ernane**, pelas correções e melhorias neste texto.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**) pela concessão da bolsa de estudos no primeiro ano do curso de mestrado.

"Aqui se encontra o perigo das escolas: de tanto ensinarem o que o passado legou - e ensinarem bem - fazem os alunos se esquecerem que o seu destino não é o passado cristalizado em saber, mas um futuro que se abre como vazio, um não saber que somente pode ser explorado com as asas do pensamento. Compreende-se então que Barthes tenha dito que, seguindo-se o tempo em que se ensina o que se sabe, deve chegar o tempo quando se ensina o que não se sabe" (Rubem Alves – Ao professor com meu carinho).

RESUMO

O Índice de Final de Livro é um dos instrumentos mais antigos utilizados para Organização e Recuperação da Informação, e a sua importância decorre do fato de ser este um dos principais pontos de acesso ao conteúdo de um livro, auxiliando o leitor na compreensão dos conceitos relevantes contidos nos manuscritos. Porém, o processo de elaboração do Índice de Final de Livro no Brasil acontece, geralmente, de maneira manual, o que demanda tempo e custos financeiros, justificando o baixo número de publicações que contêm o índice. Tendo em vista esse contexto, o objetivo desta pesquisa foi propor diretrizes para elaboração e estruturação de Índice de Final de Livro por meio da indexação semiautomática com base nas normas (ISO/AWI 999 e NBR 6034) e na literatura sobre essa temática. Para tanto, pretendeu-se: a) identificar na literatura os programas de construção de IFL automático e semiautomático, caracterizando-os, com o intuito de auxiliar na avaliação dessas ferramentas; b) sintetizar as principais recomendações das normas (NBR e ISO) e da literatura para identificar os subsídios e auxiliar na elaboração e estruturação dos Índices de Final de Livro; c) contribuir para a divulgação no Índice de Final de Livro como importante recurso de acesso à informação, a partir da análise e interpretação da literatura escolhida. Como fundamento teórico, são apresentadas abordagens na área da Biblioteconomia e Ciência da Informação, traçando temáticas sobre indexação manual, semiautomática e automática; e o relato sobre o Índice e finaliza apresentando o objeto de estudo desta dissertação o Índice de Final de Livro. Utiliza-se a base metodológica proposta por Bardin (2016) sobre Análise de conteúdo aplicada nas publicações selecionadas. A partir dos resultados da Revisão de literatura realizou-se análise de 16 publicações para elaboração de Índice de Final de Livro, resultando em duas propostas: (1) comparação entre os *softwares* que elaboram Índice de Final de Livro e (2) diretrizes para elaboração e estruturação de Índice de Final de Livro. Na primeira proposta concluiu-se que o *Topic Curation Toolkit* (TCT) atendeu a maioria dos oito critérios para avaliação de programas automáticos de indexação. Na segunda proposta elaborou-se a análise de acordo com a estrutura do Índice de Final de Livro (entrada, cabeçalho, subcabeçalho, localizadores e remissivas), obtendo-se dois resultados: O primeiro é uma figura que apresenta as diretrizes para elaboração do Índice de Final de Livro, e o segundo são as ilustrações (quadros) que indicam a estruturação necessária para o Índice de Final de Livro. Conclui-se, a partir dos resultados obtidos que, por meio da integração de princípios teóricos e metodológicos da Biblioteconomia e Ciência da Informação foi possível apresentar as diretrizes para elaboração e estruturação de Índice de Final de Livro, como também propiciar um panorama comparativo que pode servir de apoio na definição de qual *software* utilizar para construção do Índice de Final de Livro.

Palavras-chave: Índice de Final de Livros. Indexação semiautomática. Elaboração de Índice. Recuperação da Informação.

ABSTRACT

The Back-of-book indexes is one of the oldest instruments used for Organization and Information Retrieval, and its importance stems from the fact that this is one of the main points of access to the content of a book, helping the reader to understand the relevant concepts contained in the manuscripts. However, the process of elaborating the Back-of-book indexes in Brazil usually takes place manually, which requires time and financial costs, justifying the low number of publications that contain the index. In view of this context, the objective of this research was to propose guidelines for the elaboration and structuring of the Back-of-book indexes by means of semi-automatic indexing based on the standards (ISO / AWI 999 and NBR 6034) and in the literature on this theme. To this end, it was intended: a) to identify in the literature the automatic and semi-automatic IFL construction programs, characterizing them, in order to assist in the evaluation of these tools; b) synthesize the main recommendations of the standards (NBR and ISO) and the literature to identify the subsidies and assist in the elaboration and structuring of the Back-of-book indexes; c) contribute to the dissemination in the Back-of-book indexes as an important resource for accessing information, based on the analysis and interpretation of the chosen literature. As a theoretical foundation, approaches are presented in the area of Librarianship and Information Science, tracing themes about manual, semi-automatic and automatic indexing; and the report on the Index and concludes by presenting the object of study of this dissertation the Back-of-book indexes. The methodological basis proposed by Bardin (2016) on Content Analysis applied in the selected publications is used. Based on the results of the Literature Review, 16 publications were analyzed for the elaboration of the Back-of-book indexes, resulting in two proposals: (1) comparison between the software that produce the Back-of-book indexes and (2) guidelines for elaboration and structuring the Back-of-book indexes. In the first proposal, it was concluded that the Topic Curation Toolkit (TCT) met most of the eight criteria for evaluating automatic indexing programs. In the second proposal, the analysis was elaborated according to the structure of the Back-of-book indexes (entry, header, sub-header, locators and cross-references), obtaining two results: The first is a figure that presents the guidelines for the elaboration of the Index at the end of the book, and the second is the illustrations (tables) that indicate the necessary structure for the index at the end of the book. It is concluded, from the results obtained that, through the integration of theoretical and methodological principles of Librarianship and Information Science, it was possible to present the guidelines for the elaboration and structuring of the Book End Index, as well as providing a comparative panorama that can serve as support in defining which software to use to build the Back-of-book indexes.

Keywords: Back-of-book indexes. Semi-automatic indexing. Elaboration of Index. Information Retrieval.

LISTA DE FIGURAS

Figura 1 - Sistematização de conteúdos em organização e representação do conhecimento	39
Figura 2 - Fatores que influenciam a qualidade da indexação	40
Figura 3 - Exemplo de índice de citação	52
Figura 4 - Exemplo de índice KWIC	54
Figura 5 - Exemplo de índice KWOC	54
Figura 6 - Exemplo de índice KWAC.....	55
Figura 7 - Estrutura do Índice de Final de Livro.....	58
Figura 8 - Exemplo de remissivas no IFL	59
Figura 9 - Modelo contemporâneo da compreensão na leitura	65
Figura 10 - Leitores do Índice de Final de Livro	76
Figura 11 - Organização da análise: fases e atividades	80
Figura 12 - Procedimentos da metodologia: pré análise e exploração do material.....	80
Figura 13 - Publicações selecionadas para análise de conteúdo	88
Figura 14 – Esquema dos resultados da dissertação.....	110
Figura 15 – Elementos semânticos relacionados à elaboração de IFL.....	127
Figura 16 - Esquema para elaboração do índice de final de livro	136
Figura 17 - Diretrizes para estruturação do IFL no todo	148
Figura 18 - Exemplo de estrutura para o IFL.....	149
Figura 19 - Síntese dos elementos a serem considerados na elaboração de uma política de indexação	165
Figura 20 - Sugestão de estrutura para a política de indexação em Índice de Final de Livro	169

LISTA DE QUADROS

Quadro 1 - Controle de vocabulário	24
Quadro 2 - Interconexão entre concepções de análise de assunto, tipos de informação de documentos e métodos de indexação.....	63
Quadro 3 - Elementos para a leitura de livros para elaborar IFL.....	67
Quadro 4 - Responsabilidades éticas indicadas pela ASI para elaboração de IFL.....	72
Quadro 5 - Critérios para escolha das publicações.....	82
Quadro 6 - Regras nacionais e internacionais (objetivo e vigência).....	83
Quadro 7 - Artigos da revisão de literatura	85
Quadro 8 - Publicações analisadas.....	91
Quadro 9 - Critérios utilizados para avaliação do software	115
Quadro 10 - Comparativo dos softwares.....	117
Quadro 11 - Elementos semânticos relacionados à elaboração de IFL.....	124
Quadro 12 – Elementos semânticos relacionados à estruturação de IFL: partes do IFL	130
Quadro 13 – Elementos semânticos relacionados à estruturação de IFL.....	131
Quadro 14 - Formulário de auxílio para o estágio 1 “compreender do conteúdo do livro” ..	146
Quadro 15 - Detalhamento dos elementos para política de indexação	166

LISTA DE TABELAS

Tabela 1 - Processo de busca e resultado da revisão de literatura	25
Tabela 2 - Tipo de publicação	26
Tabela 3 - Quantidade de publicações por periódico	26
Tabela 4 - Vínculo dos autores	28
Tabela 5 - Quantitativo da origem dos autores.....	29
Tabela 6 – Abordagens de estudo	31
Tabela 7 - Livros selecionados	87
Tabela 8 - Apresentação dos elementos semânticos	111
Tabela 9 - Elementos semânticos identificados e selecionados	122
Tabela 10 - Elementos semânticos identificados e selecionados	129

LISTA DE GRÁFICOS

Gráfico 1 - Quantidade de publicações por ano	27
Gráfico 2 - Profissão dos autores.....	30

SUMÁRIO

1. INTRODUÇÃO	17
1.1 Justificativa e problema.....	18
1.2 Objetivos	21
1.2.1 Objetivo geral.....	21
1.2.2 Objetivos específicos	21
1.3 Estrutura da dissertação	22
2. REVISÃO DE LITERATURA	23
2.1 Revisão sistemática da literatura: procedimentos adotados.....	23
2.2 Dados quantitativos	25
2.3 Análise dos trabalhos correlatos	31
3. FUNDAMENTAÇÃO TEÓRICA	38
3.1 Indexação manual.....	38
3.1.1 Características essenciais para garantir a qualidade na indexação	39
3.2 Indexação automática e semiautomática	41
3.2.1 Tipos de indexação semiautomática e automática	43
3.2.2 Vantagens e desvantagens da indexação automática.....	44
3.3 Índice	46
3.3.1 A evolução do índice na Biblioteconomia	47
3.3.1.1 Etimologia da palavra “índice”	48
3.3.1.2 Percurso histórico	48
3.3.2 Tipos de enfoque e finalidade	51
3.4 Índice de Final de Livro.....	56
3.4.1 Estrutura do IFL	58
3.4.2 Critérios para IFL de qualidade	59
3.4.3 A análise de assunto e o controle terminológico no IFL	62
3.4.3.1 A análise de assunto.....	62
3.4.3.2 Controle terminológico	70
3.4.4 A tríade de profissionais para elaborar o índice: indexador, autor e editor	71
3.4.5 Grupos de discussão e o futuro do profissional.....	73
3.4.6 O leitor do IFL	75
4. CARACTERIZAÇÃO E DESCRIÇÃO DA METODOLOGIA	78
4.1 Caracterização da pesquisa.....	78
4.2 Delimitação do universo de pesquisa.....	78
4.3 Descrição dos procedimentos metodológicos	79
4.3.1 Pré-análise.....	81
4.3.2 Exploração do material	89
4.3.2.1 Exploração do material para Seção 1	89
4.3.2.2 Exploração do material para Seção 2	104
5. RESULTADOS	110
5.1 Resultados da Seção 1	111
5.1.1 Lista dos elementos semânticos sobre software para construção de Índice de Final de Livro.....	111
5.1.2 Tratamento e interpretação dos elementos semânticos	113
5.2 Resultados da Seção 2	121
5.2.1 Resultado da seção 2.1	121
5.2.1.1 Lista dos elementos semânticos sobre elaboração de IFL	121
5.2.1.2 Tratamento e interpretação dos elementos semânticos	123
5.2.2 Resultado da seção 2.2	128

5.2.2.1 Lista dos elementos semânticos sobre estruturação de IFL	129
5.2.2.2 Tratamento e interpretação dos elementos semânticos	130
6. DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL DE LIVRO	134
6.1 Elaboração do Índice de Final de Livro	134
6.2 Estruturação do Índice de Final de Livro	147
7. CONSIDERAÇÕES FINAIS.....	150
7.1 Limitações da pesquisa.....	153
7.2 Trabalhos futuros	154
REFERÊNCIAS	156
APÊNDICE A - SOFTWARES DE INDEXAÇÃO PARA ÍNDICE DE FINAL DE LIVRO	163
APÊNDICE B – ORIENTAÇÃO E SUGESTÃO DE ESTRUTURA PARA POLITICA DE INDEXAÇÃO EM ÍNDICE DE FINAL DE LIVRO	165

1. INTRODUÇÃO

Livro algum (na categoria de literatura geral) é tão bom que não possa tornar-se melhor graças a um índice, e livro algum é tão ruim que não possa, graças a esse acessório, fugir à pior das condenações. (Sir Edward Cook).

Neste capítulo, apresentam-se o contexto e a proposta da pesquisa, bem como as justificativas para seu desenvolvimento, o problema investigativo e os objetivos norteadores.

A crescente importância da informação tem estimulado a busca por ferramentas que permitam localizar essa informação de forma rápida e precisa. Uma dessas ferramentas é o índice, cuja finalidade é referenciar algo. Este instrumento, um dos mais antigos dispositivos de acesso à informação, tem ampla aplicação, compreendendo desde os tradicionais Índice de Final de Livro (IFL) aos índices de base de dados. A aplicação de ambos tem essência fundamentada na Organização e Representação da informação, visto que acontece a partir da descrição abreviada do texto de determinado documento e retrata sua estrutura.

O Índice é um produto da indexação, que identifica todas as palavras e/ou frases que se referem a informações que possam ser procuradas pelo leitor. A sua principal contribuição é ser um ponto de acesso ao conteúdo do livro, como mecanismo prático para Recuperação da Informação. A ferramenta proporciona, ainda, pesquisas mais efetivas em menos tempo e auxilia o usuário na compreensão de conceitos relevantes por meio da escolha e ordenação dos termos e a utilização de remissivas, já que nem sempre o título e o resumo contêm informações suficientes.

A origem do Índice é muito antiga e sua primeira menção data do séc. XXXI a.C. em Roma, quando Cícero enviou uma carta a Atícus solicitando a duas pessoas a organização de seus livros e a elaboração de Índices, ou seja, uma lista das obras que possuía. Antes da invenção da imprensa, os índices eram os registros dos títulos dos livros pertencentes a um acervo¹. A sua criação está associada à preocupação do homem em tornar mais acessível a informação registrada em documentos, o que conduziu à busca por formas de ordená-la. A partir desse momento, o Índice se transformou em objeto de estudo, e sua utilização se difundiu em nível internacional. O conceito foi formalmente introduzido no século V com a obra anônima “*Apothegmata*”, uma lista de provérbios gregos sobre tópicos teológicos (GOMES; GUSMÃO, 1983).

Após a invenção da imprensa, surgiu a necessidade de as publicações possuírem um tipo de índice que representasse o conteúdo principal dos livros, remetendo o

¹ O surgimento do Índice é reconhecido por ser a primeira tentativa de Controle Bibliográfico, conforme apresentado na seção “3.3.1.2 Percurso histórico”.

leitor para os conceitos mais relevantes tratados no documento. Esse é o contexto que motiva a criação do chamado “Índice de Final de Livro” (IFL). O produto IFL é solicitado e gerenciado, geralmente, pelo mercado editorial, situação em que se recomenda a atuação de uma equipe de profissionais, composta por: editor, autor e bibliotecário indexador. Apesar de sua indiscutível relevância, a temática relativa ao índice é pouco estudada no Brasil e, particularmente, a sua elaboração com o auxílio de tecnologias. Baseado na literatura da área, este trabalho tem como finalidade analisar as diretrizes para elaboração e estruturação de IFL, contribuir para o aumento do número de publicações que contenham índice e, auxiliar autores e profissionais de biblioteconomia na elaboração de um IFL, promovendo, dessa maneira, a divulgação de sua importância. Os problemas inicialmente apresentados, bem como a justificativa para o desenvolvimento desta pesquisa serão expostos na próxima seção.

1.1 Justificativa e problema

Desde o ano de 2015, a proponente vem ministrando, como monitora, palestras na disciplina Análise de Assunto para os estudantes do 2º período do Curso de Graduação em Biblioteconomia da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Tal experiência possibilitou refletir sobre a relevância da utilização do Índice no meio acadêmico e literário. Um dos pontos de reflexão está relacionado à sua necessidade e utilidade na era virtual, pois, segundo os discentes, com o uso da tecla de busca na caixa de texto, qualquer termo seria encontrado com facilidade. Somente após explanações relativas à aplicabilidade do Índice é que esses alunos se conscientizaram de sua importância. Assim, percebeu-se que a falta de divulgação sobre a pertinência do Índice consiste em um tópico para investigação, bem como sua construção de forma manual, pois demanda tempo e conhecimentos específicos. Essas e outras dificuldades serão tratadas, com mais vagar, a seguir.

A construção do Índice exige esforços linguísticos, cognitivos e técnicos, além do domínio de normas e critérios para criar Índices de qualidade, o que é um desafio para os profissionais da área, dado que consiste numa atividade caracterizada por um alto grau de dificuldade para os bibliotecários indexadores principiantes ou com pouco conhecimento sobre essa temática tão específica. Por se tratar de um campo com várias vertentes de atuação, há casos em que até mesmo profissionais mais experientes encontram dificuldades para elaborar esse instrumento. Por isso, a criação do índice é uma atividade que exige o estudo aprofundado das técnicas de indexação, que as rotinas de trabalho nem sempre proporcionam contato com as regras para confecção de índices, o que resulta em

adversidades associadas ao desconhecimento ou à própria complexidade inerente às normas de indexação.

Publicações são produzidas diariamente, e o quantitativo de profissionais que elaboram o IFL não acompanha o alto índice de editorações, ou seja, o número daqueles que trabalham com essa especificidade ainda não é o desejável. Wu et al (2013) argumentam que o IFL deveria ser uma obrigação dos autores, no entanto, comumente, essa tarefa é realizada por *freelancers* (profissionais autônomos), contratados pelos próprios autores, editores ou empresas independentes que gerenciam a produção de um livro. Essa situação desestimula o aumento do número de profissionais dedicados à elaboração desse recurso, pois os autores não assumem a tarefa e existe um número limitado de bibliotecários que se especializam nesta área.

Outro problema em relação à criação do índice reside no desconhecimento de sua importância por parte dos autores e editores, embora tenha sido afirmado por Nadolski e Miltschin (1979 apud Schmidmaier, 2006) que 90% dos leitores consideram que o índice seja parte necessária de um livro. Presume-se que os autores estão concentrados em redigir o conteúdo, e os editores, por serem mais pragmáticos, se empenham em evitar custos elevados na produção, não considerando que a qualidade do índice atrairá mais leitores (DARTNALL, 2008).

Todo esse cenário justifica o número baixo de publicações que contêm índice, afirmação validada mediante estudo realizado pelos autores Vahed et al (2013), os quais o levantamento dos Índices contidos na coleção em que os autores realizaram pesquisas e constataram que apenas 24,34% dos livros na área de microbiologia possuíam IFL. Além do reduzido número, os índices que existiam eram de baixa qualidade. Uma pesquisa realizada por Schmidmaier (2006) analisou os índices existentes nos livros e constatou que era difícil encontrar índices que possuíssem apresentação tipograficamente satisfatória. De acordo com o autor, os melhores índices encontrados estavam inclusos em livros sobre ciência, tecnologia, medicina e em obras sobre biblioteconomia. Assim, índices inadequados ou a falta de qualquer índice é uma realidade frequentemente encontrada no meio acadêmico.

Na revisão de literatura, observou-se que estudos atuais sobre o IFL em publicações nacionais são quase inexistentes e que as investigações internacionais estão em andamento. Assim, as autoras Csomai e Mihalcea (2007) indicam a falta de estudos sobre a elaboração automática de IFL nas áreas de Recuperação da Informação e Processamento de linguagem natural. A motivação para suscitar mais estudos sobre o tema deriva do crescente número das publicações em formato eletrônico e aos avanços na extração automática de termos, revelando a necessidade de estudos acerca da indexação de livros que considere os benefícios das técnicas de extração de termos no contexto eletrônico.

Pode-se afirmar, por meio dos resultados da Revisão de literatura que as publicações recuperadas relatam sobre práticas para elaborar e estruturar IFL. Porém, este conteúdo é disperso no documento o que dificulta sua localização, e é também muito textual, não sendo esquemático para aplicação, dessa forma, a falta de compilação das diretrizes para a construção de IFL e a ausência deste conteúdo em literatura nacional é um dos grandes problemas nessa área de estudo.

A tecnologia não alcançou o nível semântico que contemple todos os benefícios da indexação quando realizada de forma manual. Por isso, será vantajoso repensar a tecnologia, de modo a aprimorar os benefícios advindos da indexação manual já consolidada. Vale ressaltar que o processo de indexação manual não acompanhou o aumento do número de publicações e, desta forma, aliar a rapidez do processamento automático com o nível semântico proporcionado pela indexação manual é fundamental.

Conforme relatado por Fujita (1999) e Pinto (2001) e considerando a velocidade com que as informações são produzidas e publicadas, a indexação manual evidencia algumas dificuldades, quais sejam:

- 1) demanda tempo para a leitura cuidadosa e integral do livro, podendo ocorrer repetidas vezes;
- 2) exige pessoal especializado com conhecimento e domínio das técnicas de indexação, que geralmente faz parte da formação do profissional de biblioteconomia;
- 3) subjetividade na seleção dos conceitos que melhor representem o conteúdo do documento;
- 4) ocorrência de desacordos frequentes entre os indexadores possivelmente em razão da subjetividade inerente ao processo manual;
- 5) requer o domínio da temática quando da escolha e catalogação dos termos.

Todos estes indicativos resultam em altos custos de produção do índice e tempo insuficiente para atender todas as demandas podendo até inviabilizar a prática da indexação manual.

Foi identificado na revisão de literatura (capítulo 2) que os *softwares* existentes ainda não contemplam os requisitos necessários para criação de índices consistentes quando aplicada a forma totalmente automática. Os trabalhos encontrados retrataram que os *softwares* automáticos indicaram que o índice produzido não fazia relação entre os termos (utilização de remissivas) e evidenciaram problemas de contextualização como controle de sinônimos e recuperação de termos irrelevantes.

Além desses problemas, podem ser identificados outros relacionados à indexação automática, quais sejam: (1) funciona separadamente em documentos; (2) não correlaciona textos entre si ou um texto com uma visão de mundo (desconsidera o contexto); (3) vincula-se ao vocabulário e à gramática utilizados no documento indexado; (4)

não consegue tratar dados gráficos, nota de rodapé e informações contidas em tabelas; (5) não consegue trabalhar com vocabulário estrangeiro; (6) não consegue avaliar textos; (7) não consegue criar relações intertextuais; (8) não é capaz de imitar o questionamento, a resposta humana a um texto, o que acrescenta valor à indexação; (9) não traz semântica; (10) as ocorrências de termos "importantes" para referenciar no índice não ficam evidentes (SYLVA, 2002; SILVA; FUJITA, 2004). Tendo em vista todos os problemas e desafios apresentados, pode-se inferir que a elaboração do Índice baseada na indexação semiautomática é uma prática recomendada. Nesse sentido, tem-se o seguinte problema de pesquisa: **“quais são as diretrizes necessárias para a elaboração e estruturação semiautomática do Índice de final de livro?”**.

Diante das justificativas e do relato dos problemas delineados, apresentam-se os pressupostos deste estudo: a) o conhecimento sobre o processo de indexação auxilia na elaboração de índice do tipo IFL e b) a indexação semiautomática é mais eficiente na elaboração de índice porque os índices elaborados manualmente geram custos altos para a produção, assim como há perdas de semântica e de contexto em índices elaborados de forma automática.

Os pressupostos traçados direcionaram o estudo, mas não limitando a execução da pesquisa, pois a partir dessas considerações foram definidos os objetivos do estudo e serão visualizados na seção seguinte.

1.2 Objetivos

A partir do contexto e problemas relatados, são apresentados nesta subseção os objetivos desta pesquisa.

1.2.1 Objetivo geral

Propor diretrizes para elaboração e estruturação de Índice de Final de Livro por meio da indexação semiautomática com base nas normas (ISO/AWI 999 e NBR 6034) e na literatura sobre essa temática.

1.2.2 Objetivos específicos

Para atender ao objetivo geral foram delineados os seguintes objetivos específicos:

a) **identificar** na literatura os programas de construção de IFL automático e semiautomático, caracterizando-os, com o intuito de auxiliar na avaliação dessas ferramentas.

- b) **sintetizar** as principais recomendações das normas (NBR e ISO) e da literatura para identificar os subsídios e auxiliar na elaboração e estruturação dos Índices de Final de Livro;
- c) **contribuir** para a divulgação no Índice de Final de Livro como importante recurso de acesso à informação, a partir da análise e interpretação da literatura escolhida.

1.3 Estrutura da dissertação

Visando operacionalizar a investigação proposta, essa dissertação está estruturada em capítulos sendo o primeiro a *Introdução* com a discussão e contextualização do tema, problema e justificativa, bem como definição dos objetivos geral e específicos.

O segundo capítulo nomeado *Revisão de literatura*, traz reflexões teóricas sobre as publicações selecionadas nas bases de dados que tem como temática o IFL. Desta forma, apresenta em forma linear as principais ideias dos autores.

O terceiro capítulo, intitulado *Fundamentação teórica* apresenta o contexto e a prática da indexação, discutindo conceitos e aplicabilidade no IFL. Discorre ainda, sobre a indexação semiautomática e automática.

O quarto capítulo *Caracterização e descrição da metodologia* identifica a limitação e recorte utilizado para a pesquisa, bem como, a descrição das etapas metodológicas.

O quinto capítulo identificado como *Resultados*, apresenta os produtos obtidos a partir da aplicação da metodologia.

O sexto capítulo, *Diretrizes para elaboração e estruturação de IFL*, apresenta de forma estruturada os resultados obtidos no capítulo anterior, e posteriormente, a análise dos resultados.

O último capítulo, *Considerações finais*, resume os principais pontos discutidos de cada capítulo, apresenta também as limitações que foram identificadas durante a realização dos procedimentos metodológicos. E no final, discute quais lacunas precisam de destaque nessa temática de pesquisa, identificando assim os trabalhos futuros.

Além dos capítulos supramencionados, são parte integrante desta dissertação as referências e apêndices.

2. REVISÃO DE LITERATURA

“O valor de qualquer índice está em tornar rápida a consulta” (University of Chicago Press: Manual of Style)

Nesta seção são apresentados os resultados da revisão de literatura, contemplando um resumo dos procedimentos utilizados para esta etapa da pesquisa, os dados quantitativos gerados e também a análise resultante do processo de revisão.

2.1 Revisão sistemática da literatura: procedimentos adotados

A revisão sistemática da literatura é um meio de identificar, avaliar e interpretar toda a pesquisa disponível relevante para uma questão de pesquisa específica, área temática ou fenômeno de interesse. A revisão sistemática é utilizada, neste trabalho, por ter uma estratégia de pesquisa pré-definida, o que permite a integridade do estudo, em razão da diversidade de publicações que trazem argumentos condizentes e divergentes com as hipóteses do pesquisador, ampliando a percepção da pesquisa. Para este estudo, adaptou-se² as orientações da publicação de Kitchenham (2004), que foram seguidas às etapas de: 1) definição da questão de pesquisa; 2) seleção dos termos (expressão de busca) de busca; 3) identificação das bases para o levantamento bibliográfico; 4) levantamento bibliográfico e; 5) leitura exploratória para seleção dos documentos.

Para operacionalizar todo o estudo, realizou-se pesquisa bibliográfica até o mês de janeiro de 2020. Confirmou-se, por meio do teste de busca que a recuperação de publicações na temática de Índice é pouco frequente. Nesse sentido, para abranger e recuperar o número máximo de produções acadêmicas, procedeu-se à pesquisa exaustiva, ou seja, sem que fossem adotados quaisquer tipos de filtro de pesquisa (tempo, idioma, entre outros).

Na primeira etapa, definiu-se a **questão da pesquisa** com a formulação da seguinte pergunta: "quais são as principais abordagens de estudos existentes na área de IFL?". Em seguida, procedeu-se à segunda etapa, em que foi realizada a **seleção dos termos** para elaboração da estratégia de busca. Para tal, foi utilizado o vocabulário controlado do Tesouro Brasileiro de Ciência da Informação, elaborado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) para os termos em português e em inglês (PINHEIRO; FERREZ, 2014). Para o idioma em espanhol, foi consultado o *Tesouro de Biblioteconomía y Documentación*, desenvolvido pelo *Centro de Información y Documentación Científica* (MOCHÓN BEZARES; SORLI ROJO, 2002), conforme apresentado no Quadro 1.

² O foco da Revisão sistemática da literatura são as áreas da saúde e da engenharia.

Quadro 1 - Controle de Vocabulário

IDIOMA	LIGUAGEM NATURAL	VOCABULÁRIO CONTROLADO
Português	Índice controlado de livro Índice interno Índice de final de livro Índice alfabético-remissivos Índices impressos	Índices de livros
Espanhol	Índice controlado de libro Índice final de libro	Índices alfabéticos
Inglês	<i>Back-of-book indexes</i>	<i>Book indexes</i>

Fonte: elaborado pela autora.

Após o exposto, delineou-se a expressão de busca: “*Índices de livros*” OR “Índices alfabéticos” OR “*book indexes*”. Definidos os termos, o passo seguinte foi a escolha das bases de dados a serem consultadas.

Na terceira etapa, **identificação das bases para o levantamento bibliográfico**, utilizou-se o portal CAPES para pesquisar as bases de dados, nas quais foram selecionadas as principais bases interdisciplinares e específicas da BCI. As bases selecionadas foram: *Library and Information Science Abstracts* (LISA); *Information Science & Technology Abstracts* (ISTA); *Web of Science* (WoS); *Scientific Electronic Library Online* (SciELO); *SCOPUS* e *Emerald e Journals Premier* (Emerald). A expressão de busca mencionada anteriormente foi idêntica para todas as bases de dados. A partir desse levantamento, obteve-se o total de 415 publicações³.

A quarta etapa, **levantamento bibliográfico**, e a quinta etapa, **leitura exploratória para seleção dos documentos**, estão apresentadas de forma simultânea, sendo que a primeira coluna, “publicações recuperadas”, é resultado da etapa 4, e as três outras colunas retratam a análise da leitura exploratória, que são o critério de exclusão, das 415 publicações recuperadas foram realizados três critérios de exclusões, que são: 1) “não atende aos objetivos” se refere a publicações que relatavam sobre o índice, mas que abordavam temáticas sobre classificação, índice no sentido numérico, e não o IFL; a seleção para esse critério foi realizada através da leitura dos títulos e resumos; 2) “publicações repetidas”, que são publicações idênticas encontradas na mesma base, ou em bases diferentes; e 3) “texto completo não disponível” são arquivos cujo *download* do texto completo não foi possível por apresentar somente o resumo ou ser um artigo pago, conforme demonstrado na Tabela 1.

³ Conforme segunda coluna da Tabela 1.

Tabela 1 - Processo de busca e resultado da revisão de literatura

NOME DA BASE	PUBLICAÇÕES RECUPERADAS	NÃO ATENDE AOS OBJETIVOS	PUBLICAÇÕES REPETIDAS	TEXTO COMPLETO NÃO DISPONÍVEL	TOTAL
LISA	144	95	11	23	15
ISTA	101	64	7	16	14
WoS	27	13	1	8	4
SciELO	37	37	0	0	0
SCOPUS	119	82	6	13	18
Emerald	3	3	0	0	0
TOTAL	415	289	42	55	51

Fonte: adaptado de Pizzol (2014, p. 37) e Kitchenham (2004).

Observa-se, pelos resultados, que a pesquisa na base LISA obteve um maior número de publicações recuperadas, seguida da SCOPUS e ISTA. No entanto, na seleção final das publicações, o número de documentos recuperados na base SCOPUS foram mais relevantes para a pesquisa, mesmo sendo uma base interdisciplinar. Após uma leitura exploratória e a aplicação dos critérios de exclusão explicados anteriormente, foram selecionados 51 documentos que são analisados nos resultados a seguir.

Os resultados são apresentados em duas seções: na primeira mostra-se os dados quantitativos que consistem na identificação do número de publicações recuperadas, em quais periódicos, a incidência de publicações por ano e as principais abordagens de estudo nelas identificadas. A segunda seção traz a análise dos trabalhos correlatos, examinam-se os textos que abordam o tema “elaboração do índice”, e são relatados os pontos principais dos artigos.

2.2 Dados quantitativos

Após a seleção foram escolhidas 51 publicações. O maior número de publicações são artigos de divulgação em periódicos científicos, somando um total de 40 artigos. Foram recuperados 6 anais de evento e os demais documentos não continham informações sobre sua tipologia documental, conforme apresentado na Tabela 2. Considera-se que recuperar mais artigos de publicação periódica justifica-se pela necessidade de circulação rápida das informações e por conter temas inovadores.

Tabela 2 - Tipo de publicação

Tipo de publicação	Quantidade
Artigo em publicação periódica	40
Não identificado	5
Anais de evento	6
TOTAL	51

Fonte: elaborado pela autora.

Em relação aos artigos de periódicos, a fonte que mais se destacou foi o *The Indexer*, representando 42% dos itens recuperados. O *The Indexer* foi publicado pela primeira vez em 1958, e é um periódico muito importante na área de Índice, sendo publicado trimestralmente pela *Society of Indexers*, no qual se reconhece sua representação mundial no campo da indexação (THE INDEXER, 2019). Outros dois periódicos que também teve uma representatividade expressiva foram os *Journal of the American Society for Information Science* e *Technicalities*, com o total de 7% da amostra cada. Em algumas publicações não foi possível identificar o nome do periódico, embora apresentassem o número do volume, número e fascículo, o que indica se tratar de um periódico. A identificação dos periódicos está expressa na Tabela 3.

Tabela 3 - Quantidade de publicações por periódico

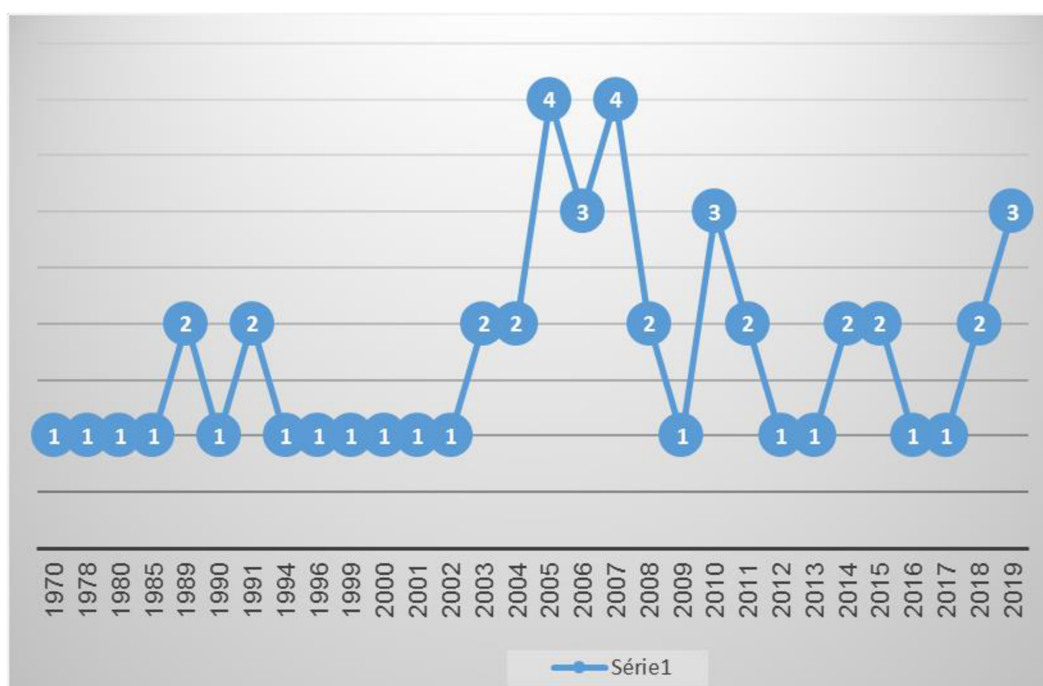
Nome do periódico	Quantidade de publicações
<i>The indexer</i>	19
<i>Não identificado</i>	5
<i>Technicalities</i>	3
<i>Journal of the American Society for Information Science</i>	3
<i>Against the Grain</i>	1
<i>Aslib Proceedings</i>	1
<i>Documentation et bibliothèques</i>	1
<i>Information Today</i>	1
<i>Information Wissenschaft & Praxis</i>	1
<i>Journal of Documentation</i>	1
<i>Journal os Scholarly Publishing</i>	1
<i>Key Words</i>	1
<i>Late Breaking Results Paper</i>	1
<i>Learned Publishing</i>	1
<i>Library Review</i>	1
<i>New Library World</i>	1
<i>Proceedings of ACL-08: HLT</i>	1
<i>Software - practice and experience</i>	1
<i>Technical Communication</i>	1
TOTAL	45*

Fonte: elaborado pela autora.

* A amostra total é 51, mas são apresentadas nessa tabela somente as publicações de periodicos que totalizam 45.

As publicações sobre Índice são esporádicas e ocorrem no máximo três vezes por ano, com exceção dos anos de 2005 e 2007, quando houve mais publicações. Salienta-se que a partir de 1999 a periodicidade de publicações sobre Índice foi mais constante, conforme apresentado no Gráfico 1. Conforme descrito anteriormente na Tabela 1, um dos critérios utilizados para excluir textos que estariam na Revisão de Literatura, foi desconsiderar arquivos cujo *download* do texto completo não foi possível. Tal situação justifica o fato da maioria das publicações anteriores ao ano de 2000 não terem todos os materiais digitalizados e catalogados em bases de dados atuais.

Gráfico 1 - Quantidade de publicações por ano



Fonte: elaborado pela autora.

Das publicações selecionadas foram identificados 69 autores, considerando que alguns documentos foram escritos em co-autoria. Dos 69 autores, 38 possuíam vínculo com Universidades⁴; 7 foram caracterizados em “outros vínculos”⁵ por estarem ligados a centros

⁴ Florida State University; Macquarie University; Moscow State Lomonosov University; Pennsylvania State University; Simmons University; Southern Polytechnic State University; Stanford University; State University Of New York; Syracuse University; Universidade de Angers; Universidade de Michigan; Universidade de Patras; Universidade de Waterloo; Universidade do Norte do Texas; Universitas Palangka Raya; Université de Lausanne; Université de Montréal; Université Paris 13; University of Wisconsin-Milwaukee; University Higher School of Economics; University of Brighton; University of British Columbia; University of Malaya; University of North Carolina; University of North Texas; University of Patras; University of Southern California; University of Strathclyde; University of Toronto Press.

de pesquisa ou outras empresas; e 5 afiliados a associações de bibliotecários⁶. Na análise dos dados, observa-se que a temática de índice é muito discutida no ambiente acadêmico, como também, em outras instituições tais como centros de pesquisas e empresas privadas, conforme apresentado na Tabela 4.

Tabela 4 - Vínculo dos autores

VÍNCULO	QUANTIDADE DE AUTORES
Universidades	38
Outros vínculos	7
Associações de bibliotecários e indexadores	5
Não informado	19
TOTAL	69

Fonte: elaborado pela autora.

Quanto ao país de origem dos autores, identificou-se uma maior concentração nos Estados Unidos, representando 35% do total. Com o quantitativo menor, os autores do Canadá e da Indonésia representam 9% e 6 %, respectivamente. Embora grande parte dos autores sejam de origem norte-americana, existe uma importante representação mundial.

Em relação aos quantitativos por continente, um total de 7 (sete) países europeus (Alemanha, Escócia, França, Grécia, Reino Unido, Rússia e Suíça) tiveram autores que publicaram sobre índice. Mas apesar de serem em maior número de países, o continente americano obteve maior quantidade de publicações relacionadas ao índice totalizando 24 publicações. No Brasil não foi identificada nenhuma publicação sobre IFL, conforme os dados estão demonstrados na Tabela 5.

⁵ Biblioteca Nacional da Alemanha; *Google*; *Harvard Library*; *News Information Center*; *Palo Alto Research Center*; presidente da ASIS & T.

⁶ *American Society Indexers e American Library Association*.

Tabela 5 - Quantitativo da origem dos autores

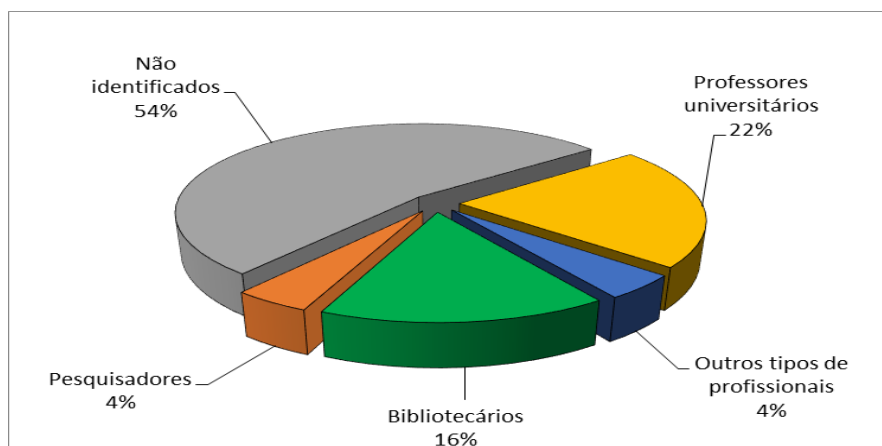
ORIGEM DOS AUTORES	QUANTIDADE DE AUTORES
Estados Unidos	24
Não localizado	23
Indonésia	4
Canadá	6
França	2
Reino Unido	2
Alemanha	1
Austrália	1
Escócia	1
Grécia	1
Malásia	1
Rússia	1
Suíça	1
TOTAL	69

Fonte: elaborado pela autora.

Em relação à ocupação profissional desses autores, a maior parte é constituída por professores universitários representando 22% do total. Consideram-se, nesse caso, professores associados, visitantes e eméritos que ministram aulas na universidade. Outra profissão identificada foi a de bibliotecário, tendo em vista que a formação em Biblioteconomia em outros países é uma especialização do profissional. Assim sendo, muito autores foram identificados como indexadores, catalogadores, taxonomistas entre outros. Para esse estudo, optou-se por nomear tais profissionais como bibliotecários, os quais totalizam 16% da amostra total.

Os pesquisadores correspondem a 4% dos autores. Professores e bibliotecários também podem ser pesquisadores. No caso desse estudo, pesquisadores são os autores que informaram ter como principal profissão o ato de realizar pesquisa. Os “outros tipos de profissionais” também representam 4% do total, portanto 1 (um) autor cientista da computação, outro especialista em linguística e o último engenheiro de software. Em análise, o índice está muito associado às práticas dos bibliotecários e pesquisas dos professores universitários, porém pode-se observar que tal temática interessa também a outros profissionais ligados à tecnologia e a linguística. No Gráfico 2 estão expostos todos os relatos desse tópico.

Gráfico 2 - Profissão dos autores



Fonte: elaborado pela autora.

Em relação ao idioma, a maioria das publicações recuperadas estavam em inglês, não havendo nenhuma recuperação de publicações em português. Pode-se inferir que no Brasil faltam estudos na área, principalmente publicação em periódicos acadêmicos. Sobre as abordagens de estudo identificadas nas publicações foram definidas a partir da leitura do resumo e leitura técnica do artigo. Após esse processo, foram verificadas sete abordagens principais, quais sejam:

1. **Teste de desempenho do índice:** estudos que comparam o índice a partir de outras ferramentas de recuperação da informação, ou mesmo em relação à eficiência de sua funcionalidade;
2. **Elaboração de índice:** relatam experiências na elaboração do índice, criando um sistema próprio de indexação automática ou semiautomática, ou apresenta a sua elaboração;
3. **Ferramenta de pesquisa:** apresentam o índice como ferramenta de pesquisa, com a finalidade de divulgar sua importância;
4. **Comentário de especialista (s):** discussão com o especialista ou grupo de estudo sobre a temática índice de forma genérica;
5. **Histórico:** apresenta a origem e antecedentes do índice;
6. **Indexação no índice:** relata melhorias na prática da indexação a partir do Índice;
7. **Revisão de Literatura:** análise de publicações com a temática índice.

A seguir é apresentada a distribuição das publicações por abordagens na Tabela 6.

Tabela 6 – Abordagens de estudo

Abordagem de estudo	Quantidade
Elaboração de Índice	16
Teste de desempenho do Índice	16
Ferramenta de pesquisa	7
Comentário de especialista (s)	4
Histórico	3
Indexação no índice	3
Revisão de literatura	2
Total	51

Fonte: elaborado pela autora.

Foram recuperados 16 documentos que abordavam sobre “elaboração de índice”. O recorte foi realizado levando em consideração os objetivos propostos neste projeto. Cada documento é apresentado na seção a seguir.

2.3 Análise dos trabalhos correlatos

Dos 51 trabalhos recuperados, conforme demonstrado na Tabela 1, somente 16 deles foram considerados para embasar a pesquisa, uma vez que eram os únicos que discorriam sobre a elaboração de índice, objeto de investigação da dissertação e, principalmente, quanto à evolução tecnológica, no intervalo temporal de 1970 a 2019. Os objetivos, os resultados alcançados e as conclusões desses documentos, foram as categorias de análises para a redação da Revisão de Literatura. Os 16 trabalhos selecionados serão apresentados, na ordem cronológica das datas de publicação.

Observando a busca bibliográfica, foi possível constatar que no Brasil há incipiência em relação a publicações referentes à elaboração do índice com o auxílio tecnológico. Porém, no âmbito internacional, os alunos do curso de biblioteconomia da *Columbia University* relatavam sobre o auxílio da tecnologia para facilitar o trabalho de indexação. Esse relato consta do trabalho intitulado “*Computer-aided production of book indexes*”, escrito por Theodore C. Hines e Jessica L. Harris em 1970 (HINES; HARRIS, 1970). Os autores referem-se à oferta de cursos de Indexação e de Sistema de Informação e, para aproveitamento dos dois cursos, os alunos desenvolveram um programa de indexação de livros com finalidade didático-educativa. O resultado do estudo foi a apresentação das atividades que o indexador realizaria para utilizar o sistema de indexação, sinalizando no texto os termos a serem representados. Os autores ficaram satisfeitos com o sistema, pois demandaria menos esforços manuais, porém mencionaram somente o funcionamento/resposta do computador, além do esforço da digitação repetitiva do indexador humano.

Em 1989, o trabalho intitulado “*Indx and findphrases, a system for generating Indexes for Ditroff documents*”, dos autores Kris K. Abe e Daniel M. Berry (ABE; BERRY, 1989), surgiu como uma das primeiras iniciativas para desenvolver um programa de indexação semiautomática, ou seja, não seria necessária a marcação dos termos indexados no texto, ao contrário da publicação de Hines e Harris. A finalidade do trabalho foi apresentar um sistema que auxiliasse o processo de determinação do termo. Os resultados foram a descrição dos procedimentos adotados para utilizar o sistema Indx, bem como a comparação desse sistema com as características existentes em outros sistemas. Como conclusão, os autores consideraram satisfatório o índice criado a partir do sistema, mas a exclusão de páginas inúteis, inseridas automaticamente, foi entendida como um item a ser melhorado no referido sistema.

O artigo francês denominado “*Nouveaux horizons en indexation automatique de Monographies*” de Sylva (2002), discorreu sobre a indexação automática para documentos extensos (indicado pela autora como *monographies*). Trabalho essencialmente norteador para desenvolvedores de *software* para extração automática, teve como propósito relatar estudos que versassem acerca da melhoria da qualidade do índice automático. A redação foi pautada em sugestões de itens que um *software* de indexação automática deveria conter, tais como: identificação de expressões relevantes; tipos de expressões procuradas (reconhecer tipograficamente as variações do tipo de palavra, por exemplo: assunto e autor); método de identificação de expressões relevantes (uma extração não somente por frequência das palavras, mas pela utilização de um tesouro embutido); homogeneização das entradas; identificação de relações semânticas e agrupamento de entradas; possibilidade de edição do índice. Em análise final, a autora indicou a dificuldade do indexador avaliar o produto (índice) que não teria sido produzido por ele mesmo, a aceitação de uma qualidade inferior do índice automático em relação ao elaborado manualmente e a necessidade de aplicação das sugestões desse estudo em um software com problemas relacionados à atividade de automação.

Com o intuito de dissertar quanto à criação de um sistema denominado ScentIndex, foram publicados dois artigos. O primeiro retratou o panorama introdutório sobre o sistema e o segundo descreveu o índice de assunto produzido pelo sistema. Escrito pelos autores Ed H. Chi, Lichan Hong, Julie Heiser e Stuart K. Card (CHI *et al.*, 2004), sendo dois especialistas da área de computação e pesquisadores do Google, e os outros dois especialistas da área de psicologia e interação humano computador. A primeira publicação “*eBooks with Indexes that Reorganize Conceptually*”, foi escrita em 2004 e a segunda, “*ScentIndex: Conceptually Reorganizing Subject Indexes for Reading*”, data de 2006 (CHI *et al.*, 2006). O sistema ScentIndex é um método que reorganiza conceitualmente grandes índices de assuntos para atender a alguma necessidade de informação, utilizando a

indexação automática. Como resultado, o primeiro trabalho relatou brevemente buscas e resultados do índice produzido pelo sistema; na segunda publicação discorreu-se sobre as técnicas e algoritmos para definir a relevância de um termo, bem como o estudo de usuário, realizado para testar o sistema. Os autores refletiram que, por meio do estudo de usuário, foi possível confirmar o funcionamento do sistema que atende a usuários experientes e novatos. Entende-se que o diferencial deste sistema é o estudo aprofundado da necessidade do usuário para estabelecer os termos relevantes.

Seguindo a temática anterior, criação de um sistema automático, em 2005 foi publicado o artigo *“Using NLP to build the hypertextual network of a back-of-the-book index”* pelos autores Touria Aït El Mekki e Adeline Nazarenko (MEKKI; NAZARENKO, 2005). O trabalho discutiu a automação total do índice (sem intervenção humana), além de propor uma rede hipertextual que auxiliasse na navegação do documento. O resultado foi a descrição do método para elaboração do sistema IndDoc. Os autores concluíram que a elaboração automática da rede hipertextual não foi satisfatória, pois a finalização resultou em um IFL sem alterações significativas. Destacaram ainda, problemas relacionados aos aspectos semânticos e linguísticos para interligar os termos, como também a definição de qual trecho do texto realmente “vale a pena” se referir como entrada, aspectos esses não atendidos pelo sistema.

O trabalho denominado *“Editing the index: developing a method”*, produzido em 2005 por Faulkner e Light (2005), retratou o desenvolvimento do método INDEX-L, criado especificamente para a edição do índice. De acordo com os autores, a edição do índice é uma tarefa “assustadora” para estagiários e indexadores recém-formados. Assim, o artigo teve por finalidade refletir sobre a temática a partir da percepção de vários indexadores. O resultado foi o relato dos indexadores sobre suas práticas no momento da edição, demonstrando as variações no comportamento de nove sujeitos pesquisados. Considerando a subjetividade do assunto, não foi possível identificar uma conclusão clara, visto que os autores expuseram habituais práticas e dificuldades na rotina de trabalho, reforçando a importância do auxílio do *software* ao relatarem que o trabalho manual demanda muito tempo.

Em uma longa carreira teórica e prática, a pesquisadora Jean Weihs, especialista na área de catalogação, publicou em 2007 o artigo intitulado *“Indexes, indexing and technical services staff part 2: cataloguers”* (WEIHS, 2007). Com o intuito de relatar a prática pessoal embasada nos argumentos de Mulvany, Weihs descreveu seu conhecimento quando da elaboração de índice e como a automação total do índice é utópica e inapropriada, causando inferioridade na qualidade desses produtos.

Os trabalhos dos pesquisadores da área de computação geraram três artigos nos anos de 2006, 2007 e 2008, relacionados à automação do índice. Andras Csomai é

engenheiro de software da empresa Google e Rada Mihalcea é professora de Ciência da Computação, conhecida por estudos no âmbito do processamento de linguagem natural, sendo premiada por trabalhar temas que vão desde a análise semântica dos textos até os *softwares* de detecção de mentiras (RADA [20--]).

O primeiro trabalho, desenvolvido em 2006, "*Creating a Testbed for the Evaluation of Automatically Generated Back-of-the-book Indexes*" (CSOMAI; MIHALCEA, 2006), teve como objetivo descrever a ação realizada na criação de um sistema (*Testbed*) para avaliar índices produzidos a partir da indexação automática. Como reflexão, os autores comprovaram que o *testbed* permitiria a avaliação comparativa de uma variedade de métodos de indexação e finalizaram indicando a criação de um método automático para elaboração de índice como objetivo de longo prazo referente ao produto dessa publicação.

O planejamento para consolidar o objetivo de longo prazo desses autores iniciou no ano de 2007, com o estudo "*Investigations in Unsupervised Back-of-the-Book Indexing*" (CSOMAI; MIHALCEA, 2007). O propósito do estudo foi apresentar o método não supervisionado, ou seja, sem a etapa de treinamento, exemplos ou modelos. No resultado foram demonstrados a lógica e os algoritmos necessários para desenvolvimento de um possível sistema. Os autores concluíram que o método seria útil e que, por meio dos dados estatísticos, seria possível identificar se houve melhoria significativa na recuperação dos termos.

Escrito pelos mesmos autores, o artigo "*Linguistically motivated features for enhanced back-of-the-book indexing*" (CSOMAI; MIHALCEA, 2008) também se baseou na ideia da automação total do índice, porém com o método supervisionado, ao contrário do artigo do ano anterior. O escopo do artigo foi explicitar o método supervisionado para elaboração do índice com melhoria de 140% se comparado com os métodos já existentes e que, geralmente, se baseavam na seleção por frequência do termo. Como resultado, foi apresentado o método supervisionado usando um conjunto de recursos com motivação linguística, que são: 1) conjunto de dados (técnicas de filtragem de dados); 2) conjunto de recursos (frases e informações, os autores utilizaram ideia do trabalho anterior) e; 3) algoritmo de aprendizado de máquinas (integração dos recursos um e dois para a estrutura aprendizado de máquina). Assim como o trabalho anterior, esse estudo também foi considerado bem-sucedido, pois possibilitou identificar melhorias em relação às técnicas tradicionais baseadas na frequência do termo. Além disso, os autores sugeriram trabalhos futuros, tais como: estudos de metodologias para indexação de índice automático e utilização dos índices extraídos automaticamente para melhorar a navegação de bibliotecas digitais. Os autores não deixaram clara a razão de desenvolverem dois artigos com métodos opostos, embora tenha sido utilizado resultado do estudo anterior para subsidiar a realização do segundo estudo.

Em 2013, Zhaohui Wu, Zhenhui Liz, Prasenjit Mitrazy e C. Lee Giles publicaram o trabalho “*Can Back-of-the-Book Indexes be Automatically Created?*”, que segue a linha de criação automática de índice com o método não supervisionado. A diferença desse artigo em relação aos dois últimos relatados anteriormente é a apresentação de uma abordagem independente do domínio para criação do índice, que resultou na explanação dos termos, principalmente no tocante ao contexto, utilizando-se o conhecimento da *web* (Wikipédia). De acordo com os testes, os autores concluíram que o método é satisfatório conseguindo “peso” (medida identificada para calcular o desempenho – algoritmo) em relação ao tempo de resposta, que leva de 5 a 8 minutos para gerar o índice. No entanto, os autores reiteraram que esse sistema não se adapta a relações semânticas do termo, referências cruzadas entre outros, e sugerem outros estudos nessa área (WU et al, 2013).

Em 2013, a autora Sylvia K. Miller publicou o artigo “*Index Appreciation: A Publisher’s Brief Guide*” (MILLER, 2013), no qual comenta sua prática de indexação e descreve as decisões básicas necessárias na criação de um índice para um livro acadêmico. A autora detalhou o indexador, o custo, o tamanho do índice e a decisão do que indexar, como montar um cronograma e, por fim, discutiu o futuro com o índice eletrônico. A autora concluiu com uma reflexão em relação à automação do índice, que em sua opinião poderia não cumprir a função de facilitar a busca do usuário. A autora advertiu, ainda, que os sistemas funcionariam satisfatoriamente somente com conexão e navegação semanticamente interligadas dos termos.

Com uma abordagem prática e didática, o artigo “*A Simple System for Constructing an End-of-Book Index*”, produzido por Ronald D. Francis e Murray Greenway (FRANCIS; GREENWAY, 2015), auxiliou autores e bibliotecários indexadores a criar um índice. Francis é professor emérito na área jurídica, e Greenway é bibliotecário jurídico. O objetivo do artigo foi delinear um método econômico que envolvesse o uso de computadores, mas não empregasse o sistema de indexação automática, o que resultou em sete breves orientações de como elaborar um índice utilizando um sistema de reconhecimento de voz no qual o personagem que estivesse criando o índice falasse em voz alta os termos e o sistema digitasse. Na conclusão os autores indicaram que as orientações estão abertas a sugestões, e não houve teste aprofundado para avaliar a efetividade da proposta.

Em conferência internacional realizada em 2018 os autores Bolshakova e Ivanov (2018) relataram a criação de método para extrair e selecionar termos relevantes no texto científico educacional e assim gerar o IFL por meio da indexação automática, cujos resultados foram considerados positivos em comparação com outras técnicas de indexação automática de IFL. Essa forma apresentou aumento considerável da precisão e revocação durante a extração de termos. Concluíram que, embora o resultado tenha sido satisfatório, é

necessária a revisão/validação do indexador. Finalizaram também pontuando que é necessário incentivar o desenvolvimento de métodos para detectar e agrupar os termos sinônimos.

A autora Provo (2019) relatou sobre o resultado do projeto *Enhanced Networked Monographs* (ENM) realizado entre 2015 a 2018, escrito no artigo denominado “*From index to network: topic maps in the Enhanced Networked Monographs project*”. O projeto ENM teve o intuito de, por meio da criação do mapa de tópicos, proporcionar acesso gratuito, via *web*, a 110 monografias acadêmicas. Como resultado foram selecionados 15.800 termos, e desses, 3.000 foram editados manualmente por uma equipe de profissionais, por haver indicativos de erros na semântica (sinônimos, abreviaturas, entre outros). O “meta-índice” criado, possui *hiperlink* para que o usuário possa acessar o *e-book* dos trabalhos completos. É possível, também, fazer filtragem da quantidade de termos que serão apresentados para o usuário. Ao final, a autora sugeriu que projetos futuros se concentrem em um único domínio do conhecimento, para assim, atenuar problemas com homógrafos e outras questões semânticas. Provo (2019) considerou que o produto é de fundamental importância para o usuário e que este projeto pode ser desmembrado em outros. Comentando sobre falhas no produto final, a autora ressaltou o esforço e participação da equipe para validar os termos selecionados, o que ainda indica a necessidade do indexador.

Pode-se classificar os trabalhos descritos em duas categorias principais: índices de forma semiautomática (HINES; HARRIS, 1970; ABE; BERRY, 1989; FRANCIS; GREENWAY, 2015; BOLSHAKOVA; IVANOV, 2018; PROVO, 2019) e índice de forma automática (CHI *et al.*, 2004; CHI *et al.*, 2006; MEKKI; NAZARENKO, 2005; CSOMAI; MIHALCEA, 2006; CSOMAI; MIHALCEA, 2007; CSOMAI; MIHALCEA, 2008; WU *et al.*, 2013). Ressalta-se, que alguns autores fizeram relatos de melhorias nos tipos de conversão ou observação no momento da elaboração do índice.

Dos autores que fizeram relatos de melhoria, Sylva (2002) recomendou melhorias para criação de índice com a indexação automática. Os autores Faulkner e Light (2005) observaram que os indexadores que utilizavam o *software* INDEX-L de indexação automática eram mais rápidos para elaborar o índice, mas alertaram sobre variações na edição do índice quando consideraram que a indexação é subjetiva. Os trabalhos dos autores Weihs (2007) e Miller (2013) relataram a prática pessoal na elaboração de índice, argumentando que a automação do índice é inapropriada, por não considerar a semântica dos documentos.

Nota-se que, embora a intenção dos autores tenha sido a de facilitar a elaboração do índice, todas essas publicações contêm variações em relação aos objetivos e resultados. Em 1970, foi identificada a inserção da tecnologia de forma simples, sendo mencionada a intervenção do indexador no trabalho. Após esse período, o índice passou a

ser observado totalmente sob a ótica tecnológica, incluindo a automação da seleção de termos e edição do índice. Acredita-se que tais publicações estivessem, de certa maneira, destoantes da realidade editorial, bem como da necessidade de atingir e aumentar o interesse pela inserção de índice nos livros.

Talvez por essa razão as publicações mais recentes tenham se caracterizado pelo cunho mais didático de orientação simples e objetiva para autores e bibliotecários, quando da criação do índice. Pois, embora abordassem a criação do índice de forma automática, observaram que os estudos finalizaram com a desmistificação da elaboração de índice totalmente automático por ser considerado um grande desafio, visto que computadores ainda não correlacionam sinônimos, polissemias e relação entre termos. Assim, a intervenção do indexador, mesmo com o auxílio da tecnologia, continua sendo necessária para explicitar a semântica e o contexto na elaboração de índices.

Tomando por base as publicações selecionadas, os primeiros trabalhos apresentaram de forma incipiente a inserção tecnológica e, com o passar dos anos, houve vários relatos de índices produzidos por meio da indexação automática. Nos últimos anos, os autores identificaram que os índices gerados totalmente de forma automática não são satisfatórios, exigindo, assim, a intervenção do indexador. Dessa forma, existem estudos recentes sobre o uso da tecnologia no IFL, porém atualmente a utilização total da tecnologia (indexação automática) na elaboração de IFL não é satisfatória.

Por meio da leitura e análise dos trabalhos recuperados, avalia-se que há lacunas de estudo que precisam ser investigadas em relação à indexação semiautomática na elaboração de índice, pois se desconhece quais os papéis da tecnologia e do indexador na elaboração de tal produto, no sentido de trata-las de forma simultânea durante a produção do IFL. Dessa forma, abordar tais temáticas possibilita o surgimento de novos estudos para auxiliar os profissionais que elaboram índices.

3. FUNDAMENTAÇÃO TEÓRICA

“Os ingleses, afirmara Herr Heinrich, não compreendem a indexação. Ela é a origem de toda boa organização.” (H.G Wells: Mr. Britling Sees it Through).

Neste capítulo é apresentada a fundamentação teórica adotada para o desenvolvimento desta pesquisa. Considerando que o objeto de estudo é o IFL, procura-se abarcar o entendimento sobre o conceito e suas características.

Existem três formas de indexação para elaborar o índice: manual, semiautomática e automática a escolha pela utilização dessas, depende da preferência do indexador a partir do entendimento das facilidades e limitações que cada uma proporciona, conforme apresentado a seguir.

3.1 Indexação manual

Para Silva e Fujita (2004, p. 138) e Guimarães (2008), a indexação surgiu com a atividade de elaborar índices. Nesse trabalho, optou-se por abordar tal histórico na seção (3.3.1 A evolução do índice na biblioteconomia).

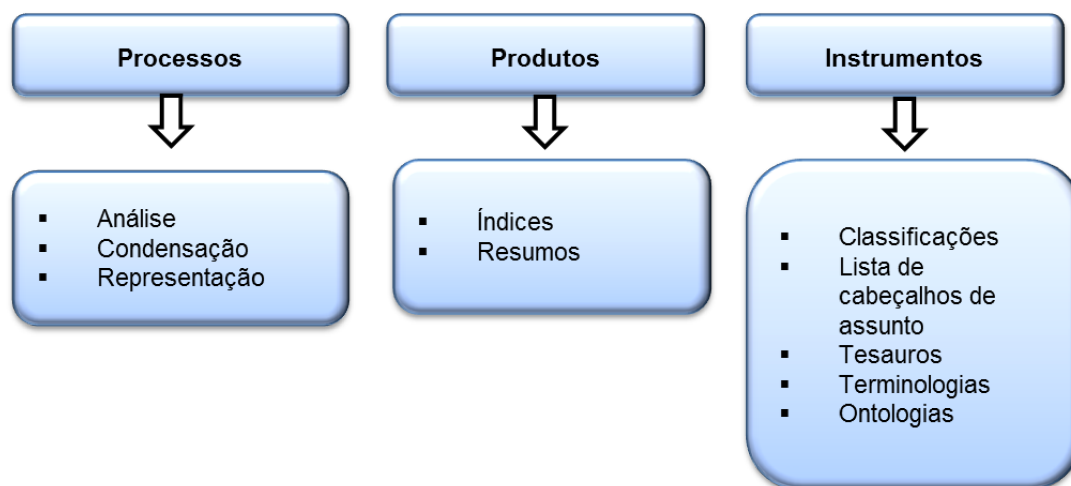
Uma das primeiras definições de indexação de que se tem registro foi elaborada pela UNISIST (1983), em que entende indexação como o processo utilizado para descrever e identificar um documento de acordo com seu assunto. Para Lima (2004, p. 65), a indexação tem estreita relação com a cognição, ao afirmar que “a indexação é o processo intelectual que envolve atividades cognitivas na compreensão do texto e na composição da representação do documento”. Observa-se que a indexação é um processo importante para identificar o conteúdo do documento, produzindo uma lista de termos que capturam a essência dos assuntos abordados neste documento, a fim de permitir sua localização em uma coleção.

A construção teórica da área de tratamento temático da informação desenvolveu-se a partir da abordagem de três aspectos que lhes são inerentes e que direcionam as pesquisas e as bases disciplinares da Organização e Representação da informação (ORI), sintetizados na Figura 1, que são:

- a) **processos:** compreendem a análise, condensação e representação. A análise é a etapa de leitura dos documentos, na qual se procura as partes de maior relevância temática; a condensação consiste na síntese do documento, a fim de extrair termos para recuperação; a representação traduz o conteúdo temático do documento para uma linguagem especializada;
- b) **produtos:** têm a finalidade de facilitar a consulta aos originais. São exemplos de produtos os índices, os resumos e os catálogos impressos;

- c) **instrumentos**: são ferramentas das linguagens documentarias para o controle terminológico dos documentos, tais como cabeçalhos de assuntos, tesouros, terminologias, ontologias, entre outros (GUIMARÃES, 2008; MARTINS; MORAES, 2015).

Figura 1 - Sistematização de conteúdos em Organização e Representação do conhecimento



Fonte: Adaptado de Guimarães, 2008.

Expostos os processos, produtos e instrumentos que integram a área de ORI, apresenta-se as características para obter uma indexação de qualidade, itens que devem ser observados durante o trabalho do indexador.

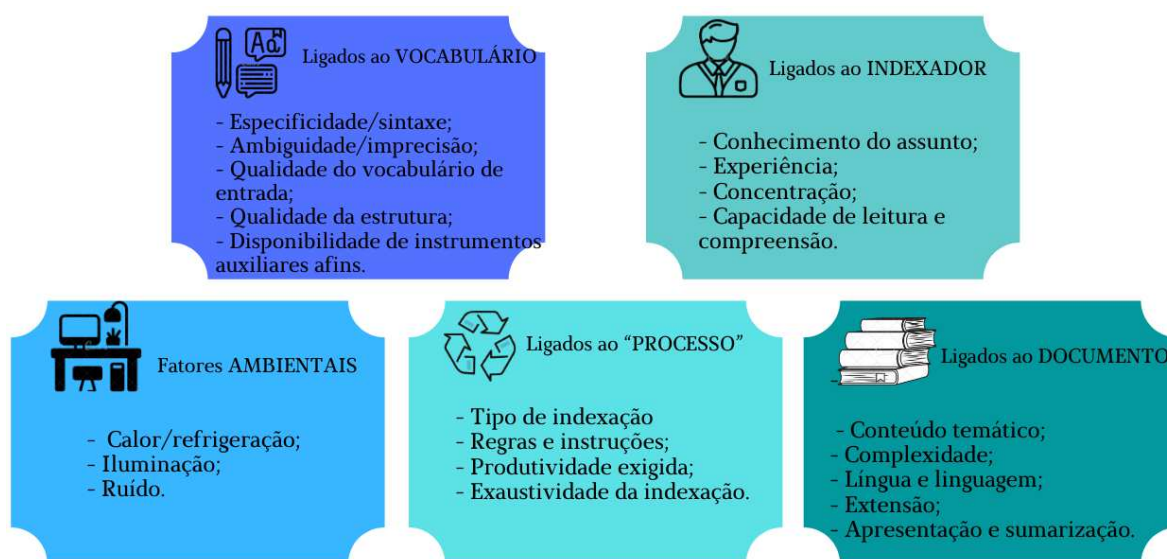
3.1.1 Características essenciais para garantir a qualidade na indexação

São identificadas na literatura sugestões que norteiam e auxiliam o profissional a garantir a qualidade na indexação, embora não existam formas para medir uma “boa indexação”. Moreiro Gonzáles (2004) destaca quatro critérios para uma boa indexação, que são: entropia, procedência dos termos, profundidade e o índice de consistência. A **entropia** está associada à identificação da maior quantidade de informações utilizando-se a menor quantidade possível de termos; é um critério difícil de alcançar, pois deve-se conseguir objetividade, neste sentido, o oposto da entropia é a redundância. Já a **procedência dos termos** surge de duas formas: a primeira é quando os termos são elaborados pelos indexadores utilizando-se da linguagem documental da própria instituição/unidade de informação, sendo necessário determinar como foram obtidos tais termos como: texto completo, epígrafe, introdução, entre outros. Geralmente, tais informações constam na política de indexação. O segundo tipo de procedência dos termos acontece quando se reutiliza a indexação feita por uma equipe de indexadores ou uma instituição externa, sendo

necessário considerar a correção dos termos para o contexto específico para o qual está sendo indexado o documento.

A **profundidade** consiste em compreender os níveis de representação do documento: se os termos devem refletir todo o texto (superestrutura⁷) ou representar o resumo dos documentos (macroestrutura parciais) e, no caso de representações mais genéricas, os conceitos dos títulos (macroestrutura global). O **índice de consistência** leva em consideração se os termos aparecem sempre da mesma forma. Diz respeito à exigência de uniformidade ou homogeneidade com a qual todos devem descrever o mesmo documento ou solicitar o mesmo assunto de formas parecidas (MOREIRO GONZALES, 2004). Os critérios para uma boa qualidade da indexação podem ser influenciados por alguns fatores internos e externos; com isto, deve-se atentar para alguns aspectos apresentados na Figura 2.

Figura 2 - Fatores que influenciam a qualidade da indexação



Fonte: elaborado pela autora com os dados de Lancaster, 2004.

Pode-se observar, a qualidade da indexação não depende somente do indexador, mas deve-se considerar vários fatores. O conhecimento de tais fatores é essencial para garantir uma indexação de qualidade e atender às necessidades informacionais do usuário.

⁷ O conceito de superestrutura e macroestrutura será abordado na seção 3.4.3.1 Análise de assunto.

3.2 Indexação automática e semiautomática

A análise automática de textos tornou-se assunto de pesquisa no final da década de 1950, quando foram desenvolvidos métodos relativamente simples para a construção de índices a partir de textos, como, por exemplo, o *Keyword in Context – KWIC* (Palavra-chave no Contexto) – e o *Keyword out of Context – KWOC* (BORGES, 2009; p. 32; BORKO; BERNIER, 1978, p. 113), definidos na seção 3.3.2 Tipos de enfoque e finalidade.

Houve quatro fatores que incentivaram o foco de pesquisa para essa área na Ciência da Informação. O primeiro deles foi o intuito de **fornecer acesso rápido** e mais completo à literatura científica, em especial no domínio da engenharia, como uma tentativa de controlar a explosão informacional. Segundo, a **maior disponibilidade de computadores** e o reconhecimento de que essas máquinas manipulam símbolos, palavras e números. O terceiro fator foi o surgimento de um novo campo de estudo denominado **linguística computacional** (aplicação da análise computacional à estrutura e ao significado da linguagem). E, por último, o quarto impulso foram pesquisas e avanços nas áreas de **inteligência artificial** e sistemas de auto-organização (“máquinas pensantes”).

O histórico da indexação automática não será aprofundado nesse estudo, por não ser o objeto principal de pesquisa; desta forma, essa seção enfatiza as definições, diferenciações entre os termos, tipos de indexação - semiautomática e automática - e principalmente as vantagens e desvantagens dos dois tipos de indexação para justificar a escolha pela indexação semiautomática nessa dissertação.

De acordo com o glossário Wellisch (2000, p. 10), a **indexação automática** é definida como “qualquer método de indexação pelo qual o texto de uma unidade documental é submetido a operações algorítmicas, a fim de extrair termos ou frases que representam assunto, tópicos ou recursos da unidade documental”. Em concordância com tal afirmação, Felipe (2012, p. 35) destaca que a indexação automática ocorre quando a “máquina realiza o processo de representação do texto com base em um sistema de regras pré-definidas através de algoritmos (comandos que constituem o software), pelos quais o documento é associado a palavras ou termos”. Em ambas definições, a indexação automática diz respeito aos algoritmos desenvolvidos e submetidos à máquina, a fim de extrair ou recuperar informações por meio de termos, frases, entre outros.

Pode-se dizer que o objetivo da indexação automática é a identificação das palavras mais representativas dos documentos, visando agilizar e auxiliar o processo intelectual de indexação mais elaborado realizado pelos profissionais da área. Além disso, esse processo auxilia na representação de textos por meio dos conceitos macroestruturais obtidos mediante a intervenção de diferentes técnicas e ferramentas originadas em outras disciplinas, principalmente a linguística e a estatística, como também a informática,

especialmente na inteligência artificial (MOREIRO GONZALES, 2004, p. 100; BORGES, 2009, p. 31).

Já a **indexação semiautomática** é definida por Felipe (2012, p.34) como o “processo misto entre processamento computacional e decisão humana. O usuário pode marcar partes do texto/documento e comandar a máquina (software) para realizar a construção do índice, com base nos termos selecionados”. Porém, assim como na indexação automática, na indexação semiautomática não se consegue solucionar os problemas da ambiguidade semântica, da polissemia e os regionalismos, exigindo que o indexador fique atento a essas variantes para minimizar esses problemas no momento da automação (FELIPE, 2012).

Observa-se na literatura variações acerca das definições dos tipos de indexação. Alguns autores consideram que a indexação automática e semiautomática são sinônimas (BORGES, 2009). Outros, no entanto, fazem diferenciação entre os termos, como é o caso de Correa e Lapa (2013), que definem a indexação auxiliada por computador como,

os programas que auxiliam no processo de armazenamento dos termos adquiridos através da indexação intelectual. Estes sistemas procuram facilitar o processo de indexação ao proporcionar através das telas de ajuda, as notas explicativas sobre o uso de um termo, os termos relacionados e permitir a atribuição de termos sem ter que digitá-los, ou mesmo consultar on-line por documentos indexados anteriormente para verificar qualquer aspecto (CORREA; LAPA, 2013, p. 258).

A indexação automática e semiautomática são definidas pelos autores, também com conceitos diferentes: para eles, a indexação automática são “programas que não precisam de validação, ou seja, os termos propostos são armazenados diretamente como descritores do documento indexado” (CORREA; LAPA, 2013, p. 258). Já a indexação semiautomática “ocorre em sistemas que indexam documentos automaticamente, mas os termos de indexação propostos, se necessário, são validados e editados por um profissional” (CORREA; LAPA, 2013, p. 258).

As definições implicam níveis diferentes de automação. No caso da indexação auxiliada pelo computador, a máquina não realiza a indexação, mas auxilia com sugestões e direcionamentos, dessa forma, a automação tem pouco protagonismo; já na indexação semiautomática a máquina realiza a indexação, mas há a possibilidade de o indexador realizar a validação dos termos; assim, computador e indexador trabalham juntos para finalizar o produto. E, finalmente, na indexação automática não há nenhuma participação do indexador: o computador sugere os termos e estes mesmos são os descritores utilizados.

3.2.1 Tipos de indexação semiautomática e automática

Para Sylva (2002) existem dois tipos de indexação semiautomática: 1) **indexação manual com software dedicado**, e a 2) **indexação assistida por computador**. Na primeira, utiliza-se um *software* dedicado para indexação de trabalhos impressos, que gerencia os aspectos mecânicos da produção de índices, como: gerenciamento de números de página, classificação alfabética, apresentação tipográfica na página entre outros. Como exemplos desse tipo de software, podemos citar o *CINDEX*, o *Macrex*, e o *SkyIndex*⁸. Eles dependem da intervenção humana para a realização de todo o trabalho intelectual de indexação, isto é, a identificação de trechos informativos, a escolha de termos úteis para sua indexação, e o agrupamento dos termos/cabeçalhos em entradas estruturadas (SYLVA, 2002). Assim, este tipo de indexação semiautomática é caracterizado pela automação apenas da parte “técnica”, no momento da elaboração do índice; a parte “intelectual” da indexação é realizada pelo indexador.

O segundo tipo, 2) *a indexação assistida por computador*, se propõe a facilitar a tarefa intelectual da indexação, oferecendo ao profissional várias ferramentas de ajuda, como: acesso *on-line* a vocabulário controlado; guias e manuais integrados; política de indexação; apresentação na tela de “grades de indexação” contendo os vários campos a serem preenchidos (tais campos servem como um auxiliar de memória, além de minimizarem o tempo e esforço de entrada); correção automática de palavras digitadas no teclado entre outros. Desta forma, o software identifica no texto as expressões que podem ser úteis para a indexação, assim, termos candidatos são propostos, mas devem ser validados pelo indexador. Exemplos de softwares que realizam essas atividades são o *SATO* e o *Sonar Bookends Professional, Indexing online e Syntactica*. Eles podem tanto compilar uma lista de palavras e termos candidatos, estruturar (até certo ponto) entradas em cabeçalhos e subcabeçalhos, definir variantes para procurar uma determinada entrada e apresentar os termos candidatos em contexto para permitir que o usuário retenha ou não uma entrada no índice para essa ocorrência do termo (SYLVA, 2002; MEKKI, NAZARENKO, 2005). Neste caso, as partes “técnica” e “intelectual” são realizadas pelo software, sendo que para a segunda é necessária uma revisão dos termos sugeridos.

Existem dois tipos de indexação automática conforme Sylva (2002), a 1) **Indexação totalmente automática**, e a 2) **Indexação por extração e atribuição**. A primeira tem o objetivo de excluir completamente a intervenção humana do processamento. Assim, o sistema realiza a localização e estruturação de termos candidatos, executa as etapas mecânicas da edição, e apresenta ao usuário um índice final. Uma indexação

⁸ Descrição no Apêndice A.

totalmente automática é realizada por alguns softwares de varredura de texto com reconhecimento óptico de caracteres, notadamente os softwares de reconhecimento de voz (como ProDEX). Por fim, os sistemas de gerenciamento de banco de dados (DBMS) também usam índices gerados automaticamente (com a possibilidade de receber comando explícito do usuário). A indexação automática destina-se a permitir que os usuários pesquisem documentos gerenciados por sistemas. O produto desse tipo de indexação é um índice com termos únicos, mesmo no caso de substantivos compostos, os termos seriam separados, embora sua recuperação possa ser realizada com a ajuda de consultas booleanas bem elaboradas. Em resumo, o que se entende por indexação automática é o suporte de todo o processo de indexação, desde a análise do texto até a apresentação dos resultados ao usuário, sem intervenção humana (SYLVA, 2002).

O segundo tipo é a *indexação de extração e atribuição*. A extração automática utiliza-se apenas dos termos explicitamente identificados no texto; o indexador, por sua vez, utilizando critérios institucionais e pessoais, seleciona no texto palavras que serão utilizadas para representar o documento. E, no caso da extração por atribuição, utiliza-se de um elemento externo ao documento, um conjunto de termos previamente definidos e normalizados (léxico) cuja complexidade pode variar deste uma lista de cabeçalhos de assunto até um tesouro ou uma ontologia. Portanto, os softwares de atribuição são mais sofisticados do que os de extração, considerando a facilidade de configuração. Porém, o desempenho não é satisfatório e a abordagem é limitada, pois dependem da terminologia exata utilizada pelo autor (SYLVA, 2002; LANCASTER, 2004, p.18).

3.2.2 Vantagens e desvantagens da indexação automática

A literatura aponta diversas vantagens e desvantagens da indexação automática. Para o indexador, essas são informações importantes para se escolher o tipo de indexação mais adequado para realizar alguma tarefa/atividade. Dessa forma, são destacadas algumas vantagens da indexação automática:

- a) realiza a leitura instantânea de todo texto;
- b) auxilia na preparação do índice;
- c) não é tendenciosa;
- d) representa maior agilidade em tarefas repetitivas mais simples;
- e) estabelecem sequências de hierarquias nas alterações das entradas;
- f) localizam as várias ocorrências de um termo;
- g) calculam automaticamente os números de página para cada termo;
- h) classificam as entradas em ordem alfabética;

- i) formatam o índice de acordo com um determinado estilo (SILVA; FUJITA, 2004; WU et al., 2013, p. 1745).

Conforme apresentado, as vantagens da indexação automática são eminentes. Mas é importante destacar os problemas advindos desse tipo de indexação, sendo o maior deles a questão semântica, “devido à particularidade de cada língua, que, em razão da sua dinâmica, determina constante modificação no significado das palavras” (FELIPE, 2012, p.35). Nesse sentido, deve-se considerar no desenvolvimento de algoritmos a “semântica e a sintaxe [da língua utilizada no] conteúdo desses documentos” (BORGES, 2009, p. 39). Para minimizar tais obstáculos existem constantes estudos na área da Biblioteconomia e da Ciência da Informação (BCI), através de estratégias como a construção de tesouros e ontologias. Porém, tais instrumentos não conseguem ainda abarcar a diversidade de domínios e contextos que são tratados nos livros. Entre as principais desvantagens relacionadas a esse problema, pode-se destacar que a indexação automática:

- a) funciona somente em documentos separadamente;
- b) não consegue fazer relações entre os textos ou entre um texto e uma visão de mundo, ou seja, não apresenta o contexto;
- c) fica vinculada ao vocabulário e à gramática da língua usada no documento indexado;
- d) não consegue lidar com os dados provenientes de gráficos, nota de rodapé, e informações contidas nas tabelas;
- e) consegue lidar com apenas uma língua por vez;
- f) não consegue avaliar textos;
- g) não consegue criar relações intertextuais;
- h) não é capaz de imitar o questionamento, a resposta humana a um texto, ou seja, não agrega valor à indexação;
- i) requer constante aprimoramento para manter-se em dia com os novos desenvolvimentos;
- j) não traz semântica incorporada;
- k) não define quais as ocorrências de termos "importantes" merecerão uma referência no índice (SILVA, 2002; SILVA; FUJITA, 2004).

Assim, conforme mencionado nesta seção, como também nos problemas relatados na Introdução (1.1 Justificativa e problema) e no resultado da revisão de literatura (2.3. Análise dos trabalhos correlatos) fica evidente que a indexação automática não constrói um IFL satisfatório, havendo perdas de semântica e contexto. Com esse relato, entende-se que o uso da indexação semiautomática é o mais indicado, por atender tanto as facilidades que a tecnologia proporciona, como extração rápida dos termos, numeração automática da página, entre outros, quanto possibilita a verificação do indexador, que analisa se os termos atendem as necessidades informacionais do leitor, e insere os termos

em um contexto, tratando as nuances de sinônimos e termos relacionados que somente o indexador é capaz de identificar.

3.3 Índice

Para compreender o objeto de estudo desta dissertação, é necessário apresentar o conceito de índice, pois esse possui diversas tipologias, as quais incluem o IFL.

Na década de 1970 a utilização do índice era primordial para busca informacional, como os volumosos catálogos/guias que continham listas de endereços e telefones, assim como outros que apresentavam os títulos dos artigos recém-publicados em periódicos. Dessa forma, antes da evolução da web, a sua importância destaca-se por sua singularidade em oferecer facilidade para que o usuário encontrasse as informações de forma agrupada e organizada por interesses, considerando que antigamente não havia sites de busca que organizavam a informação virtualmente (BORKO; BERNIER, 1978).

Com a evolução para os índices na web (catálogos on-line, índices de site, entre outros) as informações são armazenadas em arquivos de computador e pesquisadas por correlação de termos. Os computadores auxiliam na elaboração de concordâncias e na publicação de índices. Com efeito, os índices são guias úteis; eles não apenas nos direcionam para assuntos de interesse, mas também fornecem uma visão geral de um campo de assunto.

Nos Estados Unidos, o *American National Standards Institute* (ANSI) adotou um padrão para os índices, o Padrão Nacional Americano para Biblioteconomia e Ciência da Informação e Práticas Relacionadas de Publicação. Em ANSI Z39.4-1984, um índice é descrito como:

um guia sistemático para itens contidos ou conceitos derivados de uma coleção. Esses itens ou conceitos são representados por entradas organizadas em uma ordem pesquisável, como alfabética, cronológica ou numérica. Esta ordem é normalmente diferente daquela dos itens ou conceitos da própria coleção (ANSI, 1984, p. 1).

Bonura (1994, p. 5) define índice como “análise tópica sistemática organizada em ordem alfabética ou organizada por função, comando, procedimento ou tópico. É o mapa mais importante de um leitor para localizar informações no documento que é lido em um estilo de acesso aleatório”. Já para Wellisch (2000, p. 35) o índice é “um guia sistemático projetado para indicar assuntos, tópicos ou características de documentos ou suas partes, a fim de facilitar a recuperação de documentos”. Em resumo, o índice é um guia sistemático que se apresenta de forma organizada para facilitar a pesquisa do leitor, sendo assim considerado o “mapa” mais importante por indicar assuntos, tópicos ou características. Ele

tem a finalidade de localizar as informações e assim facilitar na recuperação de documentos ou informação.

Nota-se que a existência do índice é fundamental, pois a informação tende a ser dispersa em seu suporte como livro, base de dados e periódicos. Por essa razão, ele é considerado um sistema de orientação da literatura em dois aspectos, conforme citado pelos autores Borko e Bernier (1978, p. 5-6), a saber:

- a) **guia para trabalhos inteiros**: produzido para localizar os trabalhos inteiros e não tem a finalidade de diferenciar os trabalhos sobre o mesmo assunto, sendo mais amplo, a fim de orientar para o tema principal. Como exemplo, os cabeçalhos de assuntos elaborados pela biblioteca para identificar o assunto, o nome dos autores e das obras; e a classificação/sistema de notação, como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU) que agrupam e organizam logicamente categorias de obras;
- b) **guia para o conteúdo das obras**: para o leitor que deseja informações mais específicas do que a localização do livro na coleção em uma área ampla, existe o índice que identifica o conteúdo da obra, diferenciando-a das demais. Como exemplo, índices em trabalhos primários e secundários guiam pesquisadores para dados específicos dentro de obras, IFL, índices contendo as palavras-chave e índices coordenados/correlativos.

Tradicionalmente, o índice tem como função ser um sistema de orientação. Dessa forma, ele pode ser identificado em três momentos históricos: antes da invenção da imprensa, após a invenção da imprensa (a partir do século XV) e atualmente, na era tecnológica. Antes da invenção da imprensa, os índices eram os registros dos títulos dos livros encontrados nas bibliotecas dos mosteiros. Após a invenção da imprensa, são caracterizados pelo crescimento do volume da literatura e, conseqüentemente, há o aparecimento de várias listas com diversas finalidades (GOMES; GUSMÃO, 1983). E, atualmente, a inserção tecnológica facilitou a elaboração dos índices, com a criação de softwares de indexação automática que auxiliam na extração dos termos. A seguir é detalhado o histórico do índice.

3.3.1 A evolução do índice na Biblioteconomia

Nos tempos antigos, os índices não eram tão numerosos quanto são atualmente, considerando que, antes da invenção da imprensa, havia poucos livros e a leitura era uma habilidade incomum. Assim, para entender a essência desse produto, é evidenciada a etimologia da palavra, bem como seu percurso histórico.

3.3.1.1 Etimologia da palavra “índice”

A palavra *index* foi usada pela primeira vez pelos romanos para denotar descoberta, indicador, mensagem. Quando usada em relação à literatura, significa catálogo, lista, indicação ou até mesmo um título do livro. O uso do termo em inglês geralmente remete a *table of contents* ou a *literary guide* (BORKO; BERNIER, 1978; MONTEIRO, 1998), o que corrobora a ideia de Monteiro (1998, *on-line*), no relato que “o substantivo *index*, originou-se, no século XVI, da expressão *table of contents*, que, à época, significava uma lista, em ordem alfabética, das partes que formam um livro”.

Da definição em inglês pode-se inferir o equívoco, ainda hoje, no conhecimento popular em confundir “índice” com “sumário”. Mas as definições ganharam novas acepções, ficando claras as diferenças entre os dois termos. Originalmente, as palavras *index* e *table* eram usadas como sinônimos, e em meados do século XVII a primeira predomina (*index*). Atualmente, a palavra *table* (sumário) é utilizada para indicar as ordenações das seções dos livros e a palavra *index* (índice) são para os termos mais significativos contidos em uma obra (BORKO; BERNIER, 1978).

Dessa forma, essas palavras não são mais confundidas como sinônimas, pois a **ordenação** do sumário não é alfabética, indicando os níveis de cada título de um documento. Já o índice tem ordenação diversa, podendo ser alfabética, sistemática, cronológica, entre outros. Na **estrutura** do trabalho ou da obra, o sumário é localizado no início do documento (elemento pré-textual), e o índice é indicado no final do documento (elemento pós-textual), e no caso de índice de coleção de documentos é feito em volume separado. Na **elaboração** do sumário, considera-se a sequência dos títulos; o índice é elaborado pelo processo de indexação. Com a diferenciação dos termos expostos, apresenta-se o percurso histórico do índice.

3.3.1.2 Percurso histórico

Na literatura encontram-se autores que afirmam que a preocupação em localizar termos em livros sagrados (Bíblia e Alcorão) foi o estopim para o surgimento do índice. (BORKO; BERNIER, 1978; WEINBERG, 2000). Porém, outros autores indicam que os primeiros indícios apareceram em Roma e foram condicionados aos pergaminhos e às bibliotecas romanas (CISOMAI; MIHALCEA, 2007; GOMES; GUSMÃO, 1983). Neste trabalho, são apresentadas as duas vertentes de pensamentos.

Em 3100 a. C. surge o primeiro índice na forma mais antiga usada para armazenagem dos **cuneiformes da Mesopotâmia**⁹. Esses documentos eram envolvidos com uma cobertura sólida a fim de preservá-los contra a falsificação e na qual se registravam todas as informações necessárias para sua recuperação (GOMES; GUSMÃO, 1983, p. 12).

As descrições dos primeiros índices são encontradas frequentemente em **revistas teológicas**, uma vez que os índices forneciam acesso aos textos religiosos. As pessoas memorizavam a Bíblia e o Alcorão e conseguiam recitar longas passagens deles. O primeiro índice foi feito para a Bíblia, baseado no texto da escritura, elaborado por Eusébio, e dividia os Evangelhos em seções numerando-as e ordenando-as em dez Cânones (BORKO; BERNIER, 1978; GOMES; GUSMÃO, 1983; WEINBERG, 2000).

Outros autores afirmam que a primeira menção ao índice surge com os romanos indicando o índice aos **papiros**. Isso porque os índices eram anexados a rolos de papiros em bibliotecas romanas, contendo o nome do autor, o título do documento e, frequentemente, uma breve descrição que permitiria a identificação do pergaminho sem abri-lo (CISOMAI; MIHALCEA, 2007).

Em consonância com tal pensamento, as autoras Gomes e Gusmão (1983, p. 13) mencionam acontecimentos envolvendo os pergaminhos romanos, ao afirmarem que Sêneca se refere a um índice como uma lista, e Cícero, ao escrever para Áticus, pede que envie duas pessoas para restaurar seus livros e algum pergaminho para que pudesse elaborar índices. Outro exemplo é o catálogo dos fragmentos da Biblioteca de Alexandria de Calímaco, organizado em ordem alfabética dos autores, com indicação também dos assuntos mais gerais.

No século V, surge o trabalho que mais se aproximou do índice de assunto, no que se refere à estrutura e finalidade: a obra anônima *Apothegmata* apresenta uma lista de **provérbios gregos** sobre tópicos teológicos, que somente no século VI foi ordenada alfabeticamente. Na mesma época, outros autores dividiam suas obras em capítulos e seções numeradas, como, por exemplo, *Institutiones*, de Cassiodoros¹⁰, que remete não só ao número do capítulo, mas também ao título do capítulo.

Gomes e Gusmão (1983, p. 13) identificam a obra *Sacra paralela*, de John de Damascus, como a primeira tentativa em facilitar a leitura do usuário no índice, pois no início do texto aparece uma nota explicativa sobre a maneira mais fácil de encontrar a informação.

⁹ “**escrita cuneiforme** é, juntamente com os hieróglifos egípcios, o mais antigo tipo conhecido de escrita, proveniente da região da Mesopotâmia, há cerca de 3.500 anos antes de Cristo” (SILVA, 2018, *on-line*).

¹⁰ Escritor e estadista romano, conselheiro do rei ostrogodo Teodorico, o Grande, que se destacou pelos seus dotes jurídicos e literários e ocupou importantes cargos na administração pública ostrogoda da Itália.

Outro momento de identificação do índice aconteceu no século VIII: o monge, teólogo e historiador Beda junta uma relação de cerca de 40 obras à sua *Ecclesiastical History of Britain*.

Logo após a invenção da imprensa (século XV), Conrad Gessner prepara um repertório geral e europeu, denominado de “*Bibliotheca Universalis*”, relacionando cerca de 12 mil títulos de todos os livros latinos, gregos e hebraicos que eram de seu conhecimento, sendo assim considerada a maior obra nesse estilo compilada pelo homem, reconhecida também por ser a primeira tentativa de Controle Bibliográfico Universal (WELLISH, 1981). Posteriormente, publica o índice alfabético de assunto desse repertório com o nome de *Pandectarum sive partitionum universalium, libri XXI*. Gomes e Gusmão (1983, p. 13) afirmam que o “histórico da indexação teve seu início com a história da **bibliografia**”.

Por volta do século XVII, aparecem os primeiros **livros acadêmicos** com índices. Assim, em *Speed's History of Great Britaine* (1611), há um índice organizado alfabeticamente que contém os principais assuntos dessa obra. Já para índices em artigos de periódicos, William Frederick Poole é conhecido como o **primeiro inventor do índice para artigos de periódicos**, ao publicar *Poole's Index*, um tipo de índice utilizado para artigos de jornais cujas entradas são organizadas em ordem alfabética por assuntos, e os assuntos são as palavras-chave retiradas dos títulos dos artigos indexados.

Tal criação foi considerada importante, pois antecede e preconiza os atuais índices de Wilson, mas também antecipa a aprovação do *Keyword-in-Context* (KWIC), de Hans Peter Luhn. Em meados do século XIX, William Frederick Poole publica *An Alphabetical Index to Subjects Treated in the Reviews and other Periodicals*, da Universidade de Yale. As iniciativas dessas produções surgiram, porque as publicações periódicas estavam crescendo de forma acelerada, o que contribuiu para a abertura de novos horizontes no campo da indexação, havendo a necessidade de elaborar uma técnica para organização por assunto de conteúdo das publicações periódicas (BORKO; BERNIER, 1978, p. 6-8; GOMES; GUSMÃO, 1983, p. 13).

Em 1876, Charles Ammi Cutter codificou os princípios da catalogação de assuntos em *Rules for a Dictionary Catalog* (CUTTER, 1904), que foi publicado pela primeira vez em 1876 e a última edição em 1904. Sua importância se dá pelo pioneirismo na construção do método que também “mudou a lógica de organização, que até então tinha o critério sistêmico como foco principal, passando a utilizar também a ordem alfabética” (MARTINHO, GUEDES, [20--], *on-line*). No mesmo ano, houve a publicação do índice relativo da Classificação Decimal de Melvil Dewey, que, para Gomes e Gusmão (1983), foi quando as contribuições a esse campo se tornam mais expressivas.

Outro momento de destaque, conforme afirmado por Borko e Bernier (1978), foi em 1880, em que John Shaw Billings, do V.S. Army Medical Library, preparou o **primeiro**

catálogo de índice para a literatura médica. O surgimento das remissivas que ligam assuntos e outros correlatos iniciou-se em 1901. H. W. Wilson lança seu *Readers Guide to Periodical Literature*, onde cada artigo de um periódico era indexado pelo autor e pelo assunto e havia as remissivas, iniciando assim a grande época da indexação (GOMES; GUSMÃO, 1983). Posterior a tais fatos, o século XX foi marcado por ser uma

época que o conteúdo temático dos documentos começa a despertar interesse e com a “explosão da informação” há o aumento da produção bibliográfica e o conhecimento da necessidade do seu controle. A indexação começa a ser estudada mais profundamente surgindo, assim, como uma importante técnica de recuperação da informação (GOMES; GUSMÃO, 1983, p. 13-14).

Devido às mudanças que ocorreram durante os anos, é necessário abordar a diversidade dos tipos de enfoque e finalidade do índice para atender às necessidades informacionais. As variações são apresentadas no próximo tópico.

3.3.2 Tipos de enfoque e finalidade

Quanto aos tipos de enfoque, os índices são caracterizados pelos termos selecionados ou por sua finalidade. Os tipos mais comuns são os de índices de assunto, de autor e de citação. Os índices de títulos, onomástico e geográfico são apresentados brevemente, por não serem tão usuais.

Índice de assunto

É aquele que identifica as temáticas principais tratadas no documento e que representa as ideias do autor. Assunto é o foco de um trabalho para os quais a atenção e os esforços do autor foram direcionados. São os aspectos de um trabalho que contém novas ideias, explicações ou interpretações (BORKO; BERNIER, 1978). Para a ABNT (1992, p. 2) o Índice de assunto é uma “listagem alfabética ou sistemática de assuntos que indica a posição de cada assunto num documento ou numa coleção de documentos”.

Índice de autores

É considerado um índice de autor aquele que organiza as entradas por nomes dos autores citados nos documentos. Existem índices de autor pessoal e índices de autoria corporativa (indicam os nomes das organizações nas quais o trabalho foi realizado). Geralmente, os dois tipos são combinados no índice de autor.

Os índices de nomes diferem-se dos índices de autores. Os índices de autor fazem menção dos autores de vários itens no trabalho, por exemplo, os autores dos

capítulos, como ocorre no periódico. O índice de nome é um índice para os nomes das pessoas citadas no trabalho (BORKO; BERNIER, 1978; CUNHA; CAVALCANTI, 2008).

Índice de citação

A entrada do índice de citação é a indicação do documento precedida dos documentos os quais citaram o documento principal. Para exemplificar, na Figura 3 apresenta o exemplo do índice de citação, a entrada é o título do documento “Interfaces entre a ciência da informação e a ciência cognitiva”, e abaixo os documentos que o citaram “Citado por 105” ao clicar em “Citado por 105” fornece a lista de documentos que citaram o documento principal.

Figura 3 - Exemplo de Índice de citação

[HTML] Interfaces entre a ciência da informação e a ciência cognitiva
 GÂB Lima - Ciência da Informação, 2003 - SciELO Brasil
 Estudo panorâmico sobre aspectos da ciência da informação (CI) e da ciência cognitiva (CC), apontando recentes contribuições em quatro de suas possíveis interseções: categorização, indexação, recuperação da informação (RI) e interação homem-computador ...
 ☆ 77 Citado por 105 Artigos relacionados Todas as 11 versões ❧

Interfaces entre a ciência da informação e a ciência cognitiva

Pesquisar nos artigos de citação

[HTML] Ciência da informação e cognição humana: uma abordagem do processamento da informação
 DA Neves - Ciência da informação, 2006 - SciELO Brasil
 Apresenta as possibilidades de inter-relação entre a ciência da informação e os estudos relativos à cognição humana. Discute resultados de pesquisas sobre processamento da informação na psicologia cognitiva referentes ao desempenho dos indivíduos na leitura e ...
 ☆ 77 Citado por 74 Artigos relacionados Todas as 11 versões ❧

O enfoque social da segurança da informação
 JL Marciano, M Lima-Marques - Ciência da Informação, 2006 - SciELO Brasil
 O uso cada vez mais disseminado de sistemas informatizados integrados por meio de redes é um fato determinante da sociedade da informação. Este universo de conteúdos e continentes digitais está sujeito a várias ameaças que comprometem seriamente a segurança do complexo ...
 ☆ 77 Citado por 47 Artigos relacionados Todas as 12 versões ❧

[HTML] Gestão do conhecimento e a visão cognitiva dos repositórios institucionais
 GRB Fachin, J Stumm, RL Comarella... - ... em Ciência da ..., 2009 - SciELO Brasil
 Este artigo objetiva apresentar um estudo sobre os repositórios institucionais, sua relação com a gestão do conhecimento e com a visão cognitiva na formação dos mesmos. Apresenta um esquema da comunicação cognitiva entre usuários e os repositórios ...
 ☆ 77 Citado por 41 Artigos relacionados Todas as 7 versões ❧

Segurança da informação: uma abordagem social
 JLP Marciano - 2006 - repositorio.unb.br
 O uso cada vez mais disseminado de sistemas informatizados integrados por meio de redes é um fato determinante da Sociedade da Informação. Este universo de conteúdos e continentes digitais está sujeito a várias ameaças que comprometem seriamente a ...
 ☆ 77 Citado por 42 Artigos relacionados Todas as 7 versões ❧

Fonte: adaptada do Google Acadêmico, 2020.

Como mostrado anteriormente o índice de citação pode ser produzido automaticamente, não necessita do indexador, uma vez que os dados dos documentos citados estejam disponíveis em formato legível por máquina. Ele permite que os usuários iniciem uma pesquisa com documento conhecido e localizem outros documentos do mesmo assunto ou de um assunto relacionado. O índice de citações permite que um pesquisador obtenha respostas para questões como:

- a) que desenvolvimentos em um assunto específico ocorreram desde que o artigo X foi publicado?;
- b) que artigos recentes o autor do artigo X publicou?;
- c) quantos trabalhos citaram o artigo X?;
- d) quem citou o artigo X? (BORKO, BERNIER, 1978; WELLISCH, 2000).

Além dos tipos apresentados anteriormente, existe o **índice de títulos**, que é ordenado alfabeticamente, indicando os títulos das obras inseridas num documento ou numa bibliografia. Ocasionalmente, os índices de autores e de títulos são combinados num único índice. Existe também o **índice onomástico**, que contém os nomes de pessoas, locais e instituições. O **índice geográfico** inclui nomes de lugares, regiões, países e demais pontos geográficos. Já o índice topográfico inclui entradas relativas às características de uma região (CUNHA, 2008).

Os enfoques dos índices tratam-se do conteúdo que está inserido no índice. Mas eles podem ter diversas finalidades: base de dados ou tratados na web (índices unitermos, KWIC, KWOC e KWAC), organização em publicações periódicas (índice de coleções de documentos) e o tradicional IFL que é detalhado na seção 3.4.

Índice unitermos

Foi o primeiro índice não tradicional, que impulsionou a indexação coordenada, uma vez que o controle foi estabelecido por meio de uma unidade de termos. Sua importância está em ser o sistema antecessor dos termos e da recuperação de documentos computadorizada e sua aparição também coincidiu com o primeiro tesouro operacional (MOREIRO GONZALES, 2004, p. 73, tradução nossa).

Índice KWIC (Keyword-In-Context)

A sigla KWIC foi cunhada por H. P. Luhn, no artigo “*Keyword in Context Index for Technical Literature (KWIC Index)*” (LUHN, 1959), sendo definido como “um índice rotado, derivado, em sua forma mais comum, dos títulos de publicações. Cada palavra-chave que aparece num título torna-se ponto de entrada, destacada de alguma forma, realçada no centro da página” (LANCASTER, 2004, p. 54). Cada linha em um índice KWIC consiste em

três partes: a palavra-chave (que é o cabeçalho), o contexto (que funciona como uma modificação) e o código (que é a referência). Os índices KWIC têm todas as palavras significativas em ordem alfabética, de modo que cada palavra se torna um cabeçalho de índice e cada palavra é colocada adjacente à mesma sequência entre palavras-chave e seus contextos derivados de outros (BORKO; BERNIER, 1978). A seguir na Figura 4 apresenta o exemplo do índice KWIC.

Figura 4 - Exemplo de índice KWIC

	Columna de indización		
para la producción de aluminio en	Brasil*	*El consumo de energía eléctrica	3
de hierro y la economía de	Brasil*	*Las fundiciones	2
Las minas	brasileñas	de hierro	1
Las fundiciones de hierro y la	economía	española	2
Brasil* *El consumo de energía	eléctrica	para la producción de aluminio en	3
aluminio en Brasil* *El consumo de	energía	eléctrica para la producción de	3
Las	fundiciones	de hierro y la economía de Brasil	2
Las minas brasileñas de	hierro		1
Las fundiciones de	hierro	y la economía de Brasil	2
Las	minas	brasileñas de hierro	1
de energía eléctrica para la	producción	de aluminio en Brasil* *El consumo	3

Fonte: MOREIRO GONZALES, 2004, p. 72.

Índice KWOC (Key Word out of Context)

O índice KWOC é similar ao KWIC, exceto pelas palavras-chave que se tornam pontos de acesso e são repetidas fora do contexto, comumente destacadas na margem esquerda da página ou usadas como se fossem cabeçalhos de assuntos. O formato KWOC preserva o documento na íntegra; no entanto, ele não fornece o contexto adjacente conveniente que o formato KWIC faz. Um índice KWOC seria então aquele em que a palavra-chave usada como ponto de entrada não se repete no título, mas é substituída por um asterisco (*) ou outro símbolo (BORKO; BERNIER, 1978; LANCASTER, 2004). A seguir na Figura 5 apresenta o exemplo do índice KWOC.

Figura 5 - Exemplo de índice KWOC

Columna de indización		
Brasil	El consumo de energía eléctrica para la producción de aluminio en	3
	Las fundiciones de hierro y la economía de ...	2
brasileñas	Las minas de hierro	1
economía	Las fundiciones de hierro y la de Brasil	2
eléctrica	El consumo de energía para la producción de aluminio en Brasil	3
fundiciones	Las de hierro y la economía de Brasil	2
hierro	Las minas brasileñas de hierro	1
	Las fundiciones de hierro y la economía de Brasil	2
etc.		

Fonte: MOREIRO GONZALES, 2004, p. 72.

Índice KWAC (Keyword-Alongside-Context)

Alguns autores consideram que o índice KWAC é uma variante do KWOC, considerando que no KWOC o lugar do título que ocupa a palavra permutada se encontra por pontos suspensivos. No índice KWAC essa palavra é repetida no título sem qualquer omissão. Dessa forma, é obtido para permutação das palavras do título, no qual fornece o melhoramento das palavras-chave do título com palavras significativas adicionais extraídas do resumo do documento ou de seu conteúdo (MOREIRO GONZALES, 2004; FEITOSA, 2006; LIBRARIANSHIP, 2019). A seguir na Figura 6 apresenta o exemplo do índice KWAC.

Figura 6 - Exemplo de índice KWAC

Columna de indización Brasil	El consumo de energía eléctrica para la producción de aluminio en Brasil	3
	Las fundiciones de hierro y la economía de Brasil	2
brasileñas	Las minas brasileñas de hierro	1
economía	Las fundiciones de hierro y la economía de Brasil	2
etc.		

Fonte: MOREIRO GONZALES, 2004, p. 73.

Índice de coleções de documentos

Também chamado de índice corrente ou cumulativo por Feitosa (2006), são índices que contêm conjuntos de publicações periódicas. Considerado um produto importante para os pesquisadores que desejam pesquisar informações com eficiência e por atualizar informações constantes de determinados documentos. Reuni em uma só lista as entradas relativas a vários fascículos ou volumes (MOREIRO GONZALES, 2004, p. 64, tradução nossa; FEITOSA, 2006).

Os índices de coleção de documentos apresentam os mesmos objetivos e princípios dos índices dos livros, mas, devido à maior cobertura e complexidade, apresentam seus próprios problemas. Entre eles o rigor em atualizar publicações recentes com o intuito de cobertura completa, outro exemplo é acompanhar as alterações de edição relacionada a documentos individuais dentro da coleção e dessa forma atualizar o índice central. Como não é objetivo adentrar nesses problemas relacionados aos índices de coleção procede com a apresentação do índice de livro que será apresentado

separadamente, por ser o objeto de estudo desta dissertação, necessitando, assim, apresentá-lo mais detalhadamente.

3.4 Índice de Final de Livro

Em relação a sua nomenclatura, o IFL é referido na literatura a partir de diversas variações terminológicas. Dentre elas, estão: a) índice de forma geral (ARAÚJO, 2008; ABNT, 1989, 2004; COLLISON, 1971, FRANÇA *et al.*, 2014; MONTEIRO, 1998; TOMASI, 2008); b) índice de assunto (DIAS; NAVES, 2013); c) índice impresso (ROWLEY, 2002); d) Índice de Final de Livro (LANCASTER, 2004; CUNHA; CAVALCANTI, 2008); e) índices internos ou alfabético-remissivos (GOMES; GUSMÃO, 1983; FUJITA, 2003); f) índice alfabético (PIEIDADE, 1983). Neste trabalho optou-se pelo termo **Índice de Final de Livro (IFL)**, pois é o mais utilizado em publicações e reconhecido pela comunidade acadêmica.

O propósito fundamental do índice é a recuperação da informação, para atender as necessidades informacionais do leitor referente à localização de informações contidas num texto. Ele proporciona pesquisas mais assertivas em menos tempo, fornece ao leitor a fonte de informação adequada para soluções de problemas peculiares e permite que os leitores anteriores realoquem itens lembrados e novos leitores identifiquem passagens de interesse.

O IFL é um produto que apresenta a visão geral do conteúdo do livro, normalmente faz parte dos elementos pós-textuais. Consiste em uma lista dos itens relevantes abordados em um determinado livro e com indicação de quais páginas esses assuntos são tratados com relevância para o leitor. Em outras palavras, é uma lista estruturada de cabeçalhos e subcabeçalhos com indicadores que vinculam aos segmentos do livro (MEKKI; NAZARENKO, 2005; CSOMAI; MIHALCEA, 2007).

De acordo com Bonura (1994), os princípios e objetivos dos índices do livro são:

- Têm um começo e um fim;
- Se concentram em um tópico geral;
- Podem ser preparados inteiramente por uma pessoa;
- Não existe um bom livro se ele tem um índice fraco ou nenhum índice. Esses livros são incompletos e semelhantes aos livros publicados com erros, como páginas em branco onde o texto deveria estar;
- Ajuda o leitor que deve retornar ao livro mais tarde para procurar algumas informações;
- Ajuda o leitor que quer se referir a algo;
- Ajuda o leitor a determinar se o assunto desejado está no livro, em caso negativo pode dispensá-lo (BONURA, 1994, p. 14, tradução nossa).

Quanto às suas funções, Mulvany (1994, p. 5-6) cita algumas:

1. identificar e localizar a informação relevante contida no livro;
2. excluir assuntos que não oferecem informação significativa para o leitor;

3. analisar os assuntos contidos no documento, de modo a produzir cabeçalhos baseados em sua terminologia;
4. indicar relações entre os assuntos;
5. reunir informações sobre assuntos que estão espalhados pelo arranjo do documento;
6. direcionar o leitor a buscar entradas escolhidas para compor o índice, ao invés de sua linguagem;
7. sintetizar títulos e subtítulos em entradas;
8. organizar as entradas em uma ordem sistemática e útil.

Os **dois primeiros** itens exigem que o indexador julgue a diferença entre informações relevantes e irrelevantes para distinguir informação substantiva e menção passageira de um tópico, o que gera economia de tempo do leitor.

O **item 3** aborda duas funções importantes do índice. Primeiro, os assuntos devem ser identificados e analisados, e os assuntos encontrados no livro nem sempre são declarados textualmente. Por exemplo, em um livro sobre criação de cães, vários parágrafos podem ser dedicados a vários tipos de comida para cães e não havendo menção à palavra “nutrição”, no entanto, o assunto nutricional deve estar no índice. O indexador analisa o texto e identifica assuntos relevantes, sejam eles mencionados ou não. Dessa forma, utiliza-se a indexação por atribuição, ou seja, envolve-se a atribuição de termos a partir de uma fonte que não é o próprio documento. O segundo item orienta a usar a terminologia do documento. A linguagem do autor deve sempre ter preferência sobre termos atribuídos e a linguagem do texto, geralmente, é a que deve ser usada no índice.

Os **itens 4 e 5** relacionam-se com a construção da rede de inter-relacionamentos no índice, sendo que o **item 4** diz respeito à hierarquia básica da entrada de índice (cabeçalho com subcabeçalho) e indica a relação entre os assuntos. No entanto, o **item 5** “reúne informações sobre assuntos que estão espalhados pelo arranjo do documento” refere-se a um aspecto muito mais sutil dos índices. Quando os leitores procuram um termo específico, eles devem encontrar remissivas para todas as informações relevantes sobre esse termo. Identificar informações relacionadas e reuni-las em um local apropriado é um dos aspectos mais difíceis da indexação. Frequentemente, os indexadores inexperientes concentram-se nas minúcias do texto e negligencia as informações relevantes e o índice que contém informação dispersa diminui a confiança do leitor.

O **item 6** aborda o elemento crucial das remissivas no índice, especificamente a remissiva “ver” que existe para antecipar a linguagem dos leitores de índice, e concilia o texto do documento com a linguagem dos leitores. Quando faltam remissivas apropriadas, a busca pela informação é interrompida e a usabilidade do índice comprometida.

O **item 7** instrui o indexador a “sintetizar títulos e subtítulos em entradas”. A capacidade de o indexador sintetizar a entrada contribui grandemente para a integridade do

índice. Usualmente, esse esforço é simples quanto a resolver diferenças entre termos sinônimos. Outras vezes, o processo de síntese é mais complexo, quando a frase não está compreensível. Enquanto o item 7 se refere ao *design* de entradas de índice individuais, o **item 8** discute a organização de entradas de índice, e a ordem de organização mais comum para entradas de índice é alfabética.

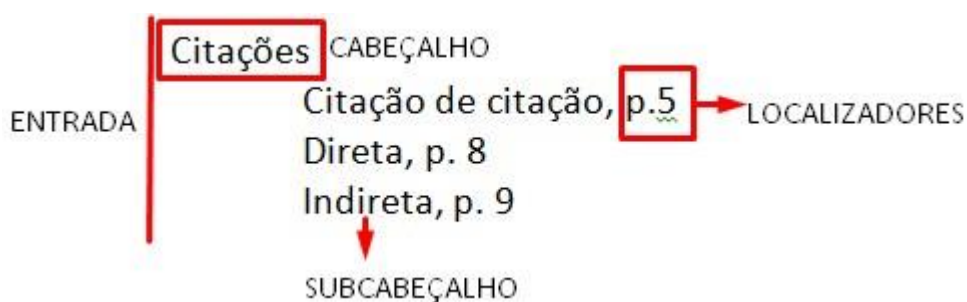
Cabe ressaltar que toda decisão feita pelo indexador necessita ser sistematicamente registrada na Política de Indexação; mesmo na elaboração do Índice, o indexador deve descrever e cumprir as delimitações que ela impõe aos profissionais (ISO, 1996; REDIGOLO, 2014). Gomes e Gusmão (1983) sugerem algumas informações que devem estar na Política de indexação do IFL (ver Apêndice B).

Enquanto a Política de indexação é o guia para o indexador, as Notas introdutórias no IFL constituem o guia para o leitor. E, estas devem fornecer orientação sobre as abordagens utilizadas que serão significativas para facilitar a busca informacional do leitor (BROWNE; JERMEY, 2007, p. 87). O IFL deve cobrir todo o assunto do Livro, as exclusões significativas devem ser registradas para conhecimento do leitor em uma nota introdutória (ISO, 1996, p. 7).

3.4.1 Estrutura do IFL

Um dos aspectos mais importantes na elaboração de um índice é sua estrutura. Ela é composta de cabeçalho principal, possivelmente acompanhado por subcabeçalhos e remissivas. A **entrada** do índice é composta pelo cabeçalho com seus localizadores, que geralmente são representados por números de página única, intervalos de números de páginas ou até mesmo números de seção, em que o número em negrito indica a referência mais importante. A organização das entradas permite a apresentação da relação semântica entre o cabeçalho e os subcabeçalhos, como a hiperonímia (vocábulo de sentido mais genérico e outro de sentido mais específico) (SYLVA, 2002; WU, et al, 2013), ilustrado no exemplo a seguir na Figura 7:

Figura 7 - Estrutura do Índice de Final de Livro



Fonte: elaborada pela autora.

A diferenciação de vários aspectos do cabeçalho se chama **subcabeçalho**. A criação do subcabeçalho é importante porque sua ausência implica que o leitor não estaria satisfeito por ter que consultar cada um dos localizadores de páginas para determinar aquelas que realmente lhe interessa (SYLVA, 2002).

Já para o leitor identificar os assuntos relacionados, como também o termo autorizado e não autorizado no livro se utilizam as **remissivas**. A remissa “ver também” faz comparações entre entradas semanticamente próximas, ampliando a pesquisa para o leitor. Para direcionar o leitor para a terminologia usada pelo autor emprega-se a remissiva “ver”, que conecta: 1) sinônimos ou expressões equivalentes; 2) termos populares a termo científico ou técnico; 3) termo antiquado a termo de uso atual; 4) sigla a nome completo da entidade (SYLVA, 2002; ABNT, 2004), conforme este exemplo na Figura 8:

Figura 8 - Exemplo de remissivas no IFL

Análise de assunto, p. 80, 153
ver linguagem documentária, p. 2-10
ver também tradução p. 15,19
 REMISSIVAS

Fonte: elaborado pela autora.

A principal função do IFL é reorganizar as informações do livro com o intuito de guiar o leitor para as páginas relevantes durante a sua consulta, dessa forma, não é apenas uma lista alfabética de termos e sua localização. Portanto, o índice de livros deve ser considerado um trabalho criativo do profissional que entende e analisa o conteúdo do livro. Ele deve ampliar a pesquisa do leitor ao apresentar os termos relacionados por meio da remissiva “ver também” e indicar os termos não utilizados no livro que, porém, podem ser pesquisados pelo usuário com a remissiva “ver” (controle de sinônimos).

Assim, a importância do índice está em como ele é organizado, ou seja, em sua estrutura. Ele deve conter todas as informações importantes do texto e, se não for organizado sistematicamente para facilitar o acesso, ele tem valor limitado para o leitor (MULVANY, 1994; DARTNALL, 2008; WU et al, 2013). Para atender à necessidade do usuário, é preciso que o índice seja estruturado com o propósito de cumprir os critérios de qualidade, conforme apresentado a seguir.

3.4.2 Critérios para IFL de qualidade

Os critérios de qualidade para o índice são apresentados na literatura pelos autores Bonura (1994) e Browne e Jerney (2007). Primeiramente apresenta-se a avaliação

dos critérios que são necessários para elaborar um IFL. E posteriormente listam-se os 15 critérios necessários para alcançar a Medalha *Wheatley* com o intuito de mostrar para os indexadores a meta que deve ser desejada durante a elaboração do IFL. Alguns aspectos devem ser avaliados para que um IFL seja considerado de qualidade, que são: precisão, profundidade na indexação, remissivas precisas, cabeçalhos lógicos, interoperabilidade e avaliação do índice.

Garantir a precisão é a característica mais importante de um bom índice. Nos *softwares* automáticos, os erros quanto à localização incorreta são menores, mas é importante certificar que as entradas e os localizadores estejam corretos. Uma verificação completa da precisão é impraticável, mas algumas verificações são necessárias.

Escolher a profundidade na indexação relaciona-se com a seleção dos números de entradas por página para compor o IFL, para essa seleção deve-se considerar o tipo de conteúdo (romance, técnico, entre outros) e o público-alvo. Existem três profundidades no IFL, que são: geral (uma a duas entradas); média (três a cinco entradas) e detalhada (no máximo dez entradas) por página (BONURA, 1994).

No contexto dessa pesquisa não se considera a indexação de documentos (coleção) e sim as páginas dos livros. A exaustividade no caso da profundidade no IFL relaciona-se com o item “detalhada” e a especificidade com a escolha do item “geral” e o intermediário é o item “média” que não se aloca totalmente em nenhum princípio de indexação.

Verificar as remissivas se estão exatamente iguais às entradas reais, pois perde-se a credibilidade quando existe remissiva que encaminha o leitor a um caminho sinuoso. Bonura (1994) afirma ser “pecado mortal da indexação” encaminhar o leitor para uma entrada que não existe. Remissivas precisas, lógicas e úteis são as marcas de um bom índice.

Selecionar os localizadores apropriados separa um indexador experiente de um novato, pois uma longa sequência de localizadores a uma única entrada constitui uma prática desagradável para os leitores, considerando uma procura de 15 páginas diferentes para encontrar uma informação específica. Qualquer assunto com mais de seis localizadores deve ser dividido em subcabeçalhos (BONURA, 1994).

Definir a interoperabilidade¹¹, pois, alguns índices de livros, serão reutilizados¹² de alguma forma e a qualidade da indexação original tem impacto significativo na facilidade de reutilização de índices. Sempre que possível, o profissional que criou o índice original deve ser o único a reformulá-lo. A interoperabilidade é importante em duas

¹¹ Interoperabilidade é a capacidade de dois sistemas trabalharem juntos, trocando informações e usando essas informações de forma eficaz (BROWNE; JERMEY, 2007)

¹² Embora a maioria dos IFL são produtos independentes.

áreas principais: 1) para índices inteiros que serão editados para uso em edições posteriores de um livro ou com versões atualizadas de serviços perdidos; 2) para partes de índices que serão reutilizados com partes do texto ou para índices inteiros que serão reutilizados em diferentes contextos.

Avaliar o índice pressupõe que esta atividade é realizada pelo indexador, mas ressalta-se a importância dos editores ou outros clientes, quando recebem o índice, avaliá-lo. Os autores Browne e Jermy (2007, p. 114) indicam nove itens que devem ser avaliados:

- 1 Verificar as entradas do índice conferindo se contem páginas relevantes;
- 2 Verificar a localização para inserir o índice no livro, garantindo que as seções relevantes do livro tenham sido incluídas no índice em tópicos apropriados e que nenhum fragmento importante esteja faltando;
- 3 Verificar o tamanho;
- 4 Verificar se as remissivas levam a entradas úteis;
- 5 Corrigir erros de ortografia e garantir que o estilo padrão foi seguido de forma consistente: uso de negrito e itálico, regras em número de página e ordenação;
- 6 Verificar se os números de página estão listados em ordem crescente;
- 7 Verificar se há listas longas de números, se houver, o editor pedirá ao indexador para criar subcabeçalhos;
- 8 Ler o índice para garantir que os cabeçalhos e subcabeçalhos descrevam seus tópicos de forma clara e sucinta.
- 9 Se houver vários índices, verifique se os termos apropriados estão em cada um.

A qualidade do IFL é discutida e estudada por diversas comunidades de indexadores¹³. Dentre alguns exemplos de reconhecimento está a Medalha *Wheatley* criada pela *Society of Indexers* e *Library Association* ao melhor IFL tendo por base o uso de 15 critérios (BONURA, 1994, p. 30), conforme identificado a seguir:

1. se houver uma nota introdutória, deve ser clara e bem expressa;
2. o índice é exato e os números de localização fornecidos no índice devem ser compatíveis com o texto;
3. itens significativos no texto devem constar no índice;
4. as entradas identificadas com a localização devem ser consistentes;
5. as entradas devem conter subcabeçalhos suficientes para evitar sequências de localização indiferenciadas (cabeçalhos com identificação de muitas páginas);
6. o índice deve ter uma estrutura de ordenação bem identificada e clara (ordem alfabética, cronológica, entre outros);

¹³ Conforme detalhado na seção “3.4.5 Grupos de discussão e o futuro do profissional”.

7. assuntos no texto devem ser representados no índice por cabeçalhos apropriados e bem escolhidos;
8. os termos são escolhidos de forma consistente;
9. existem remissivas suficientes para conectar itens relacionados no índice;
10. há remissivas para relacionar cabeçalhos desatualizados ou idiossincráticos no texto àqueles em uso atual;
11. o *layout* é claro e ajuda o usuário;
12. o índice é abrangente (embora certas limitações sobre a abrangência possam ser permitidas, se explicadas claramente);
13. o índice tem que ser adequado ao texto e não um produto para expressar as visões e os interesses do indexador;
14. se o índice se afasta das convenções, a decisão deve ser explicada na nota introdutória;
15. abreviações, acrônimos e similares devem ser explicados.

O intuito da Medalha *Wheatley* não é identificar somente a qualidade de um IFL, mas sim, um índice excelente que se destaca dos demais, trazendo algo especial que o indexador tenha desenvolvido para tornar o texto mais acessível possível ao leitor (SOCIETY, 2020). Assim, tais critérios podem ser utilizados na avaliação da prévia do índice, auxiliando na sua elaboração.

3.4.3 A análise de assunto e o controle terminológico no IFL

Nessa seção apresentam-se as a *análise de assunto* e o *controle terminológico* no IFL. Na subseção *Análise de assunto*, a ordem de apresentação é definida pelos estágios que lhe são correspondentes: compreensão do conteúdo do documento; identificação dos conceitos que representam o conteúdo, e seleção dos conceitos válidos para recuperação. Já a subseção de Controle terminológico aborda como ocorre esta etapa no IFL.

3.4.3.1 A análise de assunto

Segundo Naves (1996, p. 1), análise de assunto é “o processo de extrair conceitos que traduzam a essência de um documento”. Sendo considerada a operação base para todo procedimento de recuperação de informações (CESARINO; PINTO, 1980). Este termo possui variações terminológicas, como: análise temática, análise documentária, ou análise de conteúdo. Existem três conceitos ou pontos de vista diferentes sobre a análise de assunto: (a) concepção simplista; (b) concepção orientada para o conteúdo, e (c) concepção orientada pela demanda, conforme esquematizado no Quadro 2 e detalhado a seguir:

a) Concepção simplista: considera os assuntos como entidades simplistas absolutas, que podem derivar de uma abstração linguística do documento ou de dados que podem ser

somados. Lida com a informação explícita que é extraída do documento. De acordo com essa concepção, a indexação pode ser totalmente automatizada.

- b) Concepção orientada para o conteúdo:** envolve uma interpretação adicional do conteúdo, que vai além dos limites da estrutura léxica e gramatical, com o estabelecimento de assuntos que não estão explicitamente colocados no texto, mas que são facilmente identificados pelo indexador; envolve, portanto, uma abstração mais indireta do documento.
- c) Concepção orientada pela demanda:** considera o assunto numa perspectiva de transferência do conhecimento. Segundo essa concepção, os documentos são criados e deveriam ser tratados como instrumentos para transmissão de informações às pessoas interessadas. Ao analisar um documento, o indexador não deve se limitar a representar ou resumir apenas a informação explícita no documento. Mais do que isso, deve perguntar-se: como eu poderia tornar esse conteúdo, ou parte dele, visível para o usuário potencial? Que termos deveriam utilizar para levar esse conhecimento até o leitor interessado? (ALBRECHTSEN, 1993, p. 219; NAVES, 1996, p. 3)

Quadro 2 - Interconexão entre concepções de análise de assunto, tipos de informação de documentos e métodos de indexação

Concepções em análise de assunto e Indexação	Tipo de informação	Método de indexação
Concepção simplista	Informação explícita	Extração*
Concepção orientada para o conteúdo	Informação implícita	Atribuição‡
Concepção orientada pela demanda		

Fonte: Adaptado de Albrechtsen (1993, p. 220). Tradução nossa.

Legenda:

* palavras ou expressões que realmente ocorrem num documento são selecionadas para representar seu conteúdo temático.

‡ envolve a atribuição de termos a um documento a partir de uma fonte que não é o próprio documento (a utilização de vocabulário controlado, por exemplo).

Havendo apresentado a definição e as concepções da análise de assunto, é preciso entender os três estágios que são necessários para sua efetivação que, de acordo com Fujita (2003, p. 64) são:

- a) compreensão do conteúdo do documento;

- b) identificação dos conceitos que representam o conteúdo;
- c) seleção dos conceitos válidos para recuperação.

Estágio 1: compreensão do conteúdo

A **compreensão do conteúdo** do documento envolve a leitura do texto pelo indexador; para isto, é preciso entender a estrutura do texto. O trabalho de indexação é realizado no suporte onde se encontra a informação, dentre os suportes existentes há o **texto**, que é o foco desse trabalho. A definição de texto é geral e abstrata; para Dias e Naves (2013), trata-se de um veículo que permite a comunicação de ideias entre o sujeito que cria e dissemina informação (emissor, no caso, o autor), e o sujeito que adquire a informação (receptor, no caso, o leitor). Existem várias formas de comunicar as ideias através do texto, sendo que para esta pesquisa o foco é o texto escrito.

Para o bom desempenho da etapa de análise de assunto, é importante ter conhecimento da trilogia estrutural do texto, a saber: microestrutura, macroestrutura e superestrutura. A (1) microestrutura envolve uma estrutura superficial, representando a realidade física do texto e seus símbolos de significação, ou seja, as palavras; a (2) macroestrutura corresponde à estrutura global concebida como um tópico representativo hierárquico e coerente da unidade textual, envolvendo uma estrutura mínima da representação textual sintático-semântica; e a (3) superestrutura é a forma textual que organiza o conteúdo do texto: é a estrutura retórico-esquemática, um tipo de esquema de produção convencional para o qual o texto é adaptado, podendo ser considerado como transição entre as estruturas de superfície e de profundidade; fornece a forma global do conteúdo do texto: formas narrativa, expositiva, argumentativa, entre outros (DIAS; NAVES, 2013, p. 29).

A leitura integral do texto é fundamental para compreensão do conteúdo, e indispensável para elaboração do IFL (GOMES; GUSMÃO, 1983; UNISIST, 1983). Portanto, deve-se esclarecer o texto do autor através de ligações entre as ideias. Para que essas ligações aconteçam, é necessário entender o processamento de informações na mente do indexador. Existem dois modelos de processamento mental da informação: o *bottom-up* e o *top-down*, que parecem ocorrer simultaneamente na mente humana ao fazer a leitura de um texto. O modelo *bottom-up* pode ser definido como “ascendente, guiado por dados, indutivo, no qual a leitura é linear, com origem nas partes em direção ao todo textual” (NAVES, 2004 *apud* BORGES, 2009, p. 23). No que concerne ao modelo *top-down*, define-se como: “descendente, dedutivo, no qual se move na direção inversa, obtendo vantagem da base de conhecimento do leitor. Os dois modelos tratam-se de uma dupla ação: percepção e compreensão” (NAVES, 2004 *apud* BORGES, 2009, p. 23).

A compreensão na leitura possui três variáveis: texto, leitor e contexto; como pode ser visto na Figura 9. A variável **texto** tem “relação enquanto a forma, a estrutura textual e sobre o seu conteúdo como as regras gramaticais, de concordância sintática, semântica e a própria intenção do autor ao passar a mensagem ao interlocutor” (REDIGOLO; FUJITA, 2015, p. 360). A variável **leitor**, que nesse contexto é o indexador, diz respeito ao seu conhecimento prévio sobre o assunto e sua experiência nas áreas de assunto que esteja inserido, como também a identificação das tipologias textuais. Além dessas duas variáveis, identifica-se o **contexto** como influenciador na atividade da leitura profissional; nesta variável o leitor relaciona-se com o contexto social, psicológico e físico (REDIGOLO; FUJITA, 2015, p. 360).

Figura 9 - Modelo contemporâneo da compreensão na leitura



Fonte: adaptado de Giasson (1993, p. 21).

Após a compreensão do texto e realizada a leitura integral do mesmo, é preciso identificar os conceitos, que é o próximo estágio apresentado a seguir.

Estágio 2: identificação dos conceitos

A **identificação dos conceitos** representa o conteúdo do documento. Conceitos são “unidades do conhecimento, identificadas através de enunciados verdadeiros, sobre um item de referência e representados por uma forma verbal (termo ou palavra)” (DIAS; NAVES, 2013, p. 55). Para formar o conceito de algo se utiliza vários processos mentais, a saber:

Análise: operação mental que consiste em separar as partes, decompor, fragmentar um todo (objeto ou fenômeno) em seus elementos constituintes [...]

Síntese: operação mental que, ao contrário da análise, consiste em recompor um todo [...] a partir de seus elementos constituintes, a fim de compreendê-lo em sua totalidade.

Abstração: operação mental que consiste em isolar ou separar, para considera-lo à parte, um elemento ou parte de um todo que não é separável na realidade, a fim de distinguir o particular (acidental) do geral (essencial).

Generalização: operação mental que consiste em estender a toda uma classe de objetos ou fenômenos os elementos essenciais, gerais, universais, constatados num certo número de objetos ou fenômeno da mesma classe (DIAS; NAVES, 2013, p. 56).

É preciso considerar também o contexto no qual o documento é produzido, entendendo para que ele existe em determinado momento (DIAS; NAVES, 2013, p. 63). O contexto é definido por Cunha (2008, p. 105) como uma “palavra ou conceito que, apresentado em relação a outra palavra ou conceito, altera sua interpretação”. Para examinar o documento a fim de identificar os conceitos deve-se, além de entender o contexto, atentar para algumas partes do livro durante a leitura. Durante a leitura técnica para a catalogação, por exemplo, devem ser observados alguns capítulos com mais atenção. E, para o IFL, a leitura do livro é realizada de outra forma, como pode ser observado a seguir no Quadro 3. Existem opiniões distintas quanto às partes que devem ser indexadas no livro. Desta forma, acredita-se que o indexador deve deixar claro na Política de indexação tal decisão, e se resguardar indicando os autores que corroborem com sua decisão.

Quadro 3 - Elementos para a leitura de livros para elaborar IFL (continua)

PARTES DO LIVRO	FREQUÊNCIA COM QUE É INDEXADO	INFORMAÇÕES ADICIONAIS	FONTE (S)
DEDICATÓRIA	Não indexada	-	Mulvany (1994, p. 41)
AGRADECIMENTOS	Não indexados	-	Mulvany (1994, p. 41)
EPÍGRAFE	Não indexada	-	Mulvany (1994, p. 41)
APRESENTAÇÃO	Raramente indexada	Pode ser indexada somente se contiver informação substantiva não apresentada em outra parte do livro.	Browne e Jerney (2007, p. 38)
LISTAS DE ILUSTRAÇÕES OU TABELAS	Não indexadas	-	Mulvany (1994, p. 41)
SUMÁRIO	Não indexado	-	Mulvany (1994, p. 41)
INTRODUÇÃO	Geralmente indexada	Muitas vezes, é lida primeiro para obter uma visão geral, mas indexada por último quando o indexador sabe como seu conteúdo foi tratado em outro lugar.	Browne e Jerney (2007, p. 38); ISO (1996)
	Decisão tomada pelo indexador	O indexador deve decidir se essa parte do livro é relevante para a indexação.	Borko e Bernier (1978, p. 75)
NOTAS EXPLICATIVAS OU NOTAS DE RODAPÉ	Indexadas somente se fornecerem informações adicionais, ou seja, além das contidas no texto.	Elas não são indexadas se fornecerem apenas citações bibliográficas ou se o seu conteúdo for facilmente acessível no texto principal. As decisões dependem do espaço disponível para o IFL e a natureza do conteúdo.	Browne e Jerney (2007, p. 38); Mulvany (1994, p. 46)
ILUSTRAÇÕES	Decisão tomada pelo indexador	Indexar somente se oferecer informação adicional que não esteja no texto. Se o material da ilustração aparecer na mesma página que a discussão textual do material, geralmente não é necessário indexar a exibição individualmente.	Borko e Bernier (1978, p. 75); Mulvany (1994, p. 46)
REFERÊNCIAS E/OU BIBLIOGRAFIAS	Raramente indexadas	Somente se for índice de autores, os autores da bibliografia são indexados nas páginas em que foram citados.	Browne e Jerney (2007)
	Não indexadas	Como regra geral não se indexa o conteúdo da bibliografia.	Mulvany (1994, p. 43)
GLOSSÁRIO	Geralmente indexado	Indexar o termo como cabeçalho e a definição como subcabeçalho, desde que a definição ainda não tenha sido indexada no texto.	Browne e Jerney (2007, p. 38)
	Não indexado	Como regra geral, os glossários não são indexados.	Mulvany (1994, p. 43)

Quadro 3 - Passos para a leitura de livros para elaborar IFL (conclusão)

PARTES DO LIVRO	FREQUÊNCIA COM QUE É INDEXADO	INFORMAÇÕES ADICIONAIS	FONTE (S)
APÊNDICE(S)	Frequentemente indexado Depende do indexador e do conteúdo	Normalmente a indexação deve abranger informações extraídas do apêndice. Os termos indexados devem estar em nível mais amplo, e deve-se evitar termos específicos. Apêndices só devem ser indexados se contiverem informações adicionais que não esteja incluído no corpo do texto.	ISO (1996); ABNT (2002); Browne e Jerney (2007, p. 38) Borko e Bernier (1978, p. 75); Mulvany (1994, p. 44)
ANEXO(S)	Frequentemente indexado(s)	Normalmente indexado(s), devem abranger informações extraídas no anexo	ISO (1996); ABNT (2002)

Fonte: elaborado pela autora.

Estágio 3: seleção dos conceitos

O último estágio é a **seleção dos conceitos** válidos para recuperação. É importante que o indexador tenha claro os objetivos para os quais as informações são indexadas. Dessa forma, nem todos os conceitos identificados serão necessariamente selecionados (FUJITA, 2003). As habilidades cognitivas necessárias para uma boa seleção dos conceitos incluem a compreensão de leitura, a capacidade de conceituar e a capacidade de articular conceitos de maneira concisa e inteligível. Os autores Lancaster (2004) e Dias e Naves (2013) definem a atividade de seleção de conceitos com a determinação da *atênência*.

A tradução de atênência para o inglês é *aboutness*, havendo sinônimos como temacidade, concernência, entre outros. A pergunta fundamental para o *aboutness* é “sobre o que trata o documento?”. Há distinção entre os termos extensional (atênência) e intencional (significado): o primeiro é o assunto inerente ao documento, e o segundo é a razão pela qual o documento foi selecionado pela biblioteca ou pelo usuário (DIAS; NAVES, 2013). O significado é um dos problemas mais complexos para a linguagem documentária, relacionado à polissemia, homonímia, sinonímia e antonímia. Assim, considera-se a atênência e o significado no momento da seleção dos conceitos. Após esse trabalho intelectual são definidas as “frases de indexação” que serão inseridas no índice.

A seleção de termos desempenha um papel significativo na determinação da estrutura geral do índice. A formulação dos cabeçalhos em particular é de fundamental importância, uma vez que os cabeçalhos são os principais pontos de acesso ao índice onde o leitor realizará a primeira busca (BROWNE; JERMEY, 2007). Dessa forma, o indexador deve estruturar os termos considerando o leitor e o contexto. Para Redigolo (2014), é no estágio de seleção dos conceitos que se deve alcançar os níveis desejados de *exaustividade* e *especificidade*.

A profundidade da indexação é a “relação média geral do número de termos por documentos em um sistema documentário” (CHAUMIER, 1988, p. 63). A eficácia da análise de assunto ou sistema de indexação é controlada por dois parâmetros principais: a exaustividade e especificidade do termo, sendo que

A exaustividade da indexação é o grau de representação do assunto no texto que realmente são selecionados para o produto da indexação. Quando a indexação é exaustiva, frequentemente é atribuído um grande número de termos, incluindo assuntos periféricos. O contrário é obtido para indexação não exaustiva, na qual apenas os aspectos principais do conteúdo do documento são reconhecidos. A especificidade de termo refere-se ao grau de amplitude ou precisão dos termos (SALTON, 1988, p. 277 *apud* MULVANY, 1994. Tradução nossa).

Os IFL podem ser exaustivos e específicos. Quando um livro é catalogado para inclusão em uma biblioteca, o catalogador fica restrito à atribuição de três a cinco termos que descrevem o seu conteúdo. Em contraste, o IFL pode ter mais de dois mil termos que representam o assunto do livro. O IFL de um texto altamente especializado que será lido por leitores fluentes na linguagem técnica utilizada provavelmente será composto por vários termos específicos e precisos, enquanto um livro escrito para leitores diversos terá termos mais amplos. Mas, além de tais fatores que influenciam a profundidade no IFL, o indexador deve considerar as restrições impostas pelo cliente, o tempo disponível para indexação e o espaço designado para o índice no livro (MULVANY, 1994, p. 48-50).

3.4.3.2 Controle terminológico

O controle terminológico no IFL é diferente dos demais produtos da indexação, como por exemplo, os índices para catálogos bibliográficos. O controle utilizado no IFL não é a tradução¹⁴ para o vocabulário controlado, mas sim o controle terminológico que acontece dentro do próprio livro por meio da utilização de remissivas (*ver* e *ver também*). A remissiva “ver” tem o intuito de controlar os sinônimos em palavras que possivelmente o leitor possa pesquisar, mas com a utilização da remissiva o leitor é direcionado para o termo autorizado (termo utilizado pelo autor no livro). A remissiva “ver também” amplia a busca do leitor, apresentando além dos termos que ele procura, outros termos relacionados que o autor utilizou.

O indexador deve utilizar instrumentos (tesauro, dicionário, lista de cabeçalho de assunto, entre outros) para auxiliar no trabalho de selecionar sinônimos e relacionamentos de termos no IFL. O *tesauro* é ressaltado por Francis e Greenway (2015) como grande auxiliar neste processo. A ISO (1996) destaca a consulta com especialistas como recurso confiável de indexação. Estes instrumentos e recursos funcionam como garantia literária para que o indexador possa ter fontes confiáveis para tomada de decisões durante o controle terminológico. Para Fugmann (1994), o uso desses instrumentos é uma condição indispensável para qualidade na indexação, embora afirme que estes normalmente não são utilizados, pois demanda muito tempo para seu uso.

Deve-se ressaltar que o domínio e contexto de análise no IFL é o próprio livro, a linguagem do autor. A linguagem natural utilizada pelos autores de livros possui variações semânticas, por isso é indispensável o controle na utilização dos termos que integrarão o IFL.

¹⁴ consiste na representação de conceitos por termos de uma linguagem de indexação” (FUJITA, 2003, p. 63)

3.4.4 A tríade de profissionais para elaborar o índice: indexador, autor e editor

Comumente, o índice é elaborado pelo autor ou indexador. Borko e Bernier (1978) descrevem o indexador como o profissional que elabora o índice, sendo o autor e editor o suporte para efetivação do trabalho. Contudo, Mulvany (1994), Bonura (1994) e Browne e Jermeý (2007) apresentam o cenário com todos esses profissionais em atuação na criação do IFL.

Diferentes profissionais no protagonismo da indexação indicam pontos fortes e fracos na preparação do IFL de acordo com cada perfil. O indexador é o profissional formado em biblioteconomia que durante sua graduação obteve formação especializada de teoria e prática para analisar e indexar documentos. Os pontos fortes desse profissional para construir o IFL são decorrentes da experiência e da objetividade, o que o capacita a contabilizar o tempo gasto de produção, conhecendo o formato que geralmente agrada o editor e examinando o livro como um todo, ou seja, sendo objetivo na indexação, além de adequar a linguagem do texto à linguagem do leitor. Porém, o ponto fraco do indexador é a falta de familiaridade com o assunto – é importante destacar que essa afirmação depende do conteúdo do documento. Os livros escritos para uma audiência geral podem ser indexados por qualquer indexador qualificado. Porém, livros com assuntos especializados dependem de profissional que compreende a terminologia da área, como os livros de medicina, de economia e de biologia (MULVANY, 1994).

Autores podem escolher indexar seu próprio livro, ocasionando em pontos fortes e fracos. Obviamente, a familiaridade com o texto e com o assunto são as principais vantagens desse profissional. Porém, dois fatores que dificultam a preparação do IFL por esse profissional: 1) falta de conhecimento do processo de indexação; 2) esgotamento e muito envolvimento com o livro quando ele está pronto para a indexação. Muitos editores enviam panfletos aos autores descrevendo o processo da indexação de maneira didática, com o intuito que aprendam as técnicas da indexação. Contudo, professores que formam bibliotecários afirmam que, mesmo com disciplinas exclusivas de indexação, os alunos aprendem as regras, mas não elaboram um índice adequado.

Mulvany (1994) afirma que o profissional que elabora o índice deve possuir conhecimento sobre indexação, porém dizer que a indexação é melhor abordada pelo indexador profissional não é justo (MULVANY, 1994; BROWNE; JERMEY, 2007). Bonura (1994) ressalta que, para criar um índice, o mais adequado seria os autores com capacidade de serem objetivos sobre seu trabalho e que compreendem o que é um bom índice e que dominam as técnicas da indexação. Porém, isso é limitar ainda mais o universo de profissionais que poderiam exercer essas atividades. Dessa forma, pode-se inferir que a relação entre os diferentes profissionais é a melhor forma de criar o IFL adequado.

Nesse sentido, Mulvany (1994, p. 31, tradução nossa) reafirma que o “trabalho cooperativo entre o autor e o indexador pode ser extremamente benéfico. O autor que está disponível para consulta durante o processo da indexação pode ser um recurso técnico muito valioso para o indexador”. Raramente encontram-se autores que se especializam em indexação. Por essa razão, cada profissional deve estar ciente de seu papel na elaboração do IFL. Para a referida autora, o equilíbrio da participação do autor e do editor deve ser destacado. O ideal é que o indexador contate o autor quando necessário e na revisão final do índice, considerando que se o envio fosse em todas as fases de produção poderia prejudicar o andamento do trabalho do indexador, pois a elaboração do IFL não é linear, as escolhas do termo hoje não significam que será mantido até o final. Esse fluxo de trabalho ocasiona habilidade de gerenciar as relações. A *American Society for Indexing* (ASI) indica as responsabilidades éticas para indexadores, editores e leitores de índice. Borko e Bernier (1978, p. 225-226) parafraseiam as indicações da ASI, apresentando as atitudes que devem ser observadas para aprimoramento, conforme descrito no Quadro 4.

Quadro 4 - Responsabilidades éticas indicadas pela ASI para elaboração de IFL

INDEXADORES	EDITORES
<ul style="list-style-type: none"> - planejar, especificar e manter a qualidade do índice; - realizar o trabalho de maneira profissional; - reduzir dispersão de entradas semelhantes no índice; - fornecer remissivas no índice, em caso de ausência, somente na garantia que o pesquisador compensará esta ausência. 	<ul style="list-style-type: none"> - avaliar o resultado do índice para o usuário para evitar perda de dados, dinheiro e tempo; - indicar claramente todas as especificações do índice em publicidade, catálogos e documentação do índice; - especificar todas as precauções necessárias para o uso efetivo do índice.
LEITORES	INDEXADORES E EDITORES
<ul style="list-style-type: none"> - estudar e compreender a estrutura e organização do índice para que seja usado como pretendido; - relatar os erros para os indexadores e editores de índice, bem como a outros usuários por meio de mídia que informa sobre índices e indexação. 	<ul style="list-style-type: none"> - criar índices de forma precisa, clara e sucinta; - fornecer erratas (uma lista de erros no índice) e trazê-las à atenção do usuário por meio de inserções, acumulações, a próxima edição ou reimpressões ou no próximo índice da série.

Fonte: adaptado de Borko e Bernier (1978, p. 225-226).

O bibliotecário que deseja aprofundar na indexação e desenvolver índices deve observar as características necessárias para desenvolver o trabalho. O indexador deve apresentar as seguintes características, como indicado a seguir:

- a) conhecimento sólido dos princípios básicos da indexação;
- b) amplitude de conhecimento;
- c) entendimento no campo de assunto;
- d) conhecimento de publicação e tipografia;
- e) interesse na área;

- f) compreensão das necessidades tanto do autor quanto dos leitores;
- g) capacidade de avaliar e ajustar sistemas automáticos;
- h) habilidade em gerenciamento de projeto;
- i) conhecimento de questões de usabilidade e habilidades em testes de usuários;
- j) objetivo;
- k) motivado;
- l) observar o livro como um todo (BORKO; BERNIER , 1978; MULVANY, 1994; BONURA, 1994; BROWNE; JERMEY, 2007).

3.4.5 Grupos de discussão e o futuro do profissional

As atividades de indexação como profissão ainda são muito restritas. Dessa forma, a rede de profissionais formada por indexadores é utilizada para suporte nos trabalhos e aperfeiçoamento. Internacionalmente são identificadas “sociedades de indexadores” que reúnem os profissionais do país, a exemplo do Canadá, e sua *Société Canadienne d’Indexation* (ISC/SCI) (INDEXERS, 2020), da Alemanha, com o *Deutsches Netzwerk der Indexer* (DNI) (DEUTSCHES, [20--]), entre outros. Porém, no Brasil não foi identificado nenhum grupo nacional especialista em indexação, talvez pela formação generalista nas escolas de biblioteconomia. Contudo, é possível acompanhar os grupos¹⁵ que debatem sobre indexação e tópicos relacionados, a saber:

- a) aliaINDEXERS: iniciada pela Associação Australiana de Bibliotecas e Informações (ALIA), tem o foco de atender indexadores australianos, mas não isenta a participação de profissionais de outras nacionalidades;
- b) America On-line (AOL): agrupa as discussões sobre indexação em subtópicos, como indexação com Cindex e Macrex e outros;
- c) Asiconference: os membros podem expressar opiniões e ideias para as próximas conferências nacionais (reuniões anuais);
- d) ASI-L: fórum eletrônico gerenciado para discussão de tópicos relacionados a índices, incluindo assuntos pertinentes à ASI e seus membros;
- e) *Black Writers United*: criada para apoiar os escritores de literatura afro, facilitando o compartilhamento de notícias, eventos e informações entre os membros da comunidade de escritores;
- f) Cindexusers: fórum de discussão para os usuários do software de indexação automática Cindex, que incentiva a troca de dicas, estratégias e técnicas entre os usuários;
- g) CompuServe: o fórum de discussão para indexadores que trabalham em home office (GO WORK) e o Fórum Literário (GO LITFORUM);

¹⁵ A lista dos grupos indicando a forma de contato pode ser visualizada no site da ASI (2020).

- h) *Digital Publications Indexing*: fórum aberto para indexadores profissionais especializados em escrever índices para publicações digitais;
- i) IndexCafe: grupo no qual os indexadores, novos e experientes, podem fortalecer seus laços sociais na comunidade de indexação, seu o objetivo é diminuir os efeitos do isolamento que geralmente os profissionais vivenciam;
- j) *InDesign Indexing*: grupo com foco no compartilhamento de dicas em relação aos recursos de indexação do software *Adobe CreativeSuite InDesign* para a produção de índices;
- k) *Indexers' Discussion Group*: grupo aberto para discussão de indexadores de todos os níveis e interesses;
- l) Indexing-PNW: indexadores profissionais, estudantes e outros interessados em indexar do Noroeste do Pacífico (Estados Unidos e Canadá), aqueles de fora dessa região também podem participar;
- m) INDEX-L: criada para promover uma boa prática de indexação, fornecendo um fórum para os indexadores profissionais e aspirantes compartilharem informações e ideias relacionadas a todos os aspectos da preparação do índice;
- n) IndexPeers: indexadores que se voluntariam para revisar os índices editados entre os pares;
- o) *IndexStudents*: pessoas interessadas em discutir questões relacionadas ao aprendizado para indexar livros, revistas, bancos de dados e a web;
- p) MacIndexers: para quem indexa na plataforma Mac, o objetivo é fornecer apoio, informação, oportunidades de networking e incentivo aos indexadores do Mac;
- q) Macrex: fórum para discussão sobre todas as questões relacionadas à indexação com Macrex;
- r) *Scholarly Indexing*: fórum aberto para indexadores profissionais especializados em indexar materiais publicados por editoras acadêmicas ou publicações destinadas a acadêmicos, pesquisadores e estudantes de disciplinas acadêmicas específicas;
- s) *SKY Index Users*: é o grupo oficial dos usuários do SKY Index para obter e fornecer informações úteis sobre o programa;
- t) *WordIndexers*: discussão para pessoas que criam índices incorporados no Microsoft Word.

Browne e Jerney (2007, p. 195-201) debatem sobre situações que “prejudicam” o trabalho do indexador, bem como as perspectivas de futuro do profissional. Nesse sentido, embora o auxílio do computador tenha destaque nos últimos anos na área da indexação, inclusive com discussões acerca da substituição dos indexadores pela máquina, o uso desse equipamento para realizar indexação automática é destacado como “ameaça”, por ser mais barata que o indexador humano. Outro obstáculo é o declínio de indexação.

Browne e Jerney (2007) afirmam que o computador ainda não alcançou a qualidade do índice gerado pelo indexador. Inclusive, a maioria dos *softwares* de indexação automática tem a necessidade de mediação humana, não sendo totalmente mecânicos. Tais acontecimentos servem de reflexão para mudanças na forma de trabalho, em que a habilidade de ter o olhar clínico quanto aos índices produzidos a partir de softwares automáticos é necessária.

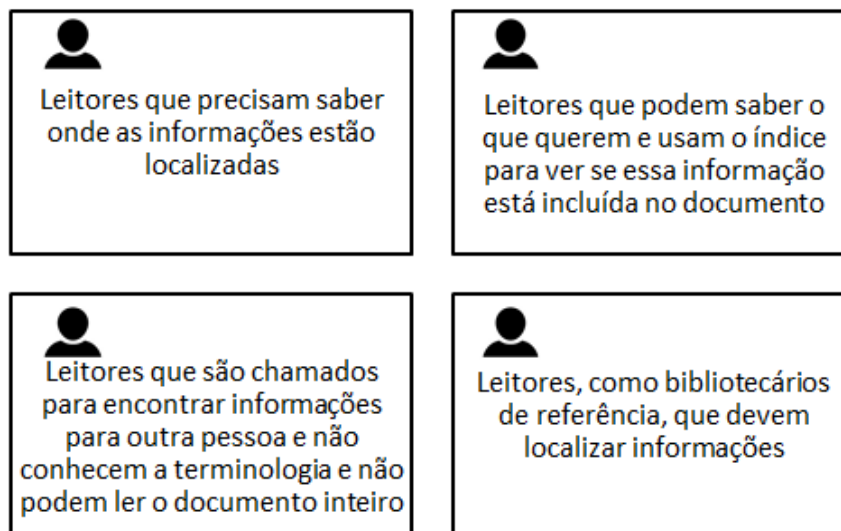
Outro aspecto a que os indexadores devem estar atentos é quanto ao declínio da indexação, da forma tradicionalmente conhecida. Os três principais suportes que o índice é utilizado são: bancos de dados bibliográficos, periódicos e livros. As duas primeiras reduziram os trabalhos de indexação, principalmente pela criação da web, pois as buscas são eletronicamente pesquisáveis, dispensando a criação de índices impressos, apesar de alguns periódicos mesmo on-line mantiveram seus índices. Mas das três áreas a indexação de livros é que mais se manteve. Os livros impressos provavelmente exigirão índices construídos pelo indexador por algum tempo, sendo a indexação de livros ser a mais provável continuar com força, já que a granularidade das informações (a importância do acesso a pequenos detalhes) torna a tarefa mais difícil de automatizar (BROWNE; JERNEY, 2007).

3.4.6 O leitor do IFL

Todas as funções do índice têm o propósito de atender ao seu leitor. Reconhecer o público que utiliza o IFL como produto é fundamental no momento de sua construção. Porém, os indexadores raramente conhecem o público ou recebem algum *feedback* deles. Os indexadores dependem de editores e autores para dizer a eles o máximo possível sobre os leitores esperados, mas também precisam tomar decisões baseadas em suposições sobre o texto. A maioria dos livros tem uma ampla gama de usuários, por isso é difícil direcionar abordagens específicas (BROWNE; JERNEY, 2007).

O usuário do índice é geralmente muito diversificado, dessa forma, ao se elaborar um índice, deve-se saber qual tipo de leitores utilizará o IFL isso fará com que o índice seja mais completo e utilizável, na Figura 10 apresenta características dos possíveis leitores desse índice.

Figura 10 - Leitores do Índice de Final de Livro



Fonte: adaptada de Bonura, 1994, p. 19, tradução nossa.

Esses leitores podem ser divididos em duas categorias: aqueles que leram o livro e aqueles que não leram. Os autores Bonura (1994) e Browne e Jerney (2007) relatam que os **leitores que leram o material** estão mais familiarizados com a terminologia e provavelmente procurarão os termos usados pelo autor. Essa é uma das razões pelas quais é importante que o indexador utilize a terminologia do autor. Contudo, para os usuários que **não leram o livro**, pode-se usar ampla variedade de termos e assim ser considerado mais difícil atender suas necessidades, pois a dificuldade está no indexador antecipar a linguagem e as expectativas de diferentes leitores, considerando que essa linguagem pode diferir da utilizada pelo autor. É importante mencionar que, para Mulvany (1994, p. 9), “a maioria dos usuários do índice não terá lido o livro em sua totalidade. Isto é particularmente verdadeiro nos livros de referência. Por sua própria natureza livros de referência são projetados para serem referidos, não para serem lidos diretamente”. Descobertas dos projetos de pesquisa de “Christine Ryan e Sandra Henselmeier (2000), Susan Olason (2000), Cecilia Wittmann (1990) e Corinne Jorgensen e Elizabeth Liddy (1996)” (BROWNE; JERNEY; 2007, p. 19). afirmam que, para os usuários, o IFL costuma ser difícil de ler e confuso de usar. A pesquisa mostrou que os usuários frequentemente:

- não compreende índices;
- não gosta de remissivas;
- quer mais termos alternativos (mas não gosta de remissivas!);
- realizam pesquisas mais amplas dos que os termos encontrados no índice;
- um melhor desempenho quando subcabeçalhos não comecem com palavras “pequenas”, como “no” e “de”;
- melhor desempenho com subcabeçalhos recuados (parágrafo);
- não leem notas introdutórias;
- precisam de ajuda para distinguir entre os cabeçalhos principais e

subcabeçalhos (BROWNE; JERMEY; 2007, p. 19).

Como solução, Browne e Jermeý (2007) sugerem algumas atitudes que os indexadores devem fazer para tornar os índices mais fáceis de usar, como:

- a) envolver os usuários no planejamento e avaliação de índices;
- b) usar variação de fonte para orientar usuários;
- c) focar menos em questões menores e considerar mais a experiência geral;
- d) oferecer treinamento de técnicas de busca nos índices;
- e) integrar tesouros para levar os usuários a termos apropriados.

Considerando a importância do índice e sua relevância para o leitor, apresentam-se os procedimentos metodológicos realizados neste trabalho com o intuito de facilitar a elaboração do índice.

4. CARACTERIZAÇÃO E DESCRIÇÃO DA METODOLOGIA

“É bastante fácil fazer um índice, como é fazer uma vassoura de retalhos, tão basta quanto palha de aveia; porém fazer um índice bem feito e que varrerá bem nos cantos não é tão fácil” (John Ruskin).

Em sintonia com as ideias apresentadas e discutidas nos capítulos anteriores, a presente pesquisa tem por base investigar e responder à questão de pesquisa: “**quais são as diretrizes necessárias para elaboração e estruturação semiautomática de Índice de Final de Livro?**”. Esse capítulo está estruturado na seguinte ordem: (4.1) Caracterização da pesquisa; (4.2) Delimitação do universo de pesquisa; (4.3) Descrição dos procedimentos metodológicos.

4.1 Caracterização da pesquisa

Tendo em vista os objetivos traçados, bem como a natureza do objeto de estudo – o IFL, esta pesquisa está caracterizada, conforme orienta Gil (2008), quanto ao nível, finalidade, tipo de pesquisa e tipo de análise:

- a) quanto ao **nível**, a pesquisa é exploratória e descritiva, pois tem o propósito de fornecer uma visão geral das diretrizes relativas à elaboração do índice, mas também pretende descrever as características do índice produzido por meio da indexação semiautomática;
- b) Quanto à **finalidade**, trata-se de uma pesquisa aplicada, visto que as diretrizes propostas se destinam a trazer consequências práticas do conhecimento anterior para aplicações futuras;
- c) Quanto ao **delineamento da pesquisa**, entende-se que seja uma pesquisa bibliográfica e documental, pois esses foram os meios utilizados para a coleta de dados;
- d) Quanto à **Análise**, entende-se que a pesquisa é qualitativa, uma vez que na reflexão das publicações utiliza-se a análise de conteúdo para auxiliar nesse processo.

Tendo em vista a caracterização da pesquisa, faz-se necessário apresentar a delimitação do universo dos dados, e assim compreender o objeto, delimitação e recorte utilizados nessa pesquisa.

4.2 Delimitação do universo de pesquisa

Ao delimitar o universo da pesquisa, é importante definir os limites da investigação. Dessa maneira, o **objeto** deste trabalho são as diretrizes para elaboração e estruturação do IFL. Conforme exposto nos objetivos, essas diretrizes exigem a consulta a fontes científico-acadêmicas. Nessa lógica, toda a base de estudo desta pesquisa é constituída pela literatura da área no que se refere a IFL. A **delimitação** diz respeito às publicações que relatam sobre a construção do IFL, excluindo-se todas as publicações que

não tinham esse foco. O **recorte** é referente à tipologia a ser utilizada, que são as normas internacionais (ISO) e as nacionais (NBR); e publicações da revisão de literatura, abordando artigos de periódicos científicos e livros que retratam as etapas necessárias para a construção do IFL.

4.3 Descrição dos procedimentos metodológicos

As primeiras tentativas de análise de conteúdo surgiram quando o homem começou a interpretar os livros sagrados, tendo sido sistematizada como método apenas na década de 1920, por Leavell (BARDIN, 2016), mas a comunidade acadêmica conheceu o método somente em 1977 com a publicação da obra "*L'Analyse de Contenu*". A obra, escrita por Bardin, descrevia o método em detalhes que servem de orientação até os dias de hoje para descrição de conteúdos. Bardin define a *análise de conteúdo* como:

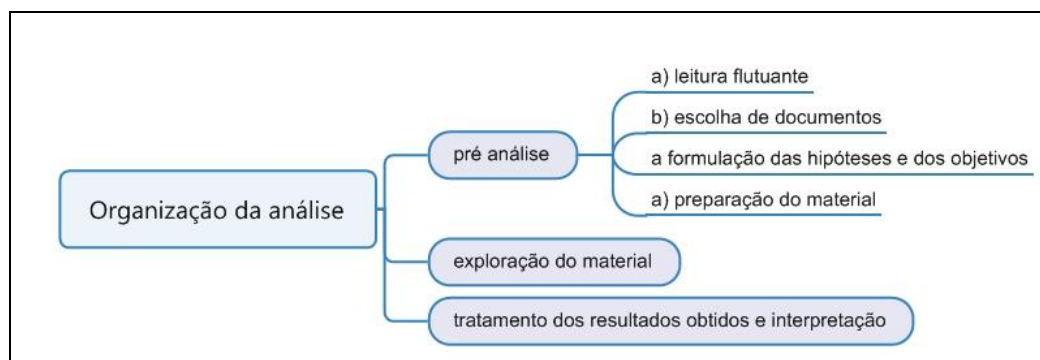
um conjunto de técnicas de análise das comunicações visando a obter, por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens (BARDIN, 2016, p. 48).

Por se tratar de uma técnica metodológica, a análise de conteúdo pode ser aplicada em diversos discursos e a todas as formas de comunicação, seja qual for a natureza do seu suporte. Nessa análise, o pesquisador busca compreender as características, estruturas ou modelos que estão por trás dos fragmentos de mensagens levados em consideração (CÂMARA, 2013). Neste estudo, o domínio de aplicação é o suporte linguístico escrito, cuja base é a literatura da área, facilitando a leitura das publicações e o delineamento das recomendações a serem selecionadas. De modo geral, os procedimentos adotados neste trabalho estão em consonância com as orientações de Bardin (2016) para a análise de conteúdo.

Para efetivação desse tipo de análise, as fases propostas por Bardin (2016) são: a organização da análise; a codificação; a categorização; a inferência, e o tratamento informático. Optou-se neste estudo por utilizar a primeira etapa "Organização das análises", uma vez que nesta etapa é possível responder às questões desta pesquisa, pois, ao explorar o material, descrevem-se os elementos que são buscados (diretrizes para IFL), assim como, ao final, são feitas análises interpretativas e críticas dos resultados obtidos.

A etapa de organização da análise compreende as formas de organizar o conteúdo para ser analisado e subdivide-se em três fases: 1) a pré análise, 2) a exploração do material, e 3) o tratamento dos resultados obtidos e interpretação, que serão detalhados a seguir na Figura 11:

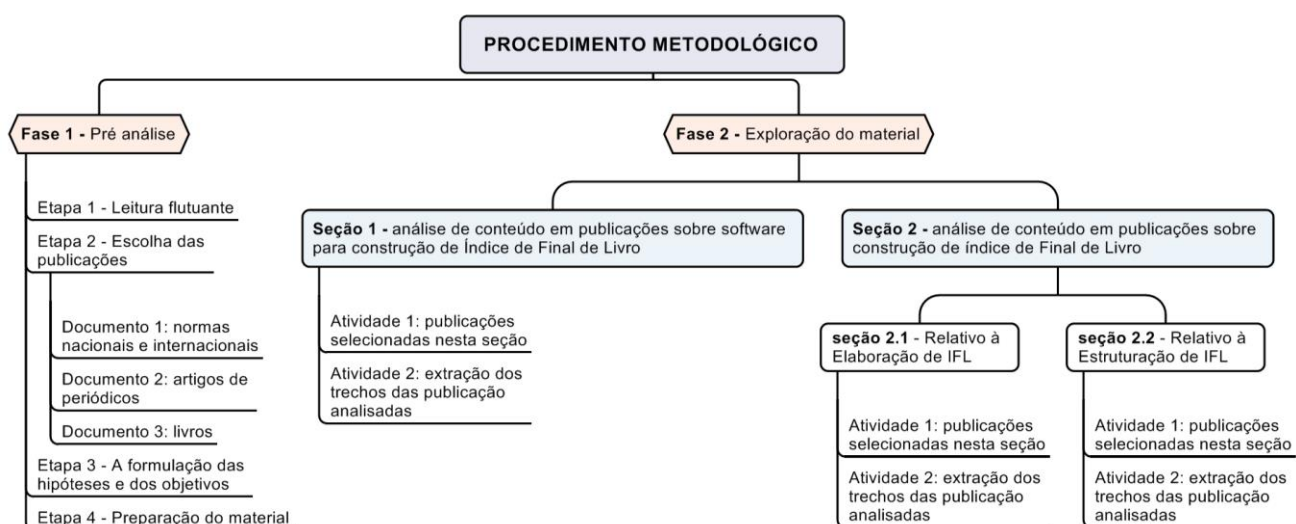
Figura 11 - Organização da análise: fases e atividades*



Fonte: elaborado pela autora.

A primeira e a segunda fase são apresentadas neste capítulo, pois as atividades previstas nestas estão relacionadas à metodologia. A terceira fase está prevista no capítulo 5, pois o tratamento e interpretação dos elementos semânticos extraídos são resultados. As fases de pré-análise e exploração do material foram subdivididas em etapas e atividades, conforme demonstrado na Figura 12.

Figura 12 - Procedimentos da metodologia: pré análise e exploração do material



Fonte: elaborado pela autora.

Em linhas gerais, o esquema da “Análise de Conteúdo” sintetiza os instrumentos adotados para realizar o trabalho prático. Na sequência, será descrita a “Pré-análise”, primeira fase do referido método.

4.3.1 Pré-análise

Neste tópico são retratadas as etapas que envolvem a fase de pré-análise, cujo intuito é identificar e selecionar as publicações que integram esta pesquisa, quais sejam: 1) leitura flutuante; 2) escolha das publicações; 3) formulação das hipóteses e dos objetivos e; 4) preparação do material.

Etapa 1 - leitura flutuante

Etapa que “consiste em estabelecer contato com os documentos a analisar e em conhecer o texto deixando-se invadir por impressões e orientações” (BARDIN, 2016, p. 126). Assim, toma-se o universo que deve ser mais claramente analisado, compreendendo as nuances de publicações análogas e vislumbrando as possíveis técnicas a serem utilizadas.

Na leitura flutuante, trabalhou-se três tipos de publicações. **Primeiro**, as regras de instituições nacionais e internacionais na área de biblioteconomia: a Norma Brasileira (NBR), e a *International Organization for Standardization* (ISO), que rege a padronização e normatização em nível mundial. O **segundo** tipo documental são os artigos de periódicos empregados na Revisão de literatura e identificados na Tabela 6. Tal escolha justifica-se por terem sido produzidos a partir de práticas dos profissionais e terem a finalidade de expor informações recentes. E, por último o **terceiro** tipo foram livros com diferentes cronologias explicitando as etapas para construção do IFL, por serem mais didáticos, facilitando a rotina de trabalho do profissional. A leitura flutuante possibilitou delimitar as tipologias documentais utilizadas nesta pesquisa. A seguir discorrer-se-á sobre a escolha das publicações, segunda etapa da análise inicial.

Etapa 2 - escolha das publicações

Determinação das publicações para análise. Nesta etapa deve-se definir o *universo empírico* que é “o conjunto de documentos tidos em conta para serem submetidos aos procedimentos analíticos. A sua constituição implica muitas vezes, escolhas, seleções e regras” (BARDIN, 2016, p. 126). As regras para a escolha das publicações foram analisadas conforme orientações de Bardin (2016) e apresentados no Quadro 5 a seguir:

Quadro 5 - critérios para escolha das publicações

REGRA	DEFINIÇÃO DO AUTOR	APLICAÇÃO NA DISSERTAÇÃO
regra da exaustividade:	seleção de todos os “elementos” para compor o corpus para aplicação do método e a exclusão de qualquer material deve ser justificada no plano do rigor;	utilizam-se buscadores confiáveis para englobar todas as publicações existentes. Para as regras de publicação de artigos, os critérios de exclusão estão dispostos no Quadro 6. Para a seleção dos livros, o critério de inclusão foi o número de citações, também identificado na Tabela 7
regra da representatividade	em caso de amostra esta deve ser representativa do universo inicial;	foi realizada uma seleção criteriosa das amostragens das publicações, e a descrição desse processo está explicitada no Quadro 6 e 7 e na Tabela 7
regra da homogeneidade	“os documentos retidos devem ser homogêneos, isto é, devem obedecer a critérios precisos de escolha e não apresentar demasiada singularidade fora desses critérios” (BARDIN, 2016, p. 128);	todas as publicações têm como assunto principal a construção do IFL.
regra de pertinência	“os documentos retidos devem ser adequados, enquanto fonte de informação, de modo a corresponderem ao objetivo que suscita a análise” (BARDIN, 2016, p. 128).	verifica-se se o assunto e a confiabilidade das publicações estão adequados, tendo por base a leitura prévia considerando a identificação das fontes da BCI

Fonte: elaborado pela autora.

Para realizar a triagem das publicações foram considerados as normas nacionais e internacionais, os artigos de periódicos e os livros, conforme relatado a seguir.

Documento 1 - normas nacionais e internacionais

Para realizar a separação das publicações, procedeu-se à seleção das regras nacionais e internacionais obtidas no *site* oficial da Associação Brasileira de Normas Técnicas (ABNT, 2018) e da *International Organization for Standardization* (ISO, [20--]). No **catálogo da ABNT**, foram aplicados os termos “índice” e “índices”, recuperando-se, respectivamente, 106 e 20 normas, porém somente uma (1) foi selecionada, tendo em vista que as demais recuperaram a palavra índice com o significado no sentido numérico de equipamentos ou não tinham como objetivo indicar características e conceitos concernentes à elaboração do IFL. No **site da ISO**, buscou-se o termo “*indexes*”, recuperando-se 508 normas e também sendo selecionada apenas (uma) 1 norma, em razão da temática específica e excluindo-se as normas canceladas. Essas informações estão representadas no Quadro 6.

Quadro 6 - Regras nacionais e internacionais (objetivo e vigência)

REGRA	OBJETIVO	VIGÊNCIA
ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 6034 : informação e documentação: Índice: apresentação. Rio de Janeiro, 2004.	Estabelece os requisitos de apresentação e os critérios básicos para a elaboração de índices.	Norma Vigente
ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 10523 : entrada de nomes de língua portuguesa em Índices. Rio de Janeiro, 1988.	Fixa as exigências para uniformizar a forma de escolha das entradas para nomes de língua portuguesa e estabelece os cabeçalhos em registros bibliográficos.	Norma cuja Vigência foi cancelada em 30 de setembro de 2003 e não é mais utilizada.
<i>INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/AWI 999: documentation - index of a publication.</i> [S.I.], 1975.	-----	Norma cuja Vigência foi cancelada. Passou pelo processo de revisão no ano de 1996.
<i>INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/AWI 999: information and documentation -- guidelines for the content, organization and presentation of indexes.</i> [S.I.], 1996.	Orienta a organização e apresentação de índices para livros, periódicos, relatórios, documentos de patentes e outros documentos escritos. Fornecer instruções para materiais não impressos, como documentos eletrônicos, filmes, gravações de som e vídeo.	Está Vigente, mas passa por processo de revisão.
<i>INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/AWI 999: information and documentation -- guidelines for the content, organization and presentation of indexes.</i> [S.I.], [20--].	-----	Documento está em fase de elaboração. Assim, não há informações sobre sua vigência

Fonte: elaborado pela autora.

Para a seleção das normas, os critérios utilizados foram: 1) a leitura do título e da descrição introdutória, para avaliar se estavam de acordo com a temática e 2) a confirmação se tais normas se encontram em vigor, ou seja, se estão ativas e devem ser cumpridas. Essas normas foram selecionadas, excluindo-se aquelas que foram canceladas ou não continham o IFL como temática. A escolha dos artigos de periódicos será comentada no próximo tópico.

Documento 2 - artigos de periódicos

Após selecionar as publicações segundo as normas da ABNT e da ISO, foi realizada a triagem dos artigos resultantes de revisão de literatura e que tinham como enfoque a construção de IFL, totalizando 16 (dezesesseis) publicações. O critério adotado

para triar essas publicações levou em consideração aqueles que relatassem sobre a construção do IFL e artigos que discorressem sobre softwares de indexação. No Quadro 7, estão descritos os artigos, o *status* e a justificativa de seleção;

Quadro 7 - Artigos da revisão de literatura (continua)

ARTIGO	STATUS DE SELEÇÃO	JUSTIFICATIVA DE SELEÇÃO
ABE, Kris; BERRY, Daniel M. Indx and findphrases: a system for generating indexes for ditroff documents. Software-practice and experience , [S.I.], v. 19, n. 1, p. 1-34, jan. 1989.	SIM	<i>Software Indx</i>
CHI, Ed. H. <i>et al.</i> eBooks with Indexes that Reorganize Conceptually. ACM , Vienna, apr. 2004.	SIM	<i>Software ScentIndex</i>
CHI, Ed. H. <i>et al.</i> ScentIndex: conceptually reorganizing subject indexes for reading. IEEE Symposium on Visual Analytics Science and Technology , Baltimore, 2006.	SIM	<i>Software ScentIndex</i>
FRANCIS, Ronald D.; GREENWAY, Murray. A simple system for constructing an end-of-book index. Journal of scholarly publishing , v. 46, n. 3, p. 290-296, apr. 2015.	SIM	Construção do ÍFL
MEKKI, Touria Ait El; NAZARENKO, Adeline. Using NLP to build the hypertextuel network of a back-of-the-book index . 2005.	SIM	<i>Software IndDoc</i>
MILLER, Sylvia K. Index appreciation: a publisher's brief guide. Against the Grain , [S.I.], v. 17, n. 4, nov. 2013.	SIM	Construção do ÍFL
PROVO, A. From index to network: topic maps in the Enhanced Networked Monographs project. The Indexer: The International Journal of Indexing , [S.I.], v. 37, n. 1, p. 13–35, 2019.	SIM	<i>Software TCT</i>
WEIHS, Jean. Indexes, indexing, and the technical services staff part 2: cataloguers. Technicalities , [S.I.], v. 27, n. 2, mar./ apr. 2007.	SIM	Construção do ÍFL
CSOMAI, András; MIHALCEA, Rada. Linguistically Motivated Features for Enhanced Back-of-the-Book Indexing. Proceedings of ACL-08: HLT , Columbus, Ohio, USA, p. 932–940, june 2008.	NÃO	Relata sobre o desenvolvimento do <i>software</i> , porém ainda não foi disponibilizado.

Quadro 7 - Artigos da revisão de literatura (conclusão)

ARTIGO	STATUS DE SELEÇÃO	JUSTIFICATIVA DE SELEÇÃO
CSOMAI, Andrés; MIHALCEA, Rada. Creating a Testbed for the evaluation of automatically generated back-of-the-book indexes . 2006.	NÃO	Estudo sobre plataforma de teste para IFL, criado por meio da automação
CSOMAI, Andrés; MIHALCEA, Rada. Investigations in Unsupervised Back-of-the-Book Indexing. [Proceedings of ACL-08: HLT], [Texas], 2007.	NÃO	Retrata o desenvolvimento do <i>software</i> , porém ainda não foi disponibilizado.
BOLSHAKOVA, E. I.; IVANOV, K. M. Term extraction for constructing subject index of educational scientific text. In: COMPUTATIONAL LINGUISTICS AND INTELLECTUAL TECHNOLOGIES: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE "DIALOGUE 2018", 2018, Moscow. Anais... Moscow: [s.n.], 2018. 11 p.	NÃO	Relata sobre o método para criação de IFL automático mas não foi desenvolvido nenhum <i>software</i> .
HINES, Theodore C.; HARRIS, Jessica L. Computer-aided production of book indexes. The Indexer , [S.l.], n. 7, v. 2, 1970.	NÃO	<i>Software</i> criado para fins didáticos. Ou seja, é de uso interno e não está disponível.
SYLVA, Lyne da. Nouveaux horizons en indexation automatique de monographies. Documentation et bibliothèques , [S.l.], v. 48, n. 4, p. 155-167, 2002.	NÃO	Relato dos pontos que precisam melhorar em <i>softwares</i> que constróem IFL. Desta forma, não se aplicam a nenhum <i>software específico</i>
WU, Zhaohui et al. Can Back-of-the-Book Indexes be Automatically Created?. CIKM'13 , San Francisco, 2013.	NÃO	Relata sobre o desenvolvimento do <i>software</i> , porém ainda não foi disponibilizado
FAULKNER, Hilary; LIGHT, Wiebke M. Editing the index: developing a method. The Indexer , [S.l.], v. 24, n. 4, oct. 2005.	NÃO	Experiência com o fórum de discussão sobre boas práticas na indexação (<i>INDEX-L</i>)

Fonte: elaborado pela autora.

Na sequência, são elencados os livros que embasaram a análise dessa pesquisa.

Documento 3 - livros

Para a escolha dos livros, foi selecionada uma amostra de livros técnicos que discorressem sobre as etapas da elaboração de construção de IFL. A seleção da amostra considerou o número de citações apresentado pelo Google acadêmico¹⁶. Os livros selecionados e o número de citações estão listados na Tabela 7:

Tabela 7 - Livros selecionados

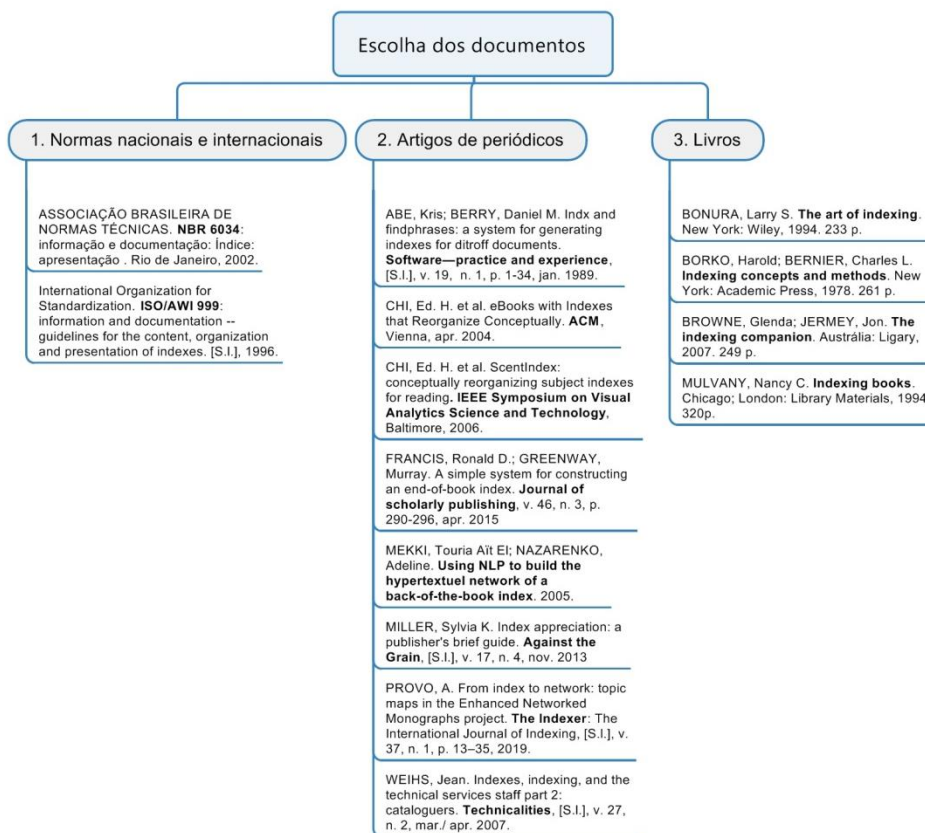
LIVRO	CITAÇÕES
BORKO, Harold; BERNIER, Charles L. Indexing concepts and methods . New York: Academic Press, 1978. 261 p.	165
MULVANY, Nancy C. Indexing books . Chicago; London: Library Materials, 1994. 320p.	162
BONURA, Larry S. The art of indexing . New York: Wiley, 1994. 233 p.	50
BROWNE, Glenda; JERMEY, Jon. The indexing companion . Austrália: Ligary, 2007. 249 p.	28

Fonte: elaborado pela autora.

Como resultado final, selecionou-se 14 (quatorze) publicações compostas de 2 (duas) normas nacionais e internacionais, 8 (oito) artigos de periódicos e 4 (quatro) livros para realizar a análise de conteúdo, conforme disposto na Figura 13 a seguir:

¹⁶ Verificação realizada pelo Google acadêmico em julho de 2019.

Figura 13 - Publicações selecionadas para análise de conteúdo



Fonte: elaborado pela autora.

Finalizada a apresentação da etapa dois, passa-se a descrever a terceira e quarta etapa que são mencionadas a seguir.

Etapa 3 - formulação das hipóteses e dos objetivos

A hipótese é entendida como uma “afirmação provisória que nos propomos a verificar (confirmar ou infirmar), recorrendo aos procedimentos de análise”, sendo concluída após a finalização de todos os procedimentos (BARDIN, 2016, online). E o objetivo “é a finalidade geral a que nos propomos (ou que é fornecida por uma instância exterior), o quadro teórico e/ou pragmático, no qual os resultados obtidos serão utilizados” (BARDIN, 2016, *online*).

As informações relativas a essa etapa foram apresentadas no capítulo Introdução. Desta forma, optou-se pela sua utilização do termo “pressupostos” ao invés de “hipóteses” por ser o utilizado nessa pesquisa. Os objetivos são apresentados na seção 1.2 desta pesquisa.

Etapa 4 - Preparação do material

Etapa em que as publicações escolhidas são preparadas e codificadas, no caso de utilização do computador, as publicações devem ser passíveis de “leitura” conforme as instruções do programa utilizado. Todas as publicações foram convertidas para Portable Document Format (PDF); no caso dos livros didáticos, foi feita a digitalização e conversão para *Optical Character Recognition* (OCR), formato esse que converte arquivos de imagens em textos.

4.3.2 Exploração do material

A exploração do material consiste na operação de codificação e decomposição das publicações selecionadas, conforme regras previamente formuladas. Nessa etapa, há a aplicação sistemática das decisões formuladas na fase anterior.

A fim de responder aos objetivos propostos nessa pesquisa, essa seção foi dividida em duas seções, que são: 1) **Seção 1**: análise de conteúdo em publicações sobre software para construção de Índice de Final de Livro e 2) **Seção 2**: análise de conteúdo em publicações sobre construção de Índice de Final de Livro. A seção 2 é dividida em duas subseções, a saber: 1) **seção 2.1**: relativo à Elaboração de IFL e 2) **seção 2.2**: relativo à Estruturação de IFL. E, em cada seção, são identificadas 2 (duas) atividades, a primeira atividade lista as publicações selecionadas e a segunda é a extração dos trechos das publicações analisadas, ou seja, destaque das frases significativas no texto. As atividades tem o intuito de identificar os elementos semânticos das publicações.

4.3.2.1 Exploração do material para Seção 1

Os softwares que elaboram IFL são analisados baseando nas informações contidas nos artigos recuperados da Revisão de literatura que relatassem sobre as características dos softwares. Dessa forma, as publicações selecionadas nessa seção são identificadas a seguir.

Atividade 01: publicações selecionadas nesta seção

- a) ABE, Kris; BERRY, Daniel M. Indx and findphrases: a system for generating indexes for ditroff documents. **Software—practice and experience**, [S.l.], v. 19, n. 1, p. 1-34, jan. 1989.
- b) CHI, Ed. H. et al. eBooks with Indexes that Reorganize Conceptually. **ACM**, Vienna, apr. 2004.

- c) MEKKI, Touria Aït El; NAZARENKO, Adeline. **Using NLP to build the hypertextuel network of a back-of-the-book index**. 2005.
- d) CHI, Ed. H. et al. ScenIndex: conceptually reorganizing subject indexes for reading. **IEEE Symposium on Visual Analytics Science and Technology**, Baltimore, 2006.
- e) PROVO, A. From index to network: topic maps in the Enhanced Networked Monographs project. **The Indexer: The International Journal of Indexing**, [S.l.], v. 37, n. 1, p. 13–35, 2019.

A decisão em analisar os *softwares* pela ótica dos artigos supracitados é justificada, pois, pressupõe-se que o relato apresentado pelo responsável do software pode ser considerado uma fonte de informação adequada, uma vez que esses profissionais seriam os mais indicados para identificar e detalhar as características.

Atividade 02: extração dos trechos das publicações analisadas

O resultado dessa atividade é apresentado no Quadro 8 a seguir visando facilitar a análise do leitor. Buscou-se selecionar textos que apresentassem informações sobre as características do software, excluindo-se informações de como utilizá-los. E, para facilitar a visualização, os trechos foram copiados de forma individual, e para cada trecho foi realizado o trabalho de **padronização dos termos**, o que significa que houve alteração de termos do texto original, com o propósito de reduzir, simplificar ou utilizar os termos escolhidos para esta pesquisa. A **numeração** indicada após cada frase de padronização tem a finalidade de indicar a fonte, considerando que nos próximos capítulos são mencionados novamente esses trechos.

Quadro 8 - Publicações analisadas (continua)

1 – <i>Software</i> Indx (ABE; BERRY, 1989)		
Trecho	Palavras/Frases destacadas	Padronização dos termos
Oferece assistência precisa na determinação de termos, permitem encontrar todas as ocorrências e algumas sugestões sobre termos e fornecem uma interface limpa na qual o texto do livro a ser indexado não é inundado com comandos de indexação. (p.2)	assistência precisa na determinação de termos permitem encontrar todas as ocorrências algumas sugestões sobre termos fornecem uma interface limpa	assistência precisa na determinação de termos - 1 permitem encontrar todas as ocorrências nos termos - 2 sugestões sobre termos - 3 fornece interface limpa - 4
Os números que seguem o cabeçalho são indicadores, que direcionam o leitor para o local no texto onde o cabeçalho é discutido. Esses números geralmente são números de página, mas também podem ser seção ou números de parágrafo. A versão atual do indx lida apenas com números de página , portanto, esses números são referidos como referências de página. (p.3)	números de página	números de página (não apresenta cáp ou seção) - 5
Como existem índices que usam ver em remissivas , nosso <i>software</i> é obrigado a fornecer a capacidade de obtê-los. (p.4)	usam ver em remissivas	possibilidade de criar remissivas - 6

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>Para demonstrar a flexibilidade do nosso <i>software</i>, a primeira metade do índice é impressa no estilo de entrada por linha e a segunda metade é impressa no estilo do parágrafo. Essa flexibilidade é alcançada, como explicado mais adiante, alterando as definições chamadas de macro criadas pelo <i>software</i> de indexação. (p.4)</p>	<p>estilo de entrada por linha</p> <p>estilo do parágrafo</p>	<p>estilo por linha - 7</p> <p>estilo por parágrafo - 8</p>
<p>A principal vantagem, acreditamos da abordagem <i>indx</i>, descrita aqui, é que o texto a ser formatado não precisa ser inundado com comandos de indexação. O formatador obtém um texto de entrada mais limpo, mais fácil de ser submetido a outras análises, e é menos provável que confunda a correspondência de padrões de um editor. (p.6)</p>	<p>texto de entrada mais limpo</p> <p>menos provável que confunda a correspondência de padrões de um editor</p>	<p>texto de entrada limpo - 9</p> <p>não confunde padrões de editor de texto - 10</p>
<p>É possível fazer pequenas alterações estéticas ou mesmo grandes mudanças reorganizacionais no texto sem ter que alterar os termos do índice, e é possível ajustar os detalhes do índice alterando os termos do índice sem alterar o texto. (p.6)</p>	<p>possível fazer pequenas alterações estéticas</p> <p>mesmo grandes mudanças reorganizacionais no texto</p> <p>sem ter que alterar os termos do índice</p>	<p>possibilidade de fazer formatações - 11</p> <p>possibilidade de fazer edições - 12</p> <p>formatar a estrutura do texto não altera os termos do índice - 13</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>Vários recursos que poderiam ser disponibilizados tornaram-se aparentes após examinar outros programas de indexação semi-automáticos e depois de usar o indx.</p> <p>1. Aumentar o número de remissivas. Embora ter uma entrada para referência cruzada geralmente seja suficiente, há momentos em que um segundo ou até mesmo um terceiro cabeçalho também deve ter referência cruzada. (...)</p> <p>3. Opções do esquema de classificação. O indx também seria mais completo se fornecesse ao indexador a opção de classificar as entradas palavra por palavra ou letra por letra. (...)</p> <p>5. Fusão de índices. O indx pode ser muito útil se partes de um índice puderem ser criadas e mescladas para formar um índice completo. Isso permitiria a criação de índices em uma base de capítulo a capítulo, quando os índices de cada capítulo são mesclados para formar o índice.</p> <p>6. Mantendo uma frase combinada no índice. Ao usar o indx para construir o índice inicial deste artigo, a opção de manter o segundo termo de uma definição de arquivo de frases combinadas em vez de excluí-lo automaticamente tornou-se desejável. Assim como o arquivo de entrada de grupo permitia que uma entrada se tornasse uma subentrada enquanto permanecia uma entrada principal, a capacidade de usar uma frase para fornecer referências de página para mais de uma frase foi considerada útil. Atualmente, se tais entradas fossem desejadas, outra frase teria que ser escolhida para produzir os mesmos números de página. Isso sobrecarregaria o indexador humano e causaria o uso de mais termos de indexação do que o necessário. (p.20)</p>	<p>Aumentar o número de remissivas</p> <p>Opções do esquema de classificação</p> <p>Fusão de índices</p> <p>Mantendo uma frase combinada no índice</p>	<p>Limitado em relação a utilização de remissivas - 14</p> <p>Não possui esquema de classificação (entradas palavra por palavra ou letra por letra) - 15</p> <p>Não tem possibilidade de criar parte de índice (exemplo: quando os índices de cada capítulo são mesclados para formar o índice) - 16</p> <p>Impossibilidade de um subcabeçalho ser um cabeçalho - 17</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
O método como um todo é simples de usar , uma vez que a difícil tarefa de selecionar frases e entradas é feita. (p.21)	simples de usar	simples de usar - 18
O programa indx foi usado com sucesso para criar o índice deste artigo. O método como um todo é simples de usar , uma vez que a difícil tarefa de selecionar frases e entradas é feita. (p.21)	simples de usar	simples de usar - 19
Além disso, deve ser mais fácil reestruturar o índice . Por exemplo, obter várias entradas principais em um grupo ou dividir uma entrada de grupos em várias entradas principais é realizado alterando uma linha nos arquivos opcionais para cada entrada envolvida, em vez de alterar cada ocorrência dos comandos de indexação para essas entradas no texto de origem. Em geral, as alterações podem ser feitas no índice sem precisar alterar o arquivo de texto. (p.21)	mais fácil reestruturar o índice	possibilidade de fazer formatações - 20
Indx ajuda o indexador humano a criar bons índices até o ponto em que as referências cruzadas no índice serão verificadas e a referência incômoda da cadeia na qual o leitor é direcionado de uma entrada para outra para outra sem ter referências de página para procurar não é permitida. É necessário que o humano passe algum tempo eliminando referências de página inúteis . Usando frases específicas que são tão curtas quanto possível reduzirá a quantidade de verificação necessária. (p.22)	ajuda o indexador humano a criar bons índices referências cruzadas no índice serão verificadas É necessário que o humano passe algum tempo eliminando referências de página inúteis	ajuda o indexador criar bons índices - 21 necessidade de verificação do indexador - 22 precisa eliminar localizadores que não fornecem informações relevantes para o leitor - 23

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>O indx executa o processo clerical da tarefa de indexação, localizando as referências de página das entradas e organizando as entradas na ordem desejada, sem confundir o texto que está sendo indexado com comandos de indexação. Este programa e suas ferramentas relacionadas fornecem muita ajuda para a pessoa que cria um índice. No entanto, a tarefa de indexação permanece altamente intensiva em termos intelectuais e humanos. (p.22)</p>	<p>localizando as referências de página das entradas</p> <p>organizando as entradas na ordem desejada</p> <p>sem confundir o texto que está sendo indexado com comandos de indexação</p> <p>fornecem muita ajuda para a pessoa que cria um índice</p> <p>a tarefa de indexação permanece altamente intensiva em termos intelectuais e humanos</p>	<p>insere os localizadores em cada entrada - 24</p> <p>organiza as entradas na ordem desejada - 25</p> <p>ajuda o indexador a criar bons índices - 26</p> <p>ajuda o indexador criar bons índices - 27</p> <p>necessidade de verificação do indexador8 - 28</p>
2. Software ScentIndex (CHI et al., 2004; 2006)		
<p>Em nosso método, usamos uma matriz de coocorrência de palavras para extrair entradas de índice que são conceitualmente relevantes para as palavras-chave que os usuários inseriram. Também aprimoramos as interações de navegação e navegação entre o índice e o e-book. Usamos uma grande variedade de técnicas de destaque para dar dicas de navegação aos usuários. (2004, p.1)</p>	<p>aprimoramos as interações de navegação e navegação entre o índice e o e-book</p> <p>dicas de navegação aos usuários</p>	<p>navegação entre o IFL e o e-book aprimorada - 1</p> <p>fornece dicas de navegação aos leitores - 2</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>Nosso [algoritmo] é baseado em similaridades conceituais de termos. Além disso, em vez de procurar gerar automaticamente novos índices para textos de livros como em, estamos interessados em formas de melhorar esses índices existentes para que eles se reorganizem para melhor atender às necessidades de informação do usuário. (2004, p.1)</p>	<p>baseado em similaridades conceituais de termos</p> <p>necessidades de informação do usuário</p>	<p>baseado em similaridades conceituais de termos - 3</p> <p>preocupa com as necessidades de informação do leitor - 4</p>
<p>Conforme mostrado, o realce automático permite que o usuário localize rapidamente a passagem que pode ser relevante para a consulta do usuário. (2004, p.3)</p>	<p>realce automático</p> <p>localize rapidamente</p>	<p>fornece realce nos termos pesquisados - 5</p> <p>leitores localizam os termos mais rapidamente - 6</p>
<p>Conforme mostrado neste cenário de uso, ao reorganizar as entradas de índice, o usuário pode restringir o número de entradas pelas quais deve pesquisar para encontrar a resposta correta. Ao inserir todas as palavras-chave relevantes, os usuários podem ver em uma única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas por várias páginas de índice diferentes. (2004, p.3)</p>	<p>reorganizar as entradas de índice</p> <p>usuário pode restringir o número de entradas</p> <p>ver em uma única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas</p>	<p>possibilidade de reorganizar as entradas de índice - 7</p> <p>leitor pode restringir o número de entradas - 8</p> <p>facilidade de visualização, pois apresenta em única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas - 9</p>
<p>Descobrimos que os usuários foram mais rápidos em concluir suas tarefas e mais precisos em suas respostas usando o ScentIndex, independentemente de seu nível de especialização no conteúdo do livro. Os usuários também preferiram a interface ScentIndex para essas tarefas. Os resultados da análise mostram que a condição da interface não teve interações com o nível de especialização. (2004, p.4)</p>	<p>usuários foram mais rápidos</p> <p>mais precisos em suas respostas</p> <p>independentemente de seu nível de especialização no conteúdo do livro</p>	<p>leitores localizam os termos mais rapidamente - 10</p> <p>leitores localizam os termos de forma mais precisa- 11</p> <p>independente do conhecimento do conteúdo do livro o leitor consegue localizar a informação- 12</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>Ao reorganizar e reduzir as entradas de índice para uma única página, os usuários podem navegar e verificar com mais eficiência as informações de interesse. Integramos essa técnica com aprimoramentos de navegação e realce na interface de navegação de e-books para permitir a rápida varredura e a varredura de passagens relevantes. Aproveitando os índices de assunto existentes, esperamos preservar a aparência dos índices de assunto em sua forma eletrônica, bem como aprimorá-los para uso real. (2004, p.4)</p>	<p>reorganizar e reduzir as entradas de índice</p> <p>verificar com mais eficiência as informações de interesse</p> <p>aprimoramentos de navegação</p> <p>realce na interface de navegação de e-books</p> <p>permitir a rápida varredura de passagens relevantes</p>	<p>possibilidade de formatação e edição- 13</p> <p>leitores localizam termos de forma mais precisa - 14</p> <p>aprimoramento de navegação- 15</p> <p>fornece realce nos termos pesquisados- 16</p> <p>permite rápida varredura nos termos relevantes - 17</p>
<p>Em seguida, nosso sistema, chamado ScentIndex, calcula quais entradas de índice são conceitualmente relacionadas e reorganiza e exhibe essas entradas de índice em uma única página. Também fornecemos várias dicas de navegação para ajudar os usuários a examinar essa lista de entradas de índice e encontrar passagens relevantes rapidamente. Comparado com a leitura regular de um livro em papel, nosso estudo mostrou que os usuários são mais eficientes e mais precisos em encontrar, comparar e compreender o material em nosso sistema.(2006, p.1)</p>	<p>reorganiza e exhibe essas entradas de índice</p> <p>dicas de navegação</p> <p>encontrar passagens relevantes rapidamente</p> <p>usuários são mais eficientes e mais precisos em encontrar, comparar e compreender o material em nosso sistema</p>	<p>possibilidade de formatação - 18</p> <p>fornece dicas de navegação - 19</p> <p>leitores localizam os termos mais rapidamente - 20</p> <p>leitores localizam os termos de forma mais precisa - 21</p>
<p>A figura ilustra que o especialista pode obter mais informações na mesma quantidade de tempo (ou a mesma informação em menos tempo). Uma meta para o sistema descrito neste artigo é permitir que os analistas iniciantes criem suas próprias curvas usando o destaque inteligente e a seleção semântica antecipada. (2006, p.1)</p>	<p>usando o destaque inteligente</p>	<p>fornece realce nos termos pesquisados - 22</p>
<p>Geralmente, há grandes listas de páginas no índice para verificar as passagens relevantes. Ajudamos os usuários a verificar essas páginas aprimorando as interações de navegação e navegação entre o índice e o e-book. Usamos uma grande variedade de técnicas de destaque para dar dicas de navegação aos usuários. (2006, p.2)</p>	<p>aprimorando as interações de navegação</p> <p>navegação entre o índice e o e-book</p> <p>dar dicas de navegação aos usuários</p>	<p>aprimoramento de navegação - 23</p> <p>navegação entre o IFL e o e-book aprimorada - 24</p> <p>fornece dicas de navegação aos leitores - 25</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>As tarefas aqui parecem fáceis por vários motivos: (1) Primeiro, usando a reorganização conceitual, os usuários têm uma alta confiança de que entradas relevantes não são omitidas, porque não dependemos apenas de correspondências exatas de palavras-chave. (2006, p.4)</p> <p>Conforme mostrado neste cenário de uso, ao reorganizar as entradas de índice, o usuário pode restringir o número de entradas pelas quais deve pesquisar para encontrar a resposta correta. Ao inserir todas as palavras-chave relevantes, os usuários podem ver em uma única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas por várias páginas de índice diferentes. (2006, p.4)</p>	<p>usuários têm uma alta confiança</p> <p>entradas relevantes não são omitidas</p> <p>reorganizar as entradas de índice</p> <p>usuário pode restringir o número de entradas pelas quais deve pesquisar</p>	<p>leitores têm uma alta confiança - 26</p> <p>entradas relevantes não são omitidas - 27</p> <p>possibilidade de formatação do índice - 28</p> <p>leitor pode restringir o número de entradas - 29</p>
<p>Para nossa agradável surpresa, especialistas e novatos puderam aproveitar a interface do ScentIndex e concluir as tarefas mais rapidamente. (2006, p.7)</p> <p>A pesquisa pós-experimento mostrou que os participantes preferiram o ScentIndex (15 de 16 indivíduos). As razões dadas para esta preferência incluem “pode pesquisar usando combinações de palavras-chave”, “clicar no número da página para navegar”, “realçar permite escanear mais rapidamente e skimming”, e “mais fácil de comparar entradas de índice porque é tudo em uma página”. discussão após os experimentos, alguns assuntos mencionaram que eles prefeririam a versão de livro em papel para leitura extensiva. (2006, p.7)</p>	<p>especialistas e novatos puderam aproveitar a interface</p> <p>concluir as tarefas mais rapidamente.</p> <p>pode pesquisar usando combinações de palavras-chave clicar no número da página para navegar</p> <p>mais fácil de comparar entradas de índice</p>	<p>independente do conhecimento do conteúdo do livro o leitor consegue localizar a informação - 30</p> <p>leitores localizam os termos mais rapidamente - 31</p> <p>possibilidade de pesquisar usando combinações de termos - 32</p> <p>Possibilidade clicar no localizar e direcionar para página - 33</p> <p>mais fácil de comparar entradas de índice - 34</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>No geral, o ScentIndex teve um desempenho melhor que o índice de assunto do papel. Assuntos usando ScentIndex foram mais rápidos em completar suas tarefas, não importa se eram especialistas ou novatos. Além disso, as respostas que eles forneceram ao usar o ScentIndex foram mais precisas do que as respostas dadas quando usaram o Índice de Assunto do Papel. Os usuários também preferiram a interface ScentIndex para essas tarefas. (2006, p.7)</p>	<p>desempenho melhor que o índice de assunto do papel mais rápidos em completar suas tarefas</p> <p>não importa se eram especialistas ou novatos</p> <p>forneceram ao usar o ScentIndex foram mais precisas</p>	<p>desempenho melhor que o Índice de Final de Livro impresso - 35</p> <p>leitores localizam os termos mais rapidamente - 36</p> <p>independente do conhecimento do conteúdo do livro o leitor consegue localizar a informação - 37</p> <p>leitores localizam os termos de forma mais precisa - 38</p>
<p>Para fazer isso, combinamos realces visuais inteligentes de texto que ajudam a direcionar a atenção dos analistas com o processamento de plano de fundo semântico analítico que filtra o índice de um livro até as entradas mais relevantes, incluindo aquelas semanticamente, mas não relacionadas textualmente. Desta forma, ampliamos o papel que os índices de assunto têm tido para os livros desde que foram inventados no século XV. O ScentIndex conceitualmente reorganiza grandes índices de assuntos de acordo com alguma necessidade de informação. Nosso estudo do usuário sugere que isso funciona. Tanto os usuários experientes quanto os novatos são mais rápidos em concluir tarefas de busca, comparação e compreensão de fatos usando o ScentIndex, e as respostas produzidas pelos usuários são mais precisas. Esperamos que isso inspire uma nova linha de pesquisa para aumentar a leitura com novas inovações. (2006, p.7)</p>	<p>realces visuais inteligentes de texto</p> <p>reorganiza grandes índices de assuntos</p>	<p>fornece realce nos termos pesquisados - 40</p> <p>possibilidade de formatação e edição - 41</p>

Quadro 8- Publicações analisadas (continua)

3 Software IndDoc (MEKKI; NAZARENKO, 2005)		
Trecho	Palavras/Frases destacadas	Padronização dos termos
Nós projetamos um método para automatizar a construção de índices. Nosso sistema IndDoc se baseia no texto do documento 1) para selecionar os termos que valem a pena mencionar no índice final e 2) para vincular cada termo aos segmentos do documento . (p.1)	selecionar os termos que valem a pena mencionar no índice final	extraí os termos ou frases mais relevantes do documento original - 1
Os métodos de segmentação são geralmente baseados na estrutura física dos documentos (tipografia, seccionamento). (p.2)		o método de segmentação utiliza as abordagens estruturais e linguísticas que são mais relevantes - 3
A abordagem de coesão lexical fornece resultados interessantes em documentos grandes e heterogêneos, mas é menos adaptada à segmentação de documentos homogêneos. As abordagens estruturais e linguísticas são mais relevantes para nossos propósitos. Nosso algoritmo de segmentação combina os dois métodos (consulte a Seção 3). (p.2)	As abordagens estruturais e linguísticas são mais relevantes	o método de segmentação utiliza as abordagens estruturais e linguísticas que são mais relevantes - 3
Todos menos um desses termos "ruins" foram realmente eliminados (menos de 0,01% de precisão). Esses números confirmam o bom desempenho do nosso algoritmo de classificação de conhecimento . (p.5)	bom desempenho do nosso algoritmo de classificação de conhecimento	bom desempenho do algoritmo de classificação de conhecimento - 4
O problema de relevância também é tratado para a sumarização de documentos, para extrair as frases mais relevantes do documento original . A pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase linguísticos ou tipográficos . Nossa medida de relevância leva esses parâmetros em consideração. (p.2)	extrair as frases mais relevantes do documento original pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase linguísticos ou tipográficos	extraí os termos ou frases mais relevantes do documento original - 5 pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase linguísticos ou tipográficos - 6
Nós adaptamos a tradicional relevância da RI para a classificação dos vários parágrafos de um documento em vez de um conjunto de documentos. (p.2)	classificação dos vários parágrafos de um documento	possibilidade de classificação - 7

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>A classificação dos termos não tem impacto direto nas funcionalidades de navegação, mas a classificação de segmentos e descritores é interdependente. (p.5)</p> <p>Propomos um método de conhecimento pobre para construir automaticamente a rede hipertextual que ajuda na navegação pelo documento. O dispositivo resultante é semelhante a um índice de final de livro. (p.5)</p> <p>A originalidade do nosso método é a grande variedade de sugestões que são levadas em conta: tipografia, estrutura lógica do documento, marcadores lingüísticos de integração linear, coesão lexical, entre outros. O impacto de cada tipo de sugestão depende do estilo do documento, mas a combinação de todos tornam nosso algoritmo de segmentação e classificação mais robusto. (p.5)</p>	<p>classificação dos termos não tem impacto direto nas funcionalidades de navegação</p> <p>método de conhecimento pobre para construir automaticamente a rede hipertextual que ajuda na navegação pelo documento</p> <p>grande variedade de sugestões que são levadas em conta: tipografia, estrutura lógica do documento, marcadores lingüísticos de integração linear, coesão lexical, entre outros.</p> <p>depende do estilo do documento</p> <p>tornam nosso algoritmo de segmentação e classificação mais robusto</p>	<p>possibilidade de formatação - 8</p> <p>simples para construir automaticamente a rede hipertextual que ajuda na navegação pelo documento - 9</p> <p>pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase lingüísticos ou tipográficos - 10</p> <p>depende do estilo do documento [as sugestões do software] - 11</p> <p>considera que o algoritmo de segmentação e classificação é mais robusto - 12</p>
<p>4 Software Topic Curation Toolkit (PROVO, 2019)</p>		
<p>Este artigo discutirá a criação do mapa de tópicos da ENM, um meta-índice desenvolvido com a combinação de muitos índices individuais de final de livro usando o Topic Curation Toolkit. (p.2)</p> <p><i>software</i> TCT cria registros de termos analisando os índices EPUB publicados e gerando um registro de cada cabeçalho do índice. (p.4)</p> <p>TCT, conforme descrito acima, agrupa e relaciona os termos, e um indexador então analisa e executa o controle de qualidade (p.8)</p>	<p>combinação de muitos índices individuais de final de livro</p> <p>gerando um registro de cada cabeçalho do índice</p> <p>agrupa e relaciona os termos indexador então analisa e executa o controle de qualidade</p>	<p>possibilidade de acumular índices - 1</p> <p>facilidade em inserir cabeçalhos no índice - 2</p> <p>possibilidade de criar remissivas - 4</p> <p>necessidade de verificação do indexador - 5</p>

Quadro 8- Publicações analisadas (continua)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>A interface editorial do TCT inclui visualizações que mostram listas alfabéticas dos termos (...). As visualizações de lista alfabética incluem opções para classificar alfabeticamente, por número de relações e pelo número de edições. (...) Um ícone de lápis ao lado de muitos dos campos indica que ele pode ser editado diretamente. Existem também botões para adicionar novos dados, um botão "reexecutar relações" e dois botões de revisão que indica se um tópico não foi revisado ou editado. Na ocorrência (ou conteúdo do livro) os botões de seta fornecem navegação sequencial entre as páginas do livro. (p.8)</p>	<p>inclui visualizações que mostram listas alfabéticas dos termos visualizações de lista alfabética incluem opções para classificar alfabeticamente, por número de relações e pelo número de edições ícone de lápis ao lado de muitos dos campos indica que ele pode ser editado diretamente botões para adicionar novos dados "reexecutar relações" e dois botões de revisão que indica se um tópico não foi revisado ou editado fornecem navegação sequencial entre as páginas do livro</p>	<p>possibilidade de classificação - 6 possibilidade de classificação - 7 possibilidade de formatação - 8 possibilidade de edição - 9 possibilidade de formatação - 10 possibilidade de edição - 11 fornece navegação sequencial entre as páginas do livro - 12</p>
<p>Outros problemas semânticos surgiram como consequências não intencionais de regras que de outra forma funcionou corretamente. (...)Também, abreviaturas causavam ambiguidade e causavam a criação de relações irrelevantes. (...)Os homógrafos também surgiram como uma questão problemática. (p.9)</p>	<p>abreviaturas causavam ambiguidade e causavam a criação de relações irrelevantes problemas semânticos surgiram homógrafos também surgiram como uma questão problemática</p>	<p>problemas com termos abreviados (criam relações irrelevantes no índice) - 13 problemas semânticos - 14 problemas semânticos - 15</p>

Quadro 8- Publicações analisadas (conclusão)

Trecho	Palavras/Frases destacadas	Padronização dos termos
<p>Algumas ocorrências se referem ao conceito de “nação”, enquanto outros referem a publicação “<i>The Nation</i>” esse problema aconteceu porque o TCT ignora artigos como “a”, entre outros. Levando em consideração os valores de autonomia do projeto e precisão, imaginamos que o usuário ao procurar sobre “The Nation” iria tornar-se frustrado. (p.13)</p> <p>Os 5.740 termos identificados como importantes foram direcionados para enriquecimento dos termos, <i>Uniform Resource Identifier</i> (URIs) em vocabulários controlados, como Wikidata, o <i>Virtual International Authority File</i> (VIAF), <i>Library of Congress Authorities</i> (LC) e <i>Application of Subject Terminology</i> (FAST). (p.14)</p> <p>a filtragem o usuário pode fazer com facetas permite controlar os limites do conjunto de pesquisa; o A visualização fornece uma visão geral do livro e, possivelmente, uma exploração. (p.18)</p>	<p>problema aconteceu porque o TCT ignora artigos como “a”</p> <p>direcionados para enriquecimento dos termos em vocabulários controlados</p> <p>permite controlar os limites do conjunto de pesquisa</p>	<p>problemas semânticos - 16</p> <p>possibilidade de utilizar vocabulário controlado - 17</p> <p>possibilidade de limitar a visualização da pesquisa - 18</p>

Fonte: elaborado pela autora.

Realizado o trabalho de extração e padronização dos trechos das publicações, passa-se para a seção 2 relacionado à construção do IFL, conforme apresentado a seguir.

4.3.2.2 Exploração do material para Seção 2

A seção 2 refere-se à busca de publicações e trechos que informam sobre a construção do IFL. A construção envolve dois processos: elaboração e estruturação; ambos serão explorados com mais vagar a seguir.

4.3.2.2.1 Seção 2.1: elaboração de IFL

A seguir são apresentados os resultados referentes a cada atividade. As publicações selecionadas nesta seção referem-se à elaboração do IFL, focando-se assim, em normas e literatura da área.

Atividade 01: publicações selecionadas nesta seção

- a) ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6034:** informação e documentação: Índice: apresentação. Rio de Janeiro, 2004.
- b) BONURA, Larry S. **The art of indexing.** New York: Wiley, 1994. 233 p.
- c) BORKO, Harold; BERNIER, Charles L. **Indexing concepts and methods.** New York: Academic Press, 1978. 261 p.
- d) BROWNE, Glenda; JERMEY, Jon. **The indexing companion.** Austrália: Ligary, 2007. 249 p.
- e) FRANCIS, Ronald D.; GREENWAY, Murray. A simple system for constructing an end-of-book index. **Journal of scholarly publishing**, v. 46, n. 3, p. 290-296, apr. 2015.
- f) INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/AWI 999: information and documentation -- guidelines for the content, organization and presentation of indexes. [S.l.], 1996.
- g) MILLER, Sylvia K. **Index appreciation:** a publisher's brief guide. Against the Grain, [S.l.], v. 17, n. 4, nov. 2013.
- h) MULVANY, Nancy C. **Indexing books.** Chicago; London: Library Materials, 1994. 320p.

Atividade 02: extração dos trechos das publicações analisadas

Nesta atividade foi realizada a extração de trechos que relatavam sobre a elaboração de IFL, destacando-se todas as frases/palavras que trouxessem informações importantes, conforme exposto a seguir.

Trecho 1 - Selecionando assuntos: **Selecionar entre todas as ideias, os assuntos e os conceitos expressos pelo autor** apenas aqueles que são os assuntos do livro, pois somente eles devem ser incluídos em um índice de assunto (BORKO; BERNIER, 1978, p. 59, 60, tradução nossa).

Trecho 2 - Os indexadores **parafraseiam assuntos especificamente para indexá-los**. As palavras usadas nas paráfrases tornam-se partes das entradas do índice. Cada palavra selecionada para indexação recebe uma paráfrase. Mesmo **generalizações devem ser parafraseadas com a máxima especificidade** e não devem ir além do que o autor **relata** (BORKO; BERNIER, 1978, p. 61, tradução nossa).

Trecho 3 - As palavras selecionadas das paráfrases de assuntos são **traduzidas em cabeçalhos de assuntos padrão**, se necessário, para **evitar a dispersão de entradas idênticas ou relacionadas no índice** (BORKO; BERNIER, 1978, p.64, tradução nossa).

Trecho 4 - Sinônimos: **A dispersão entre os sinônimos é inaceitável**. Tal dispersão é impedida (...) [usando] **referências cruzadas para orientar a busca do sinônimo para o cabeçalho do assunto escolhido** (BORKO; BERNIER, 1978, p. 65, tradução nossa).

Trecho 5 - O mais importante das **referências cruzadas do tipo ver é a indicação de sinônimos para os cabeçalhos de assuntos escolhidos** (BORKO; BERNIER, 1978, p. 69, tradução nossa).

Trecho 6 - Referências cruzadas **ver também indicam cabeçalhos relacionados** (BORKO; BERNIER, 1978, p. 69, tradução nossa).

Trecho 7 - Para **determinar a integridade, o indexador deve verificar todo o trabalho uma última vez**. Cada item adicionado exige a preparação de uma nova entrada de índice ou referência cruzada (BORKO; BERNIER, 1978, p. 77, tradução nossa).

Trecho 8 - Como as palavras selecionadas são resumos do assunto, **elas [as palavras] devem ser tão descritivas quanto possível** (BONURA, 1994, p. 52,, tradução nossa).

Trecho 9 - Quando um **assunto explica um conceito deve-se criar uma ou mais entradas no índice** para ajudar o leitor a encontrá-lo. Quando contém a **definição de um termo, crie uma ou mais entradas de índice** apontando o leitor para ele. **Crie entradas de índice quando acrônimos e abreviações forem usados. Deve-se indexar nomes de lugares ou nomes pessoais usados apenas como exemplos, se apropriado** (BONURA, 1994, p. 47, tradução nossa).

Trecho 10 - Usando subentradas: Você deve **decidir se determinados assuntos serão tratados como assuntos principais ou como subentradas**. Entradas principais que não são modificadas por **subentradas não devem ser seguidas por vários números de página** (BONURA, 1994, p. 55, tradução nossa).

Trecho 11 - Ao agrupar entradas, tenha cuidado com o **factoring**, que é o processo de dividir grupos de palavras para que os significados sejam alterados. Para reduzir o **factoring, não use um adjetivo sozinho como assunto de uma entrada principal** (BONURA, 1994, p. 64, tradução nossa).

Trecho 12 - Criando referências cruzadas: **Não devem ser usadas, a menos que realmente remetam a informações adicionais**, não apenas à mesma informação indexada sob outros cabeçalhos. Em referências cruzadas, cabeçalhos e subcabeçalhos são geralmente citados

na íntegra, com inversão e pontuação exatamente como indicado na entrada referenciada (BONURA, 1994, p. 76, tradução nossa).

Trecho 13 - **Entradas sucintas e claras**: As entradas devem ser tão sucintas quanto possível, permanecendo claras. **Evite entradas com utilização de muitos verbos e frases longas** sempre que possível. (MULVANY, 1994, p. 50, tradução nossa).

Trecho 14 - Um dos principais objetivos de marcar páginas para indexação é **indicar as possíveis entradas de índice em cada página do texto**. Os indexadores usam suas próprias anotações abreviadas para marcação de páginas. (MULVANY, 1994, p. 51, tradução nossa).

Trecho 15 - Depois que o indexador marcar as páginas, a próxima etapa é **extrair as entradas das páginas**. Seja escrevendo em fichas de índice ou digitando entradas no computador, o indexador terminará com citações com os indicadores de referência anexados ao final das entradas (MULVANY, 1994, p. 52, tradução nossa).

Trecho 16 - A escolha dos conceitos a serem incluídos nos índices depende das necessidades esperadas dos leitores e a natureza dos documentos indexados. Os **cabeçalhos devem ser específicos** quanto necessário para ajudar o usuário (ISO, 1996, p. 7, tradução nossa).

Trecho 17 - Os **termos adotados devem ser específicos, concisos e uniformes** em todo o índice, **baseando-se**, sempre que possível, [baseando-se] no próprio texto, **normalizando-se, inclusive, as variações de singular e plural**, nos termos utilizados em índices de assunto. (ABNT, 2004, p.3).

Trecho 18 - Deve-se **evitar o uso de artigos, adjetivos, conjunções entre outros. no início dos cabeçalhos** (ABNT, 2004, p. 3).

Trecho 19 - Quando as referências de um **cabeçalho são muito numerosas, é conveniente especificar o subcabeçalhos com seus indicativos próprios**. O(s) subcabeçalho(s) deve(m) ser conciso(s), evitando repetir idéias ou termos do cabeçalho (ABNT, 2004, p. 3).

Trecho 20 - A remissiva ver deve ser elaborada para:

- **termo sinônimo para termo escolhido.**
- **termo popular para termo científico ou técnico.**
- **termo antiquado para termo de uso atual.**
- **sigla para nome completo da entidade** (ABNT, 2004, p. 3).

Trecho 21 - A remissiva ver também deve ser elaborada para cabeçalhos que se relacionem com o cabeçalho proposto (ABNT, 2004, p. 4).

Trecho 22 - Os **itens devem sempre ser indexados sob o termo mais específico** que cobre o assunto. Se um termo não o cobre, uma combinação de termos pode ser usada (BROWNE; JERMEY, 2007, p. 72, tradução nossa)

Trecho 23 - **Abreviações e siglas podem ser usadas como entradas de índice se foram usadas no texto** (BROWNE; JERMEY, 2007, p. 58, tradução nossa).

Trecho 24 - Princípios de indexação

- **Use o idioma do texto;**
- **Use a ordem direta;**
- **Use substantivos e frases nominais;**
- **Digite os termos sob seu nome específico;**

- Use o plural para todos os itens que são contáveis; isso permite o uso do singular para outros significados;
- Use palavras afirmativas, a menos que o negativo seja essencial para transmitir o significado necessário (BROWNE; JERMEY, 2007, p. 58, 59, tradução nossa).

Trecho 25 - Depois de decidir indexar determinados assuntos, você deve **selecionar todas as referências substanciais a esses assuntos** (BROWNE; JERMEY, 2007, p. 54, tradução nossa)

Trecho 26 - **referências ver usadas em termos de entrada não possuem indicadores, pois informam os termos não preferenciais, os termos que possuem os indicadores são os termos preferidos** (BROWNE; JERMEY, 2007, p. 91, tradução nossa).

Trecho 27 - **referência ver também: usadas em entradas de índice que possuem indicadores; elas [referências ver também] sugerem lugares adicionais para procurar** (BROWNE; JERMEY, 2007, p. 91, tradução nossa).

Trecho 28 - Depois de inserir os termos para descrever os conceitos, **os indexadores devem avaliar suas entradas editando os termos escolhidos inicialmente para tornar o índice um todo coerente** (BROWNE; JERMEY, 2007, p. 113, tradução nossa).

Trecho 29 - A próxima etapa é **inserir os números das páginas** usando uma versão eletrônica do texto final do conjunto de páginas. (FRANCIS; GREENWAY, 2015, p. 294, tradução nossa).

Trecho 30 - É importante **incluir subentradas que identifiquem em palavras a temática que está sendo discutida na página referenciada. Caso contrário, os números de páginas que resultam [em sequência grande de números], são inúteis porque desperdiçam o tempo do pesquisador em baralhar páginas tediosas e principalmente infrutíferas** (MILLER, 2013, p. 36).

Após a apresentação dos trechos selecionados para análise, bem como o destaque das palavras que integram as diretrizes para elaboração do IFL, destaca-se a seguir sobre a estruturação.

4.3.2.2.2 Seção 2.2: estruturação de IFL

Visando atender aos objetivos propostos nesta pesquisa, esta seção pretende analisar as publicações que relatam sobre a estruturação do IFL. Para isso, são abordados as atividades e seus resultados.

Atividade 01: publicações selecionadas nesta seção:

- a) ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6034:** informação e documentação: Índice: apresentação. Rio de Janeiro, 2004.
- b) BONURA, Larry S. **The art of indexing**. New York: Wiley, 1994. 233 p.

- c) BORKO, Harold; BERNIER, Charles L. **Indexing concepts and methods**. New York: Academic Press, 1978. 261 p.
- d) BROWNE, Glenda; JERMEY, Jon. **The indexing companion**. Austrália: Ligary, 2007. 249 p.
- e) MULVANY, Nancy C. **Indexing books**. Chicago; London: Library Materials, 1994. 320p.

Atividade 02: extração dos trechos das publicações analisadas

Apresentam-se os trechos selecionados, bem como, o destaque das palavras significativas que serão insumo de análise.

Trecho 1 - Quanto à ordenação, o índice pode ser em: a) ordem alfabética; b) ordem sistemática; c) ordem cronológica; d) ordem numérica; e) ordem alfanumérica (ABNT, 2004, p. 2).

Trecho 2 - Quanto ao enfoque, o índice pode ser: a) especial, quando organizado por: - autores; - assuntos; - títulos; - pessoas e/ou entidades; - nomes geográficos; - citações; - anunciantes e matérias publicitárias (ABNT, 2004, p. 2).

Trecho 3 - Quanto ao enfoque, o índice pode ser: b) geral, quando combinadas duas ou mais das categorias indicadas na alínea a). Exemplo: Índice de autores e assuntos (ABNT, 2004, p. 2).

Trecho 4 - Localização: O índice deve ser impresso no final do documento, com paginação consecutiva ou em volume separado (ABNT, 2004, p. 2).

Trecho 5 - O índice deve ser organizado de acordo com um padrão lógico e facilmente identificável pelos usuários. Quando a forma adotada na elaboração do índice ocasionar duplicidade de interpretações, deve-se acrescentar, no início do índice, uma nota explicativa do padrão adotado e das exceções eventuais (ABNT, 2004, p. 3).

Trecho 6 - O título do índice deve definir sua função e/ou conteúdo (ABNT, 2004, p. 3).

Trecho 7 - Em índice alfabético, recomenda-se imprimir, no canto superior externo de cada página, as letras iniciais ou a primeira e última entradas da página (ABNT, 2004, p. 3).

Trecho 8 - No índice geral, as entradas de cada categoria devem ser diferenciadas graficamente e ordenadas conforme a ABNT NBR 6033 (ABNT, 2004, p. 3).

Trecho 9 - Recomenda-se a apresentação das entradas em linhas separadas, com recuo progressivo da esquerda para a direita para subcabeçalhos (ABNT, 2004, p. 3).

Trecho 10 - O cabeçalho pode ser, quando necessário, qualificado por uma expressão modificadora que lhe explicita o significado (ABNT, 2004, p. 3).

Trecho 11 - Em cabeçalhos compostos, as entradas devem ser elaboradas pelas palavras significativas, fazendo-se remissivas ou novas entradas para as palavras passíveis de serem procuradas. Exemplo: Elaboração de índice ver *Índice, elaboração de*. (ABNT, 2004, p. 3).

Trecho 12 - NOTA O recurso tipográfico (negrito, sublinhado, itálico ou outro) deve ser utilizado para destacar as expressões *ver* e *ver também* (ABNT, 2004, p. 4).

Trecho 13 - [indicativo índice deve ser apresentado por] números extremos, ligados por hífen, quando o texto abranger páginas ou seções consecutivas (ABNT, 2004, p. 4).

Trecho 14 - [indicativo índice deve ser apresentado por] números separados por vírgula, quando o texto abranger páginas ou seções não consecutivas (ABNT, 2004, p. 4).

Trecho 15 - [indicativo índice deve ser apresentado por] número do volume ou parte correspondente, seguido de página(s) ou seção(ões) consecutivas ou não em documento composto de mais de um volume ou parte (ABNT, 2004, p. 4).

Trecho 16 - Organize as subentradas em ordem alfabética. Normalmente, alfabete a palavra mais importante primeiro (BONURA, 1994, p. 58, tradução nossa).

Trecho 17 - Organizando assuntos: Sempre organize entradas por ênfase. Coloque a palavra mais importante primeiro, se a legibilidade não for afetada (BONURA, 1994, p. 62, tradução nossa).

Trecho 18 - No método *palavra por palavra*, escreva em ordem alfabética até o final da primeira palavra e depois pare. Use palavras secundárias e subsequentes somente quando dois ou mais cabeçalhos começarem com a mesma palavra ou palavras. Se houver um hífen separando as palavras, trate as palavras como elementos individuais (BONURA, 1994, p. 71, tradução nossa).

Trecho 19 - Formato: Uma característica importante do formato é a organização de subcabeçalhos em estilo de entrada por linha ou de parágrafo. O estilo de entrada por linha é: fácil de folhear e ler (...). No entanto, o estilo de entrada por linha ocupa mais espaço do que o estilo de parágrafo. Alguns índices de *estilo de parágrafo organizam os subcabeçalhos por número de página, em vez de alfabeticamente* (BORKO; BERNIER, 1978, p. 73, tradução nossa).

Trecho 20 - As entradas são digitadas em espaço duplo (...) os cabeçalhos começam na margem esquerda. Para entradas com subcabeçalhos, as vírgulas vêm antes e depois do subcabeçalho. Os subcabeçalhos são todos em letras minúsculas, exceto nomes próprios, acrônimos, iniciais e símbolos que possuem maiúsculas. Entradas de índice que são muito longas para ajustar na linha impressa são executadas na próxima linha após o recuo de seis espaços. As entradas são exibidas na próxima página com o cabeçalho repetido e colocadas em itálico e parênteses (BORKO; BERNIER, 1978, p. 78, tradução nossa).

Trecho 21 - Desambiguação: Onde a mesma entrada se refere a duas entidades diferentes, informações distintas devem ser adicionadas entre parênteses após o termo. Se aplica a termos homógrafos, locais e nomes (BROWNE; JERMEY, 2007, p. 56, tradução nossa).

Trecho 22 - Arranjo de subentradas. A escolha predominante para a ordem de subentradas é alfabética. O tipo específico de ordem alfabética é o mesmo usado para classificar os cabeçalhos principais (MULVANY, 1994, p. 55, tradução nossa).

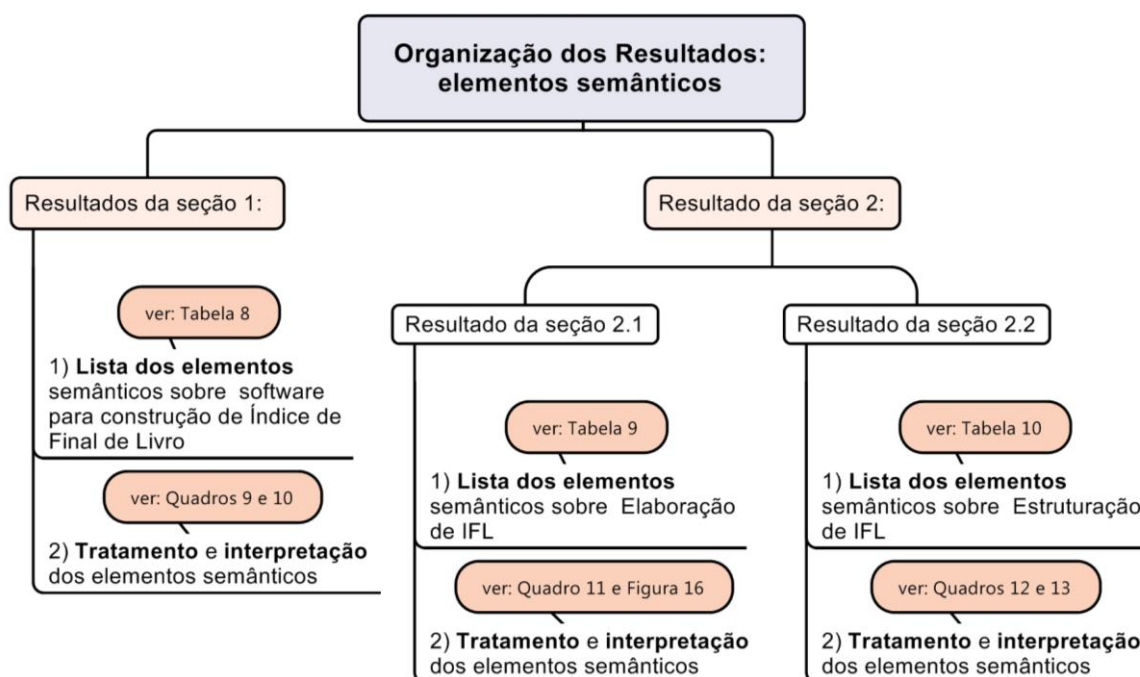
Listados os trechos e destacadas todas as frases/palavras significativas, apresenta-se, a seguir, o resultado da pesquisa que são os elementos semânticos das publicações. Os resultados também são tratados e interpretados no capítulo a seguir.

5. RESULTADOS

“Sabemos que na elaboração de um bom índice, há muito mais campo para o exercício de bom senso e competência do que comumente se supõe. Reconhecemos os méritos do compilador de tal índice e estamos dispostos a externar nosso reconhecimento pelos seus esforços (Dr. Samuel Austin Allibone: Critical Dictionary of Authors)”

Neste capítulo são apresentados os resultados obtidos a partir da aplicação do método de Análise de conteúdo descrito no capítulo 4. A lógica de apresentação dos resultados foi inspirada na dissertação de Aganette (2010). Assim, a organização tem as mesmas seções do capítulo anterior (Seção 1 e 2). Os resultados são apresentados em dois momentos: 1) lista dos elementos semânticos sobre software para construção de IFL e 2) Tratamento e interpretação dos elementos semânticos. Seguindo a estruturação das fases e atividades do método descrito, na Figura 14 a seguir são esquematizados os resultados deste capítulo.

Figura 14 - Esquema dos resultados da dissertação



Fonte: elaborado pela autora.

5.1 Resultados da Seção 1

A partir dos elementos semânticos destacados no capítulo anterior esta seção tem o objetivo de listar e apresentar os resultados relativos à softwares para construção de IFL, conforme demonstrado a seguir.

5.1.1 Lista dos elementos semânticos sobre software para construção de Índice de Final de Livro

Os elementos semânticos¹⁷ resultante do capítulo anterior foram extraídos da terceira coluna “padronização dos termos” da Tabela 8, em seguida, foram agrupados pelos artigos correspondente ao software, os elementos semânticos são organizados por ordem alfabética. Ao lado de cada elemento semântico é identificado o número que corresponde ao seu trecho de origem, conforme apresentado a seguir.

Tabela 8 - Apresentação dos elementos semânticos (continua)

ajuda o indexador a criar bons índices	21. 26. 27
assistência precisa na determinação de termos	1
estilo por linha	7
estilo por parágrafo	8
formatar a estrutura do texto não alterando os termos do índice	13
fornece interface limpa	4
impossibilidade de um subcabeçalho ser um cabeçalho	17
insere os localizadores em cada entrada	24
limitado em relação a utilização de remissivas	14
não confunde padrões de editor de texto	10
não possui esquema de classificação (entradas palavra por palavra ou letra por letra)	15
não tem possibilidade de criar parte de índice (exemplo: quando os índices de cada capítulo são mesclados para formar o índice)	16
necessidade de verificação do indexador	22. 28
números de página (não apresenta capítulo ou seção)	5
organiza as entradas na ordem desejada	25
permitem encontrar todas as ocorrências nos termos	2
possibilidade de criar remissivas	6
possibilidade de fazer edições	12
possibilidade de fazer formatações	11. 20
precisa eliminar localizadores que não fornecem informações relevantes para o leitor	23

¹⁷ O objetivo de tais elementos semânticos é utiliza-los para responder os 8 critérios que os programas de indexação devem ser avaliados. E também comparar os softwares identificando as vantagens e desvantagens de cada um.

Tabela 8 - Apresentação dos elementos semânticos (continua)

Elementos semânticos	Nº do trecho extraído do artigo (ABE, BERRY, 1989)
simples de usar	18. 19
sugestões sobre termos	3
texto de entrada limpo	9
Elementos semânticos	Nº do trecho extraído do artigo (CHI et al., 2004; 2006)
aprimoramento de navegação	15. 23
baseado em similaridades conceituais de termos	3
desempenho melhor que o índice de final de livro impresso	35
entradas relevantes não são omitidas	27
facilidade de visualização, pois apresenta em única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas	9
fornece dicas de navegação aos leitores	2. 19. 25
fornece realce nos termos pesquisados	5. 16. 22. 40
independente do conhecimento do conteúdo do livro o leitor consegue localizar a informação	12. 30. 37
leitor pode restringir o número de entradas	8. 29
leitores agradaram da interface	39
leitores localizam os termos de forma mais precisa	11. 14. 21. 38
leitores localizam os termos mais rapidamente	6. 10. 20. 31. 36.
leitores têm uma alta confiança	26
mais fácil de comparar entradas de índice	34
navegação entre o ifl e o e-book aprimorada	1. 24
permite rápida varredura nos termos relevantes	17
possibilidade clicar no localizar e direcionar para página	33
possibilidade de formatação do índice	18. 28
possibilidade de formatação e edição do índice	13. 41
possibilidade de pesquisar usando combinações de termos	32
possibilidade de reorganizar as entradas de índice	7
preocupa com as necessidades de informação do leitor	4
Elementos semânticos	Nº do trecho extraído do artigo (MEKKI; NAZARENKO, 2005)
bom desempenho do algoritmo de classificação de conhecimento	4
considera que o algoritmo de segmentação e classificação é mais robusto	12
depende do estilo do documento [as sugestões do software]	11
extrai os termos ou frases mais relevantes do documento original	1. 5
o método de segmentação utiliza as abordagens estruturais e linguísticas que são mais relevantes	3
pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase linguísticos ou tipográficos	6. 10
possibilidade de classificação	7
possibilidade de formatação	8
simples para construir automaticamente a rede hipertextual que ajuda na navegação pelo documento	9
vincula cada cabeçalho aos localizadores do documento	2

Tabela 8 - Apresentação dos elementos semânticos (conclusão)

Elementos semânticos	Nº do trecho extraído do artigo (PROVO, 2019)
facilidade em inserir cabeçalhos no índice	2
fornece navegação sequencial entre as páginas do livro	12
necessidade de verificação do indexador	5
o software retira os sufixos e pontuações dos termos	3
possibilidade de acumular índices	1
possibilidade de classificação	6. 7
possibilidade de criar remissivas	4
possibilidade de edição	9. 11
possibilidade de formatação	8. 10
possibilidade de limitar a visualização da pesquisa	18
possibilidade de utilizar vocabulário controlado	17
problemas com termos abreviados (criam relações irrelevantes no índice)	13
problemas semânticos	14. 15. 16

Fonte: elaborada pela autora.

































5.1.2 Tratamento e interpretação dos elementos semânticos

Dos *software* descritos nos artigos, nenhum foi localizado para *download* nem mesmo foram identificadas informações sobre forma de aquisição, possivelmente em razão do uso ser restrito ou devido à descontinuidade ou indisponibilidade. Entende-se que a análise das funcionalidades auxilia o profissional no momento da escolha do *software* que será utilizado para auxiliar na elaboração do IFL. São descritos no (Apêndice A - Softwares de indexação para Índice de Final de Livro) alguns softwares existentes, mas todos pagos.

Para Fetters (1986), os programas de indexação devem ser avaliados levando-se em consideração 8 critérios. O primeiro diz respeito à “**facilidade de inserir cabeçalhos no índice**”, que significa a seleção do termo e sua identificação toda vez que o mesmo aparecer no texto. O segundo critério trata-se da “**capacidade de criar remissivas**”, pois um bom índice apresenta relações entre os termos. O terceiro critério contempla a “**facilidade de edição na entrada do índice**”, sendo necessário que o *software* proporcione formas de edição das entradas e assim facilite as inclusões ou exclusões de termos sem que modifique o Índice no todo. Outro critério apontado pelo autor é a “**capacidade de classificação**”, com a possibilidade de classificar os termos de diferentes maneiras. O critério “**limitações de tamanho**” relaciona-se com a apresentação das entradas. Já o critério “**capacidade de formatação**” é a flexibilidade de organizar, estruturar e alterar a estrutura do índice. Os “**efeitos de impressão**” referem-se à possibilidade de imprimir o IFL com opções de personalização. Por fim, tem-se a “**capacidade de acumular índices**”, que aponta para a possibilidade de criar partes do índice e posteriormente, gerar um índice inteiro, exemplo: índices para capítulos e depois o índice do livro inteiro agrupando-os.

Durante a análise das publicações, buscou-se identificar trechos dos artigos que pudessem evidenciar se o *software* atendia ou não tais critérios. O resultado é apresentado no Quadro 9.

Quadro 9 - Critérios utilizados para avaliação do software

Critério	<i>Software 01: Indx (A)</i>	<i>Software 03: ScentIndex (B)</i>	<i>Software 02: IndDoc (C)</i>	<i>Software 04: Topic Curation Toolkit (TCT) (D)</i>
(1) facilidade de inserir cabeçalhos no índice (<i>ease of entering index headings</i>)	 1. 24. ¹⁸		 1. 2. 5.	 2.
(2) capacidade de criar remissivas (<i>ability to create cross-references</i>)	 6.			 4.
(3) facilidade de edição na entrada do índice (<i>ease of editing index entries</i>)	 12. 25.	 5. 7. 13. 16. 22. 40. 41.		 9. 11.
(4) capacidade de classificação (<i>sorting capabilities</i>)	 15.		 7.	 6. 7.
(5) limitações de tamanho (<i>size limitations</i>)		 8. 29.		 18.
(6) capacidade de formatação (<i>formatting capabilities</i>)	 7. 8. 11. 13. 20. 25.	 13. 41. 18. 28	 8.	 8. 10.
(7) efeitos de impressão (<i>printing effects</i>)				
(8) capacidade de acumular índices (<i>ability to cumulate indexes</i>)	 5. 16.			 1.

Fonte: elaborado pela autora.


Legenda:



¹⁸ Número extraído dos elementos semânticos das publicações, os números podem ser conferidos na Tabela 8.

Atende: de acordo com a afirmação do autor descrito na referência utilizada.

 *Não atende*: de acordo com a afirmação do autor descrito na referência utilizada.

 *Não informado*: quando o autor não mencionou nenhuma informação que pudesse responder ao critério.

As análises das publicações resultaram na identificação de vantagens e desvantagens dos *softwares*. Essas informações foram organizadas no Quadro 10, e têm como propósito contribuir com a definição dos requisitos necessários para *softwares* utilizados para construção do IFL.

Quadro 10 - Comparativo dos Softwares (continua)¹⁹

SOFTWARE	DESCRIÇÃO	VANTAGENS	DESVANTAGENS	REFERÊNCIA UTILIZADA
Indx	Programa de indexação semiautomático, que auxilia no processo de determinação dos termos, bem como design e implementação do IFL.	<ul style="list-style-type: none"> • ajuda o indexador a criar bons índices (21. 26. 27) • assistência precisa na determinação de termos (1) • sugestões sobre termos (3) • formata a estrutura do texto não alterando os termos do índice (13) • fornece interface limpa (4) • texto de entrada limpo (9) • não confunde padrões de editor de texto (10) • simples de usar (18. 19) • organiza as entradas na ordem desejada (25) <p>permite encontrar todas as ocorrências nos termos (2)</p>	<ul style="list-style-type: none"> • impossibilidade de um subcabeçalho ser um cabeçalho (17) • limitado em relação a utilização de remissivas (14) • não possui esquema de classificação (entradas palavra por palavra ou letra por letra) (15) • não tem possibilidade de criar parte de índice (exemplo: quando os índices de cada capítulo são mesclados para formar o índice) (16) • necessidade de verificação do indexador (22. 28) <p>precisa eliminar indicadores que não fornecem informações relevantes para o leitor (23)</p>	Abe e Berry (1989)

¹⁹ Números extraídos dos elementos semânticos das publicações, os números podem ser conferidos na Tabela 8.

Quadro 10 - Comparativo dos Softwares (continua)

SOFTWARE	DESCRIÇÃO	VANTAGENS	DESVANTAGENS	REFERÊNCIA UTILIZADA
ScentIndex	utiliza o método que reorganiza conceitualmente grandes índices de assuntos de acordo com alguma necessidade de informação. Desta forma, é uma nova maneira de utilizar índices de assunto em livros eletrônicos.	<ul style="list-style-type: none"> • aprimoramento de navegação (15. 23) • baseado em similaridades conceituais de termos (3) • desempenho melhor que o índice de final de livro impresso (35) • entradas relevantes não são omitidas (27) • facilidade de visualização (pois apresenta em única tela o que pode ser relevante sem precisar consultar várias entradas de índice dispersas) (9) • fornece dicas de navegação aos leitores (2. 19. 25) • independente do conhecimento do conteúdo do livro o leitor consegue localizar a informação (12. 30. 37) • leitores agradaram da interface (39) • leitores localizam os termos mais rapidamente (6. 10. 20. 31. 36) • leitores localizam termos de forma mais precisa (11. 14. 21. 38) • leitores têm uma alta confiança (26) • mais fácil de comparar entradas de índice (34) • navegação entre o IFLe o e-book aprimorada (1. 24) • permite rápida varredura nos termos relevantes (17) • possibilidade clicar no localizar e direcionar para página (33) • possibilidade de reorganizar as entradas de índice (7) <p>preocupa com as necessidades de informação do leitor (4)</p>		<p>Chi <i>et al.</i> (2004)</p> <p>Chi <i>et al.</i> (2006).</p>

Quadro 10 - Comparativo dos Softwares (conclusão)

SOFTWARE	DESCRIÇÃO	VANTAGENS	DESVANTAGENS	REFERÊNCIA UTILIZADA
IndDoc	Constrói uma rede hipertextual que ajuda na navegação de um documento, e dessa forma auxilia no momento de selecionar os termos relevantes no índice final e vincula cada cabeçalho aos indicadores do documento.	<ul style="list-style-type: none"> • bom desempenho do algoritmo de classificação de conhecimento (4) • considera que o algoritmo de segmentação e classificação é mais robusto (12) • o método de segmentação utiliza as abordagens estruturais e linguísticas que são mais relevantes (3) • pontuação de relevância baseia-se nos pesos das palavras, estrutura do documento e marcadores de ênfase lingüísticos ou tipográficos (6. 10) • depende do estilo do documento [as sugestões do software] (11) 	<ul style="list-style-type: none"> • simples para construir automaticamente a rede hipertextual que ajuda na navegação pelo documento (9) 	Mekki e Nazarenko (2005)
Topic Curation Toolkit (TCT)	é um aplicativo de <i>software</i> livre criado na estrutura da web do Django com um banco de dados relacional do Postgres.	<ul style="list-style-type: none"> • facilidade em inserir cabeçalhos no índice (2) • possibilidade de utilizar vocabulário controlado (17) • fornece navegação sequencial entre as páginas do livro (12) <p>o software retira os sufixos e pontuações dos termos (3)</p>	<ul style="list-style-type: none"> • necessidade de verificação do indexador (5) • problemas com termos abreviados (criam relações irrelevantes no índice) (13) <p>problemas semânticos (14. 15. 16)</p>	Provo (2019)

Fonte: elaborado pela autora.

Fazendo uma interseção entre os objetivos apresentados nos artigos e os subsídios que esses dispuseram para as análises, pode-se inferir, em linhas gerais, que a maioria dos autores relata mais elogios/vantagens dos *softwares*, aprofundando superficialmente em críticas ou melhorias. Em exceção, o artigo dos autores Abe e Berry (1989) foi o que apresentou informações mais detalhadas quanto às vantagens dos *software* e itens que precisam de aprimoramento. Considera-se que tal resultado decorreu do fato de que objetivo do artigo era apresentar o *design*, a implementação e a aplicação do Indx, favorecendo, assim, a identificação do que poderia ser melhorado.

Os artigos de Chi *et al.* (2004, 2006) que relatam sobre o *software* ScentIndex têm o intuito de apresentar um estudo de usuário e analisar o uso do *software*. Dessa maneira, entende-se que os artigos esclarecem sobre a usabilidade e relatam com menos intensidade sobre a elaboração. Já o artigo de Mekki e Nazarenko (2005) tem como objetivo discutir sobre os algoritmos projetados para que o *software* atendesse a diversas demandas. Entende-se que, por essa razão, esse estudo também ofereceu poucos subsídios para a elaboração do IFL no *software*. E Provo (2019) também tem a finalidade de relatar sobre a construção do IFL com o uso do *software* TCT, mas principalmente identifica os pontos de melhoria que merecem atenção após testes de usabilidade. Tais afirmativas reiteram que todas as análises e os resultados foram condicionados à opinião dos autores e à cobertura de assunto que cada artigo propôs esclarecer.

Partindo-se dessa premissa, os resultados alcançados no Quadro 9 indicam que os *softwares* Indx, IndDoc e TCT proporcionam facilidade quanto à inserção de entradas no IFL. Somente o Indx e TCT afirmam ter a opção para inserir remissivas. As edições na entrada do índice trata-se de uma funcionalidade atendida pela Indx, ScentIndex e TCT. A possibilidade de classificar os termos por ordem alfabética e outras classificações são atendidas pelo IndDoc e TCT. Para facilitar a visualização dos resultados recuperados no IFL, principalmente os criados para atender e-book, o ScentIndex e TCT atendem tal quesito. Observou-se que todos os *softwares* têm capacidade de formatação. Nenhum dos artigos informou se existe a possibilidade de realizar efeitos de impressão após a finalização do IFL. E somente o TCT tem a capacidade de acumular índices, pois o objetivo do projeto que o utilizou tinha essa finalidade.

Sobre a comparação dos *softwares*, todos afirmam que o auxílio oferecido é satisfatório. No caso do Indx, essa afirmação é mais focada no indexador; no ScentIndex, o leitor é avaliado com mais precisão; no IndDoc, o foco é no próprio programa de indexação; o TCT preocupa-se em avaliar o *software* e o leitor. Os enfoques diferenciados contribuíram para que a análise do Indx respondesse mais as respostas quanto aos critérios; no IndDoc recupera-se mais informações quanto à apresentação, visualização e navegação do

software, com foco no atendimento do usuário; e o ScenIndex oferecesse mais dados técnicos quanto a algoritmos e teste de desempenho, o que não impediu identificar características, e o TCT trouxe mais informações sobre as características do *software*, não abordando em profundidade sobre as vantagens e desvantagens.

Conclui-se que o Indx atende a boa parte dos critérios que bons programas de indexação devem conter. Embora não dispense a avaliação do indexador, como afirmado pelos autores, pois existem muitos aspectos cruciais em um índice que esse *software* ainda não atende, como limitações na utilização de remissas e de classificação das entradas. Já o *software* ScenIndex para índices em e-books é muito útil, pois auxilia o leitor durante sua busca informacional. O IndDoc atende à tarefa de criação de um IFL, mas a proposta de criar uma rede hipertextual no documento não foi atendida, limitando o indexador à elaboração de um índice de maneira “tradicional”. E o TCT foi o *software* que mais atendeu aos quesitos do índice de qualidade, e este é utilizado mais para livros digitais e por indexadores que pretendem elaborar mais de um IFL para livros diferentes. Apresentadas as análises sobre os *softwares* para construção de IFL, na próxima seção discorre-se sobre a Construção de IFL com indexação manual.

5.2 Resultados da Seção 2

A partir dos elementos semânticos destacados no capítulo anterior esta seção tem o objetivo de listar e apresentar os resultados relativos construção de IFL, que é subdividido em seção 2.1: relativo à elaboração do IFL e seção 2.2 relativo à estruturação do IFL. A elaboração e estruturação possivelmente acontecem de forma simultânea durante a construção do IFL. Mas, para facilitar a visualização e o entendimento deste trabalho, essas atividades são apresentadas separadamente, pressupondo que a elaboração acontece primeiro e, com os termos escolhidos, parte-se para a estruturação do IFL.

5.2.1 Resultado da seção 2.1

Na construção de IFL a elaboração é considerada uma das primeiras atividades realizadas. A seguir são apresentados os resultados extraídos das publicações analisadas.

5.2.1.1 Lista dos elementos semânticos sobre elaboração de IFL

A lista apresentada a seguir na Tabela 9 é composta por 44 elementos semânticos identificados e extraídos dos trechos das publicações e organizados alfabeticamente (trechos apresentandos na seção 4.3.2.2.1 – atividade 2). Os números

acrescidos após cada elemento semântico correspondem à sua identificação ao trecho da publicação.

Tabela 9 - Elementos semânticos identificados e selecionados (continua)

Elementos semânticos selecionados	Identificação dos 30 trechos dos artigos selecionados
a dispersão entre os sinônimos é inaceitável	4
abreviações e siglas podem ser usadas como entradas de índice se foram usadas no texto	23
assunto que explica um conceito exige a criação de uma ou mais entradas de índice	9
cabeçalhos que são muito numerosos, é conveniente especificá-los por subcabeçalhos com seus indicativos próprios	19
cabeçalhos devem ser específicos	16
crie entradas de índice quando acrônimos e abreviações forem usados	9
decidir se determinados assuntos serão tratados como assuntos principais ou como subentradas	10
definição de um termo, crie uma ou mais entradas de índice	9
determinar a integridade, o indexador deve verificar todo o trabalho uma última vez	7
digite os termos sob seu nome específico	24
elas [as palavras] devem ser tão descritivas quanto possível	8
entradas sucintas e claras	13
evitar a dispersão de entradas idênticas ou relacionadas no índice	3
evitar o uso de artigos, adjetivos, conjunções entre outros. no início dos cabeçalhos	18
evite entradas com utilização de muitos verbos e longos	13
generalizações devem ser parafraseadas com a máxima especificidade	2
incluir subentradas que identifiquem em palavras a temática que está sendo discutida na página referenciada	30
indexar referências a nomes de lugares ou nomes pessoais usados apenas como exemplos, se apropriado	9
indicar as possíveis entradas de índice em cada página do texto	14
inserir os números das páginas	29
itens devem sempre ser indexados sob o termo mais específico	22
não devem ser usados [remissivas], a menos que realmente remetam a informações adicionais	12
não use um adjetivo sozinho como assunto de uma entrada principal	11
números de páginas que resultam [em sequência grande de números], são inúteis porque desperdiçam o tempo do pesquisador	30
os indexadores devem avaliar suas entradas editando os termos iniciais para tornar o índice em um todo coerente	28
parafraseie assuntos especificamente para indexá-los	2
referência ver também: usadas em entradas de índice que possuem indicadores; eles [referência ver também] sugerem lugares adicionais para procurar	27
referências ver usadas em termos de entrada não possuem indicadores, pois informam os termos não preferenciais, os termos que possuem os indicadores são os termos preferidos	26

Tabela 9 - Elementos semânticos identificados e selecionados (conclusão)

Elementos semânticos selecionados	Identificação dos 30 trechos dos artigos selecionados
referências cruzadas para orientar a busca do sinônimo para o cabeçalho do assunto escolhido	4
referências cruzadas <i>ver</i> são a indicação de sinônimos para os cabeçalhos de assuntos escolhidos	5
remissiva <i>ver</i> também deve ser elaborada para cabeçalhos que se relacionem com o cabeçalho proposto	21
[remissiva <i>ver</i> deve ser elaborada para] sigla para nome completo da entidade	20
[remissiva <i>ver</i> deve ser elaborada para] termo antiquado para termo de uso atual	20
[remissiva <i>ver</i> deve ser elaborada para] termo popular para termo científico ou técnico	20
[remissiva <i>ver</i> deve ser elaborada para] termo sinônimo para termo escolhido	20
selecionar dentre todas as ideias, assuntos e conceitos expressos pelo autor	1
selecionar todas as referências substanciais a esses assuntos	25
subcabeçalho(s) deve(m) ser conciso(s), evitando repetir idéias ou termos do cabeçalho	19
subentradas não devem ser seguidas por vários números de página	10
termos adotados devem ser específicos, concisos e uniformes	17
use o idioma do texto	24
use palavras afirmativas, a menos que o negativo seja essencial para transmitir o significado necessário	24
use substantivos e frases nominais	24
<i>ver</i> também indicam cabeçalhos relacionados	6

Fonte: elaborado pela autora.

O tratamento e interpretação dos elementos semânticos é apresentada na seção a seguir.

5.2.1.2 Tratamento e interpretação dos elementos semânticos

Para efetivação da etapa de elaboração do IFL o indexador realiza antecipadamente a leitura das partes do livro que são relevantes, conforme demonstrado no Quadro 3 dessa dissertação. O Quadro 11 explicita os agrupamentos dos elementos semânticos de acordo com a parte do IFL a qual se destina. Dessa maneira, são identificadas afirmações relacionadas aos termos, ao cabeçalho, ao subcabeçalho, às remissivas e aos localizadores. Para se chegar a esses resultados, foi necessário analisar os elementos semânticos que estavam relacionados, ou seja, que afirmavam a mesma argumentativa, e posteriormente, foram identificados os verbos necessários para efetivar as ações e os processos relatados em cada agrupamento, conforme apresentado a seguir.

Quadro 11 – Elementos semânticos relacionados à elaboração de IFL (continua)

PARTE DO IFL	AÇÕES	ELEMENTOS SEMÂNTICOS
TERMOS	Utilizar	<ul style="list-style-type: none"> ▪ termos específicos, concisos e uniformes - 17 ▪ o termo mais específico - 22 ▪ generalizações com paráfrases utilizando a máxima especificidade - 2 ▪ cabeçalhos específicos - 16 ▪ os termos sob seu nome específico - 24 ▪ assuntos parafraseados especificamente para indexá-los - 2 ▪ as palavras de forma descritiva - 8 ▪ o idioma do texto - 24 ▪ substantivos e frases nominais - 24 ▪ palavras afirmativas, a menos que o negativo seja essencial para transmitir o significado necessário - 24
	Evitar	<ul style="list-style-type: none"> ▪ entradas com utilização de muitos verbos e frases longas - 13 ▪ indexar nomes de lugares ou nomes pessoais usados apenas como exemplos, se apropriado - 9
	Selecionar	<ul style="list-style-type: none"> ▪ dentre todas as ideias, assuntos e conceitos expressos pelo autor - 1 ▪ todas as referências substanciais a esses assuntos - 25
	Revisar	<ul style="list-style-type: none"> ▪ as entradas editando os termos escolhidos inicialmente para tornar o índice em um todo coerente - 28 ▪ todo o trabalho uma última vez para determinar a integridade - 7
CABEÇALHO	Criar o cabeçalho quando	<ul style="list-style-type: none"> ▪ o assunto explica um conceito deve-se criar uma ou mais entradas no índice - 9 ▪ tem definição de um termo, deve-se criar uma ou mais entradas de índice - 9
	Deve ter cabeçalhos Evitar no início dos cabeçalhos o uso de	<ul style="list-style-type: none"> ▪ sucintos e claros - 13 ▪ de artigos, adjetivos, conjunções entre outros no início dos cabeçalhos - 18 ▪ adjetivo sozinho como assunto de uma entrada principal - 11 ▪ abreviações e siglas podem ser usadas como entradas de índice se foram usadas no texto - 23 ▪ acrônimos e abreviações - 9

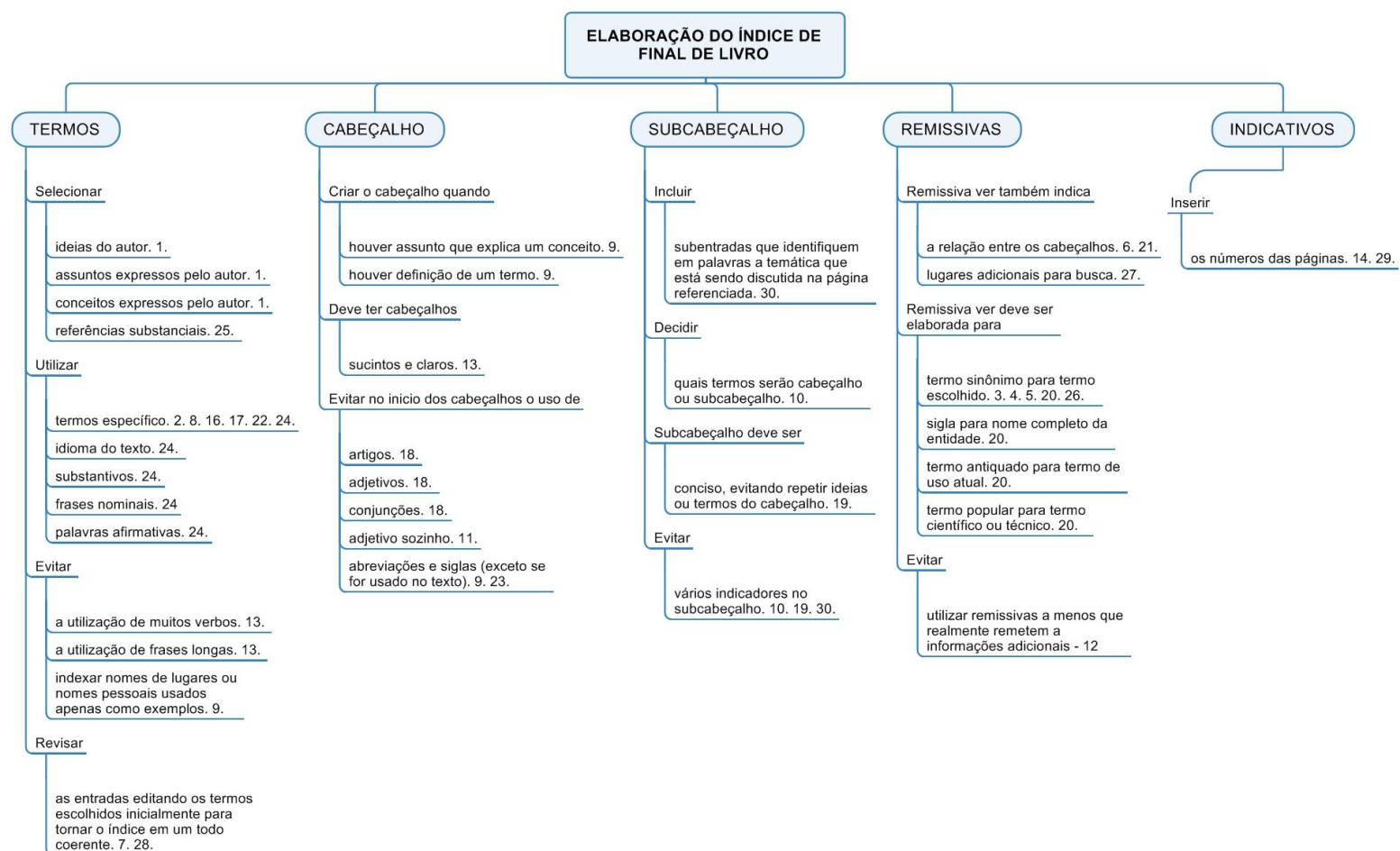
Quadro 11 - Elementos semânticos relacionados à elaboração de IFL (conclusão)

PARTE DO IFL	AÇÕES	ELEMENTOS SEMÂNTICOS
SUBCABEÇALHO	Evitar Decidir Incluir Deve ser	<ul style="list-style-type: none"> ▪ cabeçalho que são muito numerosos, é conveniente especificá-lo por subcabeçalhos com seus indicativos próprios - 19 ▪ números de páginas que resultam [em sequência grande de números], são inúteis porque desperdiçam o tempo do pesquisador - 30 ▪ vários números de página seguidos - 10 ▪ se determinados assuntos serão tratados como assuntos principais ou como subentradas - 10 ▪ subentradas que identifiquem em palavras a temática que está sendo discutida na página referenciada - 30 ▪ ser conciso(s), evitando repetir ideias ou termos do cabeçalho - 19
REMISSIVAS	Remissiva ver também Remissiva ver Evitar	<ul style="list-style-type: none"> ▪ usadas em entradas de índice que possuem indicadores; eles [referência ver também] sugerem lugares adicionais para procurar - 27 ▪ deve ser elaborada para cabeçalhos que se relacionem com o cabeçalho proposto - 21 ▪ indicam cabeçalhos relacionados - 6 ▪ orienta a busca do sinônimo para o cabeçalho do assunto escolhido - 4 ▪ são a indicação de sinônimos para os cabeçalhos de assuntos escolhidos - 5 ▪ deve ser elaborada para termo sinônimo para termo escolhido - 20 ▪ deve ser elaborada para evitar a dispersão entre os sinônimos - 4 ▪ evitar a dispersão de entradas idênticas ou relacionadas no índice - 3 ▪ usadas em termos de entrada não possuem indicadores, pois informam os termos não preferenciais, os termos que possuem os indicadores são os termos preferidos - 26 ▪ deve ser elaborada para sigla para nome completo da entidade - 20 ▪ deve ser elaborada para termo antiquado para termo de uso atual - 20 ▪ deve ser elaborada para termo popular para termo científico ou técnico - 20 ▪ deve ser usadas, a menos que realmente remetam a informações adicionais - 12
LOCALIZADORES	Inserir	<ul style="list-style-type: none"> ▪ as possíveis entradas de índice em cada página do texto - 14 ▪ os números das páginas - 29

Fonte: elaborado pela autora.

Com a finalidade de facilitar a visualização das diretrizes para elaborar o IFL, representa-se uma estrutura dos dados apresentados no quadro anterior, utilizando a ferramenta MindManager. A padronização dos termos foi realizada com o objetivo de reduzir as frases sem perder o sentido e, dessa maneira, facilitar a visualização, conforme exposto na Figura 15 a seguir.

Figura 15 – Elementos semânticos relacionados à elaboração de IFL



Fonte: elaborado pela autora.

Durante a análise das publicações, foram identificados alguns trechos com argumentação distinta em relação a fundamentar-se ou não no texto durante a leitura e seleção dos termos. Para Borko e Bernier (1978, p. 61) e para a ABNT (2004), o indexador não deve ir além do que o autor relata, ou seja, basear somente no próprio texto. Em concordância, os autores Browne e Jerney (2007), com o foco na grafia, enfatizam que todos os termos devem estar escritos conforme estão no texto, seguindo a grafia e o uso de aspas e itálicos invertidos, a menos que isso cause confusão no índice. Já para Bonura (1994, p. 53), o indexador não deve se limitar aos termos que aparecem no texto. Importante destacar que Browne e Jerney (2007) convergem em certo momento ao afirmarem que palavras usadas nas paráfrases tornam-se partes das entradas do índice, ou seja, quando se utiliza nova afirmação do sentido de um texto usando outras palavras, desta forma, alterando o texto original.

Em análise, há vantagens e desvantagens nas duas decisões. Baseando-se no texto, o indexador se assegura de que o especialista da temática é o autor, desta forma diminui a possibilidade de erros com termos técnicos. Contudo, há limitação de utilização dos termos restringindo a busca do leitor. Por outro lado, a vantagem de não se basear somente no texto é que o indexador entende o leitor e seus possíveis termos de busca, que nem sempre foram os mesmos utilizados pelo autor. Assim, as opções de busca pelo leitor são ampliadas. Quaisquer umas das decisões devem ser descritas na Política de indexação do IFL ou na Nota explicativa.

Em relação aos resultados obtidos nesta seção, infere-se que houve elementos semânticos para todas as partes integrantes do IFL, o que possibilita a elaboração completa, embasando nas normas e na literatura da área. Sobre a incidência dos elementos semânticos, observou-se que em alguns casos os autores citaram a mesma afirmação, e na maioria dos casos houve apenas uma constatação de um elemento semântico, o que não interfere na credibilidade das informações por se tratarem de publicações confiáveis. Apresentado o tratamento e a interpretação da elaboração do IFL, expõem-se na próxima seção sobre a estruturação.

5.2.2 Resultado da seção 2.2

Na construção de IFL a estruturação é a organização dos termos para o IFL. A seguir são apresentados os resultados extraídos das publicações analisadas que relatavam sobre a estruturação do IFL.

5.2.2.1 Lista dos elementos semânticos sobre estruturação de IFL

A lista a seguir na Tabela 10 é composta dos 17 elementos semânticos identificados e extraídos dos trechos apresentados e organizados alfabeticamente (trechos apresentandos na seção 4.3.2.2.1 – atividade 2).

Tabela 10 - Elementos semânticos identificados e selecionados

Elementos semânticos selecionados	Identificação dos 30 trechos dos artigos selecionados
apresentação das entradas em linhas separadas, com recuo progressivo da esquerda para a direita para subcabçalhos	9
cabeçalho pode ser, quando necessário, qualificado por uma expressão modificadora que lhe explicita o significado	10
cabeçalhos começam na margem esquerda	20
duplicidade de interpretações, deve-se acrescentar, no início do índice, uma nota explicativa do padrão adotado e das exceções eventuais	5
ênfase, o índice pode ser: a) especial, quando organizado por: - autores; - assuntos; - títulos; - pessoas e/ou entidades; - nomes geográficos; - citações; - anunciantes e matérias publicitárias	2
ênfase, o índice pode ser: b) geral, quando combinadas duas ou mais das categorias	6
entradas de cada categoria devem ser diferenciadas graficamente e ordenadas conforme a abnt nbr 6033	3
entradas são digitadas em espaço duplo	20
índice alfabético, recomenda-se imprimir, no canto superior externo de cada página, as letras iniciais ou a primeira e última entradas da página	7
[indicativo índice deve ser apresentado por] número do volume ou parte correspondente, seguido de página(s) ou seção(ões) consecutivas ou não em documento composto de mais de um volume ou parte	15
[indicativo índice deve ser apresentado por] números extremos, ligados por hífen, quando o texto abranger páginas ou seções consecutivas	13
[indicativo índice deve ser apresentado por] números separados por vírgula, quando o texto abranger páginas ou seções não consecutivas	14
O índice deve ser impresso no final do documento, com paginação consecutiva ou em volume separado	4
o recurso tipográfico (negrito, sublinhado, itálico ou outro) deve ser utilizado para destacar as expressões ver e ver também	12
ordenação, o índice pode ser em: a) ordem alfabética; b) ordem sistemática; c) ordem cronológica; d) ordem numérica; e) ordem alfanumérica	1
subcabçalhos são todos em letras minúsculas, exceto nomes próprios, acrônimos, iniciais e símbolos que possuem maiúsculas. entradas de índice que são muito longas para ajustar na linha impressa são executadas na próxima linha após o recuo de seis espaços	20
título do índice deve definir sua função e/ou conteúdo	6

Fonte: elaborado pela autora.

A partir das informações sobre os softwares para construção de IFL e considerando os elementos semânticos identificados, a próxima seção tem como propósito apresentar o tratamento desses resultados, visando facilitar a visualização e a análise com o intuito de sintetizar e avaliar os dados obtidos.

5.2.2.2 Tratamento e interpretação dos elementos semânticos

A seguir no Quadro 12, apresentam-se de forma agrupada as diretrizes sugeridas pelas publicações em relação à estruturação do IFL. A estrutura do índice deve ser organizada considerando “um padrão lógico e facilmente identificável pelos usuários” (ABNT, 2004, p. 3). Desta forma, manter clareza e consistência quanto aos estilos, à ordenação e à padronização dos termos no IFL, para economizar o tempo do leitor e, assim, atender o seu principal objetivo. Para auxiliar em tal organização apresentam-se as recomendações para estruturar as partes constituintes do IFL (título, entradas, cabeçalho, subcabeçalho, localizadores, remissivas e nota explicativa).

Quadro 12 – Elementos semânticos relacionados à estruturação de IFL: partes do IFL

PARTE DO ÍNDICE	DIRETRIZES PARA ESTRUTURAÇÃO
Título do índice	título do índice deve definir sua função e/ou conteúdo – 6²⁰
Entradas	apresentação das entradas em linhas separadas, com recuo progressivo da esquerda para a direita para subcabeçalhos – 9 entradas de cada categoria devem ser diferenciadas graficamente e ordenadas conforme a ABNT NBR 6033 - 7 entradas são digitadas em espaço duplo - 20
Cabeçalho	cabeçalho pode ser, quando necessário, qualificado por uma expressão modificadora que lhe explicita o significado – 10 cabeçalhos começam na margem esquerda – 20
Subcabeçalho	subcabeçalhos são todos em letras minúsculas, exceto nomes próprios, acrônimos, iniciais e símbolos que possuem maiúsculas. Entradas de índice que são muito longas para ajustar na linha impressa são executadas na próxima linha após o recuo de seis espaços - 20
localizadores	[indicativo índice deve ser apresentado por] número do volume ou parte correspondente, seguido de página(s) ou seção(ões) consecutivas ou não em documento composto de mais de um volume ou parte - 15 [indicativo índice deve ser apresentado por] números extremos, ligados por hífen, quando o texto abranger páginas ou seções consecutivas - 13 [indicativo índice deve ser apresentado por] números separados por vírgula, quando o texto abranger páginas ou seções não consecutivas - 14
Remissivas	o recurso tipográfico (negrito, sublinhado, itálico ou outro) deve ser utilizado para destacar as expressões <i>ver</i> e <i>ver também</i> - 12
Nota explicativa	duplicidade de interpretações, deve-se acrescentar, no início do índice, uma nota explicativa do padrão adotado e das exceções eventuais - 5

Fonte: elaborado pela autora.

O Quadro 13 demonstra quais os enfoques existentes no IFL, como também a sua localização no livro e, por último, quais as possíveis ordenações de arranjo para as entradas do IFL.

²⁰ Números após as frases são identificados na Tabela 10.

Quadro 13 – Elementos semânticos relacionados à estruturação de IFL

ELEMENTOS IMPORTANTES NO ÍNDICE	DIRETRIZES PARA ESTRUTURAÇÃO
Enfoque	enfoque, o índice pode ser: a) especial, quando organizado por: - autores; - assuntos; - títulos; - pessoas e/ou entidades; - nomes geográficos; - citações; - anunciantes e matérias publicitárias - 5 enfoque, o índice pode ser: b) geral, quando combinadas duas ou mais das categorias - 6
Localização	O índice deve ser impresso no final do documento, com paginação consecutiva ou em volume separado - 4
Ordenação	ordenação, o índice pode ser em: a) ordem alfabética; b) ordem sistemática; c) ordem cronológica; d) ordem numérica; e) ordem alfanumérica. – 1 índice alfabético, recomenda-se imprimir, no canto superior externo de cada página, as letras iniciais ou a primeira e última entradas da página - 7

Fonte: elaborado pela autora.

Na elaboração do IFL, houve argumentações divergentes na parte de estruturação. Entre elas, as que dizem respeito ao uso do singular e plural. A ABNT (2004, p. 3) indica que deve-se normalizar “as variações de singular e plural, nos termos utilizados em índices de assunto”, no entanto, para Bonura (1994, p. 55, tradução nossa) “não há problema em usar o singular ou o plural” desde que o indexador seja consistente, ou seja, do início ao fim manter o mesmo padrão decisório do uso singular e plural. Entre os dois extremos, sendo uma publicação afirmando a normalização dos termos e outro assentindo o uso de singular e plural no IFL, os autores Borko e Bernier (1978, p. 65, tradução nossa), assumem uma posição intermediária, ao relatarem que tanto “o singular como o plural podem ser usados, mas não ambos no mesmo índice, a menos que o singular tenha um significado diferente do plural”. A escolha é geralmente baseada no uso comum, ou seja, novamente atentando para a consistência dos termos. De forma mais simplista, Borko e Bernier (1978, p. 65, tradução nossa) indicam que o plural deve ser usado em “todos os itens que são contáveis; isso permite o uso do singular para outros significados”. Em síntese, é fundamental que no momento de padronizar os cabeçalhos e subcabeçalhos recorra-se à argumentativas claras quanto à decisão de utilizar plural ou singular.

Outra argumentação conflitante nas publicações é sobre a padronização das remissivas. Para a ABNT 6034 (2004, p. 4), o “recurso tipográfico (negrito, sublinhado, itálico ou outro) deve ser utilizado para destacar as expressões ver e ver também”. Ampliando as possibilidades de apresentar formatos e posicionamentos das remissivas, a autora Mulvany (1994) afirma que

Muitos querem ver e ver também em itálico. Às vezes, a remissiva "Ver também" será colocada no início de uma entrada logo após o cabeçalho principal e seus números de página. Outras vezes, o ver também será colocado no final da entrada após as subentradas (MULVANY, 1994, p. 56, tradução nossa).

Apresentado as possibilidades de disposição das remissivas, outro ponto de destaque é sobre as ordenações das palavras no IFL. Para Bonura (1994, p. 62) deve-se colocar "a palavra mais importante primeiro", assim como para a ABNT (2002, p. 4) nos "cabeçalhos compostos, as entradas devem ser elaboradas pelas palavras significativas", contudo, para Browne e Jermeý (2007, p. 58, 59) deve-se utilizar a ordem direta das palavras, ou seja, sem alterações na ordem das palavras descritas no texto. Sobre a ordenação dos subcabeçalhos, orienta-se utilizar a ordem alfabética (BONURA, 1994; MULVANY, 1994), porém no estilo paragrafo "organizam os subcabeçalhos por número de página, em vez de alfabeticamente" (BORKO; BERNIER, 1978, p. 73, tradução nossa). Borko e Bernier (1978, p. 78, tradução nossa) acrescentam em "entradas com subcabeçalhos, as vírgulas vêm antes e depois do subcabeçalho", dessa forma, apresentam três possibilidades de ordenação do subcabeçalho.

Sobre os cabeçalhos e subcabeçalhos, que têm continuação em outra página, existem duas abordagens para indicar tal ocorrência no IFL. Para a ABNT (2002, p. 3) "subdivisões que o cabeçalho estende de uma página (ou coluna) para a seguinte, o cabeçalho e, se necessário, um subcabeçalho deve (m) ser repetido(s) e acrescido(s) da palavra 'continuação' entre parênteses ou em itálico, por extenso ou de forma abreviada". Sob outra perspectiva, Borko e Bernier (1978, p. 78, tradução nossa) afirmam que as "entradas são exibidas na próxima página com o cabeçalho repetido e colocadas em itálico e parênteses". As exposições de todas as argumentações contraditórias têm o intuito de explicitar as várias possibilidades de estruturação no IFL, e assim ampliar a perspectiva de decisão do indexador.

Considera-se que a estruturação apresentou mais convergências nas argumentações e, em consequência, menos elementos semânticos foram selecionados para compor a seção. Em contrapartida, a estruturação é menos subjetiva que a elaboração e, assim, os elementos semânticos identificados são diretos e imperativos. Avalia-se que os resultados obtidos foram satisfatórios, considerando que os elementos semânticos foram claros e sucintos, porém poucos exemplos foram relatados nas publicações.

Mas, de modo geral, a estruturação envolve dois aspectos: o primeiro diz respeito ao formato, que é a "maneira como as entradas serão exibidas na impressão" (MULVANY, 1994, p. 55, tradução nossa) do IFL; o segundo é o *layout* que tem a finalidade de "facilitar o acesso ao índice" (BROWNE; JERMEY, 2007, p. 105, tradução nossa). Dessa

forma, é fundamental preocupar-se com a elaboração e estruturação, pois o primeiro contato do leitor é com a estruturação do IFL. Durante sua pesquisa, o leitor observará o enfoque, a ordenação e a localização do IFL e analisará como a estrutura das entradas está disposta. A elaboração será analisada pelo leitor posteriormente, no momento em que ele analisa as minúcias dos termos. Assim, os dois processos são essenciais para qualidade do IFL e, assim, atender às necessidades do leitor.

No próximo Capítulo, apresentam-se as Diretrizes para elaboração e estruturação de IFL.

6. DIRETRIZES PARA ELABORAÇÃO E ESTRUTURAÇÃO DE ÍNDICE DE FINAL DE LIVRO

“E em tais índices, embora pequenos comparados com os volumes que os acompanham, entrevê-se em miniatura a forma gigantesca da massa de coisas que virão.” (William Shakespeare, Troilo e Cressida)

Apresentam-se neste capítulo as diretrizes para elaborar e estruturar IFL por meio dos resultados obtidos no capítulo 5 que são informações extraídas dos trechos de publicações e normas da área e as orientações que são apresentados no decorrer da dissertação. Entende-se por elaboração a “1. Ação de elaborar, de elaborar-se. 2. Composição, **construção**, preparação” (DICIO, [20--], online, grifos nossos), e estruturação é o “modo de **organização** das partes que compõe um todo” (DICIO, [20--], online, grifos nossos). Nesse trabalho, optou-se por trabalhar com a elaboração selecionando os trechos que relatassem sobre a construção do IFL, e na estruturação, os que abordassem sobre sua organização.

A ordem das atividades descritas na prática pode acontecer de forma simultânea. Mas geralmente, a elaboração do IFL está relacionada ao processo inicial de leitura, seleção e escolha dos termos. Com os termos selecionados e padronizados, passa-se para a estruturação que é a organização destes e a finalização do produto que é o IFL.

6.1 Elaboração do Índice de Final de Livro

Para facilitar o entendimento do processo de elaboração de IFL, inicialmente é demonstrado um esquema de como elaborar o IFL (Figura 16); este esquema, além de indicar as fases e atividades envolvidas na elaboração do Índice, também aponta para formulários, quadros e outros itens apresentados nesta dissertação. Neste esquema, as etapas foram divididas em 3 fases: 1) Iniciação; 2) Execução, e 3) Finalização. As etapas 1 e 2 estão intimamente ligadas a elaboração. Na fase 3 houve apenas a indicação da estruturação sem detalhamento para fazer uma ligação com as outras atividades da construção do IFL.

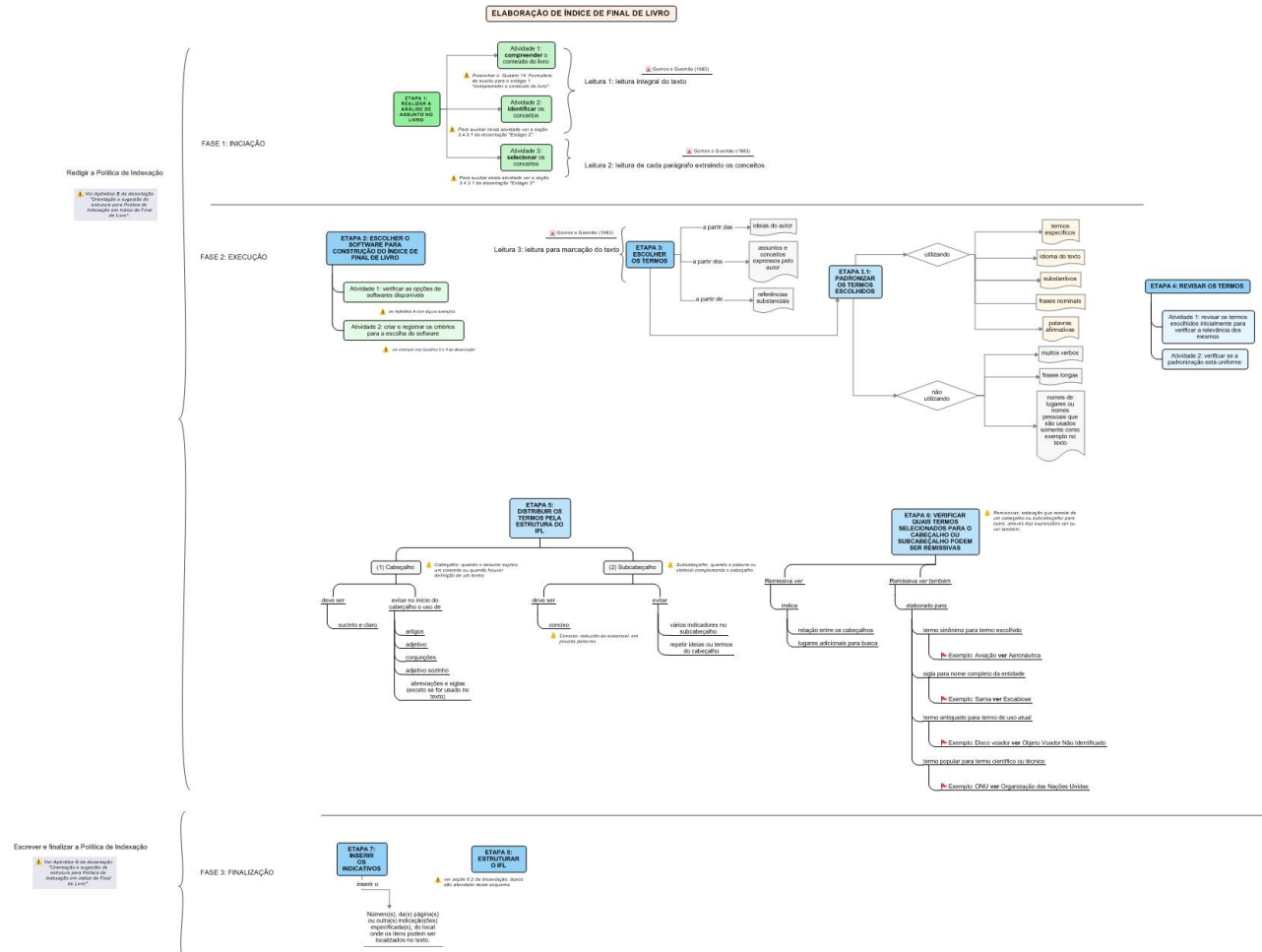
Importante destacar três itens de atenção para entendimento do esquema apresentado na Figura 16. O primeiro diz respeito à redação da Política de Indexação que orienta todo o processo de construção do IFL; este processo culmina em várias decisões que devem ser registradas nessa Política. O segundo é sobre a leitura do livro para o qual será produzido o IFL, para efetivação das etapas descritas, indicado na publicação de Gomes e Gusmão (1983, p. 24-25). A leitura do texto para o qual o IFL será criado é essencial, mas é considerada independente, por acontecer em vários momentos da

construção o que inviabiliza a sua inserção em etapa. Mas é possível descrever três principais momentos nesse processo: 1) a leitura integral do texto; 2) a leitura de cada parágrafo, e 3) a leitura para marcação do texto. Importante destacar que essas leituras podem acontecer em etapas e atividades além das indicadas no esquema.

E por último a escolha do software que está identificado na “Etapa 2” no esquema. Ressalta-se que a ordenação é relativa por existir dois casos: a) quando a instituição já assinou/trabalha com algum software, então esta etapa não existirá neste caso²¹; b) quando não existe nenhum software indicado previamente para a construção, assim o profissional deve escolher algum de sua preferência. Se for o último caso a decisão de escolher o software antes ou depois da Análise de assunto (Etapa 1 do esquema) depende do profissional. Se for um trabalho para um único livro aconselha-se que essa escolha se dê depois da análise de assunto, pois assim o profissional tem mais chance de optar por um software personalizado que já atenda as especificidades de tal. Se for para elaboração de vários IFL é preciso estudar a coleção no todo para que o software atenda todas as demandas necessárias. Sucedendo tais afirmações, apresenta-se a Figura 16 a seguir.

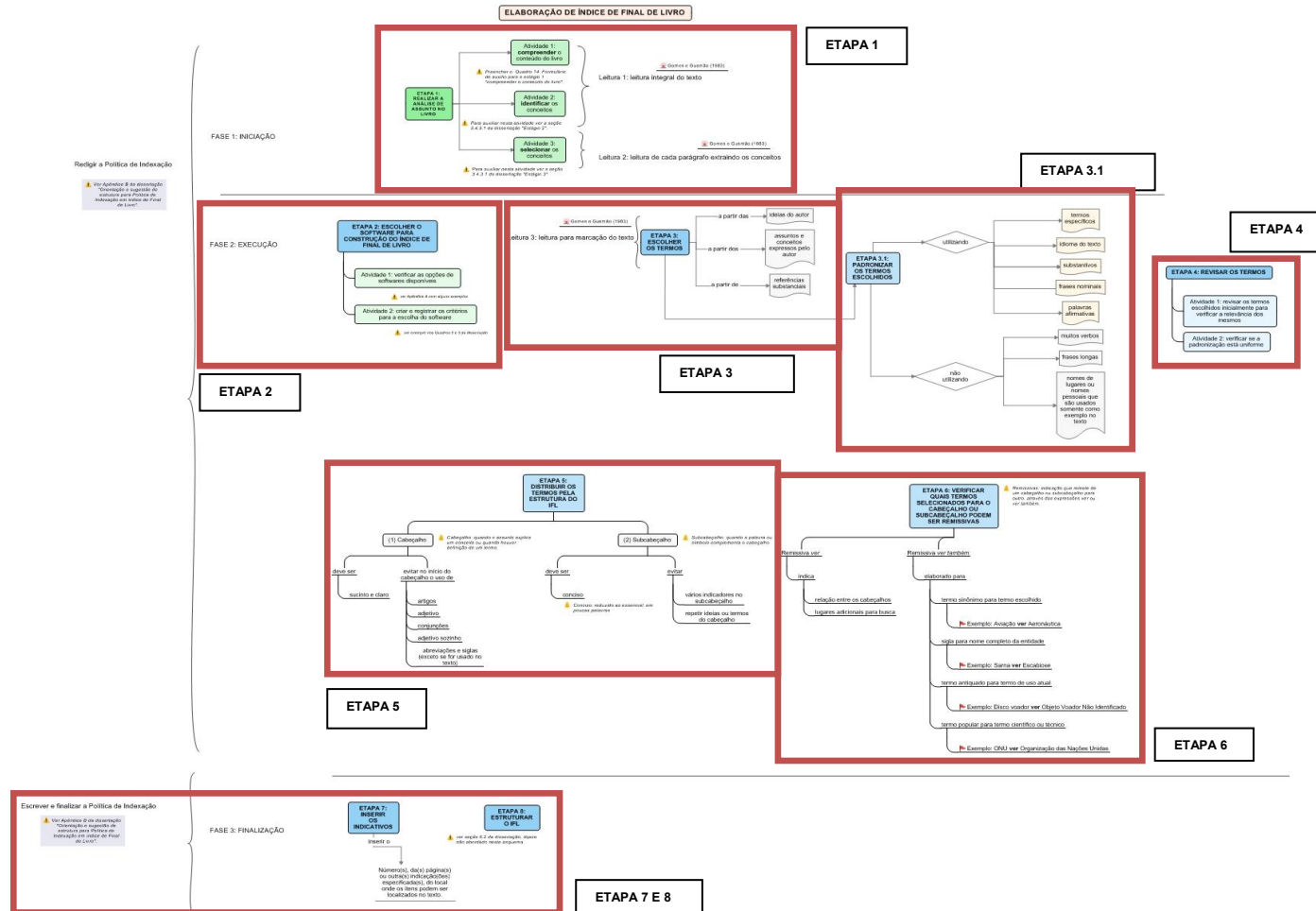
²¹ Mas pode ser realizada uma reavaliação do software já utilizado.

Figura 16 - Esquema para elaboração do Índice de Final de Livro



Fonte: elaborado pela autora.

ESQUEMA DAS ETAPAS AMPLIADAS



ETAPA 1



ETAPA 2

FASE 2: EXECUÇÃO

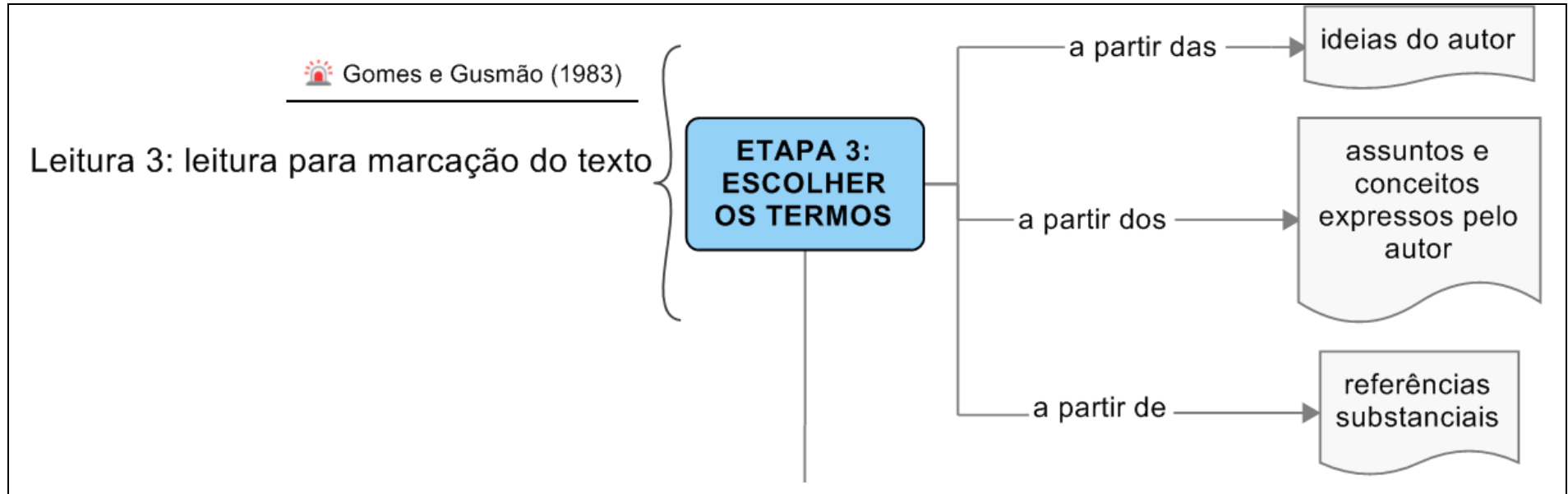
ETAPA 2: ESCOLHER O SOFTWARE PARA CONSTRUÇÃO DO ÍNDICE DE FINAL DE LIVRO

Atividade 1: verificar as opções de softwares disponíveis

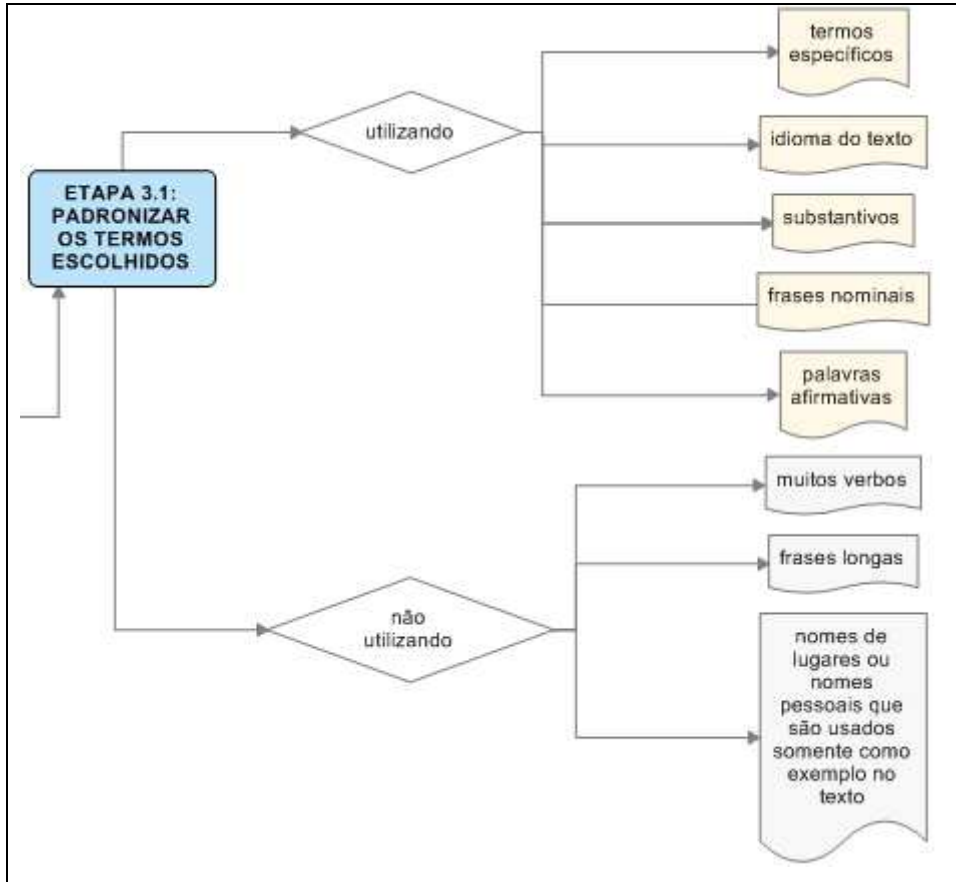
⚠ ver Apêndice A com alguns exemplos

Atividade 2: criar e registrar os critérios para a escolha do software

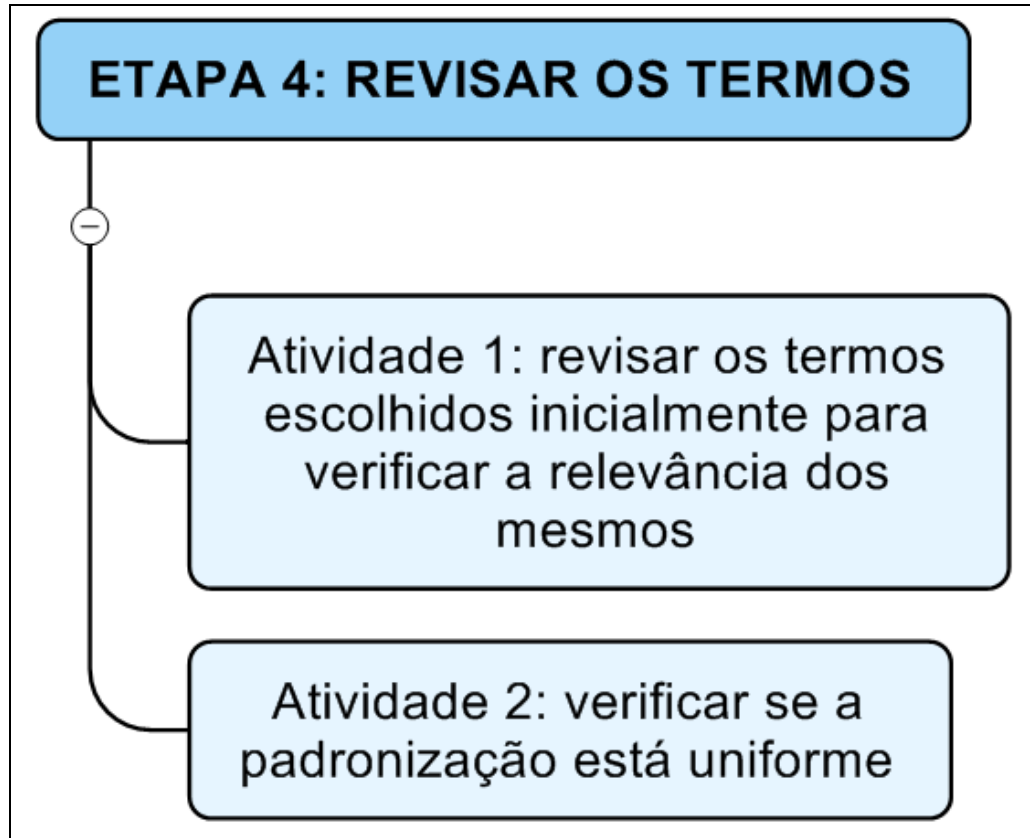
⚠ ver exemplo nos Quadros 8 e 9 da dissertação

ETAPA 3

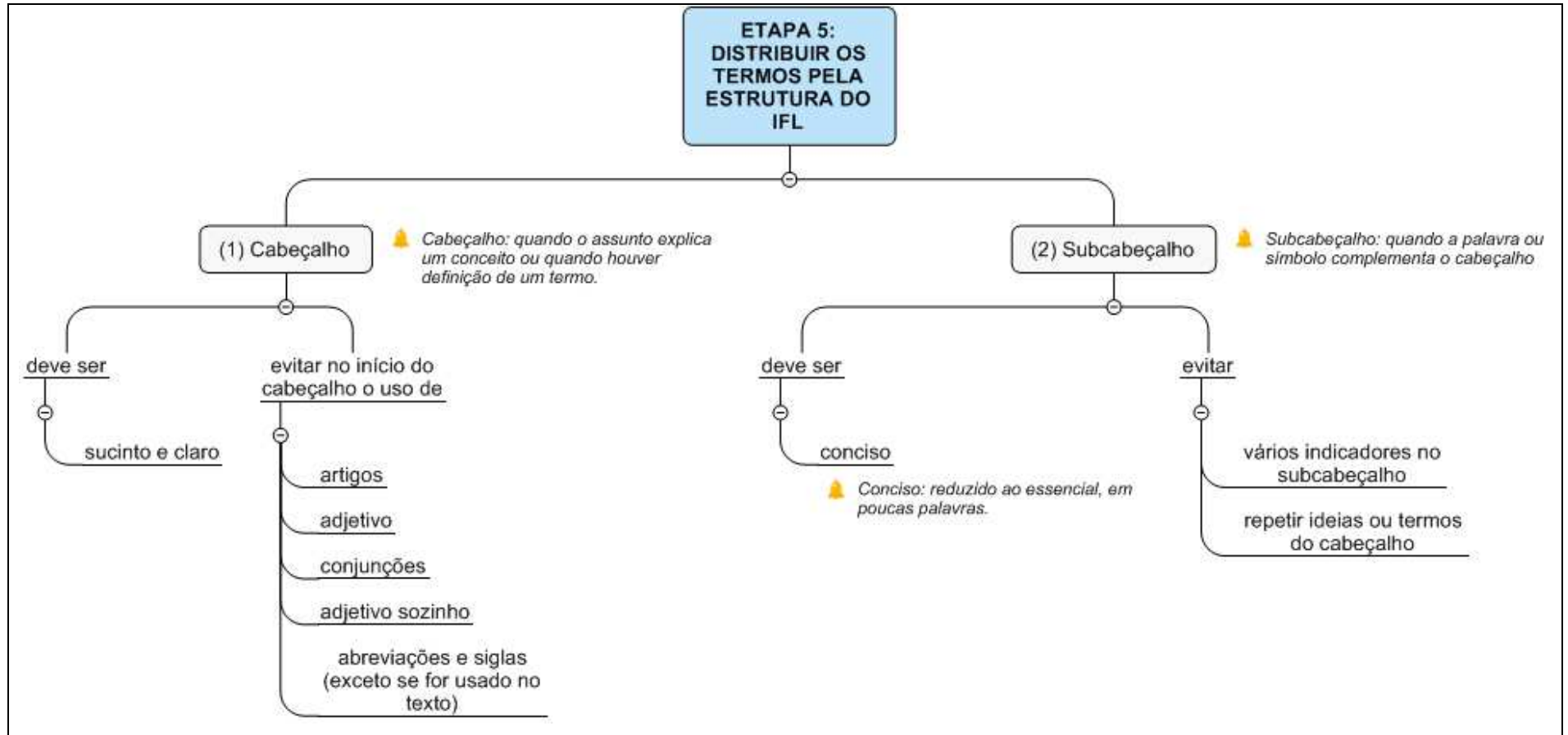
ETAPA 3.1



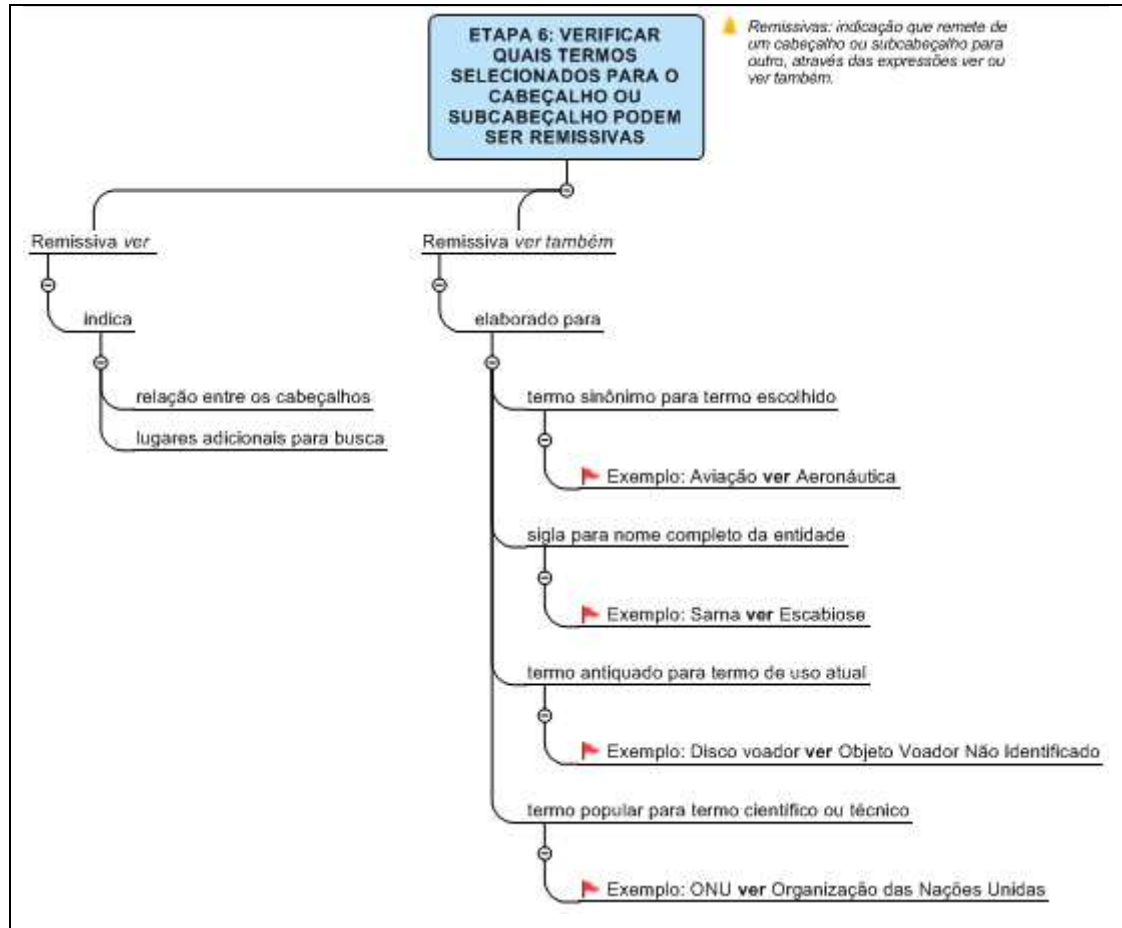
ETAPA 4

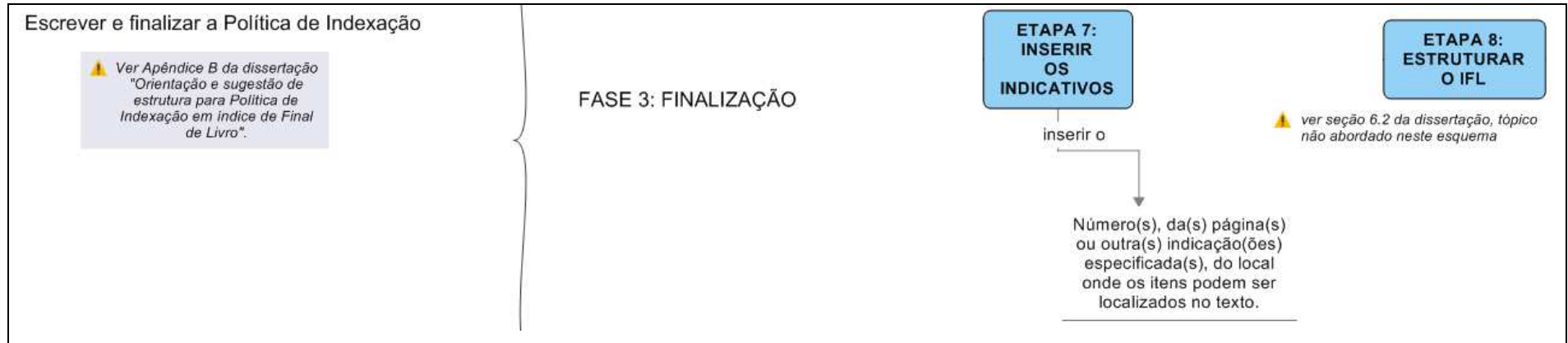


ETAPA 5



ETAPA 6



ETAPA 7 E 8

A figura anterior, na primeira etapa, destacou a necessidade de preencher um formulário este apresentado no Quadro 14 a seguir, que tem o objetivo de facilitar no estágio 1 da análise de assunto que é a compreensão do conteúdo do livro. Ao preencher este formulário o profissional deve compreender que este é um guia inicial, que possivelmente seja necessário a criação de novas perguntas para entendimento do conteúdo.

Quadro 14 - Formulário de auxílio para o estágio 1 “compreender do conteúdo do livro”

TEXTO		
1) Qual é o conteúdo do livro ?	Resposta:	
2) Qual é o objetivo do autor?	Resposta:	
3) Qual é a tipologia textual deste livro:		
	SIM	NÃO
a) pela estrutura interna:	<i>Descritivo</i> ²²	
	<i>Narrativo</i> ²³	
	<i>Dissertativo</i> ²⁴	
b) pela finalidade:	<i>Jornalístico</i>	
	<i>Jurídico</i>	
	<i>Científico</i>	
	<i>Técnico</i>	
	<i>Político</i>	
	<i>Didático</i>	
Outros? Quais?	Resposta:	
LEITORES ²⁵		
1) Qual(s) a (s) possível (s) ocupação (ões)/ profissão (ões)/cargo (s) do leitor deste livro?	Resposta:	
2) O seu Índice de Final de Livro pretende atender há:		
	SIM	NÃO
<i>Leitores que precisam saber onde as informações estão localizadas</i>		
<i>Leitores que podem saber o que querem e usam o índice para ver se essa informação está incluída no documento</i>		
<i>Leitores que são chamados para encontrar informações para outra pessoa e não conhecem a terminologia e não podem ler o documento inteiro</i>		
<i>Leitores, como bibliotecários de referência,</i>		

²² Caracterizado por descrever algo ou alguém detalhadamente, sendo possível ao leitor criar uma imagem mental do objeto ou ser descrito, de acordo com a descrição efetuada (NORMA CULTA, 2020, *online*).

²³ Caracterizado por narrar uma história, ou seja, contar uma história através de uma sequência de várias ações reais ou imaginárias (NORMA CULTA, 2020, *online*).

²⁴ Informar e esclarecer o leitor através da exposição rigorosa e clara de um determinado assunto ou tema (NORMA CULTA, 2020, *online*).

²⁵ Ver a seção: “3.4.6 O leitor do IFL”.

<i>que devem localizar informações</i>	
2.1) O que deve ser considerado durante a indexação do livro para atender aos leitores selecionados com “sim”?	<i>Resposta:</i>
CONTEXTO	
1) Quais os aspectos físicos envolvidos em seu ambiente de trabalho? <i>[estrutura física do ambiente (luz, conforto, etc.), como também os materiais que são disponibilizados para a realização do seu trabalho]</i>	<i>Resposta:</i>
2) Existe algum aspecto psicológico negativo envolvendo o trabalho? Se sim, o que posso fazer para amenizar a situação e não influenciar em minha atividade?	<i>Resposta:</i>
3) Quais são os aspectos sociocognitivos envolvidos em meu trabalho? <i>[Diz respeito aos objetivos da indexação, política de indexação, regras e procedimentos do manual de indexação, a linguagem documentária para representação e mediação da linguagem do usuário e os interesses de busca do usuário]</i>	<i>Resposta:</i>

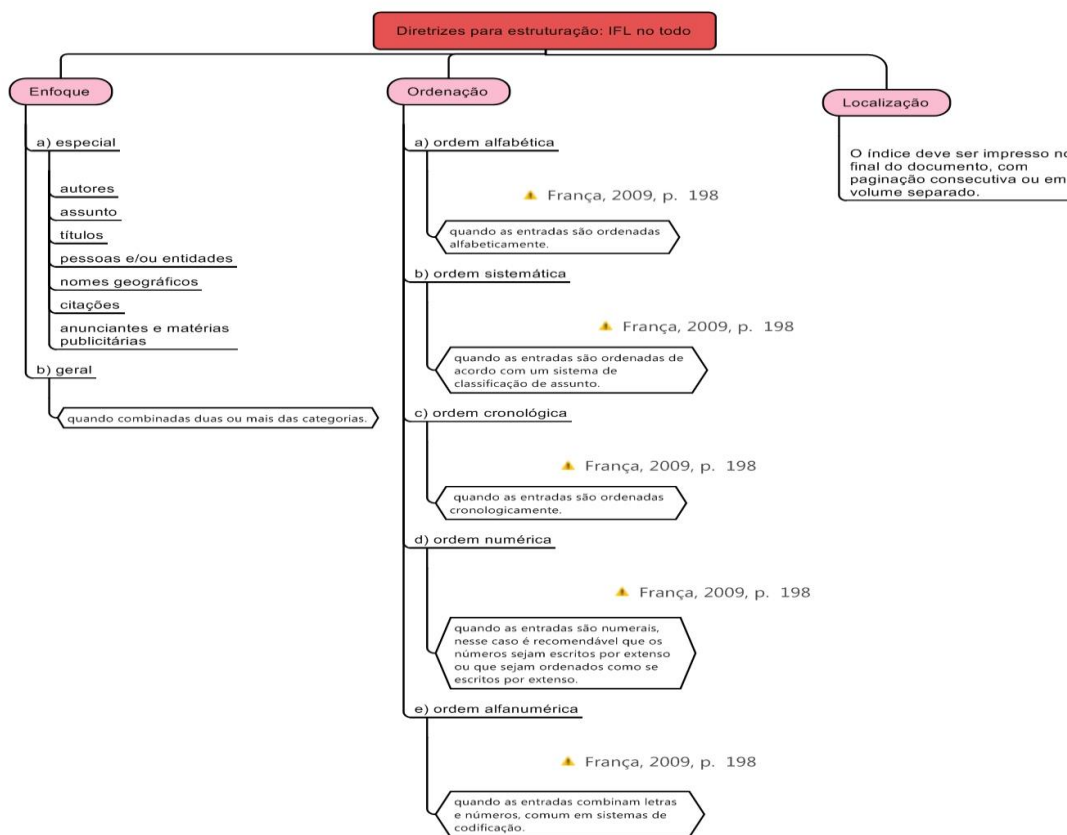
Fonte: elaborado pela autora.

Apresentado a elaboração do IFL a próxima seção aborda a atividade de estruturação.

6.2 Estruturação do Índice de Final de Livro

Esta seção aborda a estruturação do IFL, que acontece após a seleção dos termos que o integrarão, sendo necessário organiza-los para a finalização do produto. A Figura 17 a seguir tem o intuito de tornar mais intuitivo e claro o Quadro 13 desta dissertação. A estruturação do IFL como um todo precisa ser uma decisão que antecede a estruturação das partes individuais. Desta forma, deve-se escolher/identificar o Enfoque do IFL, depois, com todos os termos selecionados, é necessário verificar qual a ordenação adequada; e, por último, estruturar o IFL na localização correta do livro, como demonstrado na Figura 17 a seguir.

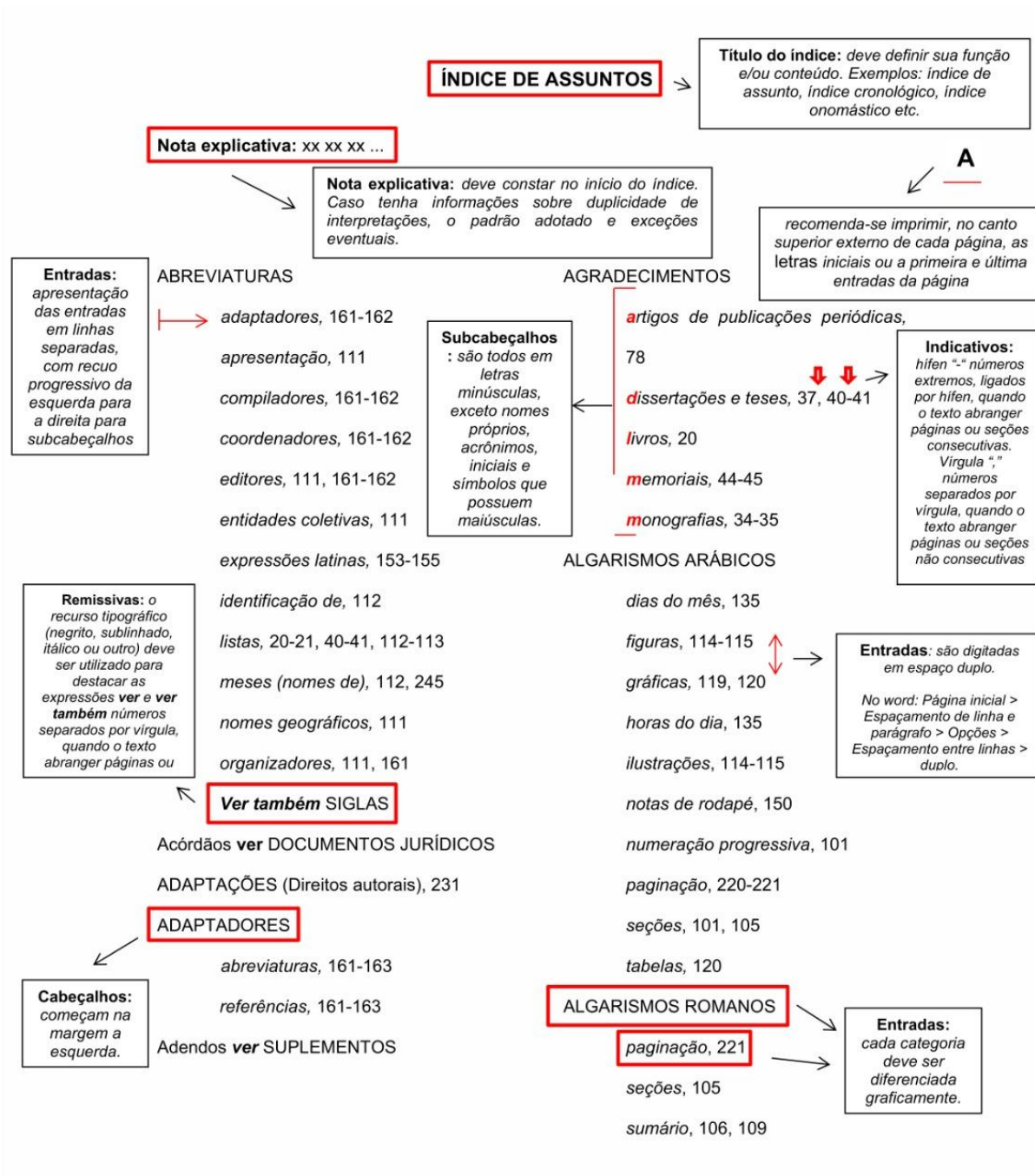
Figura 17 - Diretrizes para estruturação do IFL no todo



Fonte: elaborado pela autora.

Efetuada as decisões sobre o enfoque e ordenação, e com o IFL já na localização correta, é necessário organizar os termos selecionados na fase de elaboração. Assim, o esquema a seguir é um exemplo de IFL com orientações de como este deve ser organizado conforme as normas e publicações descritas no Quadro 12. A Figura 18 tem o propósito de tornar as diretrizes mais claras para o profissional que estiver estruturando o IFL.

Figura 18 - Exemplo de estrutura para o IFL



Fonte: adaptado de França, 2009.

No próximo capítulo é apresentada as considerações finais com reflexões sobre os resultados desta pesquisa e as sugestões de trabalhos futuros.

7. CONSIDERAÇÕES FINAIS

“Mentes curiosas nos impulsionam pra frente” (Anne with an E)

Neste capítulo, são expressas reflexões relativas ao resultado final desta pesquisa, retomando os pressupostos e objetivos levantados inicialmente e avaliando se esses foram alcançados. Evidenciando, ainda, as principais ideias discutidas e finalizando com a descrição das limitações e possibilidades de trabalhos futuros identificados no percurso da pesquisa.

O índice é uma importante fonte de informação na área de organização e representação da informação. Com sua característica sistemática, o índice possibilita o acesso a unidades de informação quando permite ao leitor navegar pelas páginas do corpo principal do livro onde pode encontrar os diferentes assuntos tratados em seu contexto. Por isso, um índice bem elaborado e estruturado, nas premissas da representação da informação, é uma ferramenta valiosa para a recuperação da informação.

O presente trabalho teve a perspectiva de apresentar, de forma didática e agrupada, as diretrizes para elaboração e estruturação do IFL por meio da indexação semiautomática, com recorte nas normas nacionais e internacionais (NBR e ISO) e na literatura que aborda essa temática. Para tanto, o uso da técnica de Análise de Conteúdo mostrou-se adequada para extrair elementos semânticos significativos na composição do universo empírico de análise. A partir da revisão de literatura, observou-se que não há artigos recentes sobre a temática de IFL no Brasil, e menos ainda sobre a utilização da tecnologia como auxiliadora no trabalho do indexador. O baixo número de publicações que contém o IFL é outro fator negativo, por se tratar de um recurso de pesquisa fundamental, principalmente em livros acadêmico-científicos. O cenário dessas lacunas foi incentivador para o desenvolvimento desta pesquisa.

Para solucionar os problemas mencionados anteriormente e as necessidades eminentes desta pesquisa partiram-se de dois pressupostos. O **primeiro** indicou que “o conhecimento sobre o processo de indexação auxilia na elaboração de índice do tipo IFL”. Tal afirmação remete ao que foi descrito nos resultados dessa dissertação identificando os aspectos relacionados às etapas de: análise de assunto e controle terminológico. A análise de assuntos é subdividida em três estágios: 1) compreensão do conteúdo; 2) identificação dos conceitos que representam o conteúdo e; 3) seleção dos conceitos válidos para recuperação. Neste estudo, aprofundou-se nesses estágios na seção “3.4.3.1 Análise de assunto”. A etapa seguinte é identificada na seção “3.4.3.2 Controle terminológico”, pois na elaboração das remissivas no IFL segue-se a etapa de controle, quando se realiza o controle de sinônimos e relacionamento dos termos. Na Figura 16 é identificado o processo

de indexação em diversas etapas do esquema para elaboração de IFL. Na fase de iniciação da primeira etapa, é necessário realizar a Análise de Assunto. No processo de seleção e padronização dos termos (etapas 3 e subtópico 3.1), a indexação está presente também. Dessa forma, entende-se que, para elaborar o IFL, é necessário ater-se aos princípios da indexação, de ambas as etapas, e ter conhecimento sobre esses princípios é fundamental.

E o **segundo** pressuposto aponta que “a indexação semiautomática é mais eficiente na elaboração de índice porque os índices elaborados manualmente geram custos altos para a produção e nos índices elaborados de forma automática há a perda de semântica e de contexto”. Esse relato é validado na revisão de literatura, na seção “2.3. Análise dos trabalhos correlatos” desta dissertação e, após analisar 16 (dezesesseis) publicações que abordavam o tema construção de IFL, é possível inferir que construir IFL manualmente, é inviável, se considerado o alto número de produções de livros atualmente. E o uso da utilização da indexação automática não é recomendável, pois não lida com questões de polissemias e semânticas das palavras. Assim, conclui-se que a melhor alternativa é elaborar o IFL por meio da indexação semiautomática. Assim, os dois pressupostos que guiaram essa pesquisa foram confirmados.

Após analisar os pressupostos levantados na pesquisa e relatar os problemas identificados, buscou-se responder a seguinte pergunta “**quais são as diretrizes necessárias para a elaboração e estruturação semiautomática do Índice de final de livro?**” para procurar a solução elabora-se o objetivo geral da pesquisa, que é “Identificar as diretrizes para a elaboração e estruturação do IFL a partir das normas (NBR e ISO) e da literatura aplicável a essa temática por meio da indexação semiautomática”. Todos os objetivos específicos foram desmembrados visando ao alcance do objetivo geral, sendo um objetivo conceitual (letra "a") e dois objetivos procedimentais (letra b e c). O primeiro objetivo específico é “identificar na literatura os programas de construção de IFL automático e semiautomático, caracterizando-os, com o intuito de auxiliar na avaliação dessas ferramentas” foi apresentado na seção “5.1 Resultados da Seção 1”, na qual descreve os procedimentos adotados do método de Análise de conteúdo, o qual possibilitou a extração dos elementos semânticos para analisar as características funcionais e estruturais dos softwares. Nesse processo foram identificadas as vantagens, desvantagens e características de cada software no Quadro 10, como também, houve a análise destes sob a abordagem dos 8 (oito) critérios que os programas de indexação devem ser validados no Quadro 9. Assim, os softwares foram caracterizados a partir dos subsídios fornecidos pelas publicações da área. Considera-se que os resultados obtidos auxiliam os indexadores no momento de decisão sobre qual programa utilizar, ao apresentar os critérios que os

programas de indexação devem atender e ao apresentar o esquema de como caracterizar e identificar as vantagens e desvantagens dos softwares.

O segundo objetivo “sintetizar as principais recomendações das normas (NBR e ISO) e da literatura para identificar os subsídios e auxiliar na elaboração e estruturação dos Índices de Final de Livro”. Para alcançar este objetivo, aplicou-se o método de Análise de Conteúdo nas publicações selecionadas identificando-se o universo empírico de análise a partir dos elementos semânticos extraídos dos trechos das publicações. As partes do IFL a que se refere (cabeçalho, subcabeçalho, entre outros) foram agrupadas obtendo dois resultados: o primeiro é a figura que apresenta as diretrizes para elaboração do IFL no Quadro 11 e na Figura 15; o segundo são os quadros que esquematizam sobre a estruturação do IFL no Quadro 12 e 13. O resultado alcançado nessa dissertação, a identificação dos elementos semânticos das publicações foi mais pormenorizado no Capítulo 6, a fim de facilitar a aplicação das diretrizes. Nesse Capítulo, na Figura 16, apresenta-se um esquema geral de como elaborar IFL, este por sua vez, indica várias seções, formulários e itens para entendimento mais amplo de efetivação da etapa. Para a estruturação a Figura 17 intenta destrinchar e esquematizar as orientações descritas no Quadro 13. E a Figura 18 apresenta um exemplo de IFL com várias caixas de orientações de como estruturar, apresentando desta forma exemplos para facilitar no entendimento do Quadro 12. Desta forma, acredita-se que o universo empírico proporcionou dados suficientes para apresentar a elaboração e estruturação do IFL, atendendo a esse objetivo descrito.

O último objetivo concerne o propósito de desenvolvimento dessa dissertação que é “contribuir para a divulgação no Índice de Final de Livro como importante recurso de acesso à informação, a partir da análise e interpretação da literatura escolhida”. O desenvolvimento deste estudo, em diversos momentos, intensificou a importância do IFL, afirmando que sua fundamental importância como recurso de acesso à informação consiste em auxiliar e guiar o leitor na localização de informações relevantes contidas nos textos. Além disso, ajuda na obtenção de pesquisas mais assertivas em menos tempo, como também determina se o assunto desejado está no livro. A partir dessa reflexão, espera-se que essa pesquisa contribua com a divulgação da importância do IFL, entendendo que o processo de conscientização vai além da escrita: é preciso ir ao encontro dos autores de livros e editores para orienta-los sobre os proveitos que o IFL pode proporcionar aos seus leitores. Mas tais ideias são concebíveis no meio profissional somente se forem comprovadas por pesquisas; dessa forma, acredita-se que esse trabalho pretendeu colaborar com esses estudos.

A fundamentação teórica apoiou-se em publicações de autores seminais e normas consolidadas que sustentaram o embasamento teórico-metodológico para o

entendimento sobre (1) indexação manual, semiautomática e automática, finalizando com o produto de tais processos que é o IFL.

Os métodos utilizados foram o da (1) revisão sistemática da literatura, para levantamento bibliográfico e o da (2) análise de conteúdo. O **levantamento bibliográfico** permitiu a compilação das principais pesquisas sobre o IFL, resultando na análise das publicações sendo algumas dessas utilizadas como amostras para aplicação da metodologia. O método da **análise de conteúdo** auxiliou no estudo das publicações conduzindo as interpretações de modo científico, sendo considerada satisfatória, bem como as fases e etapas descritas neste trabalho podem ser reaplicadas a fim de obter novos resultados com publicações mais recentes.

Com base nos resultados obtidos, acredita-se que o objetivo e a proposta da pesquisa foram alcançados, conforme representado na Figura 16, e a análise das características dos *softwares* para indexação semiautomática através dos Quadros 9 e 10. Dessa forma, esse trabalho pretende ser um guia de elaboração do IFL para autores e profissionais da informação, contendo diretrizes avaliadas a partir da literatura da área (apresentado no capítulo 6). Contribuindo, desse modo, para auxiliar bibliotecários e autores na elaboração e estruturação do IFL e impactar positivamente nos estudos na área sobre indexação no Brasil a partir das diretrizes embasadas nas normas e literatura da BCI que devem subsidiar a prática do indexador.

Pretendeu-se com esta dissertação contribuir para que os **discentes de biblioteconomia e bibliotecários** reconheçam o IFL como produto que pode ser vendido no mercado. E para que os alunos da disciplina de Análise de assunto do Curso de Biblioteconomia contem com informações condensadas sobre IFL e conheçam as discussões e pesquisadores atuantes no contexto internacional. Esperamos que ele possa auxiliar também o trabalho dos **profissionais que atuam em editoras** na elaboração do IFL e no reconhecimento de sua importância para os seus clientes (leitores). E dessa forma, aumentar o número de publicações que contem IFL, otimizando o tempo do **leitor** ao facilitar durante as buscas, tornando-as mais assertivas.

7.1 Limitações da pesquisa

A pesquisa científica, por utilizar vários métodos a fim de investigar determinado assunto possui limitações que devem ser destacadas, pois a continuidade ou atualização da pesquisa necessita visualizar os obstáculos com clareza e, dessa forma, possibilitar melhorias nas aplicações dos métodos. Existem três limitações de destaque neste trabalho, a primeira diz respeito aos insumos utilizados: as normas tendem a ser atualizadas

conforme o tempo, tornando-se vigentes ou sendo canceladas. Desta forma, alguns dos insumos utilizados neste estudo podem sofrer alterações em alguns anos. Assim como publicações são constantemente atualizadas e precisam reexaminar as afirmações destacadas nos resultados.

Como segunda limitação acrescenta-se a não localização dos softwares relatados nas publicações e analisados neste trabalho. Esse fato foi destacado na seção 5.1.2 Tratamento e interpretação dos elementos semânticos, mas é importante resgatar essa discussão para fechamento do trabalho. Não foi possível realizar nenhum *download* para testar na prática as vantagens, desvantagens e características dos softwares, utilizando somente as análises dos autores, embora esses tivessem propriedade para dissertar sobre, além de serem fontes confiáveis, os próprios desenvolvedores dos softwares criticaram os modelos desenvolvidos. Também não foi possível localizar nenhuma informação sobre a forma de aquisição, possivelmente em razão de a utilização ser restrita ou devido à descontinuidade ou indisponibilidade do software. Os softwares disponíveis hoje tem alto custo de aquisição impossibilitando a utilização dos mesmos nesse trabalho (conforme apresentado no Apêndice A). Acredita-se que a não utilização dessas ferramentas não seja empecilho para ajudar o profissional a aplicar os métodos de análise realizados nesta pesquisa no momento em que for escolher o software para o seu trabalho, e também entende-se que a aplicação dos resultados seja discutida em trabalhos mais robustos, como uma tese por exemplo.

E a última limitação é a necessidade de adotar os métodos experimental ou observacional (GIL, 2008) para visualizar como os indexadores aplicariam as diretrizes sugeridas nesta dissertação na prática, pois se acredita que existam ações que não são descritas (conhecimento tácito) e que precisariam ser registradas para aperfeiçoamento dos produtos elaborados (IFL). Assim, acrescentar essas observações à essa pesquisa seria de grande enriquecimento. Nesse sentido, os resultados obtidos com este estudo serão o pilar de observação, pois ter o embasamento teórico é fundamental para a prática, e como trabalho futuro, com a aplicação das diretrizes em estudos de caso, haveria insumos para comparação entre a teoria e a prática, aprimorando assim o produto obtido.

7.2 Trabalhos futuros

Ao longo desta pesquisa de mestrado, algumas questões enfrentadas mostraram que há a possibilidade de realizar estudos futuros, tais como o aprofundamento do IFL em Trabalhos acadêmicos (dissertações e teses e outros), já que a NBR 14724 (ABNT, 2011) prescreve o Índice como elemento opcional da estrutura dos trabalhos acadêmicos

(elementos pós-textuais). Os trabalhos acadêmicos tendem a ser extensos e, geralmente, possuem menos acessos do que os artigos advindos desses, pressupondo ser um material de pouca dinamicidade de localização das informações. O IFL poderia facilitar a leitura e localização das informações para que o leitor possa usufruir de todas as vantagens mencionadas neste trabalho. Além disso, ampliaria a visualização de quais assuntos são mais abordados em determinados documentos, pois entradas com mais subcabeçalhos demonstram que aquele cabeçalho foi mais abordado, assim como possibilitaria observar qual abordagem do autor citado no trabalho o discente abordou. E, desse modo, aumentaria o número de acessos do trabalho acadêmico atendendo seu objetivo que é o de divulgar o conhecimento produzido. O IFL no trabalho acadêmico apresenta o cuidado do pesquisador com o seu leitor. Desta forma, um estudo que fizesse observações com leitores e bancas examinadoras do uso do IFL nos trabalhos acadêmicos e o levantamento de quais foram os impactos percebidos durante a leitura seria significativo para área.

Entre as novas perspectivas nessa área de pesquisa, destacam-se a inserção do IFL em *Ebooks*. Provo (2019), em seu artigo identifica um relatório (HUMPHREYS *et al.*, 2017) que tinha por finalidade realizar teste de usabilidade para entender às expectativas do usuário em relação aos *Ebooks* produzidos pela *Digital Production Editor*. Dentre os itens investigados, analisou-se o Índice e como ele seria utilizado pelos usuários por meio dos *ebooks*. Importante destacar que o Índice utilizado nesse cenário é o metaíndice, ou seja, a elaboração de um único índice que concentra a temática de vários documentos. Constatou-se a relevância do IFL em três aspectos: 1) navegação interna através de *links* (os termos estariam destacados com *links* que redirecionaria para a parte correspondente no texto; 2) pesquisa de texto completo (os termos remeteriam aos textos completos); e 3) navegação em uma coleção. Considerando que o IFL não poderá manter em *ebooks* a forma que é apresentado em livros tradicionais, seriam necessários estudos de como seria a elaboração do IFL em materiais não convencionais, como os *ebooks*. Observa-se que esse item é importante para o usuário mesmo em *ebooks* e, por conseguinte, não pode ser dispensado com a evolução da visualização do livro. É necessário, também, que esses estudos sejam realizados no Brasil, pois existem poucos estudos sobre o tema IFL no Brasil.

Em relação aos Índices para web, ressalta-se o trabalho de Broccoli (1998) que, além de autor do artigo, é um indexador profissional que criou vários índices para sites. Em seu trabalho relata a importância do Índice para a web, assim como a importância desse produto para os livros, mencionando também a semelhança no processo de criação dos dois produtos. Embora a publicação seja antiga, destaca-se esse trabalho como norteador para trabalhos futuros, pois ainda hoje navegar com eficiência em sites é fundamental, assim como relatar como os indexadores devem elaborar e estruturar índices no ambiente da web.

REFERÊNCIAS

ABE, K.; BERRY, D. M. Indx and findphrases: a system for generating indexes for ditroff documents. **Software—practice and experience**, [S.l.], v. 19, n. 1, p. 1-34, jan. 1989.

ABNT Catálogo. 2018. Disponível em: <<https://www.abntcatalogo.com.br/>>. Acesso em: 3 out. 2018.

AGANETTE, E. C. **Taxonomias corporativas**: um estudo sobre definições e etapas de construção fundamentado na literatura publicada. 2010. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010. Disponível em: <<https://repositorio.ufmg.br/handle/1843/ECID-87BJSR>>. Acesso em: dez. 2019.

ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. **The indexer**, v. 18, n. 4, oct. 1993.

AMERICAN SOCIETY FOR INDEXING (ASI). Online discussion groups. [S.l.: s.n.], 2020. Disponível em: <<https://www.asindexing.org/reference-shelf/online-discussion-groups/>>. Acesso em: 14 abr. 2020.

ANSI Z39.4:1984: Basic Criteria for Indexe. [S.l.: s.n.]. 1984.

ARAUJO, E. BIBLIOTECA NACIONAL (BRASIL). **A construção do livro**: princípios da técnica de editoração. 2. ed. rev. e ampl. Rio de Janeiro: Lexikon: Fundação Biblioteca Nacional, 2008. 635 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14724**: Informação e documentação — Trabalhos acadêmicos — Apresentação. Rio de Janeiro, 2011.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6033**: ordem alfabética em índices de trabalhos acadêmicos. Rio de Janeiro, 1989.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6034**: informação e documentação: Índice: apresentação . Rio de Janeiro, 2004.

BARDIN, L. **Análise de conteúdo**. Tradução Luís Antero Reto, Augusto Pinheiro. São Paulo: Edições 70, 2016.

BOLSHAKOVA, E. I.; IVANOV, K. M. Term extraction for constructing subject index of educational scientific text. In: COMPUTATIONAL LINGUISTICS AND INTELLECTUAL TECHNOLOGIES: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE “DIALOGUE 2018”, 2018, Moscow. **Anais...** Moscow: [s.n.], 2018. 11 p.

BONURA, L. S. **The art of indexing**. New York: Wiley, 1994. 233 p.

BORGES, Graciane Silva Bruzina. **Indexação automática de documentos textuais: proposta de critérios essenciais**. 2009. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Disponível em

<<https://repositorio.ufmg.br/browse?type=author&value=Graciane+Silva+Bruzina>>. Acesso em: dezembro de 2019.

BORKO, H.; BERNIER, C. L. **Indexing concepts and methods**. New York: Academic Press, 1978. 261 p.

BROCCOLI, Kevin. **Indexes**: an old tool for a new medium. [S.l.: s.n.], 1998.

BROWNE, G. Classified versus specific entry in book indexes: guidelines for decision making. **The Indexer**, v. 28, n. 1, mar. 2010. Disponível em: <<http://www.webindexing.com.au/wp-content/uploads/2016/02/ClassificationVsSpecific.pdf>>. Acesso em: 14 abr. 2020.

BROWNE, G.; JERMEY, J. **The indexing companion**. Austrália: Ligary, 2007. 249 p.

CÂMARA, R. H. **Análise de conteúdo**: da teoria à prática em pesquisas sociais aplicadas às organizações. 2013. Disponível em: <<http://pepsic.bvsalud.org/pdf/gerais/v6n2/v6n2a03.pdf> 2013>. Acesso em: 02 nov. 2019.

CESARINO, M. A.; PINTO, M. C. M. F. Análise de assunto. **Revista de Biblioteconomia de Brasília**, n. 1, 32-43, jan./jun. 1980.

CHAUMIER, J. Indexação: conceito, etapas e instrumentos. **Revista Brasileira de Biblioteconomia e Documentação**, v. 21, n.1/2, 1988, p.63-79

CHI, Ed. H. *et al.* eBooks with Indexes that Reorganize Conceptually. **ACM**, Vienna, apr. 2004.

CHI, Ed. H. *et al.* ScentIndex: conceptually reorganizing subject indexes for reading. **IEEE Symposium on Visual Analytics Science and Technology**, Baltimore, 2006.

CHRISTINA, S. *et al.* Identifying the relevant indexed term related with the book domain using semantic relatedness approach. In: INTERNATIONAL CONFERENCE ON INFORMATICS AND COMPUTING (ICIC), 3. 2018, Palembang (Indonésia). **Anais eletrônicos ...** Palembang: [s.n.], 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8780562>>. Acesso em: 14 abr. 2020.

COLLISON, R. L. **Índices e indexação**: guia para a indexação de livros, e coleções de livros, periódicos, partituras musicais, discos, filmes e outros materiais, com uma seção de referencia e sugestões para leitura adicional. São Paulo: Polígono, 1971. 225p.

CORREA, R. F.; LAPA, R. C. Panorama de estudos sobre indexação automática no âmbito da ciência da informação no Brasil (1973-2012). **Ci. Inf., Brasília**, DF, v. 42 n. 2, p.255-273, maio/ago., 2013.

CSOMAI, A.; MIHALCEA, R. Creating a Testbed for the evaluation of automatically generated back-of-the-book indexes. In: INTERNATIONAL CONFERENCE ON INTELLIGENT TEXT PROCESSING AND COMPUTATIONAL LINGUISTICS, 7., 2006, Mexico. **Anais eletrônicos**. México: Springer, 2006. Disponível em: <<https://link.springer.com/book/10.1007/11671299>>. Acesso em: jul. 2020.

CSOMAI, A.; MIHALCEA, R. Investigations in unsupervised back-of-the-book indexing. In: THE FLORIDA ARTIFICIAL INTELLIGENCE RESEARCH SOCIETY (FLAIRS) CONFERENCE, 2007, Florida. **Anais eletrônicos**. Disponível em:

<<https://webcache.googleusercontent.com/search?q=cache:iRsONeJURGoJ:https://digital.library.unt.edu/ark:/67531/metadc30990/+&cd=2&hl=pt-BR&ct=clnk&gl=br>>. Acesso em: jul. 2020.

CSOMAI, A.; MIHALCEA, R. Linguistically Motivated Features for Enhanced Back-of-the-Book Indexing. **Proceedings of ACL-08: HLT**, Columbus, Ohio, USA, p. 932–940, June 2008.

CUNHA, M. B. da; CAVALCANTI, C. R. **Dicionário de biblioteconomia e arquivologia**. Brasília: Briquet de Lemos/Livros, 2008.

CUTTER, Charles A. **Rules for a dictionary catalog**. Washington: Government Printing Office, 1904.

DARTNALL, J. Are indexes worth it? Evidence from book reviews. **Learned publishing**, [S.l.], v. 21, n. 4, oct. 2008.

DEUTSCHES Netzwerk der Indexer. [S. l.: s.n.], [20--]. Disponível em: <<http://www.d-indexer.org/>>. Acesso em: 14 abr. 2020.

DIAS, E. W.; NAVES, M. M. **Análise de assunto: teoria e prática**. 2. ed. rev. Brasília: Briquet de Lemos, 2013.

DICIO. Dicionário Online de Português. Disponível em <<https://www.dicio.com.br/elaboracao/>>. Acesso em dezembro de 2019.

FAULKNER, H.; LIGHT, W. M. Editing the index: developing a method. **The Indexer**, [S.l.], v. 24, n. 4, oct. 2005.

FEITOSA, A. **Organização da informação na Web: das tags à web semântica**. Brasília: Thesaurus, 2006. 131 p.

FELIPE, Eduardo. **A importância dos metadados em bibliotecas digitais: da organização à recuperação da informação**. 2012. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Disponível em <https://repositorio.ufmg.br/bitstream/1843/ECID-943PDD/1/dissertacao_eduardo_v15.pdf>. Acesso em: dezembro de 2019.

FETTERS, L. K. INDEXIT: an economical but limited indexing program. **DATABASE**, 9, 54–56, 1986.

FRANÇA, J. L. *et al.* Índice. In: _____. **Manual para normalização de publicações técnico-científicas**. 9. ed. 1. reimp. Belo Horizonte: Editora UFMG, 2014. cap. 17

FRANCIS, R. D.; GREENWAY, M. A simple system for constructing an end-of-book index. **Journal of scholarly publishing**, v. 46, n. 3, p. 290-296, apr. 2015.

FUGMMAN, R. **Book Indexing: The Classificatory Approach**. Knowl. Org. v.21, n. 1, 1994.

FUJITA, M. S. L. A leitura do indexador: estudo de observação. **Perspect. cienc. inf.**, Belo Horizonte, v. 4, n. 1, p. 101 - 116, jan./jun.1999. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/597/366>>. Acesso em: dezembro de 2019.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003 – ISSN: 1678-765X.

GIASSON, J. **A compreensão na leitura**. Lisboa: Asa, 1993. 317 p.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Editora Atlas, 2008.

GOMES, Hagar Espanha; GUSMÃO, Heloisa Rios. **Guia prático para elaboração de índice**. Niterói, RJ: Grupo de Bibliotecários em Informação e Documentação em Ciências Sociais e Humanidades, 1983. 68p.

GOOGLE ACADÊMICO. 2020. Disponível em: <https://scholar.google.com.br/scholar?cites=5755637037214836391&as_sdt=2005&scioldt=0,5&hl=pt-BR> Acesso em: 1 abr. 2020.

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

HINES, Theodore C.; HARRIS, Jessica L. Computer-aided production of book indexes. **The Indexer**, [S.l.], n. 7, v. 2, 1970.

HUMPHREYS, Alex *et al.* Reimagining the Digital Monograph: Design Thinking to Build New Tools for Researchers. **A JSTOR Labs Report**. Nova Iorque: JSTOR, 2017. Disponível em <<https://labs.jstor.org/download/JSTORLabsMonographJune2017.pdf>>. Acesso em: dez. de 2019.

INDEXERS.CA. Toronto: [s.n.], 2020. Disponível em: <<https://indexers.ca/>>. Acesso em: 14 abr. 2020.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/AWI 999**: information and documentation -- guidelines for the content, organization and presentation of indexes. [S.l.], 1996. Disponível em: <<https://www.iso.org/standard/5446.html>>. Acesso em: 02 nov. 2019.

KITCHENHAM, Barbara. **Procedures for performing systematic reviews**. Austrália: [s.n.], 2004. Disponível em: < <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf> >. Acesso em: 4 ago. 2018.

LANCASTER, F Wilfrid. Índices pré-coordenados. In: **Indexação e resumos**: teoria e prática. 2. ed. rec. atual. Brasília: Briquet de Lemos/Livros, 2004.

LIBRARIANSHIP STUDIES & INFORMATION TECHNOLOGY. **Derived Indexing**. July 03, 2019. <https://www.librarianshipstudies.com/2016/08/derived-indexing.html>

LIMA, Gercina Â. B. O. **Mapa hipertextual (MHTX)**: um modelo para organização hipertextual de documentos. 2004. 204f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação da UFMG, Belo Horizonte. 2004.

LIPETZ, B. A. The usefulness of indexes. **Indexer**, v. 16, n. 3, p. 173-176, 1989.

LUHN, H. P. **Keyword-in-Context index for Technical Literature (KWIC Index)**, RC- 127; also in *American Documentation*, v. 11, n. 4, p. 288- 295. 1960.

MARTINHO, N. O.; GUEDES, E. F. Charles Ammi Cutter: sua contribuição para organização da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Anais eletrônicos**. João Pessoa: UFPB, 2009. Disponível em <<http://enancib.ibict.br/index.php/enancib/xenancib/paper/viewFile/3152/2278>>. Acesso em: dez. 2019.

MARTINS, G. K.; MORAES, J. B. E. Organização e representação do conhecimento: institucionalização como disciplina científica no âmbito da Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa (PB). **Anais eletrônicos** ... João Pessoa: ENANCIB, 2015. Disponível em: <<http://www.ufpb.br/evento/index.php/enancib2015/enancib2015/paper/viewFile/3162/1030>>. Acesso em: 14 abr. 2020.

MEKKI, T. Aït El; NAZARENKO, A. Using NLP to build the hypertextuel network of a back-of-the-book index. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING, 2005. [S.l.]. **Anais eletrônicos**. [S.l.: s.n], 2005. Disponível em: < <https://arxiv.org/abs/cs/0609134>>. Acesso em: jul. 2020.

MILLER, Sylvia K. Index appreciation: a publisher's brief guide. **Against the Grain**, [S.l.], v. 17, n. 4, nov. 2013.

MOCHÓN BEZARES, Gonzalo; SORLI ROJO, Ángela. **Tesauro de Biblioteconomía y Documentación**. Madrid: Consejo Superior De Investigaciones Científicas, 2002.

MONTEIRO, Gilson Vieira. Sumário ou índice? conceitos, definições e controvérsias. **Acta Cir. Bras.**, v. 13, n. 2, São Paulo. 1998.

MOREIRO GONZÁLES, José Antonio. El proceso de indización y su resultado: los índices. In: _____. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Ediciones Trea, 2004.

MULVANY, Nancy C. **Indexing books**. Chicago; London: Library Materials, 1994. 320p.

NAVES, M. M. L. **Curso de indexação**: princípios e técnicas de indexação, com vistas à recuperação da informação. Belo Horizonte: UFMG, Biblioteca Universitária, 2004. Material didático. 23p.

NAVES, Madalena Martins Lopes. Análise de assunto: Concepções **Revista de Biblioteconomia de Brasília**, v. 20, n. 2, p. 215-226, jul./dez. 1996. Disponível em <http://www.brapci.inf.br/_repositorio/2010/03/pdf_89759389ea_0008824.pdf>. Acesso em: dez. 2019.

NORMA CULTA. **Tipologia textual**: os diferentes tipos textuais. [S.l.: s.n.], 2020. Disponível em: < <https://www.normaculta.com.br/tipos-de-texto/>>. Acesso em: 1 jul. 2020.

PIEADADE, M. A. R. **Introdução a teoria da classificação**. 2. ed., rev e aum. Rio de Janeiro: Interciência, 1983.

PINHEIRO, L. V. R.; FERREZ, H. D. **Tesouro brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), 2014. Disponível em: <http://www.ibict.br/publicacoes-e-institucionais/tesouro-brasileiro-de-ciencia-da-informacao-1/copy_of_TESAUROCOMPLETOFINALCOMCAPA24102014.pdf>. Acesso em: 20 set. 2018.

PINTO, V. B. Indexação documentária: uma forma de representação do conhecimento registrado. **Perspect. cienc. inf.**, Belo Horizonte, v. 6, n. 2, p. 223 - 234, jul./dez. 2001. Disponível em <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/423/239>>. Acesso em dezembro de 2019.

PIZZOL, L. **Uso da web de dados como fonte de informação no processo de inteligência competitiva setorial**. 2014. Dissertação (Mestrado em Engenharia e Gestão do Conhecimento) - Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2014. Disponível em: <<http://btd.egc.ufsc.br/wp-content/uploads/2015/04/Leandro-Dal-Pizzol.pdf>>. Acesso em: jul. 2020.

PROVO, A. From index to network: topic maps in the Enhanced Networked Monographs project. **The Indexer: The International Journal of Indexing**, [S.l.], v. 37, n. 1, p. 13–35, 2019.

RADA Mihalcea. Recent news. [20--]. Disponível em: <<https://web.eecs.umich.edu/~mihalcea/>>. Acesso em: 15 abr. 2020.

REDIGOLO, F. M. O processo de análise de assunto na catalogação em bibliotecas universitárias para proposta de normalização: estudo de observação com protocolo verbal. **Brazilian Journal of Information Science: Research Trends**, v. 8, n. 1, 2014. Disponível em: <<http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/4248>>. Acesso em: 14 abr. 2020.

REDIGOLO, F. M.; FUJITA, M. S. L. A leitura profissional do catalogador e seu papel como mediadora da informação. **Inf. Inf.**, Londrina, v. 20, n. 3, p. 356 - 376, set./dez. 2015.

ROWLEY, J. Índices impressos e serviços de notificação corrente. In: _____. **A biblioteca eletrônica**. 2.ed. Brasília: Briquet de Lemos, 2002.

RUBI, Milena Polsinelli. Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. In: FUJITA, MSL., (org.) et al. **A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias**. Um estudo de observação do contexto sociocognitivo com protocolos verbais [online]. São Paulo: Editora UNESP; São Paulo: Cultura Acadêmica, 2009. 149 p. Disponível em: <<http://books.scielo.org/id/wcvbc/pdf/bocato-9788579830150-06.pdf>> Acesso em: 01 abr. 2020.

SALTON, G. **Automatic text processing**. Boston: Addison-Wesley Longman Publishing. 1988.

SCHMIDMAIER, D. Book indexes: experiences and expectations in the German Democratic Republic (GDR). **The Indexer**, v. 25, n. 2, oct. 2006.

SILVA, D. **Escrita cuneiforme**. [S.l.: s.n], 2018. Disponível em: <<https://www.estudopratico.com.br/escrita-cuneiforme/>>. Acesso em: 2 nov. 2019.

SILVA, M. dos R. da; FUJITA, M.a S. L. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação** [online]. v.16, n.2, p.133-161. 2004.

SOCIETY OF INDEXERS. **Wheatley Medal**. Sheffield: [s.n], 2020. Disponível em: <<https://www.indexers.org.uk/about-us/awards/wheatley-medal/>> Acesso em: 1 abr. 2020.

SYLVA, L. da. **Nouveaux horizons en indexation automatique de monographies. Documentation et bibliothèques**, [S.l.], v. 48, n. 4, p. 155-167, 2002.

THE INDEXER. Disponível em <<https://www.theindexer.org/>>. Acesso em: dez. 2019.

TOMASI, C. Como elaborar Índices em trabalho acadêmico. In: _____. **Comunicação Científica**. São Paulo: Atlas, 2008.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO): **Guia para la redacción de artículos científicos destinados a la publicación**, París: UNESCO, Programa General de Información y UNISIST, 1983.

VAHED, Z.; SHAPOURI, S.; MIRJOOD, S. H. Evaluation back-of-the-book indexes of Persian books with subject of microbiology in central library of Islamic Azad University, Lahijan branch (Iran) based on indexing international standards ISO 999. **Intl. Res. J. Appl. Basic. Sci.** v. 4, n. 8, p. 2125-2128, 2013.

WEIHS, Jean. Indexes, indexing, and the technical services staff part 2: cataloguers. **Technicalities**, [S.l.], v. 27, n. 2, mar./ apr. 2007.

WEINBERG, B. H. Book indexes in France: medieval specimens and modern practices. **The Indexer**, v. 22, n. 1, apr. 2000.

WELLISCH, Hans H. **Glossary of terminology in abstracting, classification, indexing, and thesaurus construction**. Medford, New Jersey: Information Today, 2000.

WELLISH, H. H. How to make an Index - 16th century style: Conrad Gessner on Indexes and Catalogs. **Intern. Classificat.** v. 8, n. 1, 1981.

WU, Z. et al. Can Back-of-the-Book Indexes be Automatically Created?. In: CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 13., 2013, San Francisco. **Anais...** San Francisco, 2013.

APÊNDICE A - SOFTWARES DE INDEXAÇÃO PARA ÍNDICE DE FINAL DE LIVRO²⁶ (continua)

NOME	BREVE DESCRIÇÃO	PREÇO ²⁷	LINK PARA ACESSO AO SOFTWARE
Cindex	Criado pela <i>Indexing Research</i> , o CINDEX é um software que elabora índices para livros, jornais e outras publicações. Utilizado também na preparação de glossários e listas de cabeçalho de assunto com base nos índices existentes. Auxilia em etapas como, verificação de remissivas e formatação. Está disponível para Windows e Mac.	Cindex Student 4.0 \$79 Cindex Standard 4.0 \$549 Cindex Publishers 4.0 \$799	INDEXING Research. New York: [s.n.], [20--]. Disponível em: < https://www.indexres.com/ >. Acesso em: 01 abr. 2020.
SKY Index	Software que utiliza interface semelhante à de uma planilha do Microsoft Excel. E possui facilidade de edição e localização das entradas do IFL. É um software profissional, mas tem desconto para estudantes. Está disponível para Windows.	SKY Index Professional 8.0 \$579	SKY Index Professional. Stephens City: [s.n.], 19 mar. 2020. Disponível em: < https://www.sky-software.com/ >. Acesso em: 01 abr. 2020.
Macrex	É um software criado para auxiliar indexadores que trabalham com material impresso, não tem possibilidade de realizar marcação de entrada no texto (somente em arquivos HTML e XML). O objetivo do software é ajudar os indexadores a melhorar a consistência e produtividade automatizando tarefas rotineiras (classificação, impressão, repaginação entre outros), desta forma, os indexadores dispõem de mais tempo para concentrar na redação e construção das entradas do índice. Está disponível para Windows 7, XP e Vista, a versão 8 é executada no Apple Macintoshes e Parallels.	Não informado	MACREX. [S.l.: s.n.], 19 jul. 2018. Disponível < http://www.macrex.com/ >. Acesso em: 01 abr. 2020.

²⁶ Para inglês foi utilizada a busca no google “*software for back-of-book indexes*” sendo selecionados 6 softwares relevantes. Em português foi utilizado à busca “softwares para ‘índice de final de livro’ ou ‘índice remissivo’” não sendo localizado nenhum software, somente menção ao Microsoft Word.

²⁷ Preços informados até o dia 1 de abril de 2020.

APÊNDICE A - SOFTWARES DE INDEXAÇÃO PARA ÍNDICE DE FINAL DE LIVRO (conclusão)

NOME	BREVE DESCRIÇÃO	PREÇO ²⁸	LINK PARA ACESSO AO SOFTWARE
TExtract	Cria Índice de Final de Livro para edições impressas e e-books, extrai palavras e frases de forma automática ²⁹ e apresenta uma lista em ordem alfabética em seguida, o índice deve ser exportado para formatação e revisão. No site oficial do software tem demonstração de utilização do software.	\$79,00 (para um único livro) a \$595 (modo profissional)	TEXTTRACT the back-of-book indexing system. Utrecht: [s.n.], 2019. Disponível < https://www.texyz.com/texttract/index.htm >. Acesso em: 01 abr. 2020.
WINDEX	Foi projetado por um indexador profissional. O indexador deve digitar as entradas e número das páginas, o WINDEX coloca em ordem alfabética, combina entradas e realiza a formatação. Está disponível para Windows.	Disco e manual do Windex - \$129,00	WINDEX Indexing Software from Watch City Software. [S.l.: s.n.]. Disponível em: < http://www.abbington.com/holbert/windex.htm >. Acesso em: 01 abr. 2020.
PDF Index Generator.	O PDF Index Generator analisa o livro em PDF, extrai as palavras do texto para o índice e indica sua localização no PDF. O principal objetivo do PDF Index Generator é automatizar o processo de elaboração de Índices de Final de Livros. Está disponível para Windows, Mac OS X e Linux.	\$69.95	PDF Index Generator. [S.l.: s.n.], 2020. Disponível em: < https://www.pdfindexgenerator.com/ >. Acesso em: 01 abr. 2020.
Microsoft Word	O Microsoft Word é um processador de texto que inclui recurso de indexação como o Índice. As frases e palavras devem ser marcadas e depois compiladas para criar um índice, os localizadores também são adicionados e em caso de inclusão ou exclusão podem ser editados. O Word não é recomendado para publicações em larga escala e é considerada muito complexa para leigos (BROWNE; JERMEY, 2007).	Pacote Office de R\$ 299,00 a R\$ 499,00 (por ano uso doméstico)	MICROSOFT. [S.l.: s.n.], [20--]. Disponível em: < https://www.microsoft.com/pt-br/ >. Acesso em: 01 abr. 2020.

Fonte: elaborado pela autora.

²⁸ Preços informados até o dia 1 de abril de 2020.

²⁹ Browne e Jerney (2007) informa as limitações desse tipo de ferramenta. Alguns nomes próprios são perdidos, outros não são invertidos; sinônimos não são combinados; algumas frases são quebradas de forma inadequada e outras vezes palavras isoladas são combinadas em frases.

APÊNDICE B – ORIENTAÇÃO E SUGESTÃO DE ESTRUTURA PARA POLÍTICA DE INDEXAÇÃO EM ÍNDICE DE FINAL DE LIVRO

Rubi (2012, p. 118) em seu trabalho realiza um levantamento com afirmações de vários autores sobre os elementos que devem constar na política de indexação. Nesta dissertação, destacam-se os elementos que podem ser aplicados ao IFL, conforme demonstrado na Figura 19, pois os demais estão relacionados à indexação para a base de dados e afins.

Figura 19 - Síntese dos elementos a serem considerados na elaboração de uma política de indexação

ELEMENTOS DE POLÍTICA DE INDEXAÇÃO	FONTES
Avaliação do sistema	Carneiro (1985);
Campos e sub/campos do formato MARC	Moen e Benardino (2003)
Capacidade de consulta a esmo (<i>browsing</i>)	Fosket (1973)
Capacidade de revocação e precisão do sistema	Carneiro (1985);
Cobertura de assuntos	Carneiro (1985); Guimarães (2000)
Consistência/ Uniformidade	Olson e Boll (2001); Vilchez Pardo (2002)
Detalhamento dos procedimentos relacionados à leitura documentária	Manuais INIS, AGRIS e BIREME
Economia	Vilchez Pardo (2002)
Escolha da linguagem	Carneiro (1985); Vilchez Pardo (2002)
Especificidade	Cutter (1876); Foskett (1973); Carneiro (1985); Chaumier (1986); Connell (1996); Olson e Boll (2001); Vilchez Pardo (2002)
Estratégia de busca	Carneiro (1985);
Adequação	Olson e Boll (2001)
Exaustividade	Lancaster (1968; 1993); Carneiro (1985); Chaumier (1986); Van Slype (1991); Olson e Boll (2001)
Forma de saída dos resultados	Carneiro (1985);
Formação do indexador	Fosket (1973)
Manual de indexação (elaboração/utilização)	Manuais INIS, AGRIS e BIREME; Guimarães (2000)
Seleção e aquisição de documentos-fonte	Carneiro (1985); Guimarães (2000)
Síntese	Vilchez Pardo (2002)
Uso (pelo usuário)	Vilchez Pardo (2002)

Fonte: adaptado de Rubi (2012, p. 118)

Os elementos destacados na figura anterior são explicitados com mais detalhes no Quadro 15 com o propósito de auxiliar o profissional que estiver elaborando a Política de indexação.

Quadro 15 - Detalhamento dos elementos para Política de indexação (continua)

ELEMENTOS DE POLITICA DE INDEXAÇÃO	EXPLICAÇÃO (O QUE É?)	FONTE
Cobertura de assuntos	assuntos cobertos pelo sistema (centrais e periféricos) aspectos como a conversão retrospectiva de dados e a compatibilidade de linguagem documentária entre integrantes de um mesmo sistema cooperativo	Carneiro (1985) Guimarães (2000)
Consistência/ Uniformidade	Este tópico trata da forma como um mesmo assunto deve ser analisado conceitualmente e traduzido da mesma maneira, sendo o tamanho do vocabulário utilizado e o número de conceitos representados fatores que afetam a consistência e a uniformidade	Olson e Boll (2001)
Detalhamento dos procedimentos relacionados a leitura documentária	para cada assunto haverá um cabeçalho uniforme, destacar os casos de homonímia todos os manuais apresentam uma descrição de como deve ser realizada a leitura do documento com fins à indexação, como por exemplo, recomendações sobre quais as partes do documento que devem ser lidas e quais aquelas que devem ser evitadas. Em nosso entendimento, as descrições apresentadas, ainda que possam parecer simplistas, demonstram uma preocupação com a fase inicial da indexação.	Vílchez Pardo (2002) Manuais INIS, AGRIS e BIREME
Especificidade	O princípio da entrada específica esclarece que os assuntos devem dar entrada pelo termo mais específico e pela não pela classe a que estão subordinados. Especificidade do sistema de recuperação da informação: refere-se “a extensão em que o sistema nos permite ser precisos ao especificarmos o assunto de um documento que estamos processando”. nível de abrangência em que o sistema permite especificar os conceitos identificados documento; a seleção dos conceitos no processo da indexação será influenciada pelos elementos da exaustividade e especificidade. não como um elemento específico da política de indexação, mas sim como uma questão prática a ser considerada no momento da indexação. relacionado à fase de tradução do conceito para o vocabulário controlado, diz respeito ao nível hierárquico da representação do assunto. Está dividido em três fatores: a especificidade e a co-extensividade do vocabulário; a especificidade de sua aplicação e a especificidade do termo no contexto da indexação. o cabeçalho designa um único assunto;	Cutter (1876); Foskett (1973); Carneiro (1985); Chaumier (1986); Connell (1996); Olson e Boll (2001); Vílchez Pardo (2002)

Quadro 16 - Detalhamento dos elementos para Política de indexação (conclusão)

ELEMENTOS DE POLITICA DE INDEXAÇÃO	EXPLICAÇÃO (O QUE É?)	FONTE
Exaustividade	<p>a principal decisão política diz respeito à exaustividade da indexação, a qual corresponde, grosso modo, ao número de termos atribuídos em média.</p> <p>medida de extensão em que todos os assuntos discutidos em um certo documento são reconhecidos durante a indexação e traduzidos na linguagem do sistema;</p> <p>a seleção dos conceitos no processo da indexação será influenciada pelos elementos da exaustividade e especificidade.</p> <p>A exaustividade depende fundamentalmente da política de indexação e qualidade do trabalho dos documentalistas, e especialmente de sua capacidade de julgar o que é importante e o que não é importante e seu 'olfato' para detectar os conceitos implícitos.</p> <p>número de conceitos representados no registro bibliográfico; está condicionado ao estágio de análise de assunto.</p>	<p>Lancaster (1968; 1993);</p> <p>Carneiro (1985);</p> <p>Chaumier (1986);</p> <p>Van Slype (1991);</p> <p>Olson e Boll (2001)</p>
Formação do indexador	Formação do indexador: em termos de conhecimento das áreas de assunto dos documentos; da metodologia de indexação das características da linguagem documentária e de suas habilidades práticas.	Fosket (1973)
Uso (pelo usuário)	não perder de vista o usuário, a coleção etc	Vílchez Pardo (2002)
Exaustividade	a principal decisão política diz respeito à exaustividade da indexação, a qual corresponde, grosso modo, ao número de termos atribuídos em média.	Lancaster (1968; 1993);

Fonte: adaptado de Rubi (2012, p. 118).

Ao escreverem sobre a Política de indexação para o IFL, as autoras Gomes e Gusmão (1983, p. 63) apresentam algumas informações essenciais que devem constar na Política de indexação, as autoras apresentam um “roteiro para o manual de procedimentos”, o que nesta dissertação será adaptado para ser apresentado na Política de indexação. Para as autoras essas são as informações que devem constar na Política, no mínimo:

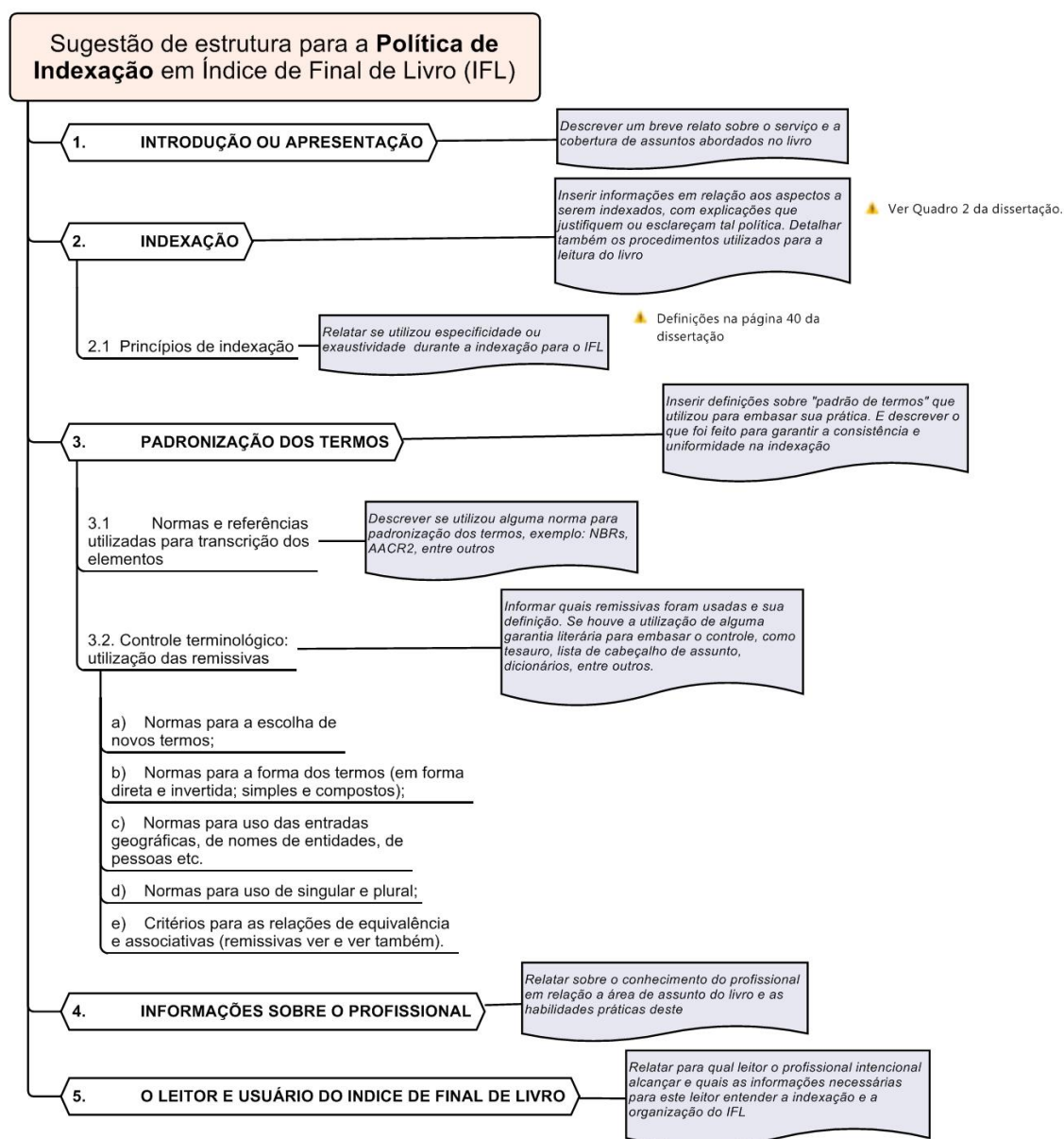
- 1) Breve relato sobre o serviço e a coleção
- 2) Aspectos a serem indexados, com explicações que justifiquem ou esclareçam tal política
- 3) Princípios de indexação
- 4) Normas para transcrição dos elementos
- 5) Regras para arquivamento
- 6) Vocabulário utilizado (com rede de remissivas e referências).

Este deverá conter, ainda:

- a) normas para a escolha de novos termos;
- b) normas para a forma dos termos (em forma direta e invertida; simples e compostos);
- c) normas para uso das entradas geográficas, de nomes de entidades, de pessoas etc.
- d) normas para uso de singular e plural;
- e) critérios para as relações de equivalência e associativas (remissivas **ver** e **ver também**) (GOMES E GUSMÃO, 1983, p. 63).

Foram compiladas as orientações descritas nas publicações de Gomes e Gusmão (1983) e Rubi (2012) o que resultou na Figura 20. O esquema tem o objetivo de direcionar o profissional que está escrevendo a Política de Indexação sobre os elementos essenciais que devem constar na Política.

Figura 20 - Sugestão de estrutura para a Política de Indexação em Índice de Final de Livro



Fonte: elaborado pela autora com dados apontados por Gomes e Gusmão (1983, p. 63) e Rubi (2012, p. 107 – 118).