

A EVOLUÇÃO DA TECNOLOGIA NA ELABORAÇÃO DO ÍNDICE DE FINAL DE LIVROS: UMA REVISÃO DE LITERATURA

Resumo: O Índice de Final de Livro é um dos instrumentos mais antigos utilizados para Organização e Recuperação da Informação, e 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. Dessa forma, este artigo tem o objetivo de apresentar estudos que discutem sobre o uso da tecnologia na elaboração do índice, a partir de 16 publicações selecionadas pelo processo de levantamento bibliográfico. Os resultados mostraram que as publicações nos primeiros anos apresentaram de forma incipiente a inserção tecnológica e, com o passar dos anos, houve vários relatos de índices produzidos por meio da indexação automática. Conclui-se que são necessários estudos sobre a utilização da indexação semiautomática para auxiliar na elaboração do índice, considerando que a intervenção do indexador, mesmo com o auxílio da tecnologia, continua sendo necessária para explicitar a semântica e o contexto na elaboração de índices, assim como a utilização da tecnologia é necessária para agilizar a finalização do índice.

Palavras-chave: índice de final de livro; indexação; tecnologia; análise de assunto; editoração de livro.

Mayara Silva Gonçalves

Mestre em Gestão e Organização do Conhecimento pela Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG).
mayara.biblio@gmail.com

Gercina Angela de Lima

Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG) com pós-doutorado em Estudos sobre metodologias para o processo de indexação de documentos textuais, pela Universidad Carlos III, Madrid, Espanha, e pela Escola de Artes e Comunicação da USP. Professora Titular no Departamento de Organização e Tratamento da Informação, na Escola de Ciência da Informação da UFMG. Bolsista de Produtividade (PQ) do CNPq, Nível 2. Bolsista do Programa Pesquisador Mineiro-PPM XII-2018 da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig).
limagercina@gmail.com

THE EVOLUTION OF TECHNOLOGY IN THE PREPARATION OF BACK-OF-BOOK INDEXES: A LITERATURE REVIEW

Abstract: The back-of-book index is one of the oldest devices used for Organizing and Retrieving Information, and its significance relies on the fact of its being one of the main access point to the content of the book, helping the reader to comprehend the relevant concepts contained in the manuscripts. However, the process of elaboration of back-of-book in Brazil happens, often, in a manual way, what demands time and financial costs, what justifies the low number of publications and contain the index. Thus, this article aims at presenting studies que discuss about the use of technology for the elaboration of the index, from 16 publications selected by means of literature review. The results have shown that the publication of the first years presented in an incipient way the technological insertion, and over the years, there were reports of indexes produced by means of automatic indexing. It is evinced that are necessary studies about the use of semiautomatic indexing to help in the elaboration of the index, considering that the indexer's intervention, even with the help of technology, is still necessary for explicating the semantics and the context in the elaboration of indexes, as well as the use of technology if needed for expedite the finalization of the index.

Keywords: back-of-book index; indexing; technology; matter analysis; book publishing.

1 INTRODUÇÃO

A crescente importância da informação tem estimulado a busca por ferramentas que permitam localizá-la 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 o tradicional 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 a sua estrutura.

O índice é um produto da indexação, que identifica todas as palavras e/ou frases que se referem às 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. Além disso, o índice proporciona pesquisas mais efetivas em menos tempo e auxilia o leitor na compreensão de conceitos relevantes contidos no documento, 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 século 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 acervo. 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”, que foi 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 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, considerando que não foram identificadas publicações nacionais sobre essa temática, conforme observado na seleção dos 51 documentos localizados nas bases de dados (apresentados na seção de Metodologia deste artigo). Nas publicações, identificou-se, também, que a tecnologia é abordada como uma ferramenta – no sentido de criação de *softwares*, sendo assim, uma análise com enfoque crítico de sua

utilização no produto “índice” de forma extensiva é pouco frequente.

Tendo em vista esse cenário, este artigo tem por objetivo sintetizar as principais discussões sobre o uso da tecnologia, a respeito de sua evolução na elaboração de índice, por meio das ideias dos autores das publicações selecionadas pelo método de revisão de literatura. Desse modo, a pergunta norteadora do artigo é “como a tecnologia tem auxiliado na elaboração do IFL no decorrer dos anos?”. Ressalta-se, também, que os resultados apresentados neste artigo, são provenientes dos resultados da dissertação de mestrado de Gonçalves (2020).

2 ÍNDICE DE FINAL DE LIVRO

O IFL é um produto que apresenta a visão geral do conteúdo do livro e normalmente faz parte dos elementos pós-textuais. Consiste em uma lista dos itens relevantes abordados em um determinado livro com indicação das páginas em que esses assuntos são tratados. Em outras palavras, é uma lista estruturada de cabeçalhos e subcabeçalhos, com seus respectivos indicativos que vinculam aos segmentos do livro (CSOMAI; MIHALCEA, 2007; MEKKI; NAZARENKO, 2006).

Para Associação Brasileira de Normas Técnicas (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2004, p. 1), o índice é a “relação de palavras ou frases, ordenadas segundo determinado critério, que localiza e remete para as informações contidas num texto”. Em relação aos princípios e objetivos dos índices do livro, de acordo com Bonura (1994, p. 14, tradução nossa) 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 retornará ao livro posteriormente 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.

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 de um documento.

a. Funções do índice

Quanto às suas funções, o índice, de acordo com Mulvany (1994, p. 5-6) tem oito funcionalidades, a saber:

1. Identificar e localizar as informações relevantes contidas 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 itens 1 e 2 exigem que o indexador julgue a diferença entre informações relevantes e irrelevantes para distinguir informação substantiva e menção passageira de um tópico, com a finalidade de otimizar o tempo de busca informacional do leitor (MULVANY, 1994, p. 5-6).

O item 3 aborda discussões importantes em relação a produção dos cabeçalhos baseados na terminologia utilizada no documento a partir da indexação. Para Sylva (2002), existem dois tipos de indexação automática, a saber: a) indexação totalmente automática, e a b) 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 leitor um índice final. Uma indexação 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 leitor). A indexação automática destina-se a permitir que os leitores 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 leitor, sem intervenção humana (SYLVA, 2002).

O segundo tipo é a indexação por 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. A indexação por extração, seja manual ou automática, é o tipo de indexação que alguns profissionais consideram como a mais apropriada, conforme Mulvany (1994, p. 5-6), pois é embasado totalmente na terminologia do documento. Tendo em vista a argumentação de que, durante o processo de elaboração do índice a linguagem do autor deve ter preferência sobre termos atribuídos, considera-se que a linguagem do texto é a mais recomendável para compor os termos do índice. Porém, o desempenho não é satisfatório e a abordagem é limitada, pois dependem da terminologia exata utilizada pelo autor. E, no caso da indexação 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. Porém, o desempenho não é satisfatório e a abordagem é limitada, pois dependem da terminologia exata utilizada pelo autor. No caso, da indexação por atribuição, de forma manual, deve-se atentar a outra argumentativa dos profissionais ao utilizar preferencialmente este tipo de indexação. Considerando o argumento de que os assuntos encontrados no livro nem sempre são declarados textualmente. Por exemplo, em um livro sobre criação de cães, vários parágrafos podem relatar tipos de comida para cães e não haver menção à palavra “nutrição”, no entanto, o assunto nutricional deve estar no índice, ou seja, uma sintetização dos conceitos em frases de indexação. O indexador analisa o texto e identifica assuntos relevantes, sejam eles mencionados explicitamente ou não. Dessa forma, utiliza-se a indexação por atribuição, ou seja, essa indexação envolve a atribuição de termos a partir de uma fonte que não é o próprio documento (LANCASTER, 2004; SYLVA, 2002).

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 (utilização das remissivas). No entanto, o item 5 “reúne informações sobre assuntos que estão espalhados pelo arranjo do documento” e 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 negligenciam as informações relevantes e o índice que contém informação dispersa diminui a confiança do leitor (MULVANY, 1994).

O item 6 aborda o elemento crucial das remissivas no índice, especificamente a remissiva “*ver*” que existe para antecipar a linguagem dos leitores durante sua busca no 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 é comprometida (MULVANY, 1994).

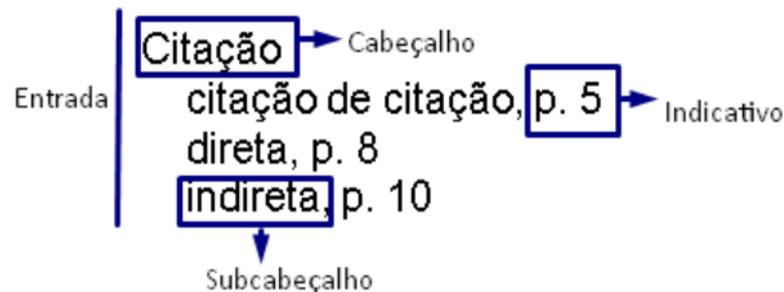
O item 7 instrui o indexador a “sintetizar títulos e subtítulos em entradas”. A capacidade de o indexador sintetizar a entrada contribui grandemente para a integridade do índice. Geralmente, esse esforço é tão simples quanto resolver diferenças entre termos sinônimos. Outras vezes, o processo de síntese é mais complexo, quando a frase não está compreensível.

Enquanto o item 7 se refere ao *design* de entradas de índice, o item 8 discute a organização de entradas do índice, que pode ter a ordenação dos seguintes tipos: a) alfabética; b) sistemática; c) cronológica; d) numérica e e) alfanumérica, sendo a ordenação alfabética a mais usual para entradas de índice, (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2004; MULVANY, 1994).

b. Estrutura do índice

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 indicativos, 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 1:

Figura 1: Estrutura do índice



Fonte: elaborado pelas autoras (2022).

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 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2004; SYLVA, 2002), conforme este exemplo na Figura 2:

Figura 2: Exemplo de remissivas no índice

Análise de assunto, p. 80, 153

ver linguagem documentária, p. 2-10

ver também tradução p. 15,19

REMISSIVAS

Fonte: elaborado pelas autoras (2022).

3 METODOLOGIA

Neste artigo, adotou-se a revisão de literatura como metodologia, que é um meio de identificar, avaliar e interpretar toda a produção disponível relevante para uma questão de pesquisa específica, área temática ou fenômeno de interesse. A revisão é utilizada, neste trabalho, por apresentar uma estratégia de pesquisa predefinida, 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-se¹ as orientações da publicação de Kitchenham (2004), que foram seguidas às etapas de: 1) definição da objeto de análise; 2) seleção dos termos (expressão 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 como objeto de estudo a identificação das principais abordagens de estudos existentes à respeito do 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.

Quadro 1: Controle de Vocabulário

Idioma	Ligagem natural	Vocabulário controlado
Português	índice controlado de livro índice interno índice de final de livro índice alfabético-remissivos Índices impressos	índice de livros
Espanhol	índice controlado de libro índice final de libro	índice alfabéticos
Inglês		<i>Book indexes</i>

Fonte: elaborado pelas autoras (2022).

¹ O foco da revisão da literatura são as áreas da saúde e da engenharia.

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 Biblioteconomia e Ciência da Informação (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 publicações.

A quarta e a quinta etapa – levantamento bibliográfico e leitura exploratória para seleção dos documentos, respectivamente, estão apresentadas de forma simultânea na Tabela 1. Sendo que a primeira coluna, “publicações recuperadas”, é resultado do levantamento bibliográfico – etapa 4 que resultou na recuperação de 415 documentos, e as três últimas colunas dizem respeito à leitura exploratória para seleção dos documentos, ou seja, houve três critérios de exclusão a fim de selecionar publicações que estivessem alinhadas aos objetivos propostos inicialmente. Tais objetivos são: 1) “não atende aos objetivos” – se refere a publicações que relatavam sobre o índice no sentido numérico (ex. índices de doenças cancerígenas, índices geográficos, entre outros); assim foi realizada a leitura dos títulos e resumos para averiguar o objetivo dos documentos; 2) “publicações repetidas”, 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 estava disponível por apresentar somente o resumo ou ser um artigo pago, conforme demonstrado na 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: adaptada 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, obteve-se maior número de documentos selecionados na base SCOPUS. Após a leitura exploratória e a aplicação dos critérios de exclusão, explicados anteriormente, foram selecionados 51 documentos que são analisados quantitativamente nos resultados a seguir.

4 RESULTADOS

Os resultados são apresentados em duas seções: na primeira, apresenta-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. Na segunda seção, identifica-se os dados qualitativos, examina-se os textos que abordam o tema “elaboração do índice”, e relata-se os pontos principais dos artigos.

a. Dados quantitativos

Em relação aos resultados quantitativos, após a seleção, foram identificadas 51 publicações. O maior número de publicações são artigos de divulgação em periódicos científicos, somando um total de 40 artigos. Foram recuperados seis 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: elaborada pelas autoras (2022).

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

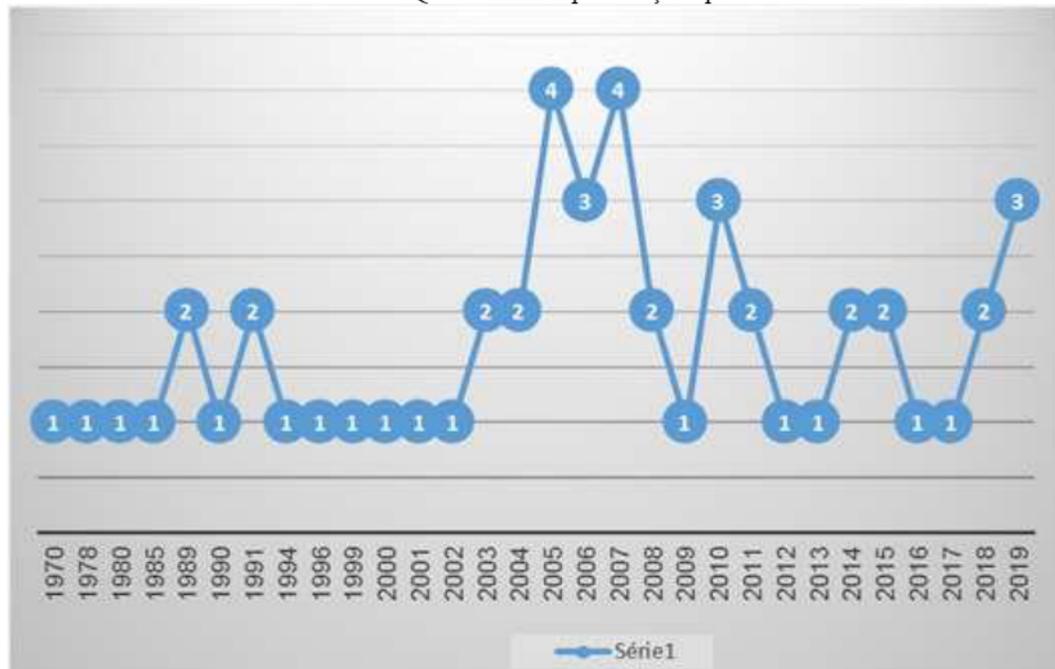
Tabela 3: Quantidade de publicações por periódico

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

Fonte: elaborada pelas autoras (2022).

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. Além disso, de acordo com a descrição da 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 pelas autoras (2022).

Das publicações selecionadas, foram identificados 69 autores, considerando que alguns documentos foram escritos em coautoria. Dos 69 autores, 38 possuíam vínculo com Universidades;² 7 foram caracterizados em “outros vínculos”³ por estarem ligados a centros de pesquisa ou outras empresas; e 5 afiliados a associações de bibliotecários.⁴ 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.

² 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.

³ Biblioteca Nacional da Alemanha; Google; Harvard Library; News Information Center; Palo Alto Research Center; presidente da ASIS & T.

⁴ American Society Indexers e American Library Association.

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

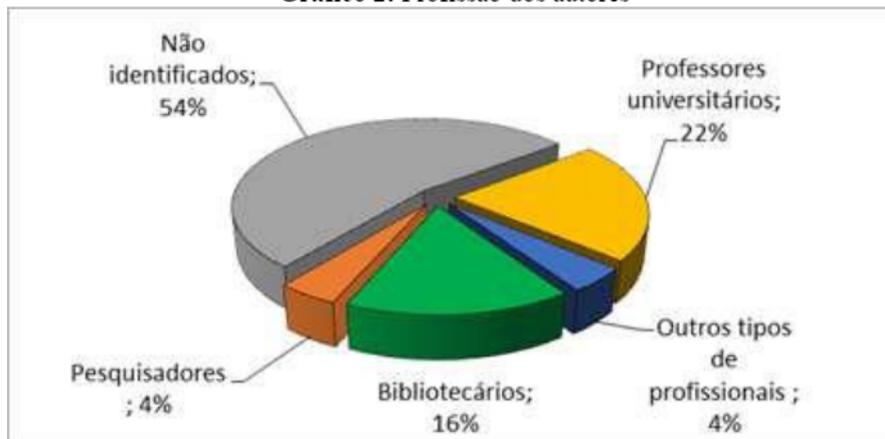
Fonte: elaborado pelas autoras (2022).

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 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 à linguística. No Gráfico 2, estão expostos todos os relatos desse tópico.

Gráfico 2: Profissão dos autores



Fonte: elaborado pelas autoras (2022).

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. 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 outros produtos da 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 as ações necessárias para a sua elaboração;
3. **Fonte de pesquisa:** apresentam o índice como fonte 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 5.

Tabela 5: Abordagens de estudo

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

Fonte: elaborado pelas autoras (2022).

Assim, a partir dos dados quantitativos dos 51 documentos, foi possível observar as principais pesquisas / estudos referentes ao índice, em relação ao tipo de publicação, os principais periódicos que tratam sobre a temática, o número de publicações durante os anos, a vinculação e profissão dos autores que se interessam sobre o tema e as principais abordagens de pesquisa identificadas nos documentos selecionados.

Na próxima seção apresenta-se a análise qualitativa dos 16 documentos das 51 publicações selecionadas, que abordam especificamente a “elaboração de índice” (apresentado na tabela anterior), com a finalidade de verificar a tecnologia neste cenário.

b. Dados qualitativos

Nesta seção, são apresentados os resultados dos 16 documentos que discorriam sobre a elaboração de índice, objeto de investigação do artigo e, principalmente, quanto à evolução tecnológica, no intervalo temporal de 1970 a 2019, apresentados na ordem cronológica das datas de publicação.

A partir do resultado da 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* abordam em seu artigo sobre o auxílio da tecnologia para facilitar o trabalho de indexação. Esse relato consta do trabalho intitulado “*Computer-aided production of book indexes*” (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.

Em 1989, o trabalho intitulado *Indx and findphrases, a system for generating Indexes for Ditroff documents* (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 não relevantes, 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. O relato dos autores foi pautado 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 de o 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 (CHI et al., 2004, 2006). O primeiro retratou o panorama

introdutório sobre o sistema e o segundo descreveu o índice de assunto produzido pelo sistema. A primeira publicação *eBooks with Indexes that Reorganize Conceptually*, foi escrita em 2004 e a segunda, *ScenIndex: Conceptually Reorganizing Subject Indexes for Reading*, data de 2006 (CHI et al., 2006). O sistema *ScenIndex* 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 (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, gerando produtos de qualidade inferior.

As publicações 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 (PRESIDENT, 2009).

O primeiro trabalho, publicado 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. Os resultados demonstraram 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 de 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.

Já em 2013, o artigo *Can Back-of-the-Book Indexes be Automatically Created?* (WU et al., 2013), 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, remissivas entre outros, e sugerem outros estudos nessa área (WU et al., 2013).

Ainda 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 leitor. 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 (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 profissional 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 um 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 leitor possa acessar o *e-book* dos trabalhos completos. É possível, também, fazer filtragem da quantidade de termos que foram apresentados para o leitor. 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 leitor e que este projeto pode ser desmembrado em outros. Comentando sobre falhas no produto final, a autora ressaltou o esforço e a participação da equipe para validar os termos selecionados, o que ainda indica a necessidade do indexador.

5 CONSIDERAÇÕES FINAIS

Considera-se que por meio da revisão de literatura realizada pode-se alcançar o objetivo proposto, pois possibilitou evidenciar a evolução da tecnologia na elaboração de índice. Apresentação da análise quantitativa dos 51 documentos e da análise qualitativa dos 16 desses que tratavam, especificamente, sobre a elaboração do IFL, permitiram classificar as publicações descritas em duas categorias principais: índices de forma semiautomática (ABE; BERRY, 1989; BOLSHAKOVA; IVANOV, 2018; FRANCIS; GREENWAY, 2015; HINES; HARRIS, 1970; PROVO, 2019) e índices de forma automática (CHI et al., 2004, 2006; CSOMAI; MIHALCEA, 2006, 2007, 2008; MEKKI; NAZARENKO, 2005; 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, ressalta o trabalho de Sylva (2002) que recomendou melhorias para criação de índice com a indexação automática. Já 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. As publicações 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, atendendo a uma necessidade no mercado editorial, bem como do esforço 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.

Assim, retomamos a pergunta inicial deste artigo “como a tecnologia tem auxiliado na elaboração do IFL no decorrer dos anos?”. Observa-se que 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, pode se dizer 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.

Conclui-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 do indexador e da tecnologia na elaboração de tal produto, no sentido de tratá-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.

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.

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

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.

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

CHI, E. H. *et al.* ScentIndex: conceptually reorganizing subject indexes for reading. **IEEE Symposium on Visual Analytics Science and Technology**, Baltimore, 2006. p. 159-166.

CHI, Ed. H. *et al.* eBooks with indexes that reorganize conceptually. **ACM**, Vienna, Apr. 2004.

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, **Anais [...]**. México: Springer, 2006. p. 19-25.

CSOMAI, A.; MIHALCEA, R. Investigations in unsupervised back-of-the-book indexing. *In: FLAIRS Conference*, May 2007, [Texas]. **Anais [...]**. [Texas: *S. n.*], 2007.

CSOMAI, A.; MIHALCEA, R. Linguistically motivated features for enhanced back-of-the-book indexing. **Association for Computational Linguistics**, Columbus, Ohio, USA, p. 932-940, June 2008.

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

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

GOMES, H. E.; GUSMÃO, H. R. **Guia prático para elaboração de índice**. Niterói, RJ: Grupo de Bibliotecários em Informação e Documentação em Ciências Sociais e Humanidades, 1983.

GONÇALVES, M. O. S. G. **Diretrizes para elaboração e estruturação de Índice de Final de Livro por meio de indexação semiautomática**. Orientadora: Gercina Ângela de Lima. 2020. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2020. Disponível em: <https://repositorio.ufmg.br/handle/1843/34192>. Acesso em: 22 fev. 2021.

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

KITCHENHAM, B. **Procedures for performing systematic reviews**. Austrália: [s. n.], 2004. Disponível em: <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>. Acesso em: 18 mar. 2022.

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

MEKKI, T. A. E.; NAZARENKO, A. **Using NLP to build the hypertextuel network of a back-of-the-book index**. [S. l.: s. n.], Sep. 2006.

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

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

MULVANY, N. C. **Indexing books**. Chicago; London: Library Materials, 1994.

PINHEIRO, L. V. R.; FERREZ, H. D. **Tesouro brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília, DF: Instituto Brasileiro de Informação em Ciência e Tecnologia, 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.

PIZZOL, L. D. **Uso da web de dados como fonte de informação no processo de inteligência competitiva setorial**. 2014. Dissertação (Engenharia e Gestão do Conhecimento) – Universidade Federal de Santa Catarina, Florianópolis, 2014. Disponível em: <https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/130931/332814.pdf?sequence=1&isAllowed=y>. Acesso em: 18 mar. 2022.

PRESIDENT Honors Outstanding Early-Career Scientists. [*S. l.: s. n.*], jul. 2009. Disponível em: https://www.nsf.gov/news/news_summ.jsp?cntn_id=115171. Acesso em: 02 nov. 2019.

PROVO, A. From index to network: topic maps in the enhanced networked monographs project. **The Indexer**, [*s. l.*], v. 37, n. 1, p. 13-35, 2019.

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. [*S. l.: s. n.*], [20--]. Disponível em: <https://www.theindexer.org/>. Acesso em: 03 nov. 2019.

WEIHS, J. Indexes, indexing, and the technical services staff part 2: cataloguers. Technicalities, **The Indexer: the International Journal of Indexing**, [*s. l.*], v. 27, n. 2, Mar./Apr. 2007.

WU, Z. *et al.* Can back-of-the-book indexes be automatically created?. **CIKM '13: Proceedings of the 22nd ACM international conference on Information & Knowledge Management**, San Francisco, p. 1745-1750, oct. 2013.