



Perspectivas de investigação

Referentes teóricos internacionais na temática exaustividade e especificidade na indexação: uma análise de domínio

José Augusto Chaves Guimarães

Universidade Estadual Paulista
Brasil · guima@marilia.unesp.br

Isadora Victorino Evangelista

Universidade Estadual Paulista
Brasil · isadora.biblio@marilia.unesp.br

Resumo: Considerando a importância da exaustividade e da especificidade na indexação para a qualidade da recuperação e uso da informação observa-se, paradoxalmente, que esse tema vem sendo ainda pouco estudado na Ciência da Informação e, em muitos casos, abordado, como mera medida de recuperação da informação. A vista disso, buscou-se analisar como essa temática vem sendo tratada na literatura internacional da área, em busca dos referentes teóricos que subsidiam essas discussões. Para tanto, e valendo-se do suporte metodológico da análise de domínio, trabalhou-se com três das onze abordagens propostas por Hjørland (2002): estudos epistemológicos, estudos de indexação e estudos bibliométrico. Para tanto, relaizou-se uma busca na base Scopus a partir dos termos exhaustivity e specificity nas áreas de maior conteúdo informacional do documento (título, resumo e palavras-chave), obtendo-se um total de doze artigos, sobre o qual realizou-se um estudo de citações para se chegar a uma rede que evidenciasse os referentes teóricos que norteiam essa produção. Os resultados revelaram, além da baixa produção científica nessa área, uma forte predominância anglo-americana e a forte presença de referentes oriundos da Ciência da Computação, o que sinaliza para uma efetiva necessidade de verticalização de estudos, nessa temática no âmbito da Ciência da Informação.

Palavras-chave: Exaustividade; Especificidade; Indexação; Análise de Domínio.

Abstract: Considering the importance of exhaustivity and specificity in indexing for the quality of retrieval and use of information, it is observed, paradoxically, that this topic has not been widely studied in Information Science and in many cases, it is addressed as a mere measure of information retrieval. Provided this, we aim to analyze how this topic has been addressed in the international literature of the area, in search of theoretical influences that support these discussions. For this, and drawing on the methodological support of domain analysis, we worked with three of the eleven approaches proposed by Hjørland (2002): epistemological studies, indexing studies and bibliometric studies. Therefore, we searched the Scopus database for the terms exhaustivity and specificity in the areas of highest information content of the document (title, abstract and keywords), retrieving a total of twelve articles on which we studied citations for reaching a network that showed the theoretical influences that guide this production. The results revealed, in addition to the low scientific production in this area, a strong Anglo-American dominance and a strong presence of influences coming from computer science, which points to a need for effective verticalization of the studies in this matter within Information Science.

Keywords: Exhaustivity; Specificity; Indexing; Domain Analysis.

Introdução

A representação do conhecimento, enquanto área da Ciência da Informação assume, na literatura internacional, duas dimensões que integram e se complementam: descritiva e temática.

Mais particularmente, no âmbito da representação temática, temos a Indexação, que se caracteriza por ser a representação do documento por intermédio de termos representativos do assunto do documento, mediante uma linguagem documentária, para futura elaboração de índices que subsidiarão todo o processo de recuperação da informação.

No entanto, para que essa indexação se faça de maneira a efetivamente estabelecer uma ponte entre a informação produzida e o uso e apropriação dessa informação pelo usuário, necessários se tornam alguns requisitos a serem considerados, requisitos esses que influenciarão diretamente na qualidade do produto final da recuperação. Dentre essas características, importante espaço ocupam a exaustividade, que diz respeito à extensão com que o tema de um documento será traduzido por termos e a especificidade, que diz respeito ao nível de precisão e particularidade que cada um dos termos irá ter, com relação ao seu documento.

Em que pese a importância dessas da exaustividade e da especificidade para os estudos de indexação, ambas são ainda pouco estudadas na literatura da área, sendo por muitas vezes entendidos apenas como forma de mensuração, ou melhor, quantidade de termos a serem atribuídos.

Dessa forma, realizou-se uma pesquisa na base de dados SCOPUS para verificar em que medida e com que concepções esses temas são abordagens pela literatura internacional de Ciência da informação. Para tanto, recuperaram-se todos os artigos que continham os termos *exhaustivity* e *specificity* em suas áreas de maior conteúdo informacional (título, resumos e palavras-chave). A esse corpus, aplicou-se uma análise bibliométrica a partir de suas referências, como subsídio à construção de uma rede social que evidenciasse os referentes teóricos da área.

A exaustividade e a especificidade no contexto da indexação

A representação do conhecimento tem por objetivo resgatar, da forma mais clara e precisa possível, o conteúdo do documento, de forma a que a informação neles contida possa ser recuperada com rapidez e eficiência, correspondendo o máximo possível ao que o usuário da biblioteca está buscando naquele momento.

De acordo com Brascher e Café (2008) a representação do conhecimento é um conjunto de processos que "modelam" o conhecimento, tendo em vista suas futuras representações. Para tanto, utiliza-se de sistemas que organizam o conhecimento e representam as informações valendo-se dos conceitos e relações semânticas dos documentos.

Esse conjunto de etapas pode ser dividido em duas modalidades: a representação descritiva, que se utiliza de processos como a catalogação, de instrumentos como códigos de catalogação e tem por produtos os catálogos, visando a representar o documento a partir do seu formato físico; e a representação temática, valendo-se de processos como classificação ou indexação, de instrumentos como sistemas de classificação e tesouros, e gerando produtos como números de classificação e índices, que se propõe a representar o documento a partir do seu conteúdo e de seus assuntos.

Trazendo o foco para a representação temática, destaca Souza (2013) que esta se dá a partir do processo de análise de assunto e possui como objetivo permitir a compreensão do assunto do documento. Para tanto, é necessário que essa representação esteja abrigada por processos que identifiquem, de maneira condizente, o assunto que aquele documento trata.

Um dos processos da representação temática é a indexação, que tem por objetivo a construção de índices e que possui, na exaustividade e na especificidade, características que permitem sua avaliação.

A história da indexação se inicia na Mesopotâmia, quando os documentos eram guardados em envelopes com fins de preservação e para que estes não fossem quebrados sem necessidade, buscou-se colocar selos dizendo sobre o que os

envelopes diziam. A indexação, da forma como nós conhecemos hoje, se originou através de cabeçalhos de assunto ou sumários, em documentos como, por exemplo, a bíblia (Witty, 1973).

Knight (1968) nos relata que a origem da indexação nos remonta às tabelas de conteúdo que poderiam ser encontradas em manuscritos medievais e que a sua forma de arranjo como utilizamos hoje, se deu início na Europa a partir do século XV.

Para Collins (1972), o principal propósito da elaboração de índices é "(...) colocar ao alcance de qualquer leitor, sem delongas, todas as informações contidas no volume" (Collins, 1972, p. 17) e, ainda, "(...) a indexação, com efeito, não é um processo mecânico: para ter utilidade, requer reflexão e ponderação em todas as fases de seu desenvolvimento" (Collins, 1972, p. 17).

É importante se ter em mente que, um leitor difere do outro e um bom profissional consegue visualizar o assunto de maneira ampla e não tão restrita, como por exemplo, o autor indexaria.

De acordo com outros autores, como Witty (1973), um elemento crucial nos índices é o arranjo de entradas, com uma ordem sistemática e pré-estabelecida. Esse arranjo depende de alguns preceitos, como sistemas de classificação ou, algumas vezes, cronologicamente ou numericamente, embora hoje os sistemas de classificação têm-se demonstrado a maneira mais eficiente de se organizar o conhecimento.

Segundo Law (1970) "o propósito do índice é reorganizar a informação de um livro (...) para trazê-la para fora, da maneira como ela é. A informação deve ser quebrada e "remontada" em partes que compõe uma a outra" (Law, 1970, p. 46). Para o autor, essas partes que compõem devem ser arranjadas de forma a serem consultadas facilmente depois.

Ainda de acordo com o autor, o índice é a forma como o indexador vê o plano de informações dado pelo autor e a forma como essa informação vai ser reapresentada, deve ser a mais conveniente possível, para que a busca seja facilitada (Law, 1970).

Para Souza e Fujita (2014), a indexação possui dois focos principais: um no documento, nos assuntos que são abordados por ele e outro, voltado para as necessidades informacionais que aquele público usuário possui, que pode ser explorada através da formulação de perguntas.

Considerando que a indexação elege conceitos que melhor representem os conteúdos de um documento, Iglesias Maturana (1992) destaca a necessidade de esse trabalho ser realizado da maneira mais correta possível, quando se faz necessário utilizar por alguma *lógica* ou *pauta de seleção* visando a garantir a qualidade de seus resultados. Nesse contexto, inserem-se a Exaustividade e a Especificidade, foco deste artigo.

De acordo com Maturana (1992), a exaustividade na indexação caracteriza-se por ser a capacidade de extensão da abordagem dos diferentes temas do documento, que são identificados pelo indexador na hora do processo de análise e que são, posteriormente, traduzidos para uma linguagem documentária e identificados como termos - pontos de acesso. Ainda de acordo com a autora, "(...) se todos estes termos são indexados, no momento da busca, o documento pode ser recuperado por qualquer destes termos, ou a combinação deles." (Maturana, 1992, p. 83, tradução nossa). Essa capacidade de recuperar o documento nos permite dizer que um alto nível de exaustividade, nos assegura um alto nível de revocação, que se caracteriza por ser "(...) a capacidade do sistema para recuperar documentos relevantes em uma busca, em relação ao total de documentos recuperados. É a capacidade de reter o relevante." (Maturana, 1992, p. 84, tradução nossa).

Para Bravo (2011), é papel da exaustividade medir a quantidade de termos e a eleição daqueles considerados significativos e que contenham informações adequadas para as necessidades informacionais dos usuários. Quando a exaustividade é baixa, automaticamente a recuperação de documentos também vai ser e, portanto, diminuirá a taxa de resposta enquanto aumentará o silêncio.

Um item bibliográfico pode ser composto por uma série de assuntos, e sobre isso, Rueda (2006) ressalta que o "(...) o indexador tem que decidir quantos conceitos deveriam ser representados, mediante cabeçalhos de assunto atribuídos ao

documento. Se se contemplam e incluem todos os documentos, a recuperação será exaustiva” (Rueda, 2006, p. 46, tradução nossa).

No cenário americano, Taylor (2009) nos traz que a exaustividade caracteriza-se por ser o número de conceitos que serão atribuídos na análise de assunto e que essa é uma decisão tomada a partir de uma política de indexação. O autor ainda ressalta que a exaustividade afeta tanto a capacidade de recuperação de documentos relevantes – revocação – como a precisão dos documentos recuperados.

Em outra vertente, temos a especificidade, que se caracteriza por ser a “capacidade da linguagem de descrever de forma específica, precisa, os diferentes temas tratados nos documentos” (Maturana, 1992, p. 86). Podemos dizer, portanto, que afeta diretamente a precisão dos documentos encontrados.

Conforme argumentado por Taylor (2009), a especificidade diz respeito ao nível de precisão semântica dada a determinado vocabulário controlado e suas representações e esse nível deve ser estipulado no momento da criação do vocabulário controlado. Como exemplo de como essa precisão pode ser dada, o autor cita um exemplo: “gatos” não são tão específicos quanto “raças de gatos”, que por sua vez, não é tão específico quanto “gatos siameses”.

Considerada por Given e Olson (2003), como uma estratégia para uma organização da informação efetiva, a especificidade é classificada pelas referidas autoras como intrínseca ao vocabulário controlado, mas também, como um fator de aplicação do vocabulário, na hora de indexação. Para elas, a especificidade diz respeito ao número de nível hierárquico estabelecido e deve satisfazer as necessidades informacionais dos usuários, como também afeta diretamente a precisão – quanto maior o nível de especificidade, maior será a precisão.

Para Rubi (2009), a especificidade está associada ao “(...) nível de abrangência que a biblioteca e a linguagem documentária permitem especificar os conceitos identificados documento” (Rubi, 2009, p. 85). Para a autora, ainda, a especificidade está ligada à capacidade de precisão, ou seja, embora se recupere um número menor de documentos, são exatamente aqueles que o usuário busca.

Diante dessas assertivas, podemos afirmar, portanto, que a exaustividade diz respeito à extensão que será dada na quantidade de termos atribuídos a um documento e está diretamente ligada ao número de documentos que serão recuperados na hora da busca, enquanto que a especificidade diz respeito ao nível de profundidade que será estabelecido na escolha dos termos, diretamente ligada à precisão dos documentos recuperados.

Metodologia

A presente tem por base metodológica a análise de domínio (Hjørland & Albrechtsen, 1995), e, valendo-se das onze abordagens de análise de domínio propostas por Hjørland (2002), insere-se especificamente na confluência das abordagens 3, 5 e 6, relativas, respectivamente, aos estudos de indexação (por se referir à exaustividade e à especificidade), aos estudos bibliométricos (por trabalhar com redes de citação) e aos estudos epistemológicos (por se referir à busca de referentes teóricos), em especial no que se refere às abordagens abordagem dos estudos em indexação e recuperação (Hjørland, 2002)

Para o referido autor, ainda não há uma definição clara da análise de domínio na Ciência da Informação, mas esse conceito insere-se em um “paradigma social, concebendo a Ciência da Informação como uma das ciências sociais e também uma abordagem funcionalista, para entender as funções implícitas e explícitas da informação” (Hjørland & Albrechtsen, 1995, p. 400), buscando dessa forma, analisar não apenas o indivíduo, mas o grupo como um todo.

Para Tennis (2003), o ato de analisar um domínio auxilia na compreensão sobre o objeto daquele domínio e sua principal preocupação é não somente teórica, mas prática, buscando operacionalizar a análise de domínio. Desse modo, o autor propõe dois eixos - Áreas de Modulação e Graus de Especialização.

O primeiro, diz respeito ao nível de extensão com que um domínio é coberto, ou seja, indo diretamente ao conceito de exaustividade na indexação. De acordo com o

autor, as Áreas de modulação possuem dois parâmetros – a extensão e os nomes que são utilizados nesse domínio, podendo nesse último, relacionarmos com os termos definidos pelo processo de indexação. Esses parâmetros podem definir quão vasto um domínio é, o que pode ou não pode ser incluso no domínio e ainda, como esse domínio pode ser denominado (Tennis, 2003).

O segundo eixo, Graus de Especialização, qualifica e dizem quais são as intensões daquele domínio, o que podemos comparar com o conceito de especificidade. Ele também possui dois parâmetros: o foco daquele domínio, aumentando sua intenção e diminuindo sua extensão e ainda, as intersecções, que estabelecem relações com outros domínios e o que se difere nestes domínios.

Portanto, quando nos propomos a fazer uma análise bibliométrica dos autores que escreveram sobre Exaustividade e Especificidade, podemos dizer que estamos a nos comprometer a fazer uma análise sobre o domínio da indexação.

A pesquisa propôs-se a criar uma rede social onde pudessem ser vistas as relações que os autores citados em artigos da área da Ciência da Informação, mais especificamente nas temáticas de Exaustividade e Especificidade na representação temática do conhecimento, possuíam. Para tanto, foi realizada uma busca a partir da base de dados Scopus. A escolha da base de dados dá-se por ser uma das maiores bases de dados internacionais, abrigando periódicos científicos, livros e anais de eventos. A base de dados também é de fácil manuseio e a autora já está familiarizada, devido a pesquisas anteriores.

De forma a apresentar o cenário latino, também se fez necessária a busca em bases de dados latinas, no entanto, a recuperação não demonstrou-se efetiva para os dados da pesquisa. Na base de dados Latindex, a recuperação é apenas de periódicos, não sendo possível buscar artigos separadamente; na base de dados Red, não houve recuperação de nenhum documento e na base de dados Dialnet, os artigos recuperados não correspondiam ao tema da pesquisa.

Dessa maneira, utilizou-se para busca apenas a base de dados Scopus, a partir dos termos exaustividade e especificidade e, dos documentos recuperados, foram selecionados apenas aqueles que além de artigos científicos, fossem publicados em periódicos científicos da área. Não foi estabelecido um período para busca, recuperando assim artigos indexados desde o início na base de dados. A lista dos artigos recuperados, um total de 12 artigos, encontra-se no Apêndice 1.

Nos artigos que continham mais de um autor foi contada uma incidência para cada um deles e foram deletados da contagem, associações e autocitações, uma forma de se manter imparcial e assim, caracterizar da melhor maneira o universo da literatura no tema.

Após essa coleta, o total de artigos citados com as devidas exclusões foi um total de 295 autores e para estabelecer os principais, viu-se necessário aplicar a Lei de Elitismo de Price, que estipula que se x é o número total de contribuintes \sqrt{x} corresponde à elite da área estudada (Price, 1963). Para Alvarado (2009, p. 70):

(...) esta lei estabelece que a raiz quadrada de todos os autores produtores de literatura em determinado campo produzirá, quando menos, a metade de todos os artigos publicados pela população de autores estudados.

Portanto, se 295 é o número total de autores coletados, a raiz quadrada resulta em 17.17, sendo necessário arredondar o número para 17. Foram utilizados para a rede social, deste modo, todos os autores que foram citados pelo menos três vezes nos artigos.

Para a elaboração da rede social, foi utilizado o software Pajek, por ser um *software open source*, sendo, portanto de fácil acesso e ainda, de fácil utilização e ainda, já ter sido utilizada pela autora em seu trabalho de conclusão de curso, causando certa familiaridade.

Análise dos resultados

Os autores identificados na coleta de dados e um breve resumo de seus currículos são apresentados a seguir. Os dados foram coletados a partir da base de dados Scopus, o Google Acadêmico e os sites das instituições.

Michael J. Nelson: o autor tem como instituição sede a Western University, localizada no Canadá. Suas principais áreas de interesse são Ciência da Computação e Ciências Sociais e possui artigos publicados na área de bibliometria e frequência de termos na indexação.

Jean M. Tague: o autor é oriundo do Reino Unido, tendo como instituição a University of Northumbria. Como o autor anterior, suas principais áreas são também a Ciência da Computação e as Ciências Sociais, embora sua área de publicação seja bem diversa.

Cyril W. Cleverdon: o autor também vem do Reino Unido, possuindo como instituição Cranfield Institute of Technology. O principal assunto tratado pelo autor é a recuperação da informação e a indexação.

Jane Fedorowicz: a autora vem de uma instituição americana, a Bentley University. Possui como principais áreas de interesse o compartilhamento de informações entre organizações, *e-government*, controles internos e sistemas empresariais. Suas principais publicações preocupam-se com a informação no ambiente empresarial.

Gerard Salton: o autor vem da universidade americana Cornell University, do departamento de Ciências da Computação e tem suas principais publicações dentro da área de tecnologia da informação.

Karen Sparck-Jones: a autora, também vinda do Reino Unido, possui como instituição University of Cambridge. Também com formação na Ciência da Computação, preocupa-se com as questões de recuperação da informação e automatização das bibliotecas.

Jaana Kekalainen: a autora, de origem institucional finlandesa, é pesquisadora sênior da University of Tampere. Tem como principais áreas de interesse recuperação da informação, informação focalizada e recuperação de informação interativa. Atenta-se, principalmente, com métodos automatizados de recuperação da informação.

Robert R. Korfhage: o autor foi professor emérito da University of Pittsburgh, nos Estados Unidos. Embora sua formação seja em Matemática, suas áreas de atuação dão-se principalmente, em recuperação da informação e tópicos relacionados.

Karen Markey: a autora é professora da University of Michigan, nos Estados Unidos. Suas áreas de atuação são classificação, catalogação e buscas online. Suas publicações mais citadas são sobre *interindexer consistency* e recuperação temática em catálogos.

Micheline Beaulieu: a autora é professora no Reino Unido, na University of Sheffield, no departamento de Ciência da Informação e já atuou como bibliotecária em bibliotecas universitárias. Preocupa-se principalmente com questões de recuperação da informação.

Nicholas J. Belkin: o autor atua, atualmente, na Rutgers University, nos Estados Unidos. Formado na área, dá aula no departamento de Biblioteconomia e Ciência da Informação. Seus principais campos de pesquisa são comportamento informacional e recuperação da informação interativa.

José Marie Griffiths: o autor tem como instituição a Bryant University, localizada também nos Estados Unidos. O autor possui pesquisas na área de comportamento informacional e perspectivas da área.

Kalervo Jarvelin: o autor é professor da University of Tampere, localizada na Finlândia. Tem se preocupado com a recuperação informacional, busca informacional e informetria. Tem publicações focadas em métodos matemáticos e automatizadas, para melhor recuperação informacional.

Michael Keen: O autor possui como instituição a Aberystwyth University, localizada no Reino Unido. Suas pesquisas tem foco na Ciência da Computação, com temas relacionados com sistemas de recuperação informacional e buscas online.

Stephen Robertson: O autor é professor do departamento de Ciência da Computação, na University College London, embora tenha trabalhado por muito tempo, na City University, do departamento de Ciência da Informação. Os campos de

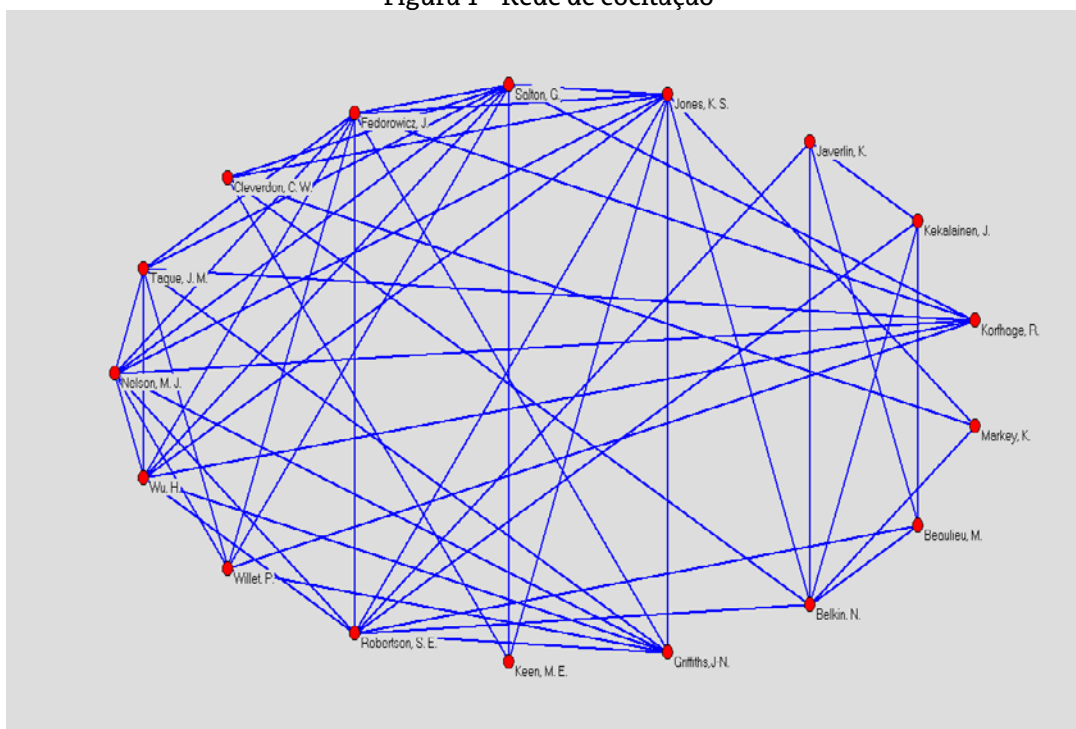
pesquisa que atua, são teorias e modelos para recuperação informacional e o design de sistemas para recuperação da informação.

Peter Willett: o autor também vem da University of Sheffield e dá aula no departamento de Ciência da Informação. Como sua área de origem é a química, seu foco de pesquisa é a aplicação do que ele chama de *chemoinformatics*, embora também possua pesquisas na recuperação da informação e na bibliometria.

Hao Wu: o autor vem da Yunnan University, universidade chinesa, e dá aula de Ciências da Computação. Atua em áreas como processos de informação e computação distribuída. Tem trabalhos principalmente na área de web semântica.

A seguir, apresenta-se a rede de cocitação realizada pela pesquisa e as relações que esses autores possuem entre si, a partir dos artigos analisados.

Figura 1 – Rede de cocitação



Fonte: Elaborado pela autora

Na rede, os círculos vermelhos indicam os autores, enquanto as linhas azuis indicam as relações. Quanto mais linhas, mais relações esse autor possui, estando envolvido assim, de maneira mais clara na temática. Podemos perceber essa forte incidência, nos seguintes autores: Salton, Jones, Robertson, Wu e Nelson, com oito relações cada. É importante destacar ainda, que os autores Kekalainen, Korfhage, Jarvelin e Beauileu, foram citados em apenas um documento.

O que podemos perceber, também, é que grande parte dos autores recuperados pela análise, possuem como área de pesquisa a ciência da computação ou a matemática, com poucos atuando na recuperação da informação e indexação, enquanto áreas da organização do conhecimento.

Devemos nos atentar ao fato da predominância da institucionalidade dos autores, com a grande maioria vinda do Reino Unido e dos Estados Unidos (6 autores cada), com nenhum representante sul-americano na lista. Esse fato pode ocorrer devido às barreiras linguísticas encontradas na tradução dos textos ou ainda, a falta de cultura dos pesquisadores tratarem sobre o tema e ainda, publicarem em revistas indexadas pela base de dados em que a pesquisa foi realizada.

Cumpramos destacar também que os autores Nelson, Tague e Cleverdon, correspondem a elite do grupo, que juntos, obtêm 29% das citações recebidas, embora apenas Nelson esteja no grupo de maiores relações. Essa discrepância pode ser dada

pelo fato da maioria dos artigos, possuírem uma quantidade pequena de referenciados do grupo estabelecido após a aplicação da Lei de Price.

Conclusões

O presente artigo buscou demonstrar quais autores tratam sobre a temática da exaustividade e da especificidade, no contexto dos artigos indexados pela Scopus, por meio de uma análise de domínio, partindo de três das abordagens propostas por Hjørland (2002), nomeadamente a análise dos estudos sobre a indexação, os estudos bibliométricos e os estudos históricos, mais especificamente, no que tange aos estudos espistemológicos, representando essa relação mediante uma rede social de análise de cocitação.

Vê-se que as publicações são centradas em autores oriundos de países como Estados Unidos e Reino Unido, nações de língua inglesa, não tendo nenhum correspondente da América Latina. Constata-se portanto, que a barreira linguística é ainda, muito grande. Pode-se perceber que grande parte dos autores, possui sua base nas Ciências da Computação, acabando por pensar nos processos de exaustividade e especificidade, enquanto processos matemáticos e automatizados, excluindo o fator humano da indexação e o que isso pode acarretar. Dessa forma, vê-se necessária uma maior verticalização de pesquisa no contexto da Ciência da Informação.

O principal objetivo do trabalho, portanto, foi comprovado, demonstrando que a falta de cultura em publicações no âmbito da indexação, pensando nas características da exaustividade e especificidade enquanto tarefas que necessitam do fator humano, considerando a seleção de assuntos mais adequados para maior recuperação e não apenas, a incidência destes, é inerente.

Conclui-se, portanto, que uma maior pesquisa é necessária no tema, por parte de autores que tratem da organização do conhecimento enquanto área principal, de maneira a estabelecer influências teóricas da área nessas questões.

Agradecimento: Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo pelo apoio e financiamento dessa pesquisa, referente ao processo 2015/06278-6.

Referências

Collison, R. L. (1972). Índices e indexação: guia para indexação de livros e coleções de livros, periódicos, partituras musicais, discos, filmes e outros materiais. Tradução por Antônio Agenor Birquet de Lemos. São Paulo: Polígono, 223 p.

Fujita, M. S. L., Rubi, M. P. & Boccato, V. R. C. (2009) O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. DataGramZero - Revista de Ciência da Informação, 10: 2.

Given, L. M. & Olson, H. A. (2003) Knowledge organization in research: A conceptual model for organizing data. Library and Information Science Research, 25, p. 157 – 176.

Hjørland, B. & Albrechtsen, H. (1995) Toward a new horizon in information science: domain analysis. Journal of the American Society for Information Science, 46: 6, p. 400-425.

Hjørland, B. (2002) Domain analysis in information science: eleven approaches – traditional as well as innovative. Journal of Documentation, 58: 4, p. 422-462.

Law, M. D. (1970) Introduction to book indexing. The Indexer, 7: 2, p. 46 – 48.

Knight, G. N. (1968) Book indexing in Great Britain: a brief history. The Indexer, 6: 1, p. 14 – 18.

Maturana, M. T. I. (1992) Analysis y recuperación de información documental. Santiago: Ediciones Instituto Profesional de Santiago, 237 p.

Rueda, L. V. (2006) La indización y el acceso por materias em los catálogos de bibliotecas: el desafio de la recuperación de información. Guadalajara: Universidad de Guadalajara, 224 p.

Sousa, B. P. & Fujita, M. S. L. (2014) Análise de assunto no processo de indexação: um percurso entre teoria e norma. *Infomação e Sociedade*, 24: 1, p. 19-34.

Sousa, B. P. (2013) Representação temática da informação documentária e sua contextualização em biblioteca. *Revista brasileira de biblioteconomia e documentação*, 9: 2, p. 132-146.

Taylor, A.G.; Joudrey, D. N. (2009) *The organization of information*. Londres: Libraries Unlimited, 512 p.

Taylor, A. G. (1999) *The organization of information*. Estados Unidos: Libraries Unlimited, 280 p.

Tennis, J. T. (2003) Two Axes of Domains for Domain Analysis. *Knowledge Organization*, 30: 3/4, p. 191-195.

Witty, F. J. (1973) The beginnings of indexing and abstracting: some notes towards a history of indexing and abstracting in antiquity and the middle ages. *The Indexer*, 8: 4, p. 193 – 198.

Apêndice 1

Ajiferuke, I., Kun, L. & Wolfram, D. (2011) Who are the research disciples of an author? Examining publication recitation and oeuvre citation exhaustivity. *Journal of Informetrics*, 5, p. 293 – 302.

Chen, X. (2008). The influence of existing consistency measures on the relationship between indexing consistence and exhaustivity. En: Arsenault, C.; Tennis, J. T. (ed.). *Cultural and Identity in Knowledge Organization*. Würzburg: ERGON, p. 144 - 150.

Chenglin, W. et al. (2012) A new DSMT combination rule in open frame of discernment and its application. *Science China Information Sciences*, 55: 3, p. 551 – 557.

Given, L. M. & Olson, H. A. (2003) Knowledge organization in research: A conceptual model for organizing data. *Library and Information Science Research*, 25, p. 157 – 176.

Hughes, A. V. (2011) Inter-indexer consistency in graphic materials indexing at the National Library of Wales. *Journal of documentation*, 67: 1, p. 9 – 32.

Jones, K. S. (1972) A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28: 1, p. 11 – 21.

Jones, K. S. (1973) Does indexing exhaustivity matter?. *Journal of the american society for information science*, 24: 5.

Nelson, M. J. & Tague, J. M. (1985) Split size-rank models for the distribution of index terms. *Journal of the american society for information science*, 36: 5.

Schabas, A. (1982) Postcoordinate Retrieval: A Comparison of Two Indexing Languages. *Journal of the American Society for Information Science*, 33: 1, p. 32 – 36.

Vakkari, P. et al. (2004) Query exhaustivity, relevance feedback and search success in automatic and interactive query expansion. *Journal of documentation*, 60: 2, p. 109 – 127.

Wolfram, D. (1996) Inter-Record Linkage Structure in a Hypertext Bibliographic Retrieval System. *Journal of the American Society for Information Science*, 46: 10, p. 765 – 774.

Wolfram, D.; Zhang, J. (2002) An investigation of the influence of indexing exhaustivity and term distributions on a document space. *Journal of the american society for information science and technology*, 53: 11, p. 943 – 952.