



Organização e representação do conhecimento: bibliometria temática em artigos de periódicos brasileiros¹

Cibele Araújo Camargo Marques dos Santos

Resumo: Objetiva identificar a temática da produção científica na área de Organização e Representação do Conhecimento no Brasil utilizando a bibliometria temática que apresenta limites metodológicos mesmo quando realizada em bases de dados bibliográficas, devido à dispersão decorrente da variabilidade de assuntos abordados nas palavras-chaves de um conjunto de artigos de periódicos, além de ocorrerem inconsistências, uso de sinônimos, variações linguísticas, e critérios diferenciados de nivelamento do tratamento da informação pelos autores. Realizou-se estudo exploratório com 150 artigos da BRAPCI sobre Organização da Informação e temas afins, publicados entre 1996 e 2013. Através da técnica de leitura *skimming* foram identificados os assuntos dos artigos, usados como nós temáticos para análise métrica de similaridade. A análise exploratória dos nós identificou 756 itens e a análise de cluster trouxe como resultado um diagrama com nós por similaridade de palavras, similaridade de codificação e mapa de nós comparados por número de itens codificados que permite visualizar os temas estudados. A análise bibliométrica identificou tópicos de pesquisa na área, as correlações temáticas tratadas nos artigos, indicando relações mais ou menos fortes, que podem ser demandas sociais de pesquisa, evolução e tendências informacionais contemporâneas.

Palavras-chave: Análise bibliométrica, Artigos de periódicos, Organização e Representação do conhecimento.

1 INTRODUÇÃO

Este estudo tem por objetivo identificar a temática da produção científica na área de Organização e Representação do Conhecimento no Brasil utilizando a bibliometria para obter parâmetros metodológicos. Esta identificação foi realizada através de estudo exploratório na BRAPCI, base de dados de revistas brasileiras de Ciência da Informação. Segundo o site da BRAPCI (2014), a base amplia o espaço da pesquisa, facilitando a visão de conjunto da produção na área e atende aos objetivos aqui propostos.

¹ Apresentado no XVI Congresso Brasileiro de Biblioteconomia e Documentação em 24 de julho de 2015.



Os recursos bibliométricos permitem medidas de atividade, impacto, colaboração e dinâmica da comunicação de resultados de pesquisa que são representados por indicadores, analisando de forma quantitativa os dados de publicação e citação (SANTIN, BRAMBILLA, STUMPF, 2013).

Porém, a bibliometria temática, que abre a possibilidade para análise qualitativa da produção científica, apresenta limites metodológicos quando realizada em bases de dados bibliográficas, mesmo que estas bases apresentem métricas.

Nos estudos métricos, as abordagens para análise temática da produção científica trazem questões em relação aos métodos mais adequados de representar um domínio disciplinar a partir de uma estruturação temática, no que se refere aos níveis de subjetividade, validade e atualização do modelo de classificação utilizado (LIBERATORE e HERRERO-SOLANA, 2013).

A dispersão decorrente da variabilidade e granularidade de assuntos abordados em um conjunto de artigos de periódicos é grande quando analisamos as palavras-chaves de cada artigo, além de ocorrerem inconsistências, uso de sinônimos, variações linguísticas, bem como critérios diferenciados de nivelamento do tratamento da informação pelos autores, principalmente em áreas interdisciplinares, o que torna importante a discussão sobre a bibliometria temática e gera a necessidade de analisar os procedimentos adotados.

2 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

A visibilidade de uma ciência relaciona-se ao seu vocabulário terminológico e, no caso da Organização e Representação do Conhecimento, a harmonização dos conceitos ainda está sendo realizada. Segundo Lara (2011), o reconhecimento pelos pares da terminologia de uma área caracteriza sua produção técnica e científica e a fixação terminológica leva à visibilidade. Assim, entendemos que as análises bibliométricas temáticas podem contribuir para identificar e fortalecer o vocabulário de uma área.

Os estudos de Organização e Representação do Conhecimento (ORC) apresentam diferentes vertentes teóricas e formas de observar seus problemas e suas práticas. Lara



(2011) enfatiza a necessidade de explicitar termos em sistemas de conceitos coerentes e considera que a dispersão terminológica dificulta a comunicação das propostas existentes.

Esse subcampo tem sua origem na Biblioteconomia, na Documentação e na Ciência da Informação. Deriva da área conhecida em inglês como “*Knowledge Organization*” que de acordo com Hjørland (2008) está entre vários campos da atualidade que desempenham um papel nos ambientes contemporâneos e futuros de comunicação e compartilhamento. Parece ter sido utilizado pela primeira vez no início do século XX por Charles Cutter e Ernest Cushing Richardson (1900) e consolidado por W. C. Sayers Berwick e Henry Bliss. O livro de Bliss publicado em 1929 e intitulado *The organization of knowledge and the system of the sciences* trouxe importante contribuição para a área. Estes autores argumentavam que a classificação de livros e a representação de documentos nas bibliografias devia ser baseada na organização do conhecimento.

A *Knowledge Organization* trata de atividades como descrição de documentos, indexação e classificação realizadas em bibliotecas, bases de dados, arquivos e outras instituições. Estas atividades são realizadas por bibliotecários, arquivistas, indexadores e por profissionais da computação. Dedicar-se, por um lado, à natureza e qualidade dos processos de organização do conhecimento e por outro, aos sistemas de organização do conhecimento usados para organizar documentos, representações e conceitos documentários. Hjørland (2008) considera ainda que o termo tem um significado mais específico quando se refere à representação da informação nas instituições de memória, mas a área também atua em um sentido mais geral quando se refere à dimensão social do conhecimento.

Giunchiglia, Dutta e Maltese (2014), lembram que entre a comunidade da Ciência da Informação, a *Knowledge Organization* desenvolveu diversas soluções bem sucedidas na organização e recuperação da informação, mas que estas soluções, limitadas aos suportes e a pesquisa nas propriedades dos documentos, podem ser aprimoradas pelo uso da inteligência artificial e web semântica baseadas na representação do conhecimento, pelo uso de ferramentas e técnicas como as ontologias.



As diversas abordagens teóricas que compõem a Organização e Representação do Conhecimento foram classificadas segundo Broughton et al (2005) e Hjørland (2008) de acordo com o Quadro 1.

Abordagens teóricas	Definições
Sistemas de classificação usados em bibliotecas, bases de dados e catálogos, incluindo CDD, CDU e outros.	Do ponto de vista concreto, refere-se a documentos organizados nas instituições de memórias, mas pode, de forma mais abstrata ser a base para classificação do conhecimento técnico científico.
Abordagem analítico facetada	Ideias e conceitos. Remove a base empírica da organização de documentos e introduz princípios lógicos e semânticos.
Recuperação da informação	Informação como foco, palavras, co-relações de palavras, relações documento-palavras.
Abordagem orientada ao usuário	Estruturas individuais e cognitivas.
Abordagens bibliométricas	Padrões de documentos e citações entre documentos.
Análise de domínio	Conhecimento é substituído por alegações de conhecimento documentado ou obras. Os conhecimentos que estão organizados não são verdades eternas, mas justificados a partir de uma ou outra perspectiva epistemológica.
Outras abordagens	Abordagens semióticas, abordagens crítico-hermenêutica, abordagens analítico-discursiva, ênfase na representação de documentos, tipologia e descrição de documentos, linguagens de marcação, arquitetura de documentos e outras.

Quadro 1 – Abordagens teóricas da Organização do Conhecimento

Fonte: adaptado de Broughton et al (2005) e Hjørland (2008), tradução nossa.

Dahlberg (2006), analisando de forma breve a lista de conteúdo dos artigos publicados na revista *Knowledge Organization* até 1998, observou que os assuntos tratados estavam de acordo com a proposta do *Classification System for Knowledge Organization Literature*, desenvolvida, 1993, para o índice dos conteúdos publicados em cada volume da revista. Este sistema de classificação apresenta dez classes, sendo elas, de acordo com nossa tradução: Classe 0 Divisões por forma; Classe 1 Fundamentos teóricos e problemas gerais; Classe 2 Sistemas de Classificação e Tesouros, estrutura e construção; Classe 3 Classificação e indexação, métodos; Classe 4 Sobre Sistemas de



Classificação Universal e Tesouros; Classe 5 Sobre Sistemas de Classificação de Objetos Específicos (Taxonomias); Classe 6 Sobre Sistemas de Classificação e Tesouros de Objetos Específicos; Classe 7 Representação do Conhecimento por Linguagem e Terminologia; Classe 8 Classificação e Indexação Aplicada e Classe 9 Ambiente da Organização do Conhecimento.

Estas classes e suas subclasses, que também foram indicadas como grupos de conceitos da área por Fujita (2008), permitem visualizar os temas da pesquisa desenvolvida na área de Organização e Representação do Conhecimento em nível internacional e nacional.

No caso da produção científica brasileira em Organização e Representação do Conhecimento (ORC), são utilizados também outros termos como “Organização da Informação” e “Organização do Conhecimento”, para representar conceitos referentes às práticas de organização da informação, traz ainda esses termos utilizados em conjunto “Organização da Informação do Conhecimento” para representar o mesmo conceito citado. Além disso, na literatura estes termos “Organização da Informação” e “Organização do Conhecimento” aparecem para representar assuntos diferentes, aparecendo “Organização do Conhecimento” para conceitos mais teóricos que aplicados, ou ligados à área da tecnologia da informação.

No Brasil, em 2006, o GT2, um dos grupos de trabalho da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB) que era denominado “Representação do conhecimento, Indexação e Teoria da Classificação” passou a ser “Organização e Representação do Conhecimento”, assim como o capítulo brasileiro da *International Society for Knowledge Organization* (ISKO) optou por utilizar também esta denominação (LARA, 2011; FUJITA, 2008).

Apesar de que nestas associações, o macro termo esteja definido, os conceitos e termos usados nas práticas, pesquisas e produção científicas ainda são diversificados, o que reflete nos procedimentos e resultados dos estudos métricos temáticos em Organização e Representação do Conhecimento.



3 BIBLIOMETRIA TEMÁTICA

Os estudos métricos quando analisam os temas de pesquisa de uma área utilizam procedimentos de classificação por assuntos gerais ou macro temas, bem como por assuntos específicos, com o uso de linguagem natural ou de linguagens documentárias, através da indexação já existente na base de dados de coleta dos dados, e sendo necessário às vezes, reindexação dos dados coletados, de modo que estes estudos normalmente não se referem ao processo de indexação dos assuntos de forma detalhada com a indicação da metodologia e critérios, mostram a lista de termos utilizada, se a mesma pertence a um tesouro ou vocabulário controlado, não discutindo as categorias presentes. Porém, foram identificados alguns artigos que discutiram questões ligadas à linguagem, a padronização terminológica e ao uso de grandes áreas temáticas nos procedimentos bibliométricos (SANTOS, 2015).

Desta forma, considera-se conveniente analisar a influência e contribuição das linguagens e terminologias utilizadas nestes estudos métricos em relação aos procedimentos, métodos e resultados das bibliometrias temáticas. Pecegueiro (2002) em análise temática dos artigos de periódicos brasileiros em Ciência da Informação da década de 90 estudou artigos de seis periódicos, indexando-os com os dez temas principais da área utilizados por Teixeira em sua dissertação de mestrado de 1997, e obteve como assuntos mais publicados: Organização e Gerência de Atividades de Informação; Estudo de Usuário, Transferência, Uso da Informação e Uso da Biblioteca; e Entrada, Tratamento, Armazenamento, Recuperação e DSI. A autora considerou o seu estudo como análise temática, não se referindo ao estudo como bibliometria, convém também observar que seus resultados apontam a ORC, mesmo com terminologia mais geral como um dos temas mais pesquisados.

Em trabalho de revisão da literatura na área de Organização e Representação do Conhecimento, Alvarenga e Silva (2010, p.50), consideram que “as análises bibliométricas podem ser coadjuvantes das análises de conteúdo” e classificaram seu trabalho como análise cientométrica por terem analisado o comportamento da literatura científica no campo. Outro estudo de Gonzalez Alcaide, et al (2008) que realizou análise



bibliométrica e temática, com os artigos da revista *Anales de Documentación*, recorreu à identificação das palavras-chaves mais frequentes e mostrou a rede de relações entre os assuntos nos artigos.

No Brasil, para investigar os principais temas de pesquisa na área de Ciência da Informação na primeira década deste século (2000 a 2009), Liberatore e Herrero-Solana (2013) realizaram análise de conteúdo da produção científica publicada em quatro das principais revistas nacionais, utilizando como metodologia a co-ocorrência de palavras-chave e a representação dos resultados de sua caracterização temática por meio de análise de redes sociais.

Outros trabalhos brasileiros nomeiam de forma diferente suas análises bibliométricas temáticas: mapeamento usado por Autran e Albuquerque (2002) enquanto Kobashi e Santos (2006) denominam “cartografia temática e de redes sociais por meio de técnicas bibliométricas”.

Embora, não haja consenso quanto à terminologia para este tipo de pesquisa, a bibliometria temática permite analisar e mapear a produção científica de uma área do conhecimento, trazendo indicadores temáticos que possam qualificar a pesquisa e mostrar pontos fortes e oportunidades de investigação para desenvolvimento de um campo.

4 METODOLOGIA

Foi realizado um estudo exploratório com análise bibliométrica temática, a partir da captura de tela do resultado da pesquisa com 150 artigos na BRAPCI sobre Organização da Informação, Organização do Conhecimento e temas afins, publicados entre 1996 e 2013. Foi necessário utilizar o recurso de captura de tela para coleta de dados, já que a BRAPCI, não possui opção para exportação de dados. Estes temas foram utilizados na estratégia de busca realizada na base em maio de 2014.

Através da técnica de leitura *skimming* (rápida passada com foco em informações salientes do texto) identificou-se uma lista de termos, encontrados nas palavras-chaves e resumos, representativos dos assuntos dos artigos. Estes termos foram usados como nós temáticos para análise métrica de similaridade, que é um método estatístico utilizado para



o cálculo de correlação entre os itens. Foi utilizado o programa NVivo, da QSR International para pesquisa qualitativa e análise de dados estruturados e não estruturados. Embora a BRAPCI apresente em sua nuvem de tags as palavras-chaves dos artigos dos resultados de busca, o *skimming* dos artigos permitiu controlar a ocorrência dos assuntos e diminuir a dispersão.

5 RESULTADOS

Foram identificados e utilizados 32 termos como nós temáticos no programa NVivo para analisar as telas capturadas da BRAPCI. Estes termos, localizados nas palavras-chaves e resumos dos artigos e representados pela linguagem natural, mostraram-se pertinentes à temática da área de Organização e Representação do Conhecimento apontadas tanto nas abordagens teóricas do Quadro 1 deste artigo, como das classes desenvolvidas por Dahlberg (2006), pertencentes ao *Classification System for Knowledge Organization Literature*.

A análise exploratória dos nós identificou 756 itens e a análise de cluster trouxe como resultado um diagrama com nós por similaridade de palavras (Figura 1), onde os nós que apresentam maior grau de similaridade com base na ocorrência e a frequência de palavras nos artigos são mostrados agrupados e aqueles com menor grau de semelhança na ocorrência e frequência de palavras são exibidos mais afastados.

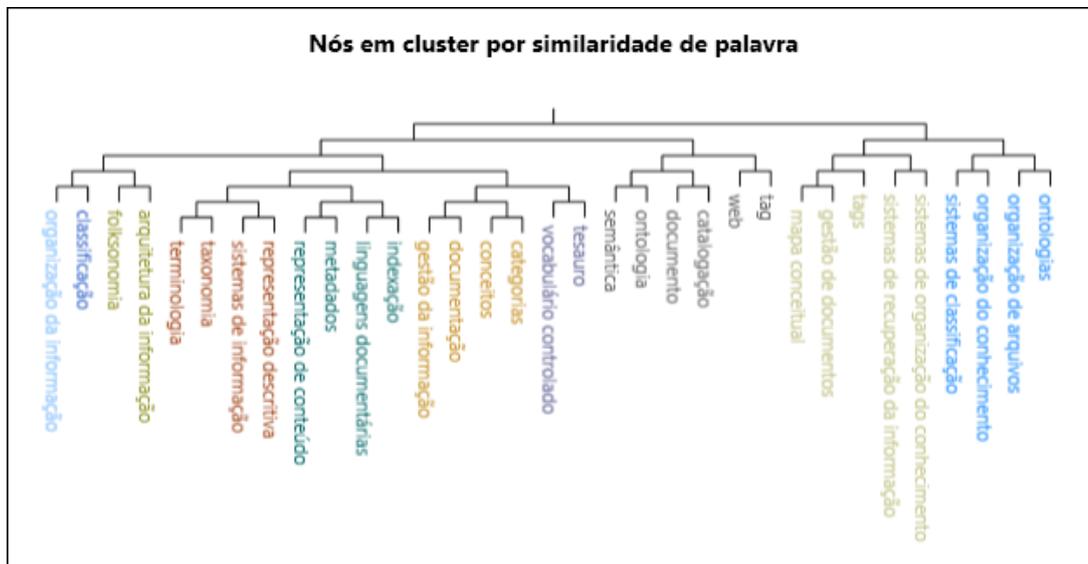


Figura 1 – Diagrama de nós por similaridade de palavra
 Fonte: gerado no NVivo pela autora

A similaridade semântica dos termos também pode ser observada nos clusters, com algumas duplas com grande proximidade de significado, apresentando relacionamento de quase sinônimos como “tesauro e vocabulário controlado” ou relacionamentos associativos do tipo disciplina ou campo e objeto como “organização do conhecimento e sistemas de classificação”, “terminologia e taxonomia”, “arquitetura da informação e folksonomia”; relacionamento de ação e objetivo como no caso de “sistemas de organização do conhecimento” e “sistemas de recuperação da informação”, relacionamento de ação e produto ou objeto como “tag e web”, “catalogação e documento”, “indexação e linguagens documentárias”, “representação descritiva e sistemas de informação”, “metadados e representação de conteúdo”, entre outros.

Outra análise de cluster realizada foi similaridade de codificação, onde os nós que foram codificados de forma similar foram agrupados no diagrama, se codificados de forma diferente apareceram mais distantes no diagrama. As diferenças de relações entre os temas nos dois diagramas foram pequenas, apenas ficando mais visível as relações de proximidade ou distância semântica entre os grupos de quatro ou mais termos, como por exemplo, a proximidade entre os termos “sistemas de organização do conhecimento, sistemas de recuperação da informação, tags, gestão de documentos e mapa conceitual”.

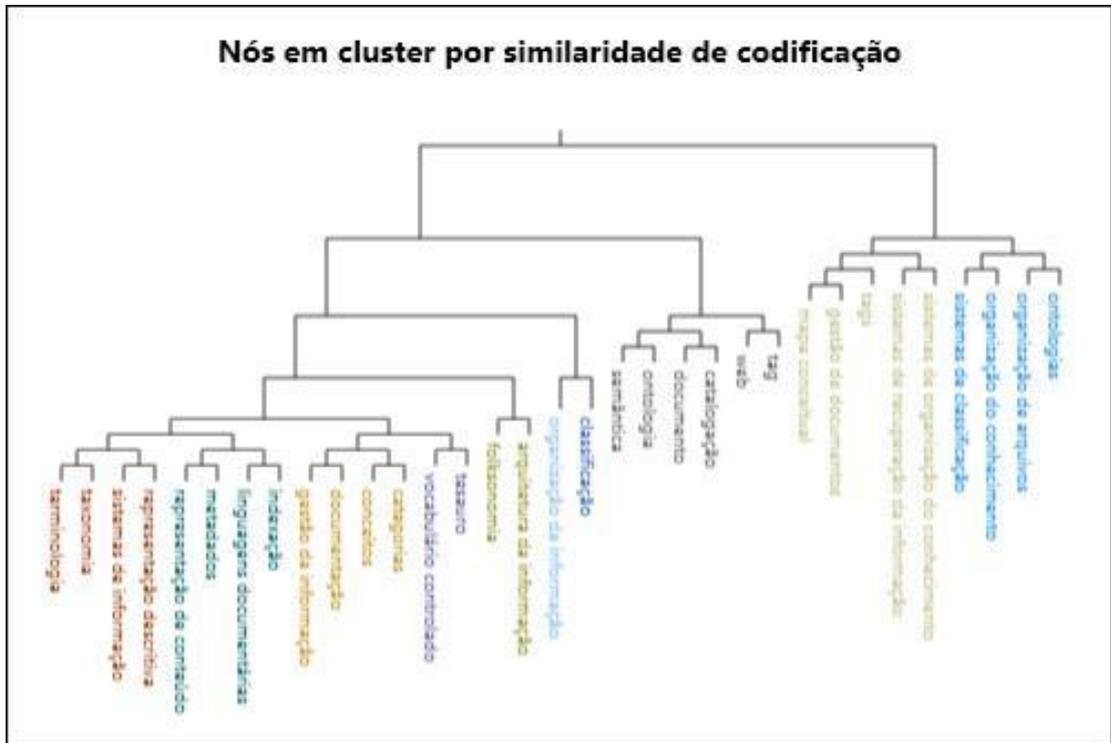


Figura 2 – Diagrama de nós por similaridade de codificação
 Fonte: gerado no NVivo pela autora

Na Figura 3 é apresentado o mapa de nós comparados por número de itens codificados que permite identificar e comparar a quantidade de referências por nó e visualizar os temas de destaque e os temas com menos ocorrência.

Nós comparados por número de itens codificados

organização da informação	arquitetura da informação	taxonomia	indexação	representação descritiva	terminologia	metadados	ontologia	catalogação	tag	ontologias
classificação	folksonomia	linguagens documentais	representação de conceitos	categorias	vocabulário controlado	tesouro	documento	tags	sistemas de classificação	sistemas de organização de conhecimento
		sistemas de informação	conceitos	documentação	gestão da informação		web	mapa conceitual	organização de	gestão de

Figura 3 – Nós comparados por número de itens codificados
 Fonte: gerado pelo NVivo pela autora



Os termos com maior ocorrência foram “organização da informação e classificação”, seguidos de “arquitetura da informação e folksonomia” sendo que a quantidade de termos pode ser visualizada pelo tamanho e cor da ocorrência.

Nos diagramas gerados, os tópicos que se mantiveram próximos apresentaram correlação lógico-semântica e de contexto de investigação, formaram duplas temáticas ou grupos temáticos relacionados, reforçando a importância da análise temática de uma área para melhor conhecimento de suas especificidades como indicado por Santin, Brambilla e Stumpf (2013), o que pode ser realizado pela bibliometria temática para o mapeamento de um campo de pesquisa.

6 CONSIDERAÇÕES FINAIS

A análise bibliométrica temática identificou tópicos de pesquisa na área de Organização e Representação do Conhecimento, de acordo com as abordagens teóricas e classificação do campo, a partir da indicação das correlações temáticas tratadas nos artigos pesquisados, indicando relações mais ou menos fortes, podendo ser indicativo de demandas sociais de pesquisa, evolução e tendências informacionais contemporâneas.

A metodologia utilizada mostrou-se adequada para o mapeamento temático, com potencial para continuidade, novas análises e aplicações.

A identificação do tipo de relações lógico-semânticas entre os termos pode ser útil para o desenvolvimento de classificações para uma área de conhecimento, bem como para o desenvolvimento de vocabulários controlados para a organização da informação, permitindo visualizar as relações existentes, as ocorrências de termos e fornecendo a garantia literária da produção científica para a construção de uma linguagem documentária.



Knowledge Organization and Representation: thematic bibliometrics in Brazilian journals articles

Abstract: It's aim to identify the scientific production in Knowledge Organization and Representation area in Brazil using thematic bibliometrics that facing methodological limits even when held in bibliographic databases, due to the dispersion arising from the variability in the subjects addressed in key words of f journal articles, and inconsistencies occur, in the use of synonyms, language variations, and different criteria leveling the information processing by the authors. Held exploratory study with 150 articles of BRAPCI on Information Organization and related issues published between 1996 and 2013. Through the skimming reading technique the subjects of articles were identified, and used as theme for metric similarity analysis. The exploratory analysis identified 756 items and cluster analysis brought as a result a diagram for similarity of words, coding similarity and map that was compared for a number of coded items that lets to view the topics studied. A bibliometric analysis identified research topics in the area, thematic correlations addressed in the articles, indicating more or less strong relationships, which can be social demands research, evolution and contemporary informational trends.

Keywords: Bibliometric analysis, Journal articles, Knowledge organization and representation.

REFERÊNCIAS

ALVARENGA, L., SILVA, D. L. da. Organização e Representação do Conhecimento na Ciência da Informação: revisão de literatura. **Pesquisa Brasileira em Ciência da Informação**, v.3, n.1, p.47-84, jan./dez. 2010.

AUTRAN, M. de M. M., ALBUQUERQUE, M. E. B. C. de. Mapeamento do periódico Informação & Sociedade, estudos: dez anos de trajetória. **Informação e Sociedade – Estudos**, v.12, n.1, 2002.

BRAPCI. **Bases de dados referenciais em Ciência da Informação**. Curitiba, Universidade Federal do Paraná, 2014. Disponível em: <<http://www.brapci.ufpr.br/>>. Acesso em 26 maio 2014.

BROUGHTON, V., HANSSON, J., HJØRLAND, B.; LÓPEZ-HUERTAS, M. J. Knowledge organisation: Report of working group 7. In: KAJBERG, L., LØRRING, L., eds. **European curriculum reflections on education in library and information science**, Copenhagen, Royal School of Library and Information Science, 2005. p.



Disponível em: <http://www.iva.dk/bh/Lifeboat_KO/knowledgeorg_chapter%207.pdf>. Acesso em 30 de novembro de 2015.

DAHLBERG, I. Knowledge organization: a new Science. **Knowledge Organization**, v.33, n. 1, 2006.

FUJITA, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do Enancib no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.1, n.1, 2008.

GIUNCHIGLIA, F., DUTTA, B., MALTESE, V. From Knowledge Organization to Knowledge Representation. **Knowledge Organization**, v. 41, n. 1, 2014. p.44-56.

HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, v.35, n. 2/3, 2008. Disponível em: <http://is.muni.cz/el/1421/podzim2008/VIKMA13/um/6341865/Hjorland_-_What_is_Knowledge_Organization.pdf>. Acesso em 30 de novembro de 2015.

GONZALEZ ALCAIDE, G., ALONSO ARROYO, A., VALDERRAMA ZURIÁN., J. C., BENAVENT, R. A. Uma década de investigaciones en Anales de Documentación (1998-2007): aproximación bibliométrica e temática. **Anales de Documentación**, n.11, 2008.

KOBASHI, N. Y., SANTOS, R. N. M dos. Institucionalização da pesquisa científica no Brasil: cartografia temática e de redes sociais por meio de técnicas bibliométricas. **Transinformação**, v. 18, n.1, p.27-36, jan./abr. de 2006.

LARA, M. L. G. de. Conceitos de organização e representação do conhecimento na ótica das reflexões do Grupo Temma. **Informação & Informação**, Londrina, v.16, n. esp. p.92-121, jan./jun. 2011.

LIBERATORE, G.; HERRERO-SOLANA, V. Caracterización temática de la investigación em Ciência de la Información en Brasil en el período 2000-2009. **Transinformação**, Campinas, v.25, n.3, p. 225-235, set./dez. 2013.

PECEGUEIRO, C. M. P. de A. Temática dos artigos de periódicos brasileiros na área da Ciência da Informação na década de 90. **Transinformação**, v.14, n.2, p.117-131, jul./dez. 2002.

SANTIN, D. M.; BRAMBILLA, S. D. S.; STUMPF, I. R. C. Produção científica em Neurociências da UFRGS indexada na Web of Science: 2000-2009. **Liinc em Revista**, Rio de Janeiro, v. 9, n. 1, p. 66-84, maio 2013.

SANTOS, C. A. C. M. dos. Organização e representação do conhecimento: contribuições aos estudos métricos. In: GUIMARÃES, J. A. C; DODEBEI, V., orgs. **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil: FUNDEPE, 2015. 835 p. (Série: Estudos Avançados em Organização e Representação do Conhecimento, v. 3). Disponível em: <<http://isko-brasil.org.br/wp->



content/uploads/2015/09/Organiza%C3%A7%C3%A3o-do-Conhecimento-e-Diversidade-Cultural-ISKO-BRASIL-2015.pdf>. Acesso em: 30 de novembro 2015.

Agência financiadora

FAPESP Projeto 2013 04109-7

Informações do autor

Cibele Araújo Camargo Marques dos Santos

Profa. do Departamento de Informação e Cultura da Escola de Comunicações e Artes da Universidade de São Paulo.

E.mail: cibeleac@usp.br

