

EIXO TEMÁTICO:
Organização e Representação da Informação e do Conhecimento

O PROCESSO DE ANÁLISE DE DESCRITORES EM PERIÓDICOS CIENTÍFICOS ELETRÔNICOS DA ÁREA DE CIÊNCIA DA INFORMAÇÃO

THE ANALYSIS OF DESCRIPTORS FROM SCIENTIFIC ELECTRONIC JOURNALS OF INFORMATION SCIENCE

Jaqueline Ribeiro Rocha - jackeflorrosa@gmail.com
Marinara Ferreira Santos - marinara.ferreira.santos@gmail.com
Brígida Maria Nogueira Cervantes - brigidacervantes@gmail.com
Ana Esmeralda Carelli - carelliana@gmail.com

Resumo: Representar apropriadamente as informações contidas nos documentos é condição fundamental para a recuperação eficaz destas informações e o processo de análise de assunto é essencial para o êxito dessa atividade. Neste sentido, mensurar os descritores que caracterizam o conhecimento científico é de suma importância aos periódicos científicos eletrônicos. Neste trabalho investigou-se o processo do uso de descritores em artigos dos periódicos científicos eletrônicos *JASIST* e *Scientometrics*, do ano de 2013, para a organização e recuperação temática da informação. Como percurso metodológico, utilizou-se o estudo exploratório-descritivo, com abordagem quanti-qualitativa e os dados coletados registrados em base de dados desenvolvida a partir do MS Access. Os principais resultados obtidos demonstraram que a quantidade de autores responsáveis pela elaboração do artigo, não influenciou no número de descritores encontrados nos artigos, uma vez que, diversos artigos elaborados por vários autores não apresentavam a atribuição de descritores. Enquanto que, outros artigos com uma única autoria, possuíam diversos descritores. Concluiu-se que os periódicos estudados são relevantes e possuem influência para a área da Ciência da Informação, contudo observou-se ausência de critérios e rigor no cumprimento das diretrizes propostas para publicação. Além de aventar negligência dos próprios autores dos artigos, quanto a indicação de descritores, fator este que implica diretamente na recuperação e acesso aos conteúdos temáticos publicados.

Palavras-chave: Análise de descritores. Periódicos Científicos Eletrônicos. Tratamento Temático da Informação. Ciência da Informação.

Abstract: An accurate representation of documents is fundamental for the effectiveness of information retrieval, being the subject analysis process a key step for it. In this sense, measuring the descriptors that represent scientific knowledge is quintessential to electronic journals. This paper analyzes the use of descriptors for the organization and retrieval of academic papers from 2013 published in two scientific journals: *JASIST* and *Scientometrics*. Following an exploratory-descriptive study with a mixed-method approach, data collected were recorded in a database developed in MS Access. Main findings showed that the amount of authors did not influence the number of descriptors found in articles, since several articles written by various authors did not present the allocation descriptors. Meanwhile, other articles with a single author had several descriptors. The study concludes that the two journals under study are relevant for the Information Science field; however, results noted the absence of criteria and rigor especially with regard to the compliance of the journals'

guidelines. Results suggest some negligence from authors with respect to their articles descriptors, which directly affects the retrieval and access of thematic content.

Keywords: Descriptor Analysis. Electronic Scientific Journals. Information Representation. Information Science

1 INTRODUÇÃO

A temática que envolve a organização dos documentos eletrônicos, em especial a organização dos periódicos científicos eletrônicos, constitui-se como elemento específico da Organização da Informação, a qual analisa e descreve as relações que ocorrem dentro do conhecimento registrado e socializado, bem como na sua organização no espaço da *Web*, na recuperação e acesso.

A organização da informação foi, desde sempre, uma preocupação da humanidade. Na medida em que os volumes de informações têm-se multiplicado constantemente, as exigências relativas à sua organização e acessibilidade crescem nas mesmas proporções. Em face da relevância e a confiabilidade do periódico científico para as Ciências, e para Ciência da Informação, em específico, torna-se necessário sua avaliação para mensurar a adequação dos metadados que caracterizam o acesso ao conhecimento científico, pois possui importância para o tratamento e a recuperação da informação.

Em geral, os autores são os que atribuem às palavras-chave de seus artigos, no entanto, nem sempre as estabelecem com o intuito de maximizar o potencial deste elemento para recuperação da informação. Corroborando com o exposto, Gil-Leiva e Alonso-Arroyo (2005, tradução nossa) enfatizam a importância de estudar as palavras-chave proporcionadas por autores em seus artigos científicos e os descritores atribuídos a esses artigos, em bases de dados para, averiguar o papel que exercem as palavras-chave no processo de indexação desses artigos. Assim, a questão de investigação que se coloca é: Os periódicos científicos apresentam palavras-chave que facilitam a recuperação dos conteúdos temáticos, por serem pertinentes e relevantes?

Como a proponente deste projeto está vinculada ao Departamento de Ciência da Informação e ao curso de Pós-Graduação *Stricto Sensu* em Ciência da Informação da Universidade Estadual de Londrina, bem como, vinculada a linha de pesquisa “Organização da Informação e do conhecimento”, a realização deste estudo, contribuiu para: ampliar e fortalecer esta linha de pesquisa.

2 PERIÓDICOS CIENTÍFICOS ELETRÔNICOS

Desde o surgimento das publicações eletrônicas, a literatura especializada e as discussões no âmbito da Biblioteconomia e Ciência da Informação (BCI) vêm estudando a organização e a representação da informação no ambiente da Web e destacam que os principais problemas relativos ao uso desses recursos informacionais eletrônicos estão intrinsecamente relacionados à representação temática da informação, ocasionando a recuperação ineficaz do usuário no momento da consulta do documento. Representar efetivamente os conteúdos dos textos e, conseqüentemente, auxiliar o usuário na sua busca por conteúdos significativos sempre foi um desafio para os profissionais da informação.

A publicação científica tornou-se um instrumento imprescindível para criação, promoção, organização e fortalecimento do ciclo e difusão do conhecimento. Podemos considerar que sua contribuição social é um dos fatores que mais influencia o ritmo de produção do conhecimento. Este fenômeno ocorreu devido à necessidade percebida pelos próprios pesquisadores sobre as fontes de informações disponíveis para o domínio da literatura de sua área, e dos meios existentes para difusão dos resultados de suas próprias pesquisas. Nessa trajetória, as tecnologias da informação trouxeram um novo desafio para as áreas de estudo da Ciência da Informação (CI), tais como a da Comunicação Científica, da Organização da Informação, envolvendo a mudança de hábitos e práticas em processos de produção, organização e disseminação do conhecimento nos meios digitais (BUFREM et al., 2007).

O periódico ocupa espaço primordial na comunicação científica, ao longo do tempo, tornou-se um dos maiores veículos de disseminação do conhecimento da Ciência. Nesse contexto, os periódicos científicos eletrônicos merecem uma atenção especial em relação ao meio em que são disponibilizados e a forma de acesso a essas informações, que são extremamente importantes. Critérios claros e consistentes no processo de organização e no tratamento dos conteúdos desses documentos são extremamente importantes para recuperar as informações. A transmissão da informação é condição essencial para a produção do conhecimento e para eficácia dessa conversão, um dos principais requisitos é que as informações sejam reunidas, armazenadas e representadas de acordo com uma organização que possibilite efetivamente a sua recuperação.

Os periódicos científicos eletrônicos são geralmente indexados utilizando-se de uma breve descrição, tais como resumos e índices, na tentativa de disponibilizar esse conteúdo informacional para sua comunidade usuária. Na área da CI, os textos são representados em resumos e índices (autor, título e assunto). A representação dos resumos é composta por paráfrases, indicando a proximidade e conformidade com o texto original. A representação dos índices - aqui se referindo aos índices de assuntos - se realiza por tradução, formulados por meio de linguagem que não se confunde com a linguagem do texto, ainda que apresentem a mesma forma. Os índices podem ser constituídos por meio de palavras-chave, notações classificatórias ou símbolos de um sistema de classificação, cabeçalhos de assunto ou unidades conceituais, como a de um tesouro (PEREIRA; BUFREM, 2005).

Nessa perspectiva, a adoção da linguagem e os instrumentos documentários que irão representar as palavras-chave e os descritores atribuídos aos conteúdos dos periódicos são fundamentais para o êxito desse processo. Guedes e Dias (2010) explicitam que a representação do conhecimento é uma manifestação do pensamento, um meio de expressão e uma linguagem na qual se pode dizer coisas sobre a realidade e o mundo.

As definições de palavra-chave e descritor estão conectadas a essa ideia. Cunha e Cavalcanti (2008, p. 274) definem a palavra-chave como uma palavra “[...] significativa encontrada no título de um documento, no resumo ou no texto. Essa palavra (ou grupo de palavras) caracteriza o conteúdo temático do item e é usada em catálogos e índices de assuntos.” Por sua vez, descritor é definido como o “Elemento de uma linguagem documentária, que pode ser empregado para representar um texto em sistema de informação. Traduz os conceitos (os assuntos) contidos no texto.” (CUNHA; CAVALCANTI, 2008, p. 119).

Representar apropriadamente as informações contidas nos documentos é condição fundamental para a recuperação eficaz das informações e as palavras-chave atribuídas pelos autores são elementos essenciais para o êxito dessa atividade. Para atribuir conceitos aos documentos é necessário o uso de uma linguagem que represente com fidedignidade o conteúdo informacional registrado no suporte, representando-o para fins de recuperação e acesso. Para Naves (2001, p. 189), “O aumento da produção científica aliado à crescente interdisciplinaridade entre as áreas do conhecimento, que vem ocorrendo nas últimas décadas, tem tornado o trabalho de organização e tratamento da informação cada vez mais árduo

e complexo.” Pode-se considerar que a recuperação eficaz das informações depende, exclusivamente, da organização e da representação consistente de tais conteúdos, bem como da sua preservação e da perspectiva de seu uso.

Nesse contexto, os processos de análise, síntese e interpretação são extremamente importantes para o desempenho deste processo. Dessa forma, torna-se necessário adotar regras de padronização para a organização e representação das informações, estabelecendo quais os tipos de instrumentos documentários são potencialmente facilitadores para o acesso eficaz dessas informações, sem esquecer, das necessidades e realidades contextuais do usuário. Portanto, a palavra-chave é um metadado presente tanto nos artigos, como nas bases de dados, com a função de facilitar o acesso à informação. Tendo em vista que este é o objeto deste estudo, averiguar os descritores como elemento representativo do conteúdo dos artigos de periódicos é imprescindível.

Para o desenvolvimento deste estudo exploratório-descritivo, com abordagem quanti-qualitativa, foram escolhidos dois periódicos internacionais: *Journal of the American Society for Information Science and Technology (JASIST)*, e *Scientometrics*. O primeiro periódico citado aborda as áreas de Ciência da Informação e Tecnologia, em que possui publicações desde 1950, atualmente as submissões podem ser efetuadas *online*, e sua periodicidade é mensal. Já o segundo periódico, da área de Cientometria, possui seu enfoque nos fatores quantitativos, da qual se caracteriza pelos estudos matemáticos estatísticos, tendo periodicidade trimestral.

Para efeito deste estudo, foi considerado o ano de 2013 para a coleta de artigos publicados no período, na intenção de averiguar o estado atual da atribuição de descritores, nos periódicos escolhidos para o estudo. Além de 2013 ser compatível com a finalização do projeto de pesquisa, em questão. Como critério de coleta de dados nos periódicos selecionados estabeleceu-se que: - no periódico *Scientometrics*, seriam coletadas apenas as palavras-chave atribuídas aos artigos da seção: *Original paper*. Já do periódico *JASIST*, foram extraídas as palavras-chaves atribuídas aos artigos da seção: *Research Article*.

A base de dados, desenvolvida a partir do *software Access*, foi nomeada como JAMAR, dispõe dos seguintes metadados para preenchimento de dados: Código (gerado automaticamente para cada artigo); referência do artigo (realizada no ato da coleta); título; autores; número de autores; resumo; tipologia de resumo

(com opção de seleção de indicativo ou informativo); palavras-chave; número de palavras-chave; periódico (com opção de seleção de *Scientometrics* ou *JASIST*). Com a coleta destas informações, tornou-se possível obter resultados pertinentes e relevantes, descritos no próximo tópico.

3 RESULTADOS

No total, obteve-se 384 artigos. Destes 216 são da *Scientometrics* e 168 pertencem a *JASIST*. Num primeiro momento, diante de tantos artigos e tendo em vista que há autores com mais de uma publicação em ambos os periódicos, buscou-se averiguar se havia padronização de indicação de autoria.

No periódico *Scientometrics*, pareceu-nos não haver rigor na padronização de indicação de autoria, ao perceber que um mesmo autor, em dois artigos distintos, encontrava-se descrito com grafias diferentes, sendo um artigo com o seu prenome abreviado, enquanto que no outro artigo o nome estava descrito de modo completo. Em seguida, no quadro 1, expõe-se a menção de autoria, como um exemplo.

Quadro 1 – Indicação de autoria Leo Egghe.

CÓDIGO NA BASE	AUTORES	TÍTULO	PALAVRAS-CHAVE	Nº DE PALAVRAS-CHAVE	REVISTA
8	<i>L. Egghe</i>	Theoretical justification of the central area indices and the central interval indices	Central area index. Central interval index. h-index. Hirsch index.	4	<i>Scientometrics</i>
11	<i>Leo Egghe, Raf Guns, Ronald Rousseau</i>	Measuring co-authors' contribution to an article's visibility	Visibility indicators. Power law model h-index. Citation analysis. Co-authorship.	4	<i>Scientometrics</i>

102	<i>L. Egghe</i>	A rationale for the relation between the citer h-index and the classical h-index of a researcher	Citer. h-index. Hirsch index.	3	<i>Scientometrics</i>
231	<i>L. Egghe</i>	Theory of the topical coverage of multiple databases	Informetrics.	1	<i>JASIST</i>

Fonte: Elaborado pelas autoras com base no banco de dados JAMAR.

Já no periódico *JASIST* foi possível verificar que havia padronização, pois os autores com mais de um artigo publicado na revista, são descritos da mesma forma. O quadro 2 demonstra a menção de autoria, como um exemplo.

Quadro 2 – Indicação de autoria Loet Leydesdorff

CÓDIGO NA BASE	AUTORES	TÍTULO	PALAVRAS-CHAVE	Nº DE PALAVRAS-CHAVE	REVISTA
3	<i>Loet Leydesdorff</i>	An evaluation of impacts in “Nanoscience & nanotechnology”: steps towards standards for citation analysis	Citation Impact. Evaluation. Nanotechnology. Statistics. Standards.	6	<i>Scientometrics</i>
16	Staša Milojević, <i>Loet Leydesdorff</i>	Information metrics (iMetrics): a research specialty with a socio-cognitive identity?	Scientometrics. Bibliometrics. Informetrics.	3	<i>Scientometrics</i>
66	<i>Loet Leydesdorff</i> , Stephen Carley, Ismael Rafols	Global maps of science based on the new Web-of-Science categories	Map. Overlay. Web of Science. Subject categories. Interdisciplinarity.	5	<i>Scientometrics</i>
229	<i>Loet Leydesdorff</i> , Ping Zhou, Lutz Bornmann	How can journal impact factors be normalized across fields of science? An assessment in terms of percentile ranks and fractional counts	Journals. Citation indexes. Information science.	3	<i>JASIST</i>

307	Loet Leydesdorff, Øivind Strand	The Swedish system of innovation: Regional synergies in a knowledge-based economy	Innovation.	1	JASIST
333	Loet Leydesdorff, Filippo Radicchi, Lutz Bornmann, Claudio Castellano, Wouter de Nooy	Field-normalized impact factors (IFs): A comparison of rescaling and fractionally counted I _{fs}	Evaluation. Journals.	2	JASIST
335	Fred Y. Ye, Susan S. Yu, Loet Leydesdorff.	The Triple Helix of university-industry-government relations at the country level and its dynamic evolution under the pressures of globalization	Measurement.	1	JASIST
353	Loet Leydesdorff, Ismael Rafols, Chaomei Chen	Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal-journal citations	Knowledge representation.	1	JASIST

Fonte: Elaborado pelas autoras com base no banco de dados JAMAR.

Na *Scientometrics*, mesmo com regras previamente definidas quanto a quantidade de palavras-chave a serem atribuídas, a maior quantidade registrada foi 12, sendo que o disposto nas regras de submissões de trabalho seja entre 4 a 6 palavras-chave. Enquanto na *JASIST*, não há qualquer delimitação quanto a quantidade de palavras-chave atribuídas, observou-se em média o número de 5 palavras-chave.

O número de autores por artigos, não demonstrou discrepância entre os periódicos, sendo de 12 autores o maior número registrado no *JASIST*, já no *Scientometrics* o número foi de 11 autores como coautores em um artigo.

Contudo, buscou-se verificar se o número de autores influencia na quantidade de palavras-chave atribuídas. Para esta análise elaborou-se um quadro para cada periódico. Em que o campo X equivale ao número de autores, o Y representa o

número de palavras-chave, e o campo em branco indica a quantidade de artigos. Sendo demonstrado ainda o total de artigos que utilizou Y quantia de palavras-chave. O quadro 3 apresenta-se o Scientometrics, primeiro periódico analisado.

Quadro 3 – Número de autores vs número de palavras-chave - *Scientometrics*.

X	1	2	3	4	5	6	7	8	9	10	TOTAL
Y											
0	1	1	-	1	-	-	-	-	-	-	3
1	-	-	-	-	-	-	-	-	-	-	0
2	2	-	-	-	-	-	-	-	-	-	2
3	3	7	4	-	2	-	-	-	-	-	16
4	14	13	12	12	3	1	1	-	-	-	56
5	10	30	14	9	6	3	-	-	1	1	74
6	6	12	16	6	1	-	1	-	-	-	42
7	4	2	3	-	2	-	-	-	-	-	11
8	1	1	3	-	-	-	-	-	-	-	5
9	-	1	1	-	2	-	-	-	-	-	4
10	-	-	1	-	-	-	-	-	-	-	1
11	1	-	-	-	-	-	-	-	-	-	1
12	-	-	1	-	-	-	-	-	-	-	1
											216

Fonte: Elaborado pelas autoras com base no banco de dados JAMAR.

Com base no quadro 3, observou-se que em três artigos não houve a atribuição de palavras-chave, variando estes artigos entre 1 e 4 autores. Não há registros de artigos com apenas uma palavras-chave. A partir da quantidade número 2 é que as palavras-chave aparecem, da qual são dois artigos, ambos de autorias individuais.

Conforme citado, anteriormente, o número de palavras-chave para a

Scientometrics situou-se entre 4 a 6. Vale salientar que o maior número quantitativo registrado de indicação de palavras-chave está entre esta recomendação do periódico. Ficando 4 palavras-chave com total de 56 artigos, variando estes entre 1 a 7 autores; 5 palavras-chave com total de 74 artigos variando entre 1 a 10 autores; e 6 palavras-chave com total de 42 artigos, com variação de 1 a 7 autores.

O periódico *JASIST*, até o período da pesquisa, não estabelecia regras explícitas quanto ao número de palavras-chave recomendadas na submissão de artigos para a publicação. Portanto, de acordo com o quadro abaixo, as variações foram bem distintas em comparação aos do periódico, anteriormente, analisado.

Quadro 3 – Número de autores vs número de palavras-chave *JASIST*.

X	1	2	3	4	5	6	7	8	9	10	11	12	TOTAL
0	2	3	2	1	-	-	-	-	-	-	-	-	8
1	5	9	10	2	3	1	-	-	-	-	-	1	31
2	4	7	10	6	3	-	-	-	-	-	-	-	30
3	18	27	23	19	2	7	2	-	-	-	-	-	98
4	-	-	-	-	-	-	-	-	-	-	-	-	0
5	-	1	-	-	-	-	-	-	-	-	-	-	1
													168

Fonte: Elaborado pelas autoras com base no banco de dados JAMAR.

Observou-se que na *JASIST* oito artigos não possuem indicação de palavras-chave, variando entre 1 a 4 autores. Tendo 31 artigos contendo uma palavras-chave com uma variação de 1 a 6 autores; 30 artigos com 2 palavras-chave variando entre 1 a 5 autores; 98 artigos com 3 palavras-chave variando entre 1 e 7 autores; não se constou nenhum registro com 4 palavras-chave, por fim, 1 artigo com 5 palavras-chave, e possui 2 autores.

Diversas vezes, a mesma quantidade de autores, de artigos distintos que se utilizou de zero palavra-chave é a que também atribuiu 1, 2, 3, 4, ..., 11 palavras-chave. Neste contexto, portanto, evidenciou-se que em ambos os periódicos o número de autores não exerceu qualquer influência na quantidade de palavras-chave atribuídas.

4 CONSIDERAÇÕES FINAIS

Com a realização desta pesquisa, pode-se observar que há necessidade de se estabelecer critérios mais claros para a atribuição de palavras-chave, tendo como foco o resultado final, ou seja, para que se possa ter acesso aos conteúdos relevantes e pertinentes contidos nos artigos de periódicos científicos.

Quando um periódico define critérios de submissão, contendo diretrizes para que tenha atribuição de palavras-chave nas publicações, é imprescindível seu cumprimento. Pois nos resultados, ainda que o artigo possuísse mais de um autor, não houve atribuição sequer de uma palavra-chave, sendo assim os preceitos adotados pelo periódico parece que não são respeitados no momento de submissão.

Considerando que os periódicos científicos estudados, por sua natureza, abrangência e relevância, exercem influência na comunidade de produtores e comunicadores científicos da CI. Considerou-se que havia certa ausência de critérios e rigor em ambos os periódicos, no ano analisado, bem como dos próprios autores dos trabalhos quanto a indicação de descritores, fator este que implica diretamente na recuperação e acesso aos conteúdos temáticos publicados.

Contudo, uma vez que palavras-chave são fundamentais para as análises quantitativas e qualitativas, é primordial orientar a sua atribuição evidenciando-se o conteúdo temático da pesquisa a ser disseminada. Permitindo assim que se obtenha a representação padronizada e relevante, de modo que propicie a recuperação e o acesso aos conteúdos temáticos disponíveis nos artigos de periódicos científicos eletrônicos.

REFERÊNCIAS

BUFREM, L. S. et al. Produção científica em ciência da informação: análise temática em artigos de revistas brasileiras. **Perspect. cienc. inf.**, Belo Horizonte, v. 12, n. 1, p. 38-49, jan./abr. 2007. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/248/36>> Acesso em: 12 fev. 2016.

CUNHA, M. B. da; CAVALCANTI, C. R. de O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília, DF: Briquet Lemos Livros, 2008.

GIL LEIVA, I.; ALONSO-ARROYO, A. La relación entre las palabras clave aportadas por autores de artículos de revista y su indexación en las bases de datos. **Rev. Esp. Doc. Cient.**, 2005, v. 28, n.º 1, p. 62-79.

GUEDES, R. M.; DIAS, E. J. W. Indexação social: abordagem conceitual. **Rev. ACB.**, Florianópolis, v. 15, n. 1, p. 39-53, jan./jun., 2010.

JOURNAL of the Association for Information Science and Technology (JASIST). Disponível em: <<http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%292330-1643>>. Acesso em: 03 jun. 2015

NAVES, M. M. L. Estudo de fatores interferentes no processo de análise de assunto. **Perspect. cienc. inf.**, Belo Horizonte, v. 6, n. 2, p. 189 - 203, jul./dez. 2001. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/428/236>> . Acesso em: 02 fev. 2016

PEREIRA, E. C.; BUFREM, L. S. Princípios de organização e representação de conceitos em linguagens documentárias. **Enc.BIBLI: R.Eletrônica de Bibl.Ci.Inform.**, Florianópolis, n. 20, p. 21-37, jul./dez. 2005.

SCIENTOMETRICS. Disponível em: <<http://link.springer.com/journal/11192>>. Acesso em: 03 jun. 2015.