

A relação entre referências e menções: estudo exploratório em artigos na base de dados BRAPCI^{1,2}

The relation between references and mentions: exploratory study on articles in the BRAPCI database

Rodrigo Aquino de Carvalho

Universidade Federal do Rio Grande do Sul
racfurg@gmail.com

Sônia Caregnato

Universidade Federal do Rio Grande do Sul
sonia.caregnato@ufrgs.br

Maria de Fátima Santos Maia

Universidade Federal do Rio Grande
mafas.maia@gmail.com

Raquel Oroski

Colégio São Carlos, Caxias do Sul – RS
queloroski@gmail.com

Resumo

Apresenta uma perspectiva para os estudos de citação a partir da relação entre a lista de referências e as menções desses documentos no texto. Os objetivos são: i) identificar os padrões existentes na relação entre as referências e as menções na área da Recuperação da Informação, através da análise de artigos publicados em periódicos acadêmicos indexados na base de dados BRAPCI; e ii) comparar um ranking de citação tradicional de autores pessoais com um ranking de menções. Trata-se de um estudo cientométrico e exploratório, para o qual foram utilizados 66 artigos, publicados entre 2008 e 2012, indexados na base de dados BRAPCI e

Abstract

It presents a perspective to citation studies through the relation between references and mentions to those documents in the text. The objectives are: i) to identify the existing patterns in the relationship between references and mentions in the area of Information Retrieval through articles published in academic journals indexed in the BRAPCI database; and ii) to compare a traditional citation ranking of personal authors with a ranking of mentions. It is a scientometric and exploratory study, to which were used 66 articles, published between 2008 and 2012, indexed in the BRAPCI database and retrieved using 12 terms related to "information retrieval". The articles

recuperados a partir de 12 termos relacionados à “recuperação da informação”. Os artigos apresentaram 1468 referências, sendo que 60% (n=877) são mencionadas uma única vez, 30% (n=451) são mencionadas duas ou mais vezes e 10% (n=136) não são mencionadas. O número de menções é de 2246, não considerando 139 menções não referenciadas. O ranking elaborado pelo estudo mostrou que dos 30 autores mais citados (sete ou mais citações), 20 estão entre os mais mencionados. Conclui-se que há um núcleo de autores que são mais mencionados e que são basicamente os mais citados, e isso valida os estudos de citação a partir da lista de referências. Sugere-se a ampliação do recorte temporal desse teste, assim como a sua aplicação em outras áreas do conhecimento.

Palavras-chave: Análise de citação. Cientometria. Comunicação científica.

presented 1468 references, being 60% (n=877) mentioned once, 30% (n=451) mentioned two or more times and 10% (n=139) not mentioned. The number of mentions is 2240, not considering 139 mentions not referenced. The ranking prepared by the study showed that of the 30 most cited authors (seven or more citations), 20 are also the most frequently mentioned. The conclusion is that there is a core of authors who are most frequently mentioned and cited, and this evidence validates citation studies through the reference list. It suggests expanding the time range and the test application in other areas of knowledge.

Keywords: Citation analysis, Scientometrics, Scientific communication.

1. Introdução

Os estudos de citação são necessários para entender o comportamento da produção e do uso de informação na ciência. Utilizar outros documentos para gerar conhecimento e apresentar uma lista desse material é elemento indispensável na publicação de pesquisas.

Tradicionalmente realizados a partir da lista de referências, os estudos de citação e os seus derivados (como análises de cocitação) buscam, cada vez mais, entender o contexto da citação, no sentido de diminuir as limitações dos estudos tradicionais. Assim, pesquisas utilizando o texto completo têm surgido com mais frequência, como apresentam Ding et al (2013), Liu, Zhang e Guo (2013), Song e Kim (2013), Strotmann e Bleier (2013), Zhang, Ding e Milojevic (2013), Jeonga, Songa e Ding (2014), Silveira e Caregnato (2014), para citar alguns exemplos.

O presente trabalho tem como tema os estudos de citação e busca responder dois objetivos, que relacionam as referências bibliográficas apresentadas na lista e a menção desses documentos no texto completo. A questão norteadora é muito simples: existe a tendência de mencionar um documento referenciado mais de uma vez no texto completo? Para tanto, os objetivos são:

- i) Identificar os padrões existentes na relação entre as referências e as menções na área da Recuperação da Informação, através da análise de artigos publicados em periódicos acadêmicos indexados na base de dados BRAPCI; e
- ii) Comparar um *ranking* de citação tradicional de autores pessoais com um *ranking* de menções.

Os objetivos apresentados serão atendidos a partir de duas unidades de contagem, a referência bibliográfica, que representa o documento, e os autores pessoais, respectivamente. Assim, referências com entradas por título, entidade coletiva e jurisdição não serão aproveitadas para atender o segundo objetivo do trabalho. A contagem dos autores inclui todas as autorias das referências e isso implica que as relações de coautoria apresentem uma contagem que não reflète o número de documentos referenciados.

Vale salientar ainda que “menção” nesse trabalho é entendida como o uso efetivo, no texto de um artigo, de um documento presente na sua lista de referências, ou seja, um documento citado (ou referenciado) deve ser mencionado no texto ao menos uma vez para caracterizar a citação. Sendo esta identificada, principalmente no texto completo, pela junção da entrada da referência (p.ex. o sobrenome do autor pessoal) e da data de publicação do documento. Essa forma de nomear a relação citação/referência com citação/menção é utilizada por Silveira e Caregnato (2014), já Ding et al (2013)³ utilizam citação para uma menção e recitação para outras menções da mesma referência. Nenhum dos trabalhos pretendeu discutir questões terminológicas que envolvem a análise de citação, assim como esse texto.

2. Aspectos conceituais e trabalhos relacionados

A citação é uma das formas de fazer ciência e um indicador que ajuda na sua avaliação. Os estudos mais comuns de análise de citação utilizam os autores como elemento de contagem (documentos e periódicos são outros exemplos), indicando que os mais citados são os mais reconhecidos, independente da forma como o contexto do documento citante aborda essa relação. Reitz (c2014) e Cunha e Cavalcanti (2008) definem a citação como uma menção à opinião ou argumento de uma autoridade sobre algum assunto. Nesse sentido, a citação deve, também, indicar o “caminho” para a verificação dessa menção: a referência bibliográfica. Assim, pode-se afirmar que a citação é algo mais amplo que a referência bibliográfica, mas é medida principalmente a partir dessa indicação.

Romancini (2010) diz que a dimensão básica e fundamental da citação é relacionar dois participantes de um circuito científico, “o par ‘autor (texto) citado/autor (texto) citante” (ROMANCINI, 2010, p. 20). Essa relação se mostra muitas vezes política (social) e assim “forçada” e conveniente, mas também se apresenta com caráter cognitivo (epistemológico), fazendo com que as citações sejam utilizadas para comparar métodos e resultados de pesquisa, posições teóricas e de pioneirismo em determinado campo científico etc. Até a não citação é um aspecto que pode ser estudado e que mostra o caráter competitivo e político da ciência.

O certo é que, somente estudando as listas de referências é mais difícil perceber algumas facetas que caracterizam o processo de citar, como as motivações, a relação com outros documentos citados, se o contexto menção é para comparar resultados ou uma discussão teórica, por exemplo. Nesse sentido, alguns estudos merecem destaque.

Brambilla, Vanz e Stumpf (2006) mapearam, através da análise de conteúdo, as razões de 24 citações de um artigo de um pesquisador da Universidade Federal do Rio Grande do Sul – UFRGS e concluíram que todas foram focadas no uso de métodos práticos ou teóricos para a solução de problemas.

Ding et al (2013) apresentam resultados sobre a distribuição das menções (eles utilizam citação) das referências na estrutura textual dos artigos de um periódico, pois questionam que a citação pela lista de referência não deixa claro o crédito ou reconhecimento ao documento citado, além de igualar a contribuição de todos os documentos listados nas referências. Os autores (DING et al, 2013) indicam que a maior parte das menções ocorre nas seções de introdução, revisão de literatura e metodologia dos artigos e que documentos recitados são mais significativos para os documentos citantes.

Jeonga, Songa e Ding (2014) apresentam uma análise de cocitação de autores (ACA) baseado em conteúdo, mostrando que os pares formados a partir das menções são mais específicos e favorecem uma análise mais precisa, pois há maior detalhamento de subdomínios do que os resultados oriundos de uma ACA tradicional.

Silveira e Caregnato (2014) propõem um estudo aplicado para a classificação das razões que levam a citação em comunicações orais e indicam que 93% das menções expressam a dimensão conceitual da citação, ou seja, “as razões das citações [...] se concentram,

essencialmente, para fins de sustentação e exemplificação da argumentação necessária para justificar as contribuições, como também para a construção do arcabouço teórico e metodológico” (SILVEIRA; CAREGNATO, 2014, p. 3627). Esse trabalho conta apenas as menções e não faz uma comparação com dados de lista de referência.

Zhao, Cappello e Johnston (2017) apresentam um estudo que, assim como esse, compara a relação quantitativa de referências e menções (mas não usa esse termo). O objetivo do texto é ponderar a contagem de citações, identificando documentos não essenciais em artigos citantes, mas a partir da localização da menção no texto, já que concluem que essa abordagem é mais eficiente do que a análise de recitação (menção) e remoção de menções únicas (ZHAO; CAPPELLO; JOHNSTON, 2017).

É evidente que os estudos de citação a partir das referências são necessários e devem continuar a ser realizados, pois, como indicam Vanz e Caregnato (2003), possibilitam identificar características e mapear a comunicação científica, além de servirem para atividades de gestão para políticas científicas. Mas o avanço em técnicas que transpõem essa tendência é uma realidade, ainda que muitos estudos sejam de caráter metodológico (JEONGA; SONGA; DING, 2014; SILVEIRA; CAREGNATO, 2014; ZHAO; CAPPELLO; JOHNSTON, 2017) e experimental, inclusive esse.

Assim, o trabalho quer mostrar se há diferenças quantitativas e significativas entre referências e menções e ampliar as possibilidades para *rankings* de citações feitos a partir das referências, como também, deixar à disposição da comunidade dados empíricos para futuras comparações.

3. Procedimentos metodológicos

O trabalho se caracteriza por ser empírico, cientométrico e exploratório. O corpus do estudo foi constituído de 66 artigos publicados entre 2008 e 2012 (cinco anos), recuperados na base BRACPI, nos meses de fevereiro e março de 2015. . Documentos publicados entre 2013 e 2014 não foram incluídos, pois apresentaram poucos resultados, indicando que números atrasados de alguns periódicos ainda não estavam na base.

As expressões que seguem foram utilizadas, como frase (entre aspas), para realizar as buscas com ocorrência dos termos no “título” e nas “palavras-chave”: a) RECUPERAÇÃO DA INFORMAÇÃO; b) RECUPERAÇÃO DE INFORMAÇÃO; c) RECUPERAÇÃO INTELIGENTE DA INFORMAÇÃO; d) RECUPERAÇÃO INTELIGENTE DE INFORMAÇÃO; e) RECUPERAÇÃO DE DADOS; f) ARMAZENAMENTO DE DADOS; g) ARMAZENAGEM DE DADOS; h) COLETA DE DADOS; i) BUSCA DE INFORMAÇÃO; j) BUSCA DA INFORMAÇÃO; k) INFORMATION RETRIEVAL; e, l) RETRIEVAL OF INFORMATION.

Definido o *corpus* foi gerada uma planilha de dados com os indicadores de caracterização e os primeiros resultados da pesquisa. A matriz foi formada por 66 casos (artigos) e as seguintes variáveis: ano de publicação, periódico, tipo de artigo, metadado (título e palavra-chave) de ocorrência do termo de busca, número de autores, número de referências (separando as com entradas de autoria pessoal das outras e identificando as que tinham apenas um autor das que tinham dois ou mais), número total de menções no corpo do texto, número de referências com uma menção, número de referências com duas menções, número de referências com três ou mais menções, referências não mencionadas e menções não referenciadas. A tabela 1 apresenta os principais indicadores de caracterização do *corpus*.

Tabela 1. Corpus da pesquisa, por ano de publicação e título do periódico

REVISTAS	ANO DE PUBLICAÇÃO					Total por revista (%)
	2008	2009	2010	2011	2012	
Informação e Sociedade	01	-	03	01	-	05 7.58
AtoZ: novas práticas em informação e conhecimento	-	-	-	01	-	01 1.52
Biblionline	-	-	03	-	03	06 9.09
Biblos	-	02	01	01	-	04 6.06
Brazilian Journal of Information Science	-	01	01	-	-	02 3.03
Ciência da Informação	01	01	-	-	-	02 3.03
DataGramaZero	-	03	-	03	01	07 10.61
Em Questão	-	-	-	01	-	01 1.52
Encontros Bibli	03	-	-	02	02	07 10.61
InCID: Revista de Ciência da Informação e Documentação	-	-	-	02	01	03 4.55
Informação & Informação	01	01	01	01	-	04 6.06
Perspectivas em Ciência da Informação	01	02	-	03	01	07 10.61
Perspectivas em Gestão & Conhecimento	-	-	-	01	-	01 1.52
Pontodeacesso	-	-	01	-	02	03 4.55
Revista Brasileira de Biblioteconomia e Documentação	01	-	-	-	-	01 1.52
Revista Digital de Biblioteconomia e Ciência da Informação	-	-	01	-	-	01 1.52
Revista ACB	02	01	01	-	01	05 7.58
Revista Ibero-Americana de Ciência da Informação	-	-	-	01	-	01 1.52
TransInformação	-	-	03	01	01	05 7.58
Total por ano (%)	10 15.15	11 16.67	15 22.73	18 27.27	12 18.18	66 100.00

Fonte: Dados da pesquisa, 2015

Depois da recuperação dos artigos, as referências foram manualmente colocadas em arquivos de texto apenas com os autores (ou entradas das referências) e a data de publicação. A partir desses dados foi realizada a contagem manual das menções e elaborado o *ranking* de citação de autores do *corpus* (299 autores com duas ou mais citações). Não foi possível localizar a fonte de duas referências elaboradas em coautoria, com a indicação “et al”, e assim apenas o primeiro autor entrou no *ranking*. Os dados foram normalizados com a verificação de autores com o mesmo sobrenome e iniciais próximas. A figura 1 exemplifica a forma de organização dos dados das referências.

A contagem de menções dos primeiros e únicos autores, das referências com entradas por título, entidade coletiva e jurisdição atendem o primeiro objetivo específico do trabalho e os resultados são apresentados na Tabela 2 e na Figura 2.

Figura 1. **Exemplo de organização dos dados retirados de cada artigo**

Autoria	Menção
1. AUTOR, A	02
AUTOR, B	
AUTOR, C	
AUTOR, D	
AUTOR, E	
2. AUTOR, C	01
3. ENTIDADE COLETIVA A	01
4. JURISDIÇÃO A	06
Total de menções no ARTIGO A	10

Fonte: Dados da pesquisa, 2015.

Além do indicador tradicional de citação de autores e do número de menções de cada autor, mais dois indicadores foram elaborados para ajudar na comparação, como apresentado no Quadro 1, que considera as primeiras 30 posições do *ranking* elaborado. Segue abaixo a descrição dos quatro indicadores:

- Número de citações: ocorrência do autor na lista de referências (com e sem autocitação);
- Número de menções: ocorrência do autor no corpo do texto (a menção é pelo documento, que mencionado duas vezes e com dois autores, p.ex., acarreta duas menções para cada autor);
- Número de artigos citantes: documentos que citam os autores, no caso desse trabalho são os 66 artigos do corpus (esse dado ajuda a entender a distribuição das citações e menções, uma vez que há casos de autores muito citados ou mencionados, mas em poucos artigos citantes); e
- Pico de menção de um único documento referenciado por autor: o maior número de menção de um único documento de um autor em todo o corpus (informação que ajuda a entender a distribuição das menções nos artigos citantes).

Foram medidos os coeficientes de correlação de Spearman⁴ e Pearson⁵ para as quatro categorias no software SPSS⁶, considerando nível de significância igual a 0,05. Esses testes servem para testar a significância da relação entre variáveis intervalares (BECKER, 2015).

4. Resultados e discussão

Os resultados iniciais mostram que existe o comportamento de mencionar mais de uma vez um documento referenciado (Tabela 2), ainda que a maioria das referências seja mencionada uma única vez (n=877), dados proporcionalmente parecidos com os apresentados por Zhao, Cappello e Johnston (2017, p.60). As 1468 referências geraram 2240 menções, com média de

33,94, um pouco abaixo dos dados apresentados por Silveira e Caregnato (2014). Mas é interessante observar que, de 451 referências, as mencionadas duas ou mais vezes correspondem à 1363 menções.

Os dados discrepantes são as “menções não referenciadas” (n=63) e as “referências não mencionadas” (n=139), mas esses resultados não parecem ser uma tendência, uma vez que apenas a mediana das referências não mencionadas foi diferente de zero. A tabela abaixo apresenta os principais dados descritivos das variáveis relativas às referências.

Tabela 2. Distribuição do número de referências e menções no corpus da pesquisa (N=66)

CATEGORIAS	Totais	Medidas de centralidade			Medidas de dispersão		
		Média	Mediana	Moda	Desvio padrão	Mínimo	Máximo
Número de referências	1468	22.24	19	17	11.85	01	54
Referências com entrada de autoria pessoal*	1348	20.74	19	14	10.99	01	53
Referências com entrada por título, entidade coletiva ou jurisdição	120	1.82	00	00	3.16	00	16
Referências com UMA menção**	877	13.29	11.50	06	8.16	00	38
Referências com DUAS menções	239	3.62	03	02	3.19	00	13
Referências com TRÊS ou MAIS menções	212	3.21	2.50	00	3.07	00	13
Referências NÃO MENCIONADAS	139	2.11	01	00	4.06	00	21
Menções no texto completo***	2240	33.94	29.50	18	21.21	02	91
Menções não referenciadas	63	0.95	00	00	1.88	00	10

Obs: * referências utilizadas para elaborar os rankings do segundo objetivo do trabalho | ** uma referência foi desconsiderada para a contagem das menções | *** não inclui as menções não referenciadas.

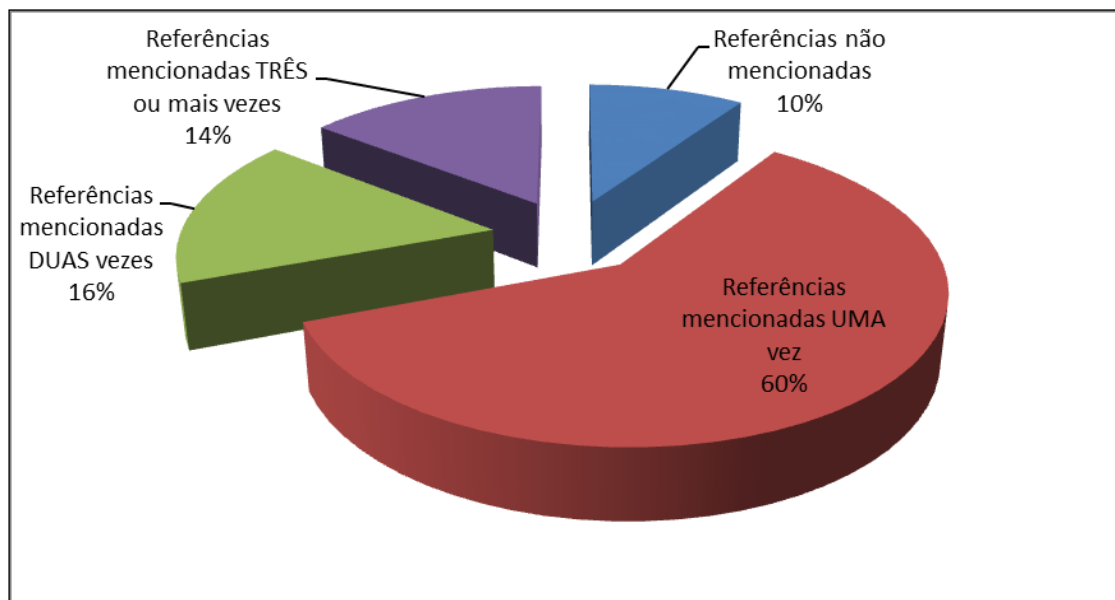
Fonte: Dados da pesquisa, 2015

A Figura 2 mostra que 30% das referências foram mencionadas mais de uma vez. Parece necessário estudar outros contextos (áreas de conhecimento e tipos de publicação) para verificar se isso é um comportamento. Voos e Dagaev⁷ (1976 apud DING et al, 2013) sugerem que o número de menções de uma referência indica a relevância para o assunto do documento.

Os 10% de referências não mencionadas se apresentam como uma prova empírica que Romancini (2010) reivindica quando comenta a afirmação de Wouters sobre muitas referências não serem citadas. Essa proporção pode ser um caso isolado, mas indica que os editores precisam estar atentos para diminuir essa prática, ainda mais que os dados

descritivos apresentaram um desvio padrão de 4,06 e um valor absoluto máximo igual a 21. As boas práticas na publicação de pesquisas são fundamentais, além de ser uma questão ética. A responsabilidade é principalmente dos autores, mas também de editores e pareceristas.

Figura 2. Proporção entre referências mencionadas



Fonte: dados da pesquisa, 2015

Para atender o segundo objetivo do trabalho foi gerado um ranking com os pesquisadores mais citados e mais mencionados. A essas duas categorias juntaram-se “número de artigos citantes” e “pico de menção de um único documento referenciado”, como salientado na seção dos procedimentos metodológicos.

O ranking ficou com 299 autores (com duas ou mais citações). A Tabela 3 apresenta os coeficientes de correlação de Spearman e Pearson das variáveis. Ding e Cronin (2011) utilizaram o coeficiente de Spearman para mostrar a relação de três *rankings* em um estudo de citação e o coeficiente de Pearson foi realizado por ser o mais comum e robusto. A análise apresentou valor de significância igual à zero para todos os pares de correlação em ambos os coeficientes, o que, segundo Field (2009), é aceitável visto que precisa ser $< 0,05$.

Em relação às categorias, “número de citação” apresentou correlação alta com “número de artigos citantes” e “número de menções” (as três mais importantes). A categoria “pico de menções” apresentou correlação baixa com “número de citação” e “número de artigos citantes”, mas alta com “número de menções”. O comportamento foi parecido com os dois

coeficientes utilizados e mostrou que a variável “pico de menção” é a mais afastada, ainda que diretamente relacionada com o “número de menções”.

Tabela 3. Distribuição das taxas de coeficiente de correlação de Spearman e de Pearson

CATEGORIAS	Nº Citação	Nº Artigos citantes	Nº Menções	Pico de menções	
Nº Citação	1,000	0,858	0,873	0,521	Correlação de Pearson
Nº Artigos citantes	0,742	1,000	0,682	0,388	
Nº Menções	0,735	0,577	1,000	0,784	
Pico de menções	0,412	0,341	0,860	1,000	
Correlação de Spearman					
O valor de significância de todos os pares da correlação foi igual a zero (0,000), ou seja, menor que 0,05.					

Fonte: dados da pesquisa, 2017

A Tabela 4 apresenta os dados descritivos do *ranking* de autores considerando as três principais variáveis, e os valores de menções se destacam, indicando um padrão que merece estudos mais apropriados, levando em consideração os autores como unidade de contagem.

Tabela 4. Medidas descritivas das variáveis: citação, artigos citantes e menção (n=299)

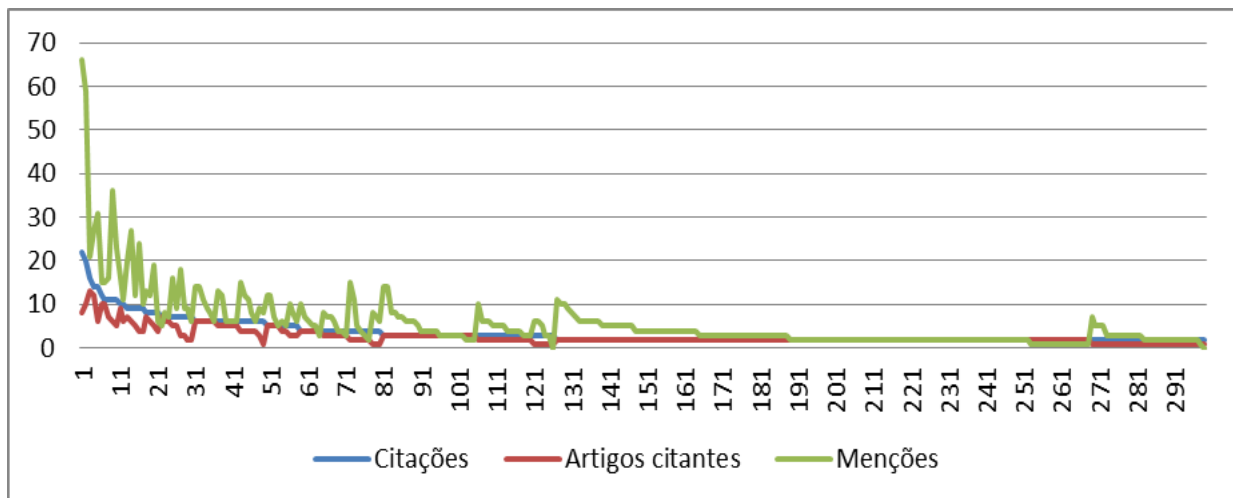
Medidas	Nº Citação	Nº Artigos citantes	Nº Menções
Média	34.55	26.82	56.19
Mediana	20	20	30
Moda	02	02	02
Desvio padrão	2.75	1.78	6.87
Mínimo	02	01	00
Máximo	22	13	66

Fonte: Dados da pesquisa, 2015

Os valores “mínimos” baixos da Tabela 4 indicam a necessidade da escolha mais criteriosa de periódicos para futuros estudos. Não fazia sentido para esse trabalho remover os documentos que apresentam esses dados, uma vez que o ponto de partida era a base de dados BRAPCI e não os periódicos. Tratar os dados como *outliers* também não pareceu apropriado, uma vez que não são duvidosos (BECKER, 2015).

A Figura 3 mostra uma distribuição de frequência das três principais categorias relacionadas aos autores. A menção apresenta os maiores valores, e há picos, mesmo quando a citação e o número de documentos citantes são baixos. As referências não mencionadas deixam a linha da menção abaixo da linha dos artigos citantes, mostrando um ruído nos dados.

Figura 3. Distribuição comparativa das variáveis: citação, artigos citantes e menção



Fonte: Dados da pesquisa, 2017

No Quadro 1 são apresentados os autores mais citados do ranking elaborado (primeira coluna de nomes), com sete ou mais citações. As demais colunas apresentam os resultados das outras categorias. Os números apresentados após o nome dos autores indicam a posição em cada categoria. Assim, o autor “Saracevic [3,1,9,42]” foi o terceiro mais citado, o primeiro em número de artigos citantes, o nono em número de menções e 42º na categoria pico de menção.

Optou-se por colocar os autores nas 30 primeiras posições de cada categoria, ao invés de uma única lista de autores e suas posições em cada categoria, para evidenciar as diferenças. Autores que aparecem na coluna da categoria “pico de menção”, como “Moura [134,117,63,2]”, “Taylor [135,118,41,6]”, “Almeida Júnior [270,269,68,10]”, “Gadamer [136,119,42,16]” e “Stambaugh [137,120,48,17]” acabariam não aparecendo no quadro e isso diminuiria a percepção do alcance das menções. Na categoria menções, autores como “Heidegger [73,96,19,14]”, “Cox [81,69,22,4]” e “Hall [82,70,23,5]” exemplificam essa percepção.

Quadro 1. Distribuição dos rankings de autores pessoais (30 posições)

Número de citações (referências) – sem e com autocitação		Número de artigos citantes		Número de menções		Pico de menção de um único documento referenciado			
01	<u>Hjorland [1,7,1,1]</u>	22	22	Saracevic [3,1,9,42]	13	<u>Hjorland [1,7,1,1]</u>	66	Hjorland [1,7,1,1]	19
02	<u>Kuhlthau [2,3,2,8]</u>	20	20	Lancaster [4,2,5,19]	12	<u>Kuhlthau [2,3,2,8]</u>	59	Moura [134,117,63,2]	10
03	<u>Saracevic [3,1,9,42]</u>	16	16	Kuhlthau [2,3,2,8]	10	<u>Ellis [9,13,3,3]</u>	36	Ellis [9,13,3,3]	09
04	<u>Lancaster [4,2,5,19]</u>	14	14	Kobashi [6,4,16,103]	10	<u>Wilson [5,12,4,9]</u>	31	Cox [81,69,22,4]	09
05	<u>Wilson [5,12,4,9]</u>	14	14	Barreto [7,5,17,43]	10	<u>Lancaster [4,2,5,19]</u>	27	Hall [82,70,23,5]	09
06	<u>Kobashi [6,4,16,103]</u>	12	12	Souza [11,6,14,20]	09	<u>Capurro [14,15,6,7]</u>	27	Taylor [135,118,41,6]	09
07	<u>Barreto [7,5,17,43]</u>	11	11	Hjorland [1,7,1,1]	08	<u>Ingwersen [16,40,7,21]</u>	24	Capurro [14,15,6,7]	08
08	<u>Spink [8,8,13,44]</u>	11	11	Spink [8,8,13,44]	07	<u>Fujita [10,25,8,11]</u>	23	Kuhlthau [2,3,2,8]	07
09	<u>Ellis [9,13,3,3]</u>	11	11	Belkin [13,9,10,45]	07	<u>Saracevic [3,1,9,42]</u>	21	Wilson [5,12,4,9]	07
10	<u>Fujita [10,25,8,11]</u>	11	11	Alvarenga [18,10,24,12]	07	<u>Belkin [13,9,10,45]</u>	20	Almeida Júnior [270,269,68,10]	07
11	<u>Souza [11,6,14,20]</u>	10	11	Cunha [22,11,96,182]	07	<u>Ferreira [20,27,11,13]</u>	19	Fujita [10,25,8,11]	06
12	<u>Lima [12,14,32,104]</u>	10	10	Wilson [5,12,4,9]	06	<u>Järvelin [27,56,12,23]</u>	18	Alvarenga [18,10,24,12]	06
13	<u>Belkin [13,9,10,45]</u>	09	09	Ellis [9,13,3,3]	06	<u>Spink [8,8,13,44]</u>	16	Ferreira [20,27,11,13]	06
14	<u>Capurro [14,15,6,7]</u>	09	09	Lima [12,14,32,104]	06	<u>Souza [11,6,14,20]</u>	16	Heidegger [73,96,19,14]	06
15	<u>Crespo [15,26,26,60]</u>	09	09	Capurro [14,15,6,7]	06	<u>Choo [25,28,15,22]</u>	16	Rubi [74,97,35,15]	06
16	<u>Ingwersen [16,40,7,21]</u>	09	09	Jansen [19,16,27,46]	06	<u>Kobashi [6,4,16,103]</u>	15	Gadamer [136,119,42,16]	06
17	Campos [17,41,37,105]	09	09	Castells [23,17,49,107]	06	<u>Barreto [7,5,17,43]</u>	15	Stambaugh [137,120,48,17]	06
18	<u>Alvarenga [18,10,24,12]</u>	08	09	Marcondes [24,18,59,183]	06	Guarino [43,43,18,25]	15	Tibbo [139,122,67,18]	06
19	<u>Jansen [19,16,27,46]</u>	08	08	Baeza-yates [31,19,20,48]	06	Heidegger [73,96,19,14]	15	Lancaster [4,2,5,19]	05
20	<u>Ferreira [20,27,11,13]</u>	08	08	Ribeiro-neto [32,20,21,49]	06	Baeza-yates [31,19,20,48]	14	Souza [11,6,14,20]	05
21	Jardim [21,42,69,106]	08	08	Ferneda [33,21,33,61]	06	Ribeiro-neto [32,20,21,49]	14	Ingwersen [16,40,7,21]	05
22	Cunha [22,11,96,182]	07	07	Gil [35,22,46,24]	06	Cox [81,69,22,4]	14	Choo [25,28,15,22]	05
23	Castells [23,17,49,107]	07	07	Lara [32,23,50,62]	06	Hall [82,70,23,5]	14	Järvelin [27,56,12,23]	05
24	Marcondes [24,18,59,183]	07	07	Le Coadic [36,24,71,111]	06	<u>Alvarenga [18,10,24,12]</u>	13	Gil [35,22,46,24]	05
25	<u>Choo [25,28,15,22]</u>	07	07	Fujita [10,25,8,11]	05	Figueiredo [37,30,25,50]	13	Guarino [43,43,18,25]	05
26	Lévy [26,29,43,47]	07	07	Crespo [15,26,26,60]	05	<u>Crespo [15,26,26,60]</u>	12	Almeida [44,44,29,26]	05
27	<u>Järvelin [27,56,12,23]</u>	07	07	Ferreira [20,27,11,13]	05	<u>Jansen [19,16,27,46]</u>	12	Rowley [59,50,39,27]	05
28	McGuinness [28,57,44,108]	07	07	Choo [25,28,15,22]	05	Campello [38,31,28,63]	12	Buckland [65,62,54,28]	05
29	Maturana [29,94,45,109]	07	07	Lévy [26,29,43,47]	05	Almeida [44,44,29,26]	12	Sousa [78,260,55,29]	05
30	Gomes [30,95,70,110]	07	07	Figueiredo [37,30,25,50]	05	Baptista [50,36,30,51]	12	Mathes [83,71,56,30]	05

Fonte: dados da pesquisa, 2015

Os resultados mostram que dos 30 autores mais citados, 20 estão entre os mais mencionados (destaques em *itálico* e *sublinhado* nas colunas um e três). Esse resultado mostra uma evidência da importância e da confiabilidade dos estudos de citação realizados a partir das referências. Ou seja, estudar as menções de autores altamente citados parece relevante para identificar o que Zhao, Cappello e Johnston (2017) chamam de documentos essenciais, mas vale salientar que o número de documentos citantes melhora essa visualização.

O autor “Hjorland [1,7,1,1]”, p.ex., que é o mais citado no *corpus*, tem 22 citações de oito documentos citantes e 66 menções, e isso pode indicar que ele é essencial para alguns dos artigos que o citam. Mas o autor “Saracevic [3,1,9,42]” que tem 16 citações, em 13 documentos citantes, e 21 menções pode não ser essencial para muitos documentos citantes, apesar de ter um alcance maior de uso no *corpus* da pesquisa. Assim, estudar as menções dos autores em mais documentos citantes pode ser uma alternativa.

Os resultados dos autores “Moura [134,117,63,2]”, “Taylor [135,118,41,6]” e “Almeida Júnior [270,269,68,10]”, na categoria pico de menção, evidenciam a importância de relacionar as categorias, uma vez que os valores são altos, mas oriundos de apenas um documento citante. Ou seja, os autores apareceriam em um *ranking* de menções mais completo, mas sem alcance no *corpus* da pesquisa.

Os resultados dos autores “Baeza-Yates [31,19,20,48]” e “Ribeiro-Neto [32,20,21,49]” que aparecem em duas categorias no Quadro 1 (nº de artigos citantes e menção) evidenciam o cuidado que se deve ter com a unidade de contagem, pois esses valores são oriundos de um mesmo documento. Assim, fazer estudos de menções de documentos, identificar os documentos dos citados e a relação de coautoria dos autores melhor posicionados no *ranking* pode melhorar a visibilidade dos dados.

As menções podem oferecer perspectivas mais amplas ou perspectivas que as referências nunca tiveram efetivamente a função de esclarecer, independentemente da unidade contagem. Mas afinal, o que leva um autor a mencionar uma referência 19, 10 ou nove vezes, como mostram as três primeiras posições da quarta coluna do Quadro 1? Este é um questionamento importante em estudos de contexto de citação.

5. Considerações finais

Na produção brasileira sobre Recuperação da Informação (2008-2012) há o comportamento de mencionar um documento mais de uma vez, ainda que a maioria das referências seja mencionada uma única vez. As referências não mencionadas e menções não referenciadas aparecem como um complicador, que distorce levemente os dados e que pode ser resolvido com uma maior atenção das equipes que gerenciam as revistas da área da Ciência da Informação no Brasil.

Em relação aos autores pode-se afirmar que os mais mencionados são basicamente os mais citados, o que implica em uma das limitações do estudo, já que essa comparação não pode ser baseada apenas em dados absolutos. Entende-se que as categorias adicionadas, artigos citantes e pico de menção, melhoram a visualização dos dados.

As limitações tecnológicas parecem ser um impedimento para estudos com menções em larga escala, pois automatizar a extração em textos completos é mais complexo do que lidar com a lista de referências. Mas de qualquer forma, estudos com menção sugerem contexto de citação e métodos qualitativos para inferências mais precisas, o que exige *corpus* menor para análises mais qualificadas e precisas.

Limitações do trabalho

O trabalho apresenta algumas limitações que merecem destaque:

- a utilização de duas unidades de contagem (documento e autoria) impede o aprofundamento maior de algumas questões, mas se justifica pelo caráter exploratório da pesquisa; e
- o número de artigos do *corpus* é pequeno, mas o trabalho manual exigiu essa condição, e lidar com citações amplifica muito o trabalho de análise.

Sugestões para futuras pesquisas

Segue abaixo algumas sugestões para futuras pesquisas:

- um estudo teórico terminológico parece necessário para definir e consolidar alguns termos, como: menção, citação, referência e recitação;
- a ampliação do recorte temporal e um número maior de documentos pode apresentar dados mais robustos, ainda que a pesquisa seja replicada; e

- o estudo de aspectos éticos relacionados a boas práticas de produção do conhecimento, tanto pelos autores como pelos periódicos é necessário.

Espera-se que dados possam colaborar em pesquisas no campo da Cientometria, principalmente no contexto brasileiro.

Referências Bibliográficas

- BRAMBILLA, S. D. S., VANZ, S. A. De S., STUMPF, I. R. C. (2006). Mapeamento de um artigo produzido na UFRGS: razões das citações recebidas. *Encontros Bibli: R. Eletr. Bibliotecon. Ci. Inf.*, 195-209. Retrieved from <https://goo.gl/frk4Zp>
- BECKER, J. L. (2015). *Estatística Básica: Transformando Dados em Informação*. Porto Alegre: Bookman.
- CUNHA, M. B., CAVALCANTI, C. R. de O. (2008). *Dicionário de Biblioteconomia e Arquivologia*. Briquet de Lemos: Brasília, DF.
- DING, Y., CRONIN, B. (2011). Popular and/or prestigious? Measures of scholarly esteem. *Information Processing and Management*, 47, 80-86. doi: 10.1016/j.ipm.2010.01.002
- DING, Y., LIU, X., GUO, Chun, CRONIN, B. (2013) The distribution of references across texts: some implications for citation analysis. *Journal of Informetrics*, 7, 583-592. doi: 10.1016/j.joi.2013.03.003
- FIELD, A. (2009). *Descobrimo a estatística usando o SPSS*. (2nd ed). Porto Alegre: Artmed.
- JEONGA, Y., SONGA, M., DING, Y. (2014). Content-based author co-citation analysis. *Journal of Infor-metrics*, 8, 197-211. doi: 10.1016/j.joi.2013.12.001
- LIU, X., ZHANG, J., GUO, C. (2013). Full-Text Citation Analysis: A New Method to Enhance Scholarly Networks. *Journal Of The American Society For Information Science And Technology*, 64(9):1852–1863. doi: 10.1002/asi.22883
- REITZ, J. M. (2014). *ODLIS - Online Dictionary for Library and Information Science*. Retrieved from http://www.abc-clio.com/ODLIS/odlis_c.aspx
- ROMANCINI, R. (2010). O que é uma citação? Análise de citações na ciência. *Intexto*, 23(2), 20-35. Retrieved from <http://goo.gl/Jgeidh>
- SILVEIRA, M. A. A., CAREGNATO, S. E. (2014). Estudo das razões das citações na Ciência da Informação: proposta de classificação. *Proceedings of the Encontro Nacional de Pesquisa em Ciência da Informação*, Brazil, 15. Retrieved from <http://goo.gl/Q5MfKA>

- SONG, Min, KIM, S. Y. (2013). Detecting the knowledge structure of bioinformatics by mining full-text collections. *Scientometrics*, 96(1), 183–201. doi: 10.1007/s11192-012-0900-9
- STROTMANN, A., BLEIER, A. (2013). Author Name Co-Mention Analysis: Testing a Poor Man's Author Co-Citation Analysis Method. *Proceedings of the International Society of Scientometrics and Informetrics Conference*, USA, 14. Retrieved from <https://arxiv.org/abs/1309.5256>
- VANZ, S. A. de S., CAREGNATO, S. E. (2003). Estudos de Citação: uma ferramenta para entender a comunicação científica. *Em Questão*, 9(2), 295-307. Retrieved from <http://goo.gl/HdRvbB>
- ZHANG, G., DING, Y., MILOJEVIC, S. (2013). Citation Content Analysis (CCA): A Framework for Syntactic and Semantic Analysis of Citation Content. *Journal Of The American Society For Information Science And Technology*, 64(7), 1490–1503. doi: 10.1002/asi.22850
- ZHAO, D., CAPPELLO, A., JOHNSTON, L. (2017). Functions of Uni- and Multi-citations: Implications for Weighted Citation Analysis. *Journal of Data and Information Science*, 2(1), 51–69. doi: 10.1515/jdis-2017-0003

Notas

¹ 1 Base de Dados em Ciência da Informação | Acervo de Publicações Brasileiras em Ciência da Informação.

² Texto originalmente publicado nos anais do XVI Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB), no Brasil, em 2015, que passou por revisão em ampliação dos dados.

³ “[...] cited/re-cited in the body of the text [...]”, p.591.

⁴ “[...] medida padronizada da força do relacionamento entre duas variáveis que não depende das hipóteses de um teste paramétrico” (FIELD, 2009, p.643).

⁵ “O coeficiente mede o grau de dependência linear entre duas variáveis” (BECKER, 2015, p.98).

⁶ Statistical Package for the Social Sciences for Windows.

⁷ Voos, H., & Dagaev, K. S. (1976). Are all citations equal? Or Did we op. cit. your idem? *J. of Acad. Librarianship*, 1, 20–21. [Artigo não localizado na Web e esse número do periódico não consta em nenhuma Biblioteca que faz parte do Catálogo Coletivo Nacional – CCN/IBICT]