

Análise de cocitação de autores: um estudo teórico-metodológico dos indicadores de proximidade, aplicados ao GT7 da ANCIB

Maria Cláudia Cabrini Grácio*

Ely Francina Tannuri de Oliveira**

Resumo Este estudo, de natureza teórico-metodológica, objetiva analisar os indicadores de proximidade absolutos e relativos - Cosseno de Salton e Índice de Jaccard-, para análise de cocitação de autores, e comparar sua contribuição para o entendimento de um domínio. Metodologicamente, aprofunda os estudos teóricos que fundamentam a análise de cocitação de autores (ACA) e os indicadores de proximidade, passando aos estudos dos índices normalizados de Cosseno de Salton (CS) e Índice de Jaccard (IJ). Apresenta uma aplicação no GT7 da ANCIB, na área de Ciência da Informação, no período de 2007 a 2011.

Palavras-chave Análise de citação, Análise de cocitação, Indicadores de proximidade, Cosseno de Salton, Índice de Jaccard.

Authors' co-citation analysis: a theoretical- methodological study of proximity indicators applied to GT7 of ANCIB

Abstract Of a theoretical-methodological nature, this research aims at analyzing the absolute and relative proximity indicators – Salton's cosine and Jaccard's index –, for authors' co-citation analysis (ACA). It also aims at comparing the contribution of these three indicators to the understanding of a domain. Methodologically, it deepens the theoretical studies that motivate authors' co-citation analysis (ACA) and the proximity indicators, and presents the studies of the normalized index of Salton' cosine (SC) and Jaccard's index (JI). It applies ACA to the GT7 of ANCIB, in the Information Science field, from 2007 to 2011.

Keywords Citation analysis, Co-citation analysis, Proximity indicators, Salton's cosine, Jaccard's index.

* Mestre em Estatística e Doutora em Filosofia, pela Universidade Estadual de Campinas. Professor Assistente Doutor do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho. Endereço postal: Av. Hygino Muzzi Filho, 737. Campus Universitário. Marília. CEP 17.525-900. Telefone: (14) 3402-1371. E-mail: cabrini@marilia.unesp.br

** Mestre e Doutora em Educação pela Universidade Estadual Paulista-Marília. Professor Assistente Doutor do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho e Docente do Departamento de Ciência da Informação da UNESP/Marília. Endereço postal: Av. Hygino Muzzi Filho, 737. Campus Universitário. Marília. CEP 17.525-900. Telefone: (14) 3433-0359. Email: etannuri@gmail.com

Introdução

Fundamentados em recursos quantitativos como método de análise, os “estudos métricos” constituem o conjunto de conhecimentos relacionados à avaliação da informação produzida, alicerçado na sociologia da ciência, na ciência da informação, matemática, estatística e computação.

Esses estudos são de natureza teórica e conceitual quando contribuem para o avanço do conhecimento da própria temática, propondo novos conceitos e indicadores, bem como reflexões e análises relativas à área. São de natureza metodológica quando se propõem a dar sustentação aos trabalhos de caráter teórico da área onde estão aplicados (OLIVEIRA; GRÁCIO, 2011).

Com conceito alinhado a esses autores, Glänzel (2003) considera que a pesquisa bibliométrica atual é destinada a três grupos-alvo principais, que determinam tópicos e subáreas da bibliometria contemporânea, a saber: *Bibliometria para profissionais da bibliometria; Bibliometria aplicada às disciplinas científicas; Bibliometria para a política científica e gestão*. O primeiro grupo é o domínio da pesquisa bibliométrica “de base”, que busca desenvolver e debater a bibliometria como metodologia, isto é, está preocupada com o seu próprio desenvolvimento conceitual-teórico-metodológico.

O escopo deste estudo situa-se nesse primeiro grupo, “*Bibliometria para profissionais da bibliometria*”, considerando que visa contribuir, de forma analítica, para o desenvolvimento da temática “análise de cocitação de autores”, ao tratar dos indicadores necessários e subjacentes a essa metodologia, oferecendo subsídios teóricos e metodológicos de modo a fortalecer a sua compreensão e uso.

Hjørland (2002a), precursor das investigações em Análise de Domínio (AD) na área de Ciência da Informação, destaca a contribuição dos estudos de citação e cocitação, dentre as metodologias de Bibliometria, para a compreensão de um domínio, entendido como reflexo de uma comunidade discursiva e do seu papel na ciência.

Os estudos de análise de cocitação, baseados na frequência com que dois autores ou documentos são citados de forma conjunta na produção científica de uma área, evidenciam como a estrutura de conhecimento de uma área é percebida pelos pesquisadores. Tem como princípio o fato de que, quando dois documentos ou autores são citados juntos em um trabalho posterior, existe, na perspectiva do autor citante, uma proximidade de assunto entre os citados. Assim, quanto maior a frequência de cocitação, mais próxima a relação entre esses autores citados.

Desse modo, para se avaliar a interlocução entre pesquisadores e seu papel em diferentes áreas da ciência, os estudos de citações e cocitações constituem procedimentos relevantes de análise na medida em que contribuem para a visualização do processo comunicativo e interativo, bem como da estrutura subjacente de um domínio do conhecimento.

Considerando o que foi exposto, esta pesquisa objetiva realizar um estudo da contribuição dos indicadores de proximidade absolutos e normalizados para análise de cocitação de autores com vistas à compreensão do conhecimento de um domínio.

De forma mais específica, propõe-se apresentar e analisar os indicadores de proximidade absolutos e os relativos, tais como o Cosseno de Salton e Índice de Jaccard, para análise de

cocitação de autores, e comparar a contribuição destes três indicadores para entendimento de um domínio, aplicando-os ao universo dos dados do GT7 da ANCIB, no período de 2007 até 2011.

Este trabalho se justifica em razão da escassez de estudos existentes na temática, especialmente daqueles que tratam de forma elucidativa e clara dos indicadores normalizados subjacentes à análise de cocitação, no Brasil. Justifica-se ainda a escolha do universo adotado para a aplicação da análise de cocitação e seus indicadores por ser o ENANCIB, com seus grupos de trabalho, um fórum de debates e reflexões que reúne pesquisadores e socializa a produção do conhecimento novo em temas especializados da Ciência da Informação, incentivando a reflexão sobre temas e tendências acadêmicas da atualidade e avaliando a produção científica na área. O grupo de trabalho GT7 – “*Produção e Comunicação da Informação em CT&I*” - apresenta uma massa documental crescente e representativa da produção científica, especialmente pelos pesquisadores que estão na pós-graduação, no Brasil, constituindo universo representativo para a análise de domínio na temática.

Análise de domínio científico: visualização por meio dos estudos de citação e cocitação e indicadores de proximidade

Para se avaliar questões de produtividade científica, a *Análise de Domínio (A.D.)* aparece hoje como o principal respaldo teórico, utilizado pela primeira vez em 1980, na área de *Ciência da Computação* (KERR, 2003). No âmbito da Ciência da Informação, Birger Hjørland foi o primeiro a usar esse conceito em parceria com Hanne Albrechtsen, fundamentando sua teoria e metodologia (HJØRLAND, 2002b).

O conceito de domínio pode ser compreendido como “uma área de especialidade, um conjunto literário ou um grupo de pessoas trabalhando juntas em uma organização” (MAI, 2005, p.605). Sob o ponto de vista desta pesquisa, é uma área de conhecimento, atividade, interesse, em que é demarcado um determinado conhecimento com limites definidos e em que profissionais ou grupos estão articulados tanto em pensamento como em linguagem.

Para Hjørland e Albrechtsen (1995), um domínio científico é entendido como o reflexo de uma comunidade discursiva e do seu papel na ciência.

Segundo Hjørland (2002 a), são 11 as abordagens sobre Análise de Domínio, dentre elas os estudos bibliométricos. O autor destaca que o uso conjunto de mais de uma dessas abordagens enriquece a análise e a compreensão de um domínio e, ainda, que os estudos bibliométricos constituem uma abordagem consistente para analisar e caracterizar um domínio científico.

Dentre os estudos bibliométricos, Hjørland (2002a) destaca a contribuição dos estudos de citação e cocitação, especialmente no que se refere à construção de mapeamentos bibliométricos ou à visualização das áreas do conhecimento científico.

O conjunto de referências dos trabalhos científicos pode ser analisado como reflexo de uma comunidade discursiva, de modo a constituir um domínio. Seu estudo baseia-se em análises das frequências de citações, quer sejam de autores ou documentos, e das frequências de coocorrência (cocitação) entre as mesmas, possibilitando a visualização de um domínio. Entende-se visualização como a “capacidade ou ato de formar na mente imagens visuais de coisas que não estão à vista, ou a imagem daí resultante” (HOUAISS, 2001, p.2872). Assim, a visualização

permite a transformação de conceitos e fenômenos em imagens perceptíveis ou visíveis mentalmente, como também a apreensão de determinados dados e de algumas relações e das estruturas subjacentes que não estão apresentadas de forma explícita.

Uma citação é tomada como um indicador objetivo e claro da comunicação científica, que permite a identificação de grupos de cientistas e suas publicações, com a finalidade de evidenciar os pesquisadores de maior impacto de uma área, apontando seus paradigmas, procedimentos metodológicos pertinentes, bem como os pesquisadores de “vanguarda” que constroem o novo conhecimento na área. Desse modo, a análise de citação contribui para o entendimento de uma comunidade científica ao identificar os pesquisadores com maior impacto na área e dar visibilidade às referências teóricas que a sustentam, bem como seus conceitos, objetos e métodos (OLIVEIRA; GRACIO; SILVA, 2010).

A partir da década de 1960, estudos de citação dão um salto qualitativo em decorrência do avanço das tecnologias e dos primeiros índices de citação criados por Eugene Garfield. Destacam-se, nesses primórdios, os trabalhos de Solla Price, entre eles Price (1965).

A “análise de citações mapeia a comunicação científica” (VANZ; CAREGNATO, 2003, p.248), dá indicadores de como está ocorrendo a comunicação de uma área do conhecimento, contribuindo para a construção da rede de relações, e explicita a comunicação e o relacionamento entre os pesquisadores.

O estudo de cocitações, derivado da análise de citações, trata da frequência com que dois autores ou documentos são citados de forma conjunta na produção científica de uma área. Esses estudos começam pela seleção dos objetos cocitados, que podem ser documentos, autores ou periódicos, entre outras possibilidades.

Spinak (1996) registra que a análise de cocitação é uma ferramenta limitada pela seleção inicial dos autores, que deverá ser adequada e representativa, observando que aqueles citados em maior número de trabalhos, isto é, os “clássicos”, fazem parte, em geral, desta seleção.

Segundo Miguel, Moya Anegon e Herreno Solana (2008), a análise de cocitação, seja de documentos, autores, jornais, especialidades ou campos de conhecimento, produz representações válidas da estrutura intelectual de um domínio científico. Ainda segundo estes autores, sua premissa fundamental estabelece que quando dois ou mais documentos, autores ou periódicos são citados juntos, em um terceiro trabalho posterior, existe, pelo menos na perspectiva do autor citante, uma similaridade de assunto entre os citados, e que quanto maior a frequência de cocitação, mais próxima a relação entre os mesmos. Assim, na análise de cocitação, os autores citados juntos podem apresentar, pelo menos na perspectiva dos autores citantes, além da similaridade, complementaridade, sobreposição ou mesmo contraposição de ideias.

Segundo Gmür (2003), a frequência de cocitação entre dois autores determina também como a estrutura de conhecimento de uma área é percebida pelos pesquisadores. O conjunto de referências dos autores citados conjuntamente produz uma representação válida da estrutura intelectual de um domínio científico.

Outros estudiosos dão suporte teórico à questão de análise de citação e cocitação. Henry Small, um dos primeiros estudiosos em análise de cocitação, trata da frequência conjunta de documentos citados em uma literatura posterior. Segundo o autor, “quando os cientistas concordam quanto ao que constitui a literatura relevante a priori, incluindo o que é significativo naquela literatura, eles estão de fato definindo as estruturas de suas comunidades” (SMALL, 2004, p.72). E prossegue: a “estrutura da ciência é gerada por padrões de co-reconhecimento”

(SMALL, 2004, p.71). Assim, “quando documentos são cocitados, autores citantes estão atribuindo co-reconhecimento bem como criando uma associação de significados” (SMALL, 2004, p.76). O autor observa ainda que a cocitação pode ser usada para estabelecer o núcleo da literatura dentro de determinado tema ou área em particular.

Segundo White e McCain (1998), estudos relativos à análise de cocitação de autores (ACA) tiveram como precursores White e Griffith, em 1981, e sua principal função é identificar os autores influentes e mostrar suas inter-relações, a partir das citações registradas.

McCain (1990) destaca que a ACA tem como premissa analisar a estrutura intelectual de uma determinada área, campo científico ou conjunto de pesquisadores, mostrando sua estrutura social e cognitiva e seu domínio de investigação. Afirma que a ACA parte do pressuposto que pesquisadores com problemas de pesquisa semelhantes citam fontes informacionais similares e próximas. Nesse sentido, ACA pode ser definida como análise de um conjunto de atores, autores ou pesquisadores organizados estruturalmente em rede (social e cognitiva) de uma determinada comunidade científica.

Quatro passos principais em uma análise de cocitação de autores são apresentados por McCain (1990). Os dois primeiros são: construção de uma matriz com as frequências de cocitação entre os autores e a conversão desta matriz em uma matriz de índice de cocitação relativa (ou normalizada), entre eles o índice de Jaccard e o cosseno de Salton. Desse modo, os estudos de cocitação de autores baseiam-se em matrizes de valores (absolutos ou relativos) que mensuram a coocorrência entre os autores citados. São matrizes de proximidade, isto é, matrizes que indicam quão próximos ou quão distantes estão dois autores, sob o olhar dos pesquisadores citantes.

O índice de Jaccard (IJ) é definido como a frequência de cocitação de dois autores dividida pelo número total de citações recebidas por eles. Em termos matemáticos, a fórmula pode ser expressa por:

$$IJ = \frac{cocit(X, Y)}{cit(X) + cit(Y) - cit(X \cap Y)}$$

em que:

$cocit(X, Y)$ = número de trabalhos em que os autores X e Y foram citados conjuntamente (cocitados);

$cit(X)$ = número de trabalhos em que o autor X foi citado;

$cit(Y)$ = número de trabalhos em que o autor Y foi citado;

$cit(X \cap Y)$ = intersecção entre o conjunto de trabalhos em que X foi citado e o conjunto de trabalhos em que Y foi citado, o que é equivalente a $cocit(X, Y)$.

O Cosseno de Salton (CS) é definido como o quociente da frequência de cocitação de dois autores (X e Y) pela raiz quadrada do número de citações recebidas pelo autor X multiplicado pelo número de citações recebidas pelo autor Y. A formulação matemática é apresentada por:

$$CS = \frac{cocit(X,Y)}{\sqrt{cit(X).cit(Y)}}$$

em que:

$cocit(X,Y)$ = número de trabalhos em que os autores X e Y foram cocitados;

$cit(X)$ = número de trabalhos em que o autor X foi citado;

$cit(Y)$ = número de trabalhos em que o autor Y foi citado.

Os índices de cocitação relativa ou normalizada mensuram a intensidade relativa de coocorrência dos autores citados em uma literatura científica, funcionando assim como medida de proximidade entre os autores cocitados. Derivam das frequências de cocitação, por meio da conversão dos valores inteiros (observados) em valores que relativizam a intensidade de cocitação de dois autores, em relação à quantidade de citações recebidas individualmente, por cada um dos autores cocitados. Esses índices padronizam a ACA aplicada às diferentes áreas do conhecimento, ao apresentarem valores variando entre zero e um: quanto mais próximo de um, mais similares (altamente cocitados) são os dois autores; quanto mais próximo de zero, mais distante (menos intensa) é a associação entre eles, na percepção do conjunto de autores citantes. Valor zero para IJ ou CS indica ausência de cocitação entre os dois autores; valor um significa que estes foram citados exatamente nos mesmos trabalhos, o que pode ser decorrente tanto de os autores serem coautores nos trabalhos citados como de terem proximidade teórico-metodológica.

Para alguns autores, como Leydesdorff e Vaughan (2006), as matrizes de cocitação – com as frequências (absolutas) observadas – podem ser usadas na construção das redes de cocitação ou dos mapas de escalonamento multidimensional (*multidimensional scaling* – MDS), entretanto elas contêm menos informação que matrizes de índice de cocitação relativa. Small e Sweeney (1985) apontam que as análises de cocitação baseadas em índices de cocitação relativa proporcionam uma representação mais compreensiva dos agrupamentos dos referentes teóricos que formam a estrutura intelectual de uma comunidade científica, tanto em grandes como em pequenas áreas de pesquisa, eliminando a dependência do padrão (tamanho do rol de referências) de citação dos artigos das diferentes áreas.

As matrizes de índice de cocitação relativa ou normalizada têm sido, preferencialmente, utilizadas nas análises estatísticas multivariadas dos estudos de cocitação por oferecerem resultados mais refinados e estarem sustentadas em teorias de probabilidade.

Assim, embora as frequências absolutas de cocitação sejam usadas e aceitas como ferramentas úteis na mensuração da conectividade científica entre os autores cocitados, os índices relativos de cocitação – entre eles IJ e CS – podem revelar, de modo mais consistente, aspectos subjacentes da estrutura intelectual de uma área científica, que não são identificáveis por análises de cocitação baseadas em frequências absolutas.

Small (1973) foi o primeiro a apontar o índice de Jaccard com medida para relativizar/normalizar os dados da matriz de frequências de cocitação em uma análise de cocitação. Small e Sweeney (1985) utilizaram o cosseno de Salton a fim de oferecer uma medida de intensidade relativa de cocitação em uma análise de cocitação, como uma alternativa ao índice de Jaccard. Estes autores

argumentam que o Cosseno de Salton apresenta-se mais adequado para a análise de cocitação que o índice de Jaccard, no processo de relativização/normalização das frequências de cocitação, por tratar com elos entre artigos com alta citação e baixa citação de modo mais efetivo. Hamers et al. (1989) apontam que a fórmula do cosseno de Salton frequentemente produz uma medida de similaridade relativa em estudos de cocitação que é duas vezes o número daquele obtido pelo índice de Jaccard.

De modo geral, segundo os estudiosos da área, essas duas medidas (cosseno de Salton e Índice de Jaccard) apresentam comportamentos similares, uma delas com valor aproximadamente igual ao dobro da outra. Todavia, também segundo Hamers et al., em casos de cocitação envolvendo dois pesquisadores, um com alta citação e outro com baixa citação, o cosseno de Salton representa mais efetivamente o elo de intensidade de cocitação entre eles.

Na área da Ciência da Informação, diversos estudos aplicados de análise de cocitação têm sido realizados, principalmente em âmbito internacional, tanto baseados nas frequências de cocitação como baseados nos índices de cocitação relativa. Dentre eles, estão os estudos de White e McCain (1998), Moya-Anegón, Jiménez-Contreras e Moneda-Corrochano (1998), Zhao e Strotmann (2008), Ma, Daí, Ni e Li (2009), Pinto, Barquín e Gonzalez (2006), Liberatore, Herrero-Solana e Guimarães (2007), Pinto, Efrain-García, Barquín e González (2007), Pinheiro e Silva (2008), Grácio, Oliveira e Mattos (2009) e Oliveira, Grácio e Silva (2010), com o objetivo de evidenciar os pesquisadores mais destacados nas áreas estudadas e apresentar indicadores que contribuam para o delineamento do panorama da atividade científica.

A importância dos dados oriundos das matrizes de cocitação e de índice de cocitação relativa, para a visualização e compreensão das estruturas subjacentes de uma comunidade científica, tem dado origem a uma significativa literatura na área da Ciência da Informação. Alguns autores, como Ahlgren, Harneving e Rousseau (2003) e Eck e Waltman (2008), posicionam-se favoravelmente ao uso do cosseno de Salton; outros, como Leydesdorff (2008), posicionam-se pelo índice de Jaccard como medida mais representativa de similaridade para normalização nos estudos de análise de cocitação de autores.

Nesse contexto, a literatura não apresenta um posicionamento conclusivo sobre a questão apontada e, por consequência, o estabelecimento da metodologia adequada para a análise de cocitação de autores (ACA) não se encontra plenamente consolidado, demandando muitos estudos a fim de se ratificar a necessidade de conversão da frequência de cocitação em índices de cocitação relativa (VANZ, 2009).

Procedimentos metodológicos

O percurso metodológico desta pesquisa seguiu duas etapas. A primeira refere-se a estudos teóricos, a fim de contribuir para a melhor compreensão conceitual-metodológica da fundamentação da análise de cocitação de autores (ACA), que envolve a análise de citação e a de cocitação, nos mais atuais enfoques e tratamentos, bem como seus indicadores de proximidade e a importância da sua normalização, passando-se assim aos estudos de Cosseno de Salton (CS) e de Índice de Jaccard (IJ). Na segunda fase, ocorreu o levantamento dos dados a partir de súmula constituída das 110 pesquisas apresentadas no período de 2007 a 2011, no GT7 nos ENANCIBs, onde constaram o título do trabalho e o conjunto de referências.

Para cada trabalho, foi levantado o rol de referências e feito o devido tratamento: as autorias múltiplas foram desdobradas para que se contemplasse o nome de todos os autores citados; houve substituição dos traços sublineares pelo nome do próprio autor; e foram realizadas uma triagem e uma limpeza geral nas citações, que, colocadas em ordem alfabética, possibilitaram verificar aqueles citados em maior número de trabalhos.

No período analisado, foram identificados 1242 pesquisadores citados, para um total de 2003 referências, destacando-se que 884 (71,2%) pesquisadores foram citados apenas uma vez. Tendo em vista o grande volume de dados, foram considerados autores de maior impacto e visibilidade aqueles citados em pelo menos 10% dos artigos analisados (correspondente a serem citados em pelo menos 10 trabalhos), em um total de 32 pesquisadores.

Com a seleção dos 32 pesquisadores mais citados, constituiu-se o grupo alvo participante da análise de cocitação de autores desta pesquisa. Em um primeiro momento, construiu-se, no software Excel, uma matriz assimétrica 110x32 de ocorrência de citação entre autores citantes (por linha: 110 linhas) e autores citados (por coluna: 32 colunas). Para cada coluna (autor citado), foi registrado o valor um na linha correspondente ao trabalho (autor(es) citante(s)) em que ele foi citado e zero quando o autor não foi citado pelo trabalho.

A Figura 1 ilustra um recorte da matriz 110x32, na qual se observa, por exemplo, que Braga foi citada nos trabalhos de Amarante e de Barbatho, aparecendo assim o valor um na célula referente ao cruzamento da coluna de Braga com as linhas referentes aos trabalhos de Amarante, C.M.V.G. e de Barbatho, R.R.G. Os zeros presentes na coluna de Braga, G. significam que esta autora não foi citada por aqueles trabalhos (autores citantes). Por outro lado, o autor Katz, J. S. não foi citado por nenhum dos oito autores citantes ilustrados no recorte da matriz.

Figura 1: Recorte da matriz de 110 trabalhos citantes e 32 autores citados.

	C	D	E	F	G	H	I	J	K	L	M	N	O
1	autores citantes	BARRETO, A	BOURDIEU, P.	BRAGA, G. M.	BUFREM, L. S.	CAREGNATO, S. E.	GOMEZ, M.L.G.	HJORLAND, B.	KATZ, J. S.	KOBASCHI, N. Y.	LE COADIC, Y.F.A.	LETA, J.	MACIAS-CHAPULA, C. A.
2	ALVES, B.H.; OLIVEIRA, E.F.T.	0	0	0	0	0	0	1	0	0	0	0	0
3	AMARANTE, C.M.V.G.	0	0	1	0	0	0	0	0	0	0	0	0
4	ANDRETTA, P.I.S.; SILVA, E.G.; SILVEIRA, J.P.B.; RAMOS, R.C.	0	0	0	1	0	0	0	0	1	0	0	0
5	BARBATHO, R.R.G.	0	1	1	0	0	0	0	0	0	0	0	0
6	BRAMBILLA, S.D.S.; STUMPF, I.R.C.	0	0	0	0	0	0	0	0	0	0	0	0
7	BUFREM, L.S.; FREITAS, J.L.; BREDA, S.M.	0	0	0	0	0	0	0	0	0	0	0	0
8	CANCHUMANI, R.M.L.; LETA, J.; DE FIGUEIREDO, A.M.	0	0	0	0	0	0	1	0	0	0	1	0
9	CHALHUB, T.; RIBEIRO, L.V.P.	0	0	0	0	0	0	0	0	0	0	0	0

A partir dessa matriz de ocorrência (110x32), construiu-se também no software Excel, usando a função matemática “SOMARPRODUTO”, a matriz quadrada e simétrica de tamanho 32x32 autores citados, cujo resultado constituiu a frequência de cocitação entre os autores mais citados, nos trabalhos analisados. A Figura 2 apresenta um recorte da matriz 32x32 de cocitação entre esses autores, após aplicação da referida função. A título de exemplo de construção, o valor 6 de frequência de cocitação entre Kobashi, N. e Bufrem, L. foi resultado da aplicação da função SOMARPRODUTO(G2:G111;L2:L111), na qual o intervalo “G2:G111” significa a coluna de presença de citação da autora Bufrem, como pode ser visto na Figura 1, e o intervalo “L2:L111” indica a coluna de citação da autora Kobashi. nos 110 trabalhos, também conforme a Figura 1. Como resultado da aplicação da referida fórmula matemática, o valor 20 presente na célula do cruzamento da “linha” e “coluna” da autora Bufrem. constitui o número total de trabalhos em que a autora foi citada, dentre os 110 trabalhos analisados. Os demais elementos da diagonal principal apresentam raciocínio análogo.

Figura 2: Recorte da matriz de frequência de cocitação.

	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY
1	Matriz de cocitação	BARRETO, A.	BOURDIEU, P.	BRAGA, G. M.	BUFREM, L. S.	CAREGNATO, S. E.	GOMEZ, M. L. G.	HJORLAND, B.	KATZ, J. S.	KOBASCHI, N. Y.	LE COADIC, Y. F. A.	LETA, J.	MACIAS-CHAPULA, C. A.	MARTIN
2	BARRETO A. de A.	10	1	1	2	1	4	0	0	1	1	2	2	0
3	BOURDIEU, P.		10	3	1	1	4	0	0	1	1	2	0	0
4	BRAGA, G. M.			10	3	1	3	1	0	0	2	0	0	0
5	BUFREM, L. S.				20	3	1	2	0	6	4	0	4	0
6	CAREGNATO, S. E.					11	1	3	1	1	2	0	1	1
7	GOMES, M. N. G.						10	0	0	0	2	1	1	0
8	HJORLAND, B.							10	0	1	2	1	1	0
9	KATZ, J. S.								10	0	0	5	1	10

As duas matrizes com os valores do Cosseno de Salton e de Índice de Jaccard, obtidas a partir da matriz de frequência absoluta de cocitação, também foram calculadas no software Excel, a partir da aplicação das fórmulas já apresentadas no referencial teórico.

Para se analisar comparativamente os resultados das três matrizes de proximidade, sendo a primeira com valores absolutos e as duas últimas normalizadas - Cosseno de Salton (CS) e Índice de Jaccard (IJ)-, foram evidenciadas as diferenças entre os valores presentes nas mesmas e os valores mais altos, a fim de observar se as matrizes tendem a apontar para as mesmas principais associações e para os mesmos destaques em relação à proximidade entre os autores cocitados. Para melhor contextualizar as tendências presentes em cada uma das três matrizes, foi calculada

a média dos valores: para a primeira matriz, a média dos valores absolutos; e para as outras duas, a média dos valores normalizados, o que permitiu verificar a distância entre os destaques feitos e as respectivas tendências centrais (médias).

Apresentação e análise dos dados

A Tabela 1 apresenta os 32 pesquisadores citados em pelo menos 10 dos trabalhos analisados. Na quase totalidade, esse conjunto de autores é formado por pesquisadores próprios e consignados da área de Ciência da Informação, o que indica que o grupo de autores selecionados para a análise de cocitação apresenta consistência teórica, especialmente no âmbito do Brasil, aqui representado pelos 22 pesquisadores brasileiros, que ratificam a existência de um grupo consolidado do ponto de vista teórico-metodológico.

Destaca-se ainda que, dos 110 trabalhos analisados, somente 9 (8,2%) deles não fizeram menção a nenhum dos autores do grupo de 32 pesquisadores mais citados, o que sugere a representatividade do grupo selecionado para a análise de cocitação, como o núcleo da literatura científica na temática.

Tabela 1: Pesquisadores citados em pelo menos 10 trabalhos.

Ajustar coluna da tabela	Nº de trabalhos em que foi citado	Pesquisador	Nº de trabalhos em que foi citado
MUELLER, S. P. M. (Brasil)	40	ROUSSEAU, R. (Bélgica)	12
MEADOWS, A. J. (Reino Unido)	31	SPINAK, E. (Uruguai)	12
PINHEIRO, L. V. R. (Brasil)	23	TARGINO, M. das G. (Brasil)	12
BUFREM, L. S. (Brasil)	20	CAREGNATO, S. E. (Brasil)	11
MENEGHINI, R. (Brasil)	14	GÓMEZ, M. N. G. de (Brasil)	11
STUMPF, I. R. C. (Brasil)	15	PACKER, A.L. (Brasil)	11
VANZ, S. A. de S. (Brasil)	15	ZIMAN, J.M. (Reino Unido)	11
SANTOS, R. N. M. (Brasil)	14	BARRETO, A. de A. (Brasil)	10
MACIAS-CHAPULA, C. A. (México)	13	BOURDIEU, P. (França)	10
MUGNAINI, R. (Brasil)	13	BRAGA, G. M. (Brasil)	10
NORONHA, D. P. (Brasil)	13	HJØRLAND, B. (Dinamarca)	10
OLIVEIRA, M de. (Brasil)	13	KATZ, J.S. (Reino Unido)	10
PRICE, D. J. de S. (E.U.A.)	17	KOBASCHI, N. Y. (Brasil)	10
VELHO, L. (Brasil)	13	LE COADIC, Y. F. (França)	10
LETA, J. (Brasil)	12	MARTIN, B.R. (Reino Unido)	10
POBLACIÓN, D. A. (Brasil)	12	VANTI, N. A. P. (Brasil)	10

Em relação às temáticas nas quais foram citados, destacam-se: comunicação científica, periódicos científicos, financiamento de pesquisa, canais de comunicação, domínio

epistemológico, infraestrutura de pesquisa, disseminação da informação, produção, estrutura e fluxo do conhecimento, ensino superior, tendências metodológicas, incluindo análises e procedimentos da bibliometria, cientometria, webometria, como análise de citação, indicadores em CT&I, redes, redes sociais, redes eletrônicas, entre outras compreendidas pela Produção e Comunicação da Informação. Recorrendo-se aos referentes teóricos adotados, destaca-se que a análise de um domínio é o estudo de uma comunidade discursiva, expressa pelos seus pesquisadores e temáticas trabalhadas.

Apresenta-se, a seguir, a matriz de frequência de cocitação absoluta dos 32 pesquisadores mais citados, cuja média de cocitação é igual a 2,2 cocitações entre os pares de autores analisados. Destacam-se 7 valores de frequência de cocitação absoluta.

Tabela 2: Matriz com as frequências absolutas de cocitação entre os 32 autores.

Matriz de frequência de cocitação	BARRETO, A.	BOURDIEU, P.	BRAGA, G. M.	BUFREM, L. S.	CAREGNATO, S.	GOMEZ, M.L.G.	HJORLAND, B.	KATZ,	KOBASCHI, N.	LE COADIC, Y.	LETA, J.	M-CHAPULA, C.	MARTIN	MEADOWS, A.	MENEGHINI, R.	MUELLER, S.	MUGNAINI, R.	NORONHA, D.	OLIVEIRA, M. de	PACKER, A. L.	PINHEIRO, L. V.	POBLACIÓN, D.	PRICE, D.S.	ROUSSEAU, R.	SANTOS, R. N.	SPINAK, E.	STUMPF, I.C.	TARGINO, M.G.	VANTI, N.	VANZ, S.	VELHO, L.	ZIMAM, J.	
BARRETO, A.	10	1	1	2	1	4	0	0	1	1	2	2	0	2	0	5	1	2	5	0	6	3	1	0	2	1	1	1	1	1	0	2	
BOURDIEU, P.	10	3	1	1	4	0	0	1	1	2	0	0	4	2	3	1	2	2	2	2	2	2	2	1	0	1	0	3	1	1	2	2	
BRAGA, G. M.	10	3	1	3	1	0	0	2	0	0	0	6	0	6	0	1	2	1	6	1	5	2	0	2	2	2	1	1	2	4			
BUFREM, L. S.	20	3	1	2	0	6	4	0	4	0	4	0	2	1	8	3	4	5	1	4	3	2	2	7	3	2	2	3	3	1	1		
CAREGNATO, S.	11	1	3	1	1	2	0	1	1	3	1	2	1	5	2	1	3	2	2	1	2	3	4	3	1	7	0	1					
GOMES, M. N. G.	10	0	0	0	2	1	1	0	2	1	4	0	3	3	1	6	3	1	0	0	0	0	2	1	1	1	1	2					
HJORLAND, B.	10	0	1	2	1	1	0	2	0	3	0	1	0	0	1	2	4	0	6	1	0	1	3	0	1								
KATZ, J. S.	10	0	0	5	1	10	3	4	6	0	0	0	2	2	0	6	1	0	1	1	0	0	1	2	0								
KOBASCHI, N.	10	3	0	2	0	0	1	4	4	2	2	1	1	3	1	1	8	2	3	1	1	2	2	0									
LE COADIC, Y.	10	0	1	0	3	0	5	2	1	2	0	3	1	2	1	3	1	1	2	2	0	0											
LETA, J.	12	0	5	4	5	4	1	0	0	3	2	0	5	0	0	1	2	0	1	2	0	1	2	0									
M-CHAPULA, C.	13	1	3	2	3	3	1	2	1	3	0	1	2	4	5	2	1	5	1	2	0												
MARTIN, B. R.	10	3	4	6	0	0	0	2	2	0	6	1	0	1	1	0	0	1	2	0													
MEADOWS, A. J.	31	7	19	3	3	5	5	11	4	9	3	1	2	7	6	3	4	6	10														
MENEGHINI, R.	13	7	4	3	1	10	2	2	3	2	0	1	4	2	2	1	4	2	1	4	2												
MUELLER, S.	40	8	6	7	6	13	6	9	2	6	0	7	7	1	2	7	10																
MUGNAINI, R.	13	2	4	4	2	3	1	1	6	1	3	4	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NORONHA, D.	13	3	3	2	6	1	0	4	1	4	4	0	3	2	2																		
OLIVEIRA, M. de	13	1	9	7	2	0	4	1	1	3	1	3	1	3	2	2																	
PACKER, A. L.	11	1	2	1	1	0	0	4	2	1	1	3	3																				
PINHEIRO, L. V.	23	5	6	0	3	1	1	4	1	3	1	4																					
POBLACIÓN, D.	12	1	0	3	0	3	3	0	3	3	0	3	3																				
PRICE, D. S.	17	2	1	3	3	1	1	3	2	4																							
ROUSSEAU, R.	12	1	6	2	1	4	3	2	1																								
SANTOS, R. N.	14	3	3	3	2	2	2	0																									
SPINAK, E.	12	0	1	2	2	2	0																										
STUMPF, I.C.	15	5	3	7	4	3																											
TARGINO, M.G.	12	1	2	2	1																												
VANTI, N.	10	2	1	0																													
VANZ, S.	15	1	2																														
VELHO, L.	13	2																															
ZIMAM, J.	11																																

Na matriz apresentada na Tabela 2, as maiores frequências de cocitações foram 19, 13, 11 e 10, e a maior delas (19) ocorre na cocitação entre os autores Mueller e Meadows. Em seguida, a segunda maior cocitação (13) ocorre entre Mueller e Pinheiro, e com 11 ocorrências aparece a cocitação entre Meadows e Pinheiro. Ainda, há 4 ocorrências de 10 cocitações, a saber: Mueller e Ziman, Meadows e Ziman, Meneghini e Paker e entre Katz e Martin. Registre-se que a alta frequência de cocitação entre Katz e Martin se deve exclusivamente a um artigo realizado em coautoria.

Em relação à cocitação decorrente de coautorias dentro do grupo em análise, destacam-se Packer e Meneghini, coautores de cinco trabalhos. As pesquisadoras Población e Noronha são coautoras de três trabalhos.

A Tabela 3, a seguir, apresenta os valores do Cosseno de Salton, obtidos a partir dos dados de frequência absoluta de cocitação presentes na Tabela 2, para o qual se obteve índice médio de CS igual a 0,15. Destacam-se nove valores de índice de cocitação normalizado (CS) acima de 0,50.

Tabela 3: Matriz com os valores de Cosseno de Salton entre os 32 autores.

Cosseno de Salton	BARRETO, A.	BOURDIEU, P.	BRAGA, G. M.	BUFREM, L.	CAREGNATO	GOMEZ, M. N.	HJORLAND, B.	KATZ,	KOBASCHI, N.	LE COADIC, Y.	LETA, J.	MACIAS-CHAPULA	MARTIN	MEADOWS, A.	MENEGHINI, R.	MUELLER, S. P. M.	MUGNAINI, R.	NORONHA, D. P.	OLIVEIRA, M. de	PACKER, A. L.	PINHEIRO, L. V.	POBLACIÓN, D. A.	PRICE, D. J. de S.	ROUSSEAU, R.	SANTOS, R. N.	SPINAK, E.	STUMPF, I. C.	TARGINO, M. G.	VANTI, N.	VANZ, S.	VELHO, L.	ZIMAM, J.
BARRETO, A.	1	0,10	0,10	0,14	0,10	0,40	0,00	0,00	0,10	0,10	0,18	0,18	0,00	0,11	0,00	0,25	0,09	0,18	0,44	0,00	0,40	0,27	0,08	0,00	0,17	0,09	0,08	0,09	0,10	0,08	0,00	0,19
BOURDIEU, P.		1,00	0,30	0,07	0,10	0,40	0,00	0,00	0,10	0,10	0,18	0,00	0,00	0,23	0,18	0,15	0,09	0,18	0,18	0,19	0,13	0,18	0,15	0,09	0,00	0,09	0,00	0,27	0,10	0,08	0,18	0,19
BRAGA, G. M.			1,00	0,21	0,10	0,30	0,10	0,00	0,00	0,20	0,00	0,00	0,00	0,34	0,00	0,30	0,00	0,09	0,18	0,10	0,40	0,09	0,38	0,18	0,00	0,18	0,16	0,18	0,10	0,08	0,18	0,38
BUFREM, L. S.				1,00	0,20	0,07	0,14	0,00	0,42	0,28	0,00	0,25	0,00	0,08	0,06	0,28	0,19	0,25	0,31	0,07	0,19	0,19	0,11	0,13	0,42	0,19	0,12	0,13	0,21	0,17	0,06	0,07
CAREGNATO, S.					1,00	0,10	0,29	0,10	0,10	0,19	0,00	0,08	0,10	0,16	0,08	0,10	0,08	0,42	0,17	0,09	0,19	0,17	0,15	0,09	0,16	0,26	0,31	0,26	0,10	0,54	0,00	0,09
GOMES, M. N. G.						1,00	0,00	0,00	0,00	0,20	0,09	0,09	0,00	0,11	0,09	0,20	0,00	0,26	0,26	0,10	0,40	0,27	0,08	0,00	0,00	0,00	0,00	0,18	0,10	0,08	0,09	0,19
HJORLAND, B.							1,00	0,00	0,10	0,20	0,09	0,09	0,00	0,11	0,00	0,15	0,00	0,09	0,00	0,00	0,00	0,09	0,15	0,37	0,00	0,55	0,08	0,00	0,10	0,24	0,00	0,10
KATZ, J. S.								1,00	0,00	0,00	0,46	0,09	1,00	0,17	0,35	0,30	0,00	0,00	0,00	0,19	0,13	0,00	0,46	0,09	0,00	0,09	0,08	0,00	0,00	0,08	0,18	0,00
KOBASCHI, N.									1,00	0,30	0,00	0,18	0,00	0,00	0,09	0,20	0,35	0,18	0,18	0,10	0,07	0,27	0,08	0,09	0,68	0,18	0,24	0,09	0,10	0,16	0,18	0,00
LE COADIC, Y.										1,00	0,00	0,09	0,00	0,17	0,00	0,25	0,18	0,09	0,18	0,00	0,20	0,09	0,15	0,09	0,25	0,09	0,08	0,18	0,20	0,16	0,00	0,00
LETA, J.											1,00	0,00	0,37	0,21	0,40	0,18	0,08	0,00	0,00	0,26	0,12	0,00	0,35	0,00	0,00	0,08	0,15	0,00	0,09	0,15	0,16	0,00
M-CHAPULA, C.												1,00	0,09	0,15	0,15	0,13	0,23	0,08	0,15	0,08	0,17	0,00	0,07	0,16	0,30	0,40	0,14	0,08	0,44	0,07	0,15	0,00
MARTIN, B. R.													1,00	0,17	0,26	0,25	0,00	0,00	0,00	0,10	0,13	0,00	0,38	0,09	0,00	0,09	0,08	0,00	0,00	0,08	0,18	0,00
MEADOWS, A.														1,00	0,35	0,54	0,15	0,15	0,25	0,27	0,41	0,21	0,39	0,16	0,05	0,10	0,32	0,31	0,17	0,19	0,30	0,54
MENEGHINI, R.															1,00	0,31	0,31	0,23	0,08	0,84	0,12	0,16	0,20	0,16	0,00	0,08	0,29	0,16	0,18	0,07	0,31	0,17
MUELLER, S.																1,00	0,35	0,26	0,31	0,29	0,43	0,27	0,35	0,09	0,25	0,00	0,29	0,32	0,05	0,08	0,31	0,48
MUGNAINI, R.																	1,00	0,15	0,31	0,33	0,12	0,24	0,07	0,08	0,44	0,08	0,21	0,32	0,09	0,14	0,08	0,08
NORONHA, D. P.																		1,00	0,23	0,25	0,12	0,48	0,07	0,00	0,30	0,08	0,29	0,32	0,00	0,21	0,15	0,17
OLIVEIRA, M. de																			1,00	0,08	0,52	0,56	0,13	0,00	0,30	0,08	0,07	0,24	0,09	0,21	0,15	0,17
PACKER, A. L.																				1,00	0,06	0,17	0,07	0,09	0,00	0,00	0,31	0,17	0,10	0,08	0,25	0,27
PINHEIRO, L. V.																					1,00	0,30	0,30	0,00	0,17	0,06	0,05	0,24	0,07	0,16	0,06	0,25
POBLACIÓN, D. A.																						1,00	0,07	0,00	0,23	0,00	0,22	0,25	0,00	0,22	0,24	0,26
PRICE, D. de S.																							1,00	0,14	0,06	0,21	0,19	0,07	0,08	0,19	0,13	0,29
ROUSSEAU, R.																								1,00	0,08	0,50	0,15	0,08	0,37	0,22	0,16	0,09
SANTOS, R. N.																									1,00	0,23	0,21	0,23	0,17	0,14	0,15	0,00
SPINAK, E.																										1,00	0,00	0,08	0,18	0,15	0,16	0,00
STUMPF, I. C.																											1,00	0,37	0,24	0,47	0,29	0,23
TARGINO, M. G.																												1,00	0,09	0,15	0,16	0,09
VANTI, N.																													1,00	0,16	0,09	0,00
VANZ, S.																														1,00	0,07	0,16
VELHO, L.																															1,00	0,17
ZIMAM, J.																																1,00

O maior índice de Cosseno de Salton observado foi entre Katz e Martin, com valor igual a um, maior valor possível para este indicador. Isto significa que os dois autores foram citados exatamente nos mesmos trabalhos em virtude de o único trabalho citado de ambos os autores ser um trabalho em coautoria entre os dois. A análise realizada a partir das frequências absolutas de cocitação mostrou que a cocitação entre estes dois autores, embora significativa, não foi a de maior destaque.

Os autores Meneghini e Packer apresentam alto índice de CS (0,84), indicando a alta proximidade de assunto entre eles, decorrente de coautorias entre os mesmos e de trabalhos com temáticas similares. Essa cocorrência de citação dos dois autores também foi destacada na análise anterior da tabela de frequência absoluta, todavia com menor intensidade.

Observam-se Santos e Kobashi, com o terceiro maior valor de CS (0,68), indicando uma forte proximidade entre eles, porém é uma cocitação que não aparece destacada na tabela de frequência absoluta. Essa diferença de intensidade é decorrente do fato de o CS considerar (relativizar) a intensidade da cocitação em relação ao total de citação (presença) dos autores no universo em estudo e não de forma absoluta. Neste caso, a frequência de cocitação entre Santos e Kobashi, igual a 8, em relação ao total de citações recebidas por ambos, é mais significativa que a frequência de cocitação, entre Muller e Meadows, igual a 19, em relação ao total de citação recebido por estes.

Comportamento similar ocorre entre Población e Oliveira, entre Caregnato e Vanz, entre Oliveira e Pinheiro e entre Spinak e Rousseau, com índices de CS salientados, todavia com valores sem destaque na análise para frequências de valores absolutos de cocitação.

A intensidade de cocitação entre Meadows e Mueller e entre Meadows e Ziman foi destacada tanto em termos relativos (CS) como absolutos.

Por outro lado, as frequências absolutas de cocitação entre Mueller e Pinheiro, entre Mueller e Ziman e entre Meadows e Pinheiro foram destacadas, todavia sem apresentar valor notável para CS.

A seguir, a Tabela 3 apresenta os valores normalizados, utilizando o Índice Jaccard (IJ), a partir dos dados da Tabela 2, com IJ médio igual a 0,09. Destacam-se nove valores de índice de cocitação normalizado (IJ) acima de 0,30.

O Índice de Jaccard (IJ) entre Katz e Martin é o máximo valor para este índice, ou seja, igual a um, assim como na tabela para os valores de CS.

Dos nove índices de Jaccard mais altos, seis deles não apresentam correspondência de destaque nos valores de frequência absoluta, a saber: Oliveira e Población, Spinak e Hjørland, Caregnato e Vanz, Spinak e Rousseau, Oliveira e Pinheiro e Población e Noronha.

Dois índices de Jaccard significativos também não apresentam correspondência de destaque nos valores de Cosseno de Salton, a saber: Spinak e Hjørland e entre Población e Noronha.

Observa-se que os valores de IJ são menores que os índices observados para CS, com exceção destes índices para Katz e Martin, que em ambos os casos teve valor normalizado igual a um. Comparados os valores de CJ e IJ presentes nas Tabelas 3 e 4, ratifica-se a observação de Hamers et al (1989) quando afirmam que CS tende a produzir uma medida de similaridade relativa que se aproxima, em valor, a duas vezes o número daquele obtido pelo índice de Jaccard.

Comparando os valores de proximidade nas três matrizes analisadas, destacam-se somente as intensidades de cocitação, tanto em termos absolutos como em termos normalizados, entre três pares de autores, a saber: Katz e Martin, entre Mueller e Meadows e entre Ziman e Meadows.

Tabela 4: Matriz com os Índices de Jaccard entre os 32 autores

Índice de Jaccard	BARRETO, A.	BOURDIEU, P.	BRAGA, G. M.	BUFREM, L. S.	S.	GOMEZ, M. L. G.	HJORLAND, B.	KATZ,	KOBASCHI, N.	LE COADIC, Y.	LETA, J.	M-CHAPULA, C.	MARTIN	MEADOWS, A.	MENEGHINI, R.	MUELLER, S.	MUGNAINI, R.	NORONHA, D.	OLIVEIRA, M.	PACKER, A. L.	PINHEIRO, L. V.	POBLACIÓN, D.	PRICE, D. S.	ROUSSEAU, R.	SANTOS, R. N.	SPINAK, E.	STUMPF, I. R. C.	TARGINO, M.	VANTI, N.	VANZ, S.	VELHO, L.	ZIMAM, J.
BARRETO, A.	1,00	0,05	0,05	0,07	0,05	0,25	0,00	0,00	0,05	0,05	0,10	0,10	0,00	0,05	0,00	0,11	0,05	0,10	0,28	0,00	0,22	0,16	0,04	0,00	0,09	0,05	0,04	0,05	0,05	0,04	0,00	0,11
BOURDIEU, P.		1,00	0,18	0,03	0,05	0,25	0,00	0,00	0,05	0,05	0,10	0,00	0,00	0,11	0,10	0,06	0,05	0,10	0,10	0,11	0,06	0,10	0,08	0,05	0,00	0,05	0,00	0,16	0,05	0,04	0,10	0,11
BRAGA, G. M.			1,00	0,11	0,05	0,18	0,05	0,00	0,00	0,11	0,00	0,00	0,00	0,17	0,00	0,14	0,00	0,05	0,10	0,05	0,22	0,05	0,23	0,10	0,00	0,10	0,09	0,10	0,05	0,04	0,10	0,24
BUFREM, L. S.				1,00	0,11	0,03	0,07	0,00	0,25	0,15	0,00	0,14	0,00	0,04	0,03	0,15	0,10	0,14	0,18	0,03	0,10	0,10	0,06	0,07	0,26	0,10	0,06	0,07	0,11	0,09	0,03	0,03
CAREGNATO, S.					1,00	0,05	0,17	0,05	0,05	0,11	0,00	0,04	0,05	0,08	0,04	0,04	0,04	0,26	0,09	0,05	0,10	0,10	0,08	0,05	0,09	0,15	0,18	0,15	0,05	0,37	0,00	0,05
GOMES, M. N.						1,00	0,00	0,00	0,00	0,11	0,05	0,05	0,00	0,05	0,05	0,09	0,00	0,15	0,15	0,05	0,22	0,16	0,04	0,00	0,00	0,00	0,00	0,10	0,05	0,04	0,05	0,11
HJORLAND, B.							1,00	0,00	0,05	0,11	0,05	0,05	0,00	0,05	0,00	0,06	0,00	0,05	0,00	0,00	0,00	0,05	0,08	0,22	0,00	0,38	0,04	0,00	0,05	0,14	0,00	0,05
KATZ, J. S.								1,00	0,00	0,00	0,29	0,05	1,00	0,08	0,21	0,14	0,00	0,00	0,00	0,11	0,06	0,00	0,29	0,05	0,00	0,05	0,04	0,00	0,00	0,04	0,10	0,00
KOBASCHI, N.									1,00	0,18	0,00	0,10	0,00	0,00	0,05	0,09	0,21	0,10	0,10	0,05	0,03	0,16	0,04	0,05	0,50	0,10	0,14	0,05	0,05	0,09	0,10	0,00
LE COADIC, Y.										1,00	0,00	0,05	0,00	0,08	0,00	0,11	0,10	0,05	0,10	0,00	0,10	0,05	0,08	0,05	0,14	0,05	0,04	0,10	0,11	0,09	0,00	0,00
LETA, J.											1,00	0,00	0,22	0,10	0,25	0,08	0,04	0,00	0,00	0,15	0,06	0,00	0,21	0,00	0,00	0,04	0,08	0,00	0,05	0,08	0,09	0,00
M-CHAPULA, C.												1,00	0,05	0,07	0,08	0,06	0,13	0,04	0,08	0,04	0,09	0,00	0,03	0,09	0,17	0,25	0,08	0,04	0,28	0,04	0,08	0,00
MARTIN, B. R.													1,00	0,08	0,15	0,11	0,00	0,00	0,00	0,05	0,06	0,00	0,23	0,05	0,00	0,05	0,04	0,00	0,00	0,04	0,10	0,00
MEADOWS, A.														1,00	0,19	0,37	0,07	0,07	0,13	0,14	0,26	0,10	0,23	0,08	0,02	0,05	0,18	0,16	0,08	0,10	0,16	0,31
MENEGHINI, R.															1,00	0,15	0,18	0,13	0,04	0,71	0,06	0,09	0,11	0,09	0,00	0,04	0,17	0,09	0,10	0,04	0,18	0,09
MUELLER, S.																1,00	0,18	0,13	0,15	0,13	0,26	0,13	0,19	0,04	0,13	0,00	0,15	0,16	0,02	0,04	0,15	0,24
MUGNAINI, R.																	1,00	0,08	0,18	0,20	0,06	0,14	0,03	0,04	0,29	0,04	0,12	0,19	0,05	0,08	0,04	0,04
NORONHA, D.																		1,00	0,13	0,14	0,06	0,32	0,03	0,00	0,17	0,04	0,17	0,19	0,00	0,12	0,08	0,09
OLIVEIRA, M. de																			1,00	0,04	0,33	0,39	0,07	0,00	0,17	0,04	0,04	0,14	0,05	0,12	0,08	0,09
PACKER, A. L.																				1,00	0,03	0,10	0,04	0,05	0,00	0,00	0,18	0,10	0,05	0,04	0,14	0,16
PINHEIRO, L. V.																					1,00	0,17	0,18	0,00	0,09	0,03	0,03	0,13	0,03	0,09	0,03	0,13
POBLACIÓN, D.																						1,00	0,04	0,00	0,13	0,00	0,13	0,14	0,00	0,13	0,14	0,15
PRICE, D. S.																							1,00	0,07	0,03	0,12	0,10	0,04	0,04	0,10	0,07	0,17
ROUSSEAU, R.																								1,00	0,04	0,33	0,08	0,04	0,22	0,13	0,09	0,05
SANTOS, R. N.																									1,00	0,13	0,12	0,13	0,09	0,07	0,08	0,00
SPINAK, E.																										1,00	0,00	0,04	0,10	0,08	0,09	0,00
STUMPF, I. R. C.																											1,00	0,23	0,14	0,30	0,17	0,13
TARGINO, M.																												1,00	0,05	0,08	0,09	0,05
VANTI, N.																													1,00	0,09	0,05	0,00
VANZ, S.																														1,00	0,04	0,08
VELHO, L.																															1,00	0,09
ZIMAM, J.																																1,00

Finalizando, acrescenta-se que os índices normalizados diferenciam a intensidade de proximidade não discriminada pelas frequências absolutas, a exemplo da frequência absoluta de cocitação igual a cinco, observada entre três pares de autores - Mueller e Barreto, Price e Martin, Vanti e Macias-Chapula- que, quando normalizados, apresentam índices para CS e IJ distintos e que distinguem a intensidade de proximidade entre os mesmos.

Conclusões

Esta pesquisa se propôs a realizar um estudo da contribuição dos indicadores de proximidade absolutos e normalizados para análise de cocitação de autores, visando compreender um domínio do conhecimento.

Foram analisados os indicadores de proximidade absolutos e relativos, especificamente Cosseno de Salton e Índice de Jaccard, comparando sua contribuição para o entendimento da proximidade temática entre os autores basilares citados e co-reconhecidos pela comunidade citante do GT7 da ANCIB, no período em estudo.

Evidenciaram-se as diferenças entre os valores presentes nas matrizes destes três indicadores e seus valores mais altos, a fim de observar se elas tendem a apontar para as mesmas principais associações e para os mesmos destaques em relação à proximidade entre os autores cocitados.

Destacaram-se e analisaram-se 7 valores de frequência de cocitação absoluta, com variação entre 10 e 19 cocitações, sobressaindo-se o total de cocitações entre os autores Mueller e Meadows. Quanto ao Cosseno de Salton, destacam-se nove valores de índice de cocitação normalizado (CS), variando de 0,50 a 1, com os autores Katz e Martin, com valor máximo para este indicador. Em relação ao índice de Jaccard (IJ), nove índices de cocitação normalizado (IJ) foram destacados, com valores entre 0,31 e 1, também com Katz e Martin, tendo o valor máximo.

Somente três pares de autores cocitados foram destacados simultaneamente na análise dos três indicadores de proximidade analisados quanto à intensidade de cocitação, ou seja, salientaram-se tanto em termos absolutos como em termos normalizados, a saber: Katz e Martin, entre Mueller e Meadows e entre Ziman e Meadows.

Observou-se que alguns destaques nas cocitações são decorrentes de coautorias dentro do grupo em análise, porém a grande maioria por proximidade temática.

Acrescenta-se que os índices normalizados diferenciaram a intensidade de proximidade não discriminada por frequências absolutas, distinguindo a intensidade de proximidade entre os autores segundo a presença de cada autor no estudo. Esse comportamento é decorrente do fato de os indicadores normalizados, CS e IJ, exprimirem uma grandeza (intensidade de cocitação) em relação a um determinado contexto (total de citação ou presença de cada autor no universo analisado).

Finalizando, espera-se que este estudo ofereça subsídios teóricos e metodológicos, contribuindo para o fortalecimento da compreensão e uso da Análise de Cocitação de Autores, como procedimento para a Análise de Domínio (AD), ao pretender elucidar de forma teórica e metodológica os indicadores normalizados subjacentes à análise de cocitação.

Artigo recebido em 06/02/2013 e aprovado em 22/03/2013.

Referências

AHLGREN, P.; JARNEVING, B.; ROUSSEAU, R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, v. 54, n. 6, p. 550-560, 2003.

ECK, N. J. V.; WALTMAN, L. Appropriate similarity measures for author co-citation analysis. *Journal of the American Society for Information Science and Technology*, v. 59, n. 10, p. 1653-1661, 2008.

GLÄNZEL, W. *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. 2003. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.5311&rep=rep1&type=pdf>>. Acesso em: 09 jan. 2011.

GMÜR, M. Co-citation analysis and the search for invisible colleges: a methodological evaluation. *Scientometrics*, v. 57, n. 1, p. 27-57, 2003.

GRÁCIO, M. C. C; OLIVEIRA, E. F. T.; MATOS, G. I. Visibilidad de los investigadores en el tema Estudios Métricos: análisis de citación y cocitación en los periódicos del SciELO. *IBERSID*, 2009.

HAMERS, L. et al. Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula. *Information Processing & Management*, v. 25, n. 3, p. 315-318, 1989.

HJØRLAND, B. Domain analysis in information science: eleven approaches-traditional as well as innovative. *Journal of Documentation*, v. 58, n. 4, p. 422-462, 2002a.

_____. Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, v. 53, n. 4, p. 257-270, 2002b.

_____; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, v. 46, n. 6, pp. 400-425, 1995.

HOUAISS, A. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Editora Objetiva Ltda, 2001.

KERR, E. S. *Ketib: um processo de representação de informações para textos complexos*. 2003. Dissertação (Mestrado Profissional em Ciência da Computação)- Instituto de Computação da Universidade Estadual de Campinas, 2003. Disponível em: <<http://libdigi.unicamp.br/document/?code=vtls000311557>>. Acesso em: 02 fev. 2006.

LEYDESDORFF, L. On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, v. 59, n. 1, p.77-85, 2008.

_____; VAUGHAN, L. Co-occurrence matrices and their applications in information science: extending ACA to the web environment. *Journal of the American Society for Information Science and Technology*, v. 57, n. 12, p.1616-1628, 2006.

LIBERATORE, G.; HERRERO-SOLANA, V.; GUIMARÃES, J. A. C. Análise bibliométrica do periódico brasileiro Ciência da Informação durante o período 2000-2004. *Brazilian Journal of Information Science*, v. 1, n. 2, p. 3-21, July/Dec. 2007. Disponível em: <<http://www.bjis.unesp.br/pt/>>. Acesso em: 11 mar. 2009.

MA, R. et al. An author co-citation analysis of information science in China with Google Scholar search engine, 2004-2006. *Scientometrics*, v. 81, n. 1, p. 33-46, 2009.

MAI, J.-E. Analysis in indexing: document and domain centered approaches. *Information Processing and Management*, v. 41, p. 599-611, 2005.

MCCAIN, K. Mapping author intellectual space: a technical overview. *Journal of the American Society for Information Science*. v. 41, n. 66, p. 433-443, 1990.

MIGUEL, S.; MOYA-ANEGON, F.; HERRERO-SOLANA, V. A new approach to institutional domain analysis: Multilevel research fronts structure. *Scientometrics*, v. 74, n. 3, p. 331-344, 2008.

MOYA-ANEGON, F.; JIMÉNEZ-CONTRERAS, E.; MONEDA-CORROCHANO, M. D. Research fronts in library and information science in Spain (1985-1994). *Scientometrics*, v. 42, n. 2, p. 229-246, 1998.

OLIVEIRA, E. F. T; GRACIO, M. C. C.; SILVA, A.C.C. Investigadores de mayor visibilidad en Organización y Representación del Conocimiento: un estudio desde el análisis de cocitaciones. *Scire*, v. 16, p. 39-46, 2010.

_____. Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tema “estudos métricos” na base Scopus. *Perspectivas em Ciência da Informação*, v. 16, p. 16-28, 2011.

PINHEIRO, L. V.; SILVA, E. L. da. As redes cognitivas na ciência da informação brasileira: um estudo nos artigos científicos publicados nos periódicos da área. *Ciência da Informação*, v. 37, p. 75-87, 2008.

PINTO, A. L.; BARQUÍN, B. A. R.; GONZÁLEZ, J. A. M. Análisis de citación de la revista Ciência da Informação del Ibict. *Ciência da Informação*, v. 35, n. 3, p. 153-165, set./dez. 2006.

_____ et al. Indicadores científicos na literatura em bibliometria e cientometria através das redes sociais. *Brazilian Journal of Information Science*, v. 1, n. 1, p. 58-76, 2007.

PRICE, D. J. de S. Networks of scientific papers. *Science*, v. 149, p. 510-515, 1965. Disponível em: <<http://www.garfield.library.upenn.edu/papers/pricenetworks1965.pdf>>. Acesso em: 01 abr. 2013.

SMALL, H. Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*, v. 24, n. 4, p. 265-269, 1973.

_____ ; SWEENEY, E. Clustering the Science Citation Index using co-citations: 1. a comparison of methods. *Scientometrics*, v.7, n. 3-6, p.391-409, 1985.

_____. On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics*, v. 60, n.1, p. 71-79, 2004.

SPINAK, E. *Dicionário enciclopédico de bibliometria, cienciometria e informetria*. Caracas: UNESCO, 1996.

VANZ, S. A. de S.; CAREGNATO, S. E. Estudos de citação: uma ferramenta para entender a comunicação científica. *Em Questão*, v. 9, n. 2, p.295-307, jul./dez. 2003.

_____. *As redes de colaboração científica no Brasil (2004-2006)*. Tese (Doutorado em Ciência da Informação)- Universidade Federal do Rio Grande do Sul, 2009.

WHITE, H.D.; GRIFFITH, B. Author co-citation: a literature measure of intellectual structure. *Journal of the American Society for Information Science &Technology*, v. 32, n. 2, p. 163-171, 1981.

_____; MCCAIN, K.W. Visualizing a discipline: an author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science*, v. 49, n. 4, p. 327-355, 1998.

ZHAO, D. Z.; STROTMANN, A. Information science during the first decade of the web: an enriched author co-citation analysis. *Journal of the American society for Information Science and Technology*, v. 59, n. 6, p. 916-937, 2008.