

Procedimentos metodológicos

Maria Cláudia Cabrini Grácio

SciELO Books / SciELO Livros / SciELO Libros

GRÁCIO, M. C. C. Procedimentos metodológicos. In: *Análises relacionais de citação para a identificação de domínios científicos: uma aplicação no campo dos Estudos Métricos da Informação no Brasil* [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020, pp. 115-144. ISBN: 978-65-86546-12-5. Available from:

<http://books.scielo.org/id/tx83k/pdf/gracio-9786586546125-04.pdf>.

<https://doi.org/10.36311/2020.978-65-86546-12-5>.



All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

4. PROCEDIMENTOS METODOLÓGICOS

Este estudo é de natureza bibliográfica e documental, elaborada a partir de material - publicações científicas e currículos Lattes - já publicado, em formato impresso e digital disponível online, para a revisão teórica e coleta de dados. É descritiva, pois tem por objetivo principal descrever e estudar características de domínios científicos no campo dos EMI. Utilizam-se dados quantificáveis e procedimentos bibliométricos e estatísticos, com os resultados analisados e interpretados, a fim de obter elementos para a compreensão da questão em estudo, isto é, a composição de domínios científicos no campo dos Estudos Métricos da Informação no Brasil, a partir da identidade e imagem de citação de um conjunto de seus pesquisadores. Apresentam-se, a seguir, os procedimentos e recursos utilizados para o seu desenvolvimento.

4.1 Seleção dos autores do universo de pesquisa

Nos dias 7 e 8 de fevereiro de 2018, por meio da consulta à base de dados Scopus, iniciou-se o levantamento dos autores atuantes no campo científico dos Estudos Métricos da Informação, seja abordando o campo como objeto de pesquisa seja como método de análise metateórica para

alguma área do conhecimento, a fim de compor o universo de pesquisa.

Dada a importância das bases de dados Scopus e Web of Science (WoS) para a análise e mensuração da ciência gerada nos diferentes campos do conhecimento de visibilidade impacto internacional, a opção pela base Scopus, como fonte para o levantamento dos dados, decorre da sua maior cobertura multidisciplinar e geográfica dos periódicos (SANTA; HERRERO-SOLANA, 2010; LIMA; VELHO; FARIA, 2012; MOED, 2017).

Apesar de o Science Citation Index, fundado por Eugene Garfield em 1964 e publicado pelo Institute for Scientific Information (atualmente, Web of Science da Clarivate), durante muitos anos, ter sido a única base de dados com cobertura multidisciplinar de periódicos revisados por pares, a base de dados Scopus lançada em 2004 pela Elsevier, em 2007, continha cerca de 40% mais artigos do que a WoS (MOED, 2017). Além disso, em relação a outras bases de dados, entre elas a SciELO, a base Scopus permite a análise do impacto e abrangência da ciência produzida, por meio da contagem das citações recebidas e suas respectivas origens.

Embora não seja intenção realizar uma menção exaustiva da ampla literatura relativa à cobertura geográfica e temática alcançada pela base de dados Scopus, observa-se que ela tem sido constantemente objeto de análise de diversos estudos, como pode ser observado em Codina (2005), Fingerman (2005) e Jacsó (2005), outros com comparações da cobertura, acessibilidade e usabilidade em relação à base de dados WoS, como atestam os estudos de Fingerman (2006) e Jacsó (2010) e, também

comparações com o Google Scholar (BAKKALBASI et al., 2006). Além disso, para países de língua latina, a base Scopus é a melhor fonte de indexação de periódicos no mundo que contempla o maior número de áreas da ciência, com ampla representação multidisciplinar, não restringindo seu conteúdo a uma só área, como as bases de dados especializadas (LUCAS; PINTO; LARA, 2013). Consideram-se esses estudos suficientes para garantir que análises baseadas em seu uso oferecem resultados significativos para a compreensão da ciência de países da América Latina, e em especial do Brasil (MIGUEL; OLIVEIRA; GRÁCIO, 2016).

Considerando que um domínio é um corpo de conhecimento, dinâmico, sempre em desenvolvimento, nunca congelado no tempo e no espaço e dependente de uma teoria coerente e socialmente institucionalizada (HJØRLAND, 2017), optou-se neste estudo por realizar um recorte temporal relativo aos últimos 15 anos (2003 a 2017) a fim de se identificar e visualizar os domínios contemporâneos distinguíveis no campo dos EMI.

4.1.1 Seleção dos autores do universo de pesquisa: busca nos campos TITLE e KEY da Busca Avançada da Scopus

Para a formação do universo de pesquisa, inicialmente, levantaram-se na base Scopus, os autores com produção de artigos científicos em EMI, por meio do recurso BUSCA AVANÇADA (Advanced search), utilizando: os termos³⁴ de busca relativos aos EMI (APÊNDICE A) nos campos TITLE(título) e KEY(palavra-chave) e a filiação brasileira dos autores (Brasil or Brazil)

no campo AFFILCOUNTRY (país de afiliação). Restringiu-se o período de publicação para 2013 a 2017 (15 últimos anos) e a tipologia documental para artigo. Por essa busca, recuperou-se um total de 1659 artigos, os quais foram exportados da base Scopus em 2 arquivos formato CSV Excel: o primeiro com os campos autores, título do documento, ano, título da fonte, abstract, palavras-chave do autor e palavras-chave adicionais do editor; e o segundo somente com as autorias dos 1659 artigos. Explicase a opção de exportação em dois arquivos: o segundo com somente o nome dos autores, para facilitar o manuseio dos dados para a síntese dos autores por maior quantidade de artigos publicados, uma vez que a Scopus registra todos os dados em uma única coluna, com os coautores e outros metadados separados por vírgula, dificultando a separação do rol de autores do restante dos metadados; o primeiro (completo), para auxiliar na identificação dos autores na base Scopus, na segunda fase do levantamento do universo de pesquisa, quando se recupera a produção total dos autores do universo, nos casos de homônimos.

Embora a base Scopus apresente, junto com o conjunto de artigos resultante da busca, uma síntese relativa a alguns dos metadados da produção, e entre estes, a lista de autores responsáveis³⁵ por aquele corpus de artigos, com seus respectivos números de artigos produzidos, optou-se por realizar a contagem das produções de cada autor por meio do software SPSS, a fim de reunir a produção dos autores com nomes presentes na base de dados com mais de uma forma.

Para essa atividade, antes de importar o arquivo Excel no software SPSS, foi necessário gerar uma nova lista com todas as coautorias presentes nos artigos “desmembradas”, a fim de se obter os dados dos autores em formato de variável, conforme padrão do software SPSS, ou seja: uma única coluna para a variável (autoria) e uma linha para cada um dos autores de cada autoria ou coautoria presente no conjunto dos 1659 artigos recuperados. Gerou-se, assim, um arquivo Word com uma tabela com uma coluna e 5581 linhas relativas às autorias desmembradas, presentes nos 1659 artigos, o que significou uma média de ~3,4 autores por artigo. Essa tabela foi importada para um arquivo Excel, para o subsequente tratamento no SPSS.

A partir da análise realizada no SPSS, obteve-se uma tabela de distribuição de frequência dos autores do conjunto 1659 artigos, em ordem decrescente de artigos publicados.

A fim de reunir as contagens de artigos publicados pelos autores que apresentaram mais de uma forma de registro dos seus nomes, reordenou-se esta tabela de frequência em ordem alfabética dos nomes.

Após esse procedimento, reuniram-se as frequências de publicação dos autores, de forma manual. Comparou-se, então, esta tabela com a síntese oferecida pela Scopus para o metadado autor, em que se observou que na tabela gerada pelo SPSS havia mais autores com altas frequências de artigos publicados que aquela presente na síntese da Scopus. Por outro lado, todos os autores registrados com alta frequência na Scopus apareceram com igual

ou maior frequência na tabela gerada pelo SPSS e revisada manualmente.

Como resultado deste procedimento, identificaram-se 4374³⁶ autores distintos responsáveis pelos 1659 artigos. Destes autores, 3743 (~86%) foram responsáveis por um único artigo publicado.

Aplicando-se a Lei do Elitismo de Price, para o total de 4374 autores, obteve-se uma elite de ~66 autores. O rol dos 71 autores decorrentes da aplicação da Lei do Elitismo encontra-se na Tabela 1.

Tabela 1 - Autores mais produtivos a partir da busca dos artigos em EMI nos campos Title e Key-words.

Autor	# artigos	Autor	# artigos	Autor	# artigos
Notare, M. S. M. A.	31	Ensslin, S. R.	7	Erdmann, A. L.	6
Mari, J. J.	14	Gerolin, J.	7	Falavigna, A.	6
Nabout, J. C.	14	Laender, A. H. F.	7	Herculano, R. D.	6
Meneghini, R.	13	Moro, M. M.	7	Teixeira, R. K. C.	6
Grácio, M. C. C.	12	Mugnaini, R.	7	Vasconcelos	6
Leta, J.	12	Oliveira, M. C. L.	7	Barbastefano, R. G.	5
de Camargo, K. R.	11	Quirino, I. G.	7	Borges, P. P.	5
Ferreira, L. M.	11	Souza, D. O.	7	Botelho, N. M.	5
Martinelli, D. R.	11	Nahas, F. X.	7	Coeli, C. M.	5
Moreira, L. F. P.	11	Braile, D. M.	6	dos Santos, W.	5
de Faria, L. I. L.	10	Bufrem, L. S.	6	Figueira, I.	5
de Oliveira, E. F. T.	10	Colosimo, E. A.	6	Lima, L. S.	5
Martelli Júnior, H.	10	do Amaral, R. M.	6	Malhado, A. C. M.	5
Rodrigues, R. S.	10	Gandia, R. M.	6	Mendlowicz, M.	5
Rocha e Silva, M.	10	Gregolin, J. A. R.	6	Procianoy, R. S.	5
Wainer, J.	10	Machado, K. B.	6	Quintella, R. H.	5
Ensslin, L.	9	Mena-Chalco, J. P.	6	Razzouk, D.	5
Milanez, D. H.	9	Oliveira, E. A.	6	Santos, N. C. F.	5
Packer, A. L.	9	Steil, A. V.	6	Yamamoto, O. H.	5
Vanz, S. A. S.	9	Stumpf, I. R. C.	6	Amancio, D. R.	5
Bini, L. M.	8	Thomaz, S. M.	6	de Lima, D. V. M.	5
de Carvalho, M. M.	8	Barreto, M. L.	6	Ferreira, M. P.	5
Albuquerque, U. P.	7	Bressan, R. A.	6	Jabbour, C. J. C.	5
dos Santos, R. N. M.	7	Caramelli, B.	6		

Fonte: elaborada pela autora (2019).

4.1.2 Seleção dos autores do universo de pesquisa: busca no campo ABSTRACT e não TITLE e KEY em Busca Avançada da Scopus

Partindo do pressuposto que a explicitação de termos relativos aos EMI no título ou nas palavras-chave dos artigos pode evidenciar um domínio distinto daquele formado por autores que evidenciam estar familiarizados com os EMI, ao mencionarem seu uso no resumo do artigo, todavia não consideraram relevante o seu destaque no título ou nas palavras-chave, os quais identificam o foco do conhecimento produzido, realizou-se um segundo levantamento na base Scopus. Neste levantamento, usando os mesmos termos de busca relativos aos EMI presentes no Apêndice A, levantaram-se os autores com produção de artigos científicos abordando os EMI por mencioná-los no resumo, mas não no título ou palavras-chave. A busca foi realizada, por meio do recurso BUSCA AVANÇADA (Advanced search) com o termo de busca: (ABS(termos relativos aos EMI listados no Apêndice A) AND NOT (TITLE(termos em Apêndice A) OR KEY(termos em Apêndice A)) AND AFFILCOUNTRY(Brasil or Brazil). Assim como para a busca descrita na seção 4.1.2, restringiu-se o período de publicação para 2013 a 2017 e artigo para o tipo documental.

Por essa busca, recuperou-se um total de 1445 artigos, para os quais foi dado tratamento análogo ao descrito na seção 4.1.1, tanto em termos de exportação dos dados, como em termos do tratamento no Excel e no Word e análise no SPSS, com revisão manual das frequências de artigos para reunião de registros distintos de nomes de um mesmo autor. A partir

dos procedimentos realizados, observaram-se 6452 autorias relativas aos 1659 artigos, o que significou uma média de ~4,5 autores por artigo.

Após a revisão manual, obteve-se uma tabela com 5888 autores distintos responsáveis pelos 1445 artigos, com respectivos totais de artigos publicados. Destes autores, 5461 (~93%) foram responsáveis por um único artigo publicado.

A aplicação da Lei do Elitismo, para o total de 5888 autores, resultou em uma elite científica composta por ~77 autores. Todavia, considerando-se a grande dispersão de autores com 3 artigos publicados (equivalente ao total de artigos publicados pelo 77º autor, em ordem decrescente de publicações), restringiu-se o universo de autores para aqueles com pelo menos 4 artigos publicados, equivalente ao conjunto composto pelos 26 autores com maior produção, apresentados na Tabela 2.

Observa-se que há uma interseção entre as Tabelas 1 e 2, formada pela presença simultânea de seis autores, a saber: Ensslin L., Ensslin S.R., Leta J., Mari J.J., Mena-Chalco J.P. e dos Santos R.N.M.

Considera-se relevante complementar que outros 28 autores presentes na Tabela 1 (autores com maior produção, observada a partir da busca pelos termos relativos aos EMI no título e nas palavras-chave) também apareceram entre os autores responsáveis pelos 1445 artigos recuperados a partir da busca pelos termos exclusivamente no resumo das publicações, porém com produção menor que o critério adotado para constar na Tabela 2 (pelo menos 4 artigos). Por outro lado, 11 autores com

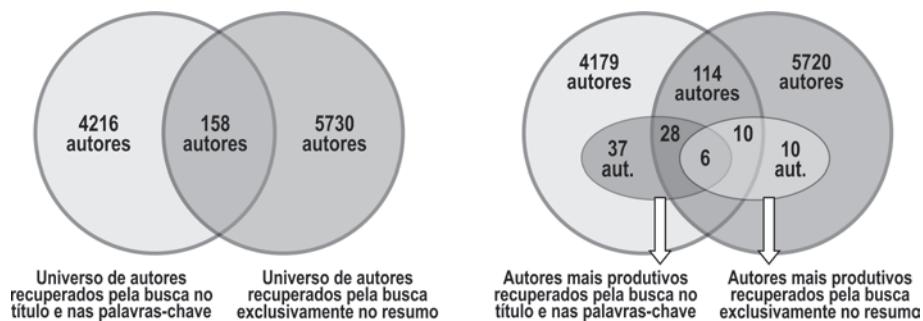
produção significativa em EMI, recuperada pela busca no resumo (presentes na Tabela 2), também encontram-se entre os autores responsáveis pelos 1659 artigos recuperados pela busca nos campos Título ou Palavras-chave das publicações (Apêndice A), porém com menos de 5 artigos publicados nesta. Constata-se, assim, uma significativa interseção entre os dois universos de autores, que pode ser visualizada no diagrama de Venn na Figura 16, em que se observa também a proporcionalidade das elites em análise em relação ao universo total de autores recuperados da Scopus.

Tabela 2 - Autores mais produtivos a partir da busca no campo Resumo (ABS).

Autor	# artigos	Autor	#artigos
Ensslin, L.	7	da Cruz, I. C. F.	4
Ensslin, S. R.	7	do Prado, M. L.	4
Leta, J.	7	dos Santos, H. M.	4
Schraiber, L. B.	6	Flores, D.	4
Dewes, H.	5	Gomes, R.	4
dos Santos, R. N. M.	5	Linhares, M. B. M.	4
Gomes, M. J.	5	Mena-Chalco, J. P.	4
Mari, J. J.	5	Moreira, T. M. M.	4
Minayo, M. C. S.	5	Padilha, M. I.	4
Vieira, N. F. C.	5	Rosa, R. F. M.	4
Andrade, A.	4	Talamini, E.	4
Cecatti, J. G.	4	Ugrinowitsch, H.	4
Costa, L. O. P.	4	Zen, P. R. G.	4

Fonte: elaborada pela autora (2019).

Figura 16 - Interseção entre os dois universos de autores e proporcionalidade das elites em relação ao universo total de autores recuperados da Scopus.



4.1.3 Universo dos autores da pesquisa

Considerando que a busca realizada a partir dos termos de busca utilizados pode deixar de fora parte da produção científica dos autores presentes nas Tabelas 1 e 2, que trata dos EMI. Como passo seguinte na direção de se estabelecer o universo de autores participantes deste estudo, realizaram-se novas buscas na base Scopus, tomando agora como princípio norteador de busca a ferramenta “busca por autor” (*Author Search*).

Nessas buscas, considerando os objetivos propostos neste estudo, adotou-se critério análogo ao presente em White (2001), a fim de estabelecer o universo final de autores brasileiros atuantes em EMI a serem analisados, a saber: autores com produção de pelo menos dois trabalhos científicos (artigo, capítulo de livro ou trabalho completo em conferência) na condição de primeiro autor, por partir do pressuposto que o primeiro autor é o responsável pela obra, incluindo o que é nela

citado. White (2001) aponta que este pressuposto não é confiável em autorias múltiplas oriundas de grandes equipes de pesquisa, frequente nas ciências naturais. O critério de ter pelo menos duas publicações é decorrente da necessidade de se observar a recitação (citação em mais de um artigo) realizada pelos autores, a qual evidencia sua identidade.

Para cada um dos autores presentes nas Tabelas 1 e 2, recuperou-se, a partir da busca na base Scopus, para o período em estudo (2003-2017): Total de documentos publicados, total de artigos como primeiro autor, total de artigos como primeiro autor na temática EMI³⁷ e total de citações recebidas neste último conjunto de artigos. Nestes levantamentos, restringiram-se as produções a: artigo em periódicos; artigos apresentados em conferências e capítulos de livro. Adotou-se este critério devido ao fato de que estas tipologias efetivamente apresentam referências bibliográficas como componente indispensável. Como resultado desta segunda fase de busca na base Scopus, o universo de autores que participam deste estudo é composto por 33 autores, sendo 32 deles oriundos da Tabela 1 e 1 oriundo da Tabela 2. A Tabela 3 apresenta o rol desses autores, organizados em ordem decrescente de artigos publicados no campo dos EMI, na condição de primeiro autor do estudo, no período de 2003 a 2017.

Tabela 3 - Universo dos 33 autores analisados com quantidade (#) de artigos publicados em EMI, na condição de primeiro autor (2003 a 2017)

Autor	# artigos busca nos títulos e palavras-chave	# documentos no período	# artigos como 1º autor	# artigos como 1º autor em EMI
Amancio, D. R.	5	40	26	24
Milanez, D. H.	9	21	12	11
Vasconcelos, S. M. R.	6	19	11	11
Meneghini, R.	13	32	10	10
Wainer, J.	10	78	23	9
Erdmann, A. L.	6	188	19	9
Mugnaini, R.	7	15	8	8
Nabout, J. C.	14	70	19	7
Teixeira, R. K. C.	6	32	10	7
Rodrigues, R. S.	10	19	7	7
Leta, J.	12	39	8	6
Herculano, R. D.	6	32	12	5
Ensslin, L.	9	58	11	5
Grácio, M. C. C.	12	16	5	5
Razzouk, D.	5	39	14	4
de Camargo, K. R.	11	83	12	4
Mari, J. J.	14	246	7	4
Oliveira, E. A.	6	75	7	4
Packer, A. L.	9	22	5	4
Barreto, M. L.	6	325	19	3
de Carvalho, M. M.	8	87	12	3
Caramelli, B.	6	138	7	3
Vanz, S. A. S.	9	12	3	3
do Prado, M. L.*	4	44	3	3
Ferreira, L. M.	11	414	18	2
Ensslin, S. R.	7	57	6	2
Mena-Chalco, J. P.	6	29	6	2
Mendlowicz, M.	5	104	5	2
Quintella, R. H.	5	15	4	2
de Oliveira, E. F. T.	10	14	3	2
Barbastefano, R. G.	5	10	2	2
Gandia, R. M.	6	7	2	2
Stumpf, I. R. C.	6	11	2	2

* único autor originalmente presente somente na Tabela 2 (busca por Resumo).

Para cada um dos 33 autores participantes deste estudo, seu conjunto de artigos relativo aos EMI, como primeiro autor, foi armazenado na ferramenta Gerenciamento de listas (*Lists*) da Base Scopus, a fim de recuperar nos procedimentos seguintes da pesquisa: seus respectivos conjuntos de referências (e autores) para a análise da identidade de citação e domínios, segundo suas proximidades de autores citados, por meio da Análise de Acoplamento Bibliográfico; conjunto de artigos citantes dos artigos publicados por cada autor (33) do estudo.

Para o conjunto de 33 autores componentes do universo de pesquisa foram realizadas análises a partir de 3 enfoques, os quais deram origem às seções 5.1, 5.2 e 5.3 do capítulo dos resultados: 1- identificação dos autores do universo de estudo; 2- identificação dos domínios em EMI, a partir da identidade de citação dos pesquisadores; 3- identificação dos domínios em EMI, a partir da imagem de citação dos pesquisadores; 4 - identificação dos domínios no campo dos EMI no Brasil, por meio da junção dos resultados obtidos a partir das perspectivas da identidade e da imagem de citação dos pesquisadores.

A fim de possibilitar uma melhor visualização dos procedimentos metodológicos adotados para o tratamento do objetivo desta pesquisa, apresenta-se o percurso desenvolvido separado em seções, as quais estão pareadas com a apresentação dos resultados.

4.2 Procedimentos para a caracterização dos pesquisadores do universo de estudo

Esta seção apresenta os procedimentos utilizados para a obtenção dos resultados apresentados na seção 5.1 do capítulo 5. Para cada um dos 33 autores componentes do estudo, recuperaram-se as áreas de atuação declaradas em seus respectivos currículos disponibilizados na Plataforma Lattes. A partir das áreas de atuação declaradas, identificou-se qual o posicionamento do autor em relação à Figura 6 do Capítulo 2, ou seja, se atua nos EMI, em áreas afins, ou em outras áreas do conhecimento.

A seguir, a fim de relativizar a produção científica em EMI destes autores, na condição de primeiro autor, em relação ao total dos seus artigos publicados também como primeiro autor, calculou-se a porcentagem relacionando estes dois indicadores.

Para cada autor, observou-se também se ele já configurava entre os autores de destaque em EMI nos estudos anteriores de Grácio e Oliveira (2012) e Oliveira (2013), relativos a três distintas esferas de disseminação do conhecimento científico - base Scopus, ENANCIB, EBBC - por meio da averiguação da presença do seu nome na Figura 4, presente no Capítulo 2.

As informações relativas a esses indicadores qualitativos foram reunidas àquelas já registradas na planilha Excel relativas aos indicadores de produção científica do período 2003 a 2017, a saber: quantidade (#) de artigos recuperados na busca por termos dos EMI; quantidade de documentos (todas as tipologias - artigos, capítulos de livro, trabalhos

em anais de eventos, editoriais, revisões, notas, entre outros) publicados; quantidade de artigos publicados como 1º autor; quantidade de artigos como 1º autor em EMI.

Os autores foram então organizados, com seus respectivos indicadores, em três grupos consonantes à Figura 6 do capítulo 2: atuantes em EMI; atuantes em áreas afins aos EMI; atuantes em outras áreas do conhecimento.

Para os indicadores com mensuração intervalar, a fim de avaliar as diferenças estatísticas entre as médias dos três grupos, foram aplicados testes ANOVA com um fator, utilizando o software SPSS, versão 21, com nível de significância igual a 0,05. Para os indicadores que apresentaram diferença estatística significativa no teste ANOVA, utilizaram-se os testes de comparações múltiplas *post hoc* de TuKey e Duncan, a fim de se identificar as diferenças significantes entre os três grupos.

4.3 Procedimentos para a identificação da identidade de citação e do acoplamento bibliográfico dos pesquisadores

A visualização dos domínios no campo dos EMI, a partir das suas proximidades teórico metodológicas, foi realizada por meio de duas etapas de análise bibliométrica subseqüentes: 1 - identificação da identidade de citação de cada um dos 33 autores componentes do universo de pesquisa; 2 - análise de acoplamento bibliográfico destes autores. Os procedimentos metodológicos adotados para a primeira etapa foram baseados em White (2001) e para a segunda em Zhao e

Strotmann (2008, 2014), conforme descrição a seguir.

Para identificar a identidade de citação de cada um dos 33 pesquisadores do universo de pesquisa, foram realizados dois procedimentos consecutivos:

a. para cada artigo presente na pasta do pesquisador analisado, armazenada no Gerenciamento de listas da Scopus, levantou-se o conjunto de autores distintos citados. Este procedimento foi realizado selecionando-se o artigo na pasta e, a seguir, por meio da ferramenta “View references”, foram visualizadas as publicações nele citadas. A partir da síntese descritiva apresentada pela Scopus, foram visualizados os autores responsáveis pelo conjunto de referenciais e respectivos totais de documentos citados. Esta síntese com os autores das referências foi copiada e registrada em arquivo Word, no formato tabela com duas colunas: 1ª com o nome dos autores citados no artigo e 2ª com o número de documentos citados do autor. Para os artigos com lista de referências muito extensas, e por este motivo com a lista dos autores mostrada parcialmente na síntese da Scopus, foi necessário completar a lista de autores (e respectivas quantidades de citações) por meio de procedimento manual análogo ao descrito na seção 4.1.1.

b. Na sequência, em um arquivo Word, reuniram-se as tabelas relativas aos autores citados em cada artigo do pesquisador analisado em uma única tabela e acrescentou-se nesta uma nova coluna a fim de se registrar o número de artigos distintos em que os autores citados

apareceram. Após este procedimento, organizou-se a tabela resultante em ordem alfabética, a fim de se verificar em quantos artigos distintos os autores citados apareceram, o que foi registrado na coluna adicionada, e registrar o total de citações acumuladas nos artigos em que foi citado. A seguir, eliminaram-se as duplicações dos nomes dos autores citados, deixando somente uma menção a cada, com o respectivo número de artigo em que foi citado e total de citações recebidas. Após esse tratamento manual, a tabela foi reorganizada em ordem decrescente do número de artigos em que o autor foi citado.

Salienta-se a necessidade de se visualizar o conjunto de autores distintos citados em cada artigo da obra de um pesquisador do universo de estudo em função da premissa principal da análise da identidade do autor: a recitação. Conforme já mencionado, a reincidência da referência (recitação) a um autor, em publicações distintas, aponta a história intelectual e o domínio científico de um pesquisador. Evidencia, ainda, a importância do autor (re)citado para a construção do lastro científico do pesquisador (WHITE, 2001). Desse modo, observar somente o total do citações recebidas por um autor, sem detectar se estas estão distribuídas na obra do pesquisador, não é suficiente para identificar sua identidade de citação.

A partir da realização dos procedimentos a) e b) para cada um dos pesquisadores do universo de pesquisa, construiu-se, então, uma tabela para cada pesquisado com os autores citados pelo menos duas

vezes em seus conjuntos de publicações analisadas. Nessas tabelas, a coluna relativa ao número de artigos em que cada autor aparece citado na obra do pesquisador constitui o seu peso na ponderação das referências presentes na obra do pesquisador A. Este critério baseia-se no procedimento metodológico presente em Zhao e Strotmann (2008, 2014): se a obra do pesquisador A contém X artigos, o conjunto de referências ponderadas deste pesquisador inclui todos os autores distintos arrolados nas listas de referências dos seus X artigos; se um autor está presente em N (com $2 \leq N \leq X$) das X listas de referências da obra pesquisador, ele aparece no conjunto de referências ponderadas do pesquisador A com peso igual a N.

Esclarece-se que, para 3 pesquisadores do universo (B. Caramelli, M. Barreto e L. M. Ferreira) foram localizadas somente as referências de um artigo na base Scopus e, desse modo, eles não participaram da análise de acoplamento bibliográfico, uma vez que não foi possível identificar suas identidades de citação (autores citados em pelo menos dois artigos).

Desse modo, as 30 tabelas construídas contêm a identidade de citação dos pesquisadores, correspondente ao conjunto de autores citados em pelo menos 2 artigos, no âmbito da ciência internacional, formada a partir dos artigos em que estes configuram o primeiro autor da publicação. Esclarece-se que pela metodologia adotada, baseada em White (2001), as autocitações foram consideradas na construção da identidade dos pesquisadores analisados, uma vez que também a recorrência da citação

aos seus próprios trabalhos aponta sua história intelectual e seu domínio científico (WHITE, 2001).

A seguir, iniciaram-se os procedimentos relativos à etapa 2 - análise de acoplamento bibliográfico dos pesquisadores-, a partir da qual buscou-se identificar domínios científicos no campo dos EMI.

Primeiramente, identificaram-se os autores citados por pelo menos 2 dos pesquisadores do universo de análise. Para isso, reuniram-se as 30 tabelas em uma única tabela organizada em ordem alfabética, a fim de se observar quais os autores cujos nomes apareciam com duplicação.

A partir do procedimento descrito, identificou-se um total de 808 autores distintos citados mais de uma vez (reincidiram na citação) por pelo menos 1 dos 30 pesquisadores do universo analisado. Destes, 80 autores foram citados por pelo menos 2 pesquisadores e são, desse modo, aqueles que acoplam os 30 pesquisadores.

A síntese com os totais de: referências, autores citados, autores citados em mais de um artigo e autores em comum com os outros pesquisadores, para o conjunto de 30 pesquisadores foi organizada em uma tabela, novamente com 3 agrupamentos consonantes à Figura 6 - área de atuação (EMI, áreas afins, outras áreas do conhecimento). A fim de avaliar as diferenças estatísticas entre as médias dos três grupos, foram aplicados testes ANOVA com um fator, utilizando o software SPSS, versão 21, com nível de significância igual a 0,05.

A seguir, por meio do software Excel, foi construída a matriz de

adjacências (*2-mode*) de dimensão 30 x 80 com os 30 pesquisadores do universo e os 80 autores citados, com as células registrando os totais de artigos distintos dos pesquisadores (pesos) em que os autores foram citados. A partir dessa matriz, foram calculadas as frequências de acoplamento bibliográfico entre os pesquisadores.

Assim, ao levar em conta, no cálculo da frequência de acoplamento bibliográfico, os pesos associados aos autores citados pelos pesquisadores analisados, reconhece-se que a recitação, ou seja, o fato de um pesquisador frequentemente citar os mesmos autores distintas vezes ao longo da sua produção científica, define em grande parte a identidade de um pesquisador (WHITE, 2001).

A fim de contribuir para a visualização dos autores (citados) em comum pelos pesquisadores (citantes), ou seja, quais são os autores que os acoplam e delineiam domínios científicos nos EMI brasileiros, foi gerada a rede de citação citante - citado, por meio do software Ucinet.

Para o cálculo da frequência, dois a dois, de acoplamento bibliográfico entre os pesquisadores, utilizou-se o procedimento adotado por Zhao e Strotmann (2008, 2014), a saber: se um autor aparece em N das listas de referências dos artigos do pesquisador A e o mesmo autor aparece em M listas de referências de um pesquisador B, então esse autor acresce (contribuirá) com o mínimo de N e M, ou seja, o menor valor de N e M, para a frequência de acoplamento bibliográfico entre os autores A e B.

Desta forma, a frequência de acoplamento bibliográfico entre dois

pesquisadores A e B é definida como a soma dos pesos (número de artigos em que foi citado) mínimos dos autores com recitação, simultaneamente, nas obras dos pesquisadores A e B.

Para a construção da matriz quadrada e simétrica 30 x 30, com as frequências de acoplamento bibliográfico entre os pesquisadores, utilizou-se a ferramenta "Affiliations: convert 2-mode to 1-mode" dentro da categoria de procedimentos "Data" do Ucinet, sobre a matriz de adjacência 30 x 80, selecionado-se como método a opção "Sums of cross-minimums". A diagonal desta matriz é vazia.

A fim de obter uma medida relativa da intensidade, as frequências absolutas de acoplamento bibliográfico entre os autores foram relativizadas pelo Cosseno de Salton. Optou-se pelo Cosseno de Salton (CS), em lugar do Índice de Jaccard (IJ), por analogia aos resultados relativos à comparação entre estes dois índices nos estudos de cocitação de autores presentes em Grácio e Oliveira (2013, 2014 e 2015b), a saber: no processo de relativização/normalização, embora CS e IJ apresentem comportamentos similares, em casos de acoplamento envolvendo dois pesquisadores, um com identidade de citação composta por muitos autores com alto peso e o outro com identidade de citação formada por poucos autores com baixa recitação, considera-se que CS representa mais efetivamente o elo de intensidade de acoplamento entre eles.

A partir dessa adequação do CS para a análise de acoplamento bibliográfico de autores, o Cosseno de Salton (CS) pode ser definido

como o quociente da frequência de acoplamento bibliográfico entre dois pesquisadores (X e Y) pela raiz quadrada da soma total dos pesos de recitação dos autores citados por X multiplicado pela soma total dos pesos de recitação dos autores citados pelo pesquisador Y. A formulação matemática pode apresentada por:

$$CS_{ABA}(X, Y) = \frac{f_{ABA}(X, Y)}{\sqrt{Pesq(X) \times Pesq(Y)}}$$

onde:

$CS_{ABA}(X, Y)$ = índice relativo por CS de acoplamento bibliográfico entre os pesquisadores X e Y;

$f_{ABA}(X, Y)$ = frequência de acoplamento bibliográfico entre os pesquisadores X e Y;

$Pesq(X)$ = soma total dos pesos (recitação) dos autores citados pelo pesquisador X;

$Pesq(Y)$ = soma total dos pesos (recitação) dos autores citados pelo pesquisador Y.

Para a matriz de acoplamento bibliográfico resultante da normalização por CS, foi gerada a rede de acoplamento bibliográfico, por meio do software Ucinet. Ainda, utilizando essa matriz de acoplamento bibliográfico normalizada, por meio do software SPSS, os pesquisadores foram agrupados pela análise de cluster, adotando método Ward de ligação

e medida de distância Euclidiana quadrada. Os agrupamentos resultantes da análise de clusters foram evidenciados na rede de acoplamento bibliográfico por meio de círculos reunindo os pesquisadores, de acordo com as proximidades apontadas no dendograma gerado no SPSS.

4.4 Procedimentos para a identificação dos criadores da imagem e da análise de cocitação entre os pesquisadores

A visualização dos domínios no campo dos EMI, a partir das suas imagens de citação foi realizada por meio de duas fases de análise subsequentes: 1 - identificação dos periódicos criadores da imagem de citação de cada um dos 33 pesquisadores do universo de pesquisa; 2 - análise de cocitação entre os pesquisadores, a partir dos periódicos em que foram citados³⁸. Os procedimentos metodológicos adotados na primeira etapa foram baseados em White (2001) e, na segunda etapa, em White (2001), Cronin e Shaw (2002), Zhao e Strotmann (2008) e Grácio e Oliveira (2013). Esclarece-se que, em relação ao artigo Zhao e Stromann (2008), cujo método é construído para a análise de Acoplamento Bibliográfico de Autores, por analogia, utilizamos a noção de pesos associados aos periódicos em que os pesquisadores foram citados no cálculo da frequência de cocitação, a fim de reconhecer que a intensidade com que um pesquisador foi citado em um periódico, define em grande parte a sua imagem de citação. Usualmente, a análise de cocitação de autores não tem levado em conta a frequência da citação (número de artigos em

que foi citado) do pesquisador por um citante, mas somente a presença ou ausência daquele pesquisador na literatura científica do citante, ou seja, em geral, a ACA é feita de forma não valorada. A opção por adotar uma matriz valorada (com pesos correspondentes à frequência de citação do pesquisador no periódico) é decorrente da busca de uma abordagem semelhante para ambos os métodos relacionais de citação - acoplamento e cocitação -, a partir da noção de identidade e imagem de citação, definidas por White (2001). Como consequência dessa opção, o método de simetrização da matriz de adjacências pesquisador x periódico (em que foi citado) é distinto daquele observado em Grácio e Oliveira (2013, 2014, 2015), em que essa matriz é não valorada. Desse modo, o método de construção da matriz de cocitação foi análogo àquele adotado para a construção da matriz de acoplamento bibliográfico, a saber: a frequência de cocitação entre dois pesquisadores A e B é definida como a soma do mínimo entre os números d e citações recebidas por A e B, nos periódicos em que foram citados simultaneamente.

Para identificar os periódicos criadores da imagem de citação de cada um dos 33 pesquisadores do universo de pesquisa, foram realizados os seguintes procedimentos consecutivos:

a. no dia 21/03/2018, para cada pesquisador, a partir dos artigos armazenados no Gerenciador de listas da Scopus, levantou-se o conjunto de periódicos em que seus artigos foram citados e suas respectivas frequências de citação. Este procedimento foi realizado,

selecionando-se todos os artigos do pesquisador em sua pasta e, a seguir, por meio da ferramenta “View cited by”, a partir da síntese descritiva apresentada pela Scopus, foram visualizados: total de citações recebidas; total de autocitações; fontes das citações (periódico, proceedings ou livro) com as respectivas frequências, excluídas as autocitações. Esses dados foram registrados em arquivos Word (após tratamento manual, para padronização em formato de tabela), sendo um para cada pesquisador. A partir desses dados, construiu-se uma tabela com a lista dos 33 pesquisadores, agrupados por área de atuação declarada no currículo Lattes, e seus respectivos valores relativos aos indicadores quantitativos de citação. A fim de avaliar as diferenças estatísticas entre as médias dos três grupos, foram aplicados testes ANOVA com um fator, utilizando o software SPSS, versão 21, com nível de significância igual a 0,05.

b. Na sequência, em um arquivo Excel, reuniram-se, em uma única tabela, as tabelas dos 33 pesquisadores, geradas em a., com os periódicos citantes dos seus artigos, acrescentando-se, nessa nova tabela, uma nova coluna a fim de se registrar o nome do pesquisador citado. Após este procedimento, organizou-se a tabela resultante em ordem alfabética, a fim de se verificar os distintos periódicos citantes e, entre estes, aqueles que citaram mais de um dos pesquisadores analisados. A seguir, eliminaram-se as duplicações dos títulos das fontes das citações, deixando somente uma menção a cada periódico, com o respectivo número de pesquisadores citados. Após esse tratamento manual, a tabela

foi reorganizada em ordem decrescente do número de pesquisadores citados pela fonte (periódico, *proceedings*, livro) citante.

A partir da realização dos procedimentos a) e b), identificou-se um total de 872 citações oriundas dos artigos de 32 dos 33 pesquisadores analisados. Após a exclusão das autocitações, resultou um total de 726 citações, decorrentes dos artigos de 31 dos pesquisadores analisados. Desse modo, a incidência geral de autocitação no corpus científico analisado foi equivalente a 17% do total das citações recebidas. As 726 citações foram oriundas de 372 fontes diferentes - periódicos, *proceedings* e livros -, majoritariamente de artigos publicados em periódicos científicos. Destas 372 fontes, 75 citaram mais de um dos pesquisadores analisados, constituindo, assim, as fontes que realizam as ligações de cocitação entre os pesquisadores. Nas demais (297 \equiv ~80%) fontes de citação, somente um dos pesquisadores analisados foi citado.

Esclarece-se que: os artigos analisados de um dos pesquisadores (Gandia, R.M.) não receberam citação até o momento da coleta dos dados; as citações aos artigos de outro pesquisador (Ensslin, S.R.) foram oriundas de autocitações; todas as citações aos artigos de dois outros pesquisadores (Ensslin, L. e Ferreira, L.M.) são oriundas de periódicos nos quais, do grupo de pesquisadores analisados, somente eles foram citados. Desse modo, esses quatro pesquisadores não participam da análise de cocitação entre os pesquisadores, a qual foi desenvolvida considerando: 29 pesquisadores e 75 periódicos que citaram simultaneamente mais de

um dos pesquisadores analisados.

A seguir, iniciaram-se os procedimentos relativos à etapa 2 - análise de cocitação entre os pesquisadores, a partir dos periódicos em que foram citados, a fim de se identificar domínios científicos no campo dos EMI a partir das suas imagens de citação.

Por meio do software Excel, foi construída a matriz de adjacências (2-mode) de dimensão 29 x 75 com os 29 pesquisadores do universo e as 75 fontes (periódicos, proceedings e livro) citantes, com os células registrando os totais de citações recebidas pelos pesquisadores (pesos) na fonte em que foram citados. A partir dessa matriz, foram calculadas as frequências de cocitação entre os pesquisadores.

A fim de contribuir para a visualização dos periódicos que cocitam os 29 pesquisadores, ou seja, quais são os periódicos que conectam os pesquisadores, ao, simultaneamente, reconhecerem as contribuições destes por meio das citações nos artigos neles publicados e delineiam domínios científicos nos EMI brasileiros, na perspectiva dos citantes, foi gerada a rede de citação periódico citante → pesquisador citado, por meio do software Ucinet.

Para o cálculo da frequência de cocitação, dois a dois, entre os pesquisadores, utilizou-se procedimento análogo àquele desenvolvido para o cálculo da frequência de acoplamento entre pesquisadores, a saber: se um periódico cita N vezes um pesquisador A e M vezes um pesquisador B, então esse periódico (contribuirá) com o mínimo entre

N e M, ou seja, o menor valor de N e M, para a frequência de cocitação entre os pesquisadores A e B.

Dessa forma, a frequência de cocitação entre dois pesquisadores A e B é definida como a soma dos pesos (número de citações) mínimos dos periódicos que citam, simultaneamente, os artigos dos pesquisadores A e B.

Para a construção da matriz quadrada e simétrica 29 x 29, com as frequências de cocitação entre os pesquisadores, utilizou-se a ferramenta "Affiliations: convert 2-mode to 1-mode" dentro da categoria de procedimentos "Data" do Ucinet, sobre a matriz de adjacência 29 x 75, selecionado-se como método a opção "Sums of cross-minimums". A diagonal desta matriz é vazia.

A fim de obter uma medida relativa da intensidade de cocitação, as frequências absolutas de cocitação entre os autores foram relativizadas pelo Cosseno de Salton, seguindo o percurso descrito em Grácio e Oliveira (2012, 2014, 2015b).

O índice relativo de cocitação por Cosseno de Salton (CS) foi, então, calculado pela equação:

$$CS_{ACA}(X, Y) = \frac{Cit_{ACA}(X, Y)}{\sqrt{Cit(X) \times Cit(Y)}}$$

onde:

$CS_{ACA}(X, Y)$ = índice relativo por CS de cocitação entre os pesquisadores X e Y;

$Cit_{ACA}(X,Y)$ = frequência de cocitação entre os pesquisadores X e Y;

$Cit(X)$ = total de citações recebidas pelo pesquisador X;

$Cit(Y)$ = total de citações recebidas pelo pesquisador Y.

Para a matriz de cocitação resultante da normalização por CS, foi gerada a rede de cocitação entre os pesquisadores, por meio do software Ucinet. Ainda, utilizando essa matriz de cocitação normalizada, por meio do software SPSS, os pesquisadores foram agrupados pela análise de cluster, adotando método Ward de ligação e medida de distância Euclidiana quadrada. Os agrupamentos resultantes da análise de clusters foram evidenciados na rede de cocitação por meio de círculos reunindo os pesquisadores, de acordo com as proximidades apontadas no dendograma gerado no SPSS.

4.5 Procedimentos para a junção dos domínios obtidos a partir das perspectivas da identidade e da imagem de citação dos pesquisadores analisados

A fim de se obter uma visualização paralela dos domínios identificados a partir das perspectivas identidade e imagem de citação, inicialmente, construiu-se uma nova figura colocando lado a lado as figuras geradas para as redes de acoplamento bibliográfico e de cocitação e os respectivos domínios identificados pela análise de cluster, a partir das matrizes geradoras dessas redes. Para analisar algumas proximidades identificadas entre pares de pesquisadores, observadas a partir desta figura, recorreu-se a informações disponibilizadas nos currículo Lattes

desses pesquisadores e também a suas produções científicas publicadas em periódicos da base Scopus.

Finalizando os procedimentos realizados para a obtenção dos resultados necessários para responder aos objetivos propostos nesta pesquisa, construiu-se um quadro de dupla classificação (identidade e imagem de citação), com uma síntese descritiva das características observadas para cada domínio e também posicionando os pesquisadores segundo os domínios que integraram em cada perspectiva.