

Análise de citação

Maria Cláudia Cabrini Grácio

SciELO Books / SciELO Livros / SciELO Libros

GRÁCIO, M. C. C. Análise de citação. 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. 77-113. ISBN: 978-65-86546-12-5. Available from:

<http://books.scielo.org/id/tx83k/pdf/gracio-9786586546125-03.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/).

3. ANÁLISE DE CITAÇÃO

Desde o século 19, a citação é parte formal do processo de construção e comunicação do conhecimento científico, em função do preceito que requer que os pesquisadores referenciem as publicações cujos conceitos e métodos inspiraram ou foram usados no desenvolvimento do seu próprio artigo. Neste contexto, distinguem-se os termos “referência” e “citação”, em que referência expressa o reconhecimento que um documento B dá a um documento A, quando o primeiro (B) utiliza conhecimento científico já elaborado no segundo documento (A). Por outro lado, a citação indica o reconhecimento do documento A recebido do documento B, pelo uso ou descrição do seu conteúdo (EGGHE; ROUSSEAU, 1990).

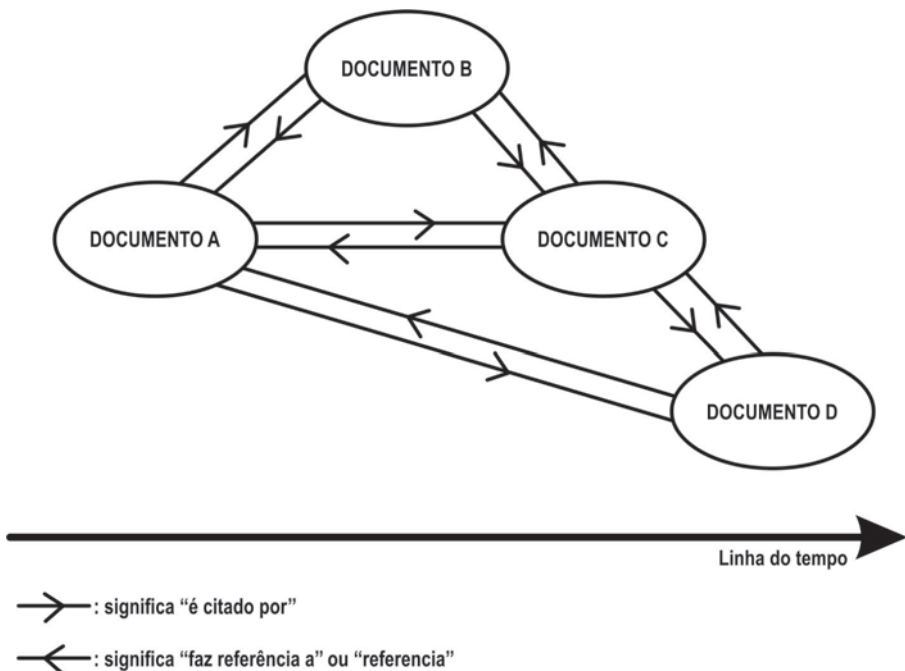
Assim, referência é um conceito sincrônico, i.e., tem um formato retrospectivo, do presente para o passado (olha para trás), e citação é um conceito diacrônico, que reflete a presença da publicação em anos consecutivos (olha para frente), o que leva a formas distintas de estudos bibliométricos (STINSON; LANCASTER, 1987; EGGHE; ROUSSEAU, 1990; GLÄNZEL, 2003).

Apesar das diferenças nas noções associadas a esses dois termos, a maioria dos autores não os utiliza de forma precisa, comutando-os de forma indistinta, o que Egghe e Rousseau (1990) consideram um

desperdício lastimável de bons termos técnicos e identificam Price (1970) como autor antecessor dessa compreensão. Apontamos que esse abuso de linguagem pode também estar presente durante o desenvolvimento deste livro, embora nos empenharemos para usá-los de forma precisa.

Essas diferentes formas de olhar a relação entre dois documentos na literatura podem ser visualizadas na Figura 7, em que se observa que A é uma referência presente nas listas de referência de B, C e D, ou seja, A é citado por B, C e D. Reciprocamente, B, C e D fazem referência (ou contém uma citação a) A. Além disso, B é citado por C e estes são citados por D, por estarem presentes em sua lista de referências.

Figura 7 - Relações de citação e referência entre documentos A, B e C.



Fonte: elaborada pela autora (2019)²⁶.

Desse modo, as referências refletem os antecedentes de um trabalho científico e as citações indicam a influência real ou impacto científico²⁷ de um trabalho, seja pelo método científico desenvolvido, seja pelos resultados obtidos, em pesquisas ulteriores (GORBEA-PORTAL, 1994).

3.1 Identidade e imagem de citação de autores

De acordo com a categorização para os indicadores métricos (Capítulo 2, seção 2.3), as análises baseadas nas referências pertencem à categoria dos indicadores de produção, ao passo que os estudos sustentados nas citações vinculam-se à categoria dos indicadores de impacto.

Nesse cenário, ao evidenciar as relações semânticas entre o seu trabalho e os documentos anteriores, assinaladas pelas referências, o autor se reporta aos fundamentos teórico-metodológicos vigentes em seu campo científico, este calcado em bases eminentemente sociais. Assim, o ato de referenciar está cingido a um contexto psicológico, social, político e histórico (ALVARENGA, 1998).

Nesse cenário, a análise das referências de um autor ou de um conjunto de produções bibliográficas em uma literatura científica permite conhecer seus principais interesses e cerne científicos. Desse modo, à medida que a obra de um pesquisador cresce, o conjunto de autores referenciados em mais de uma das suas publicações reflete sua identidade científica, em especial, a partir daqueles autores cuja reincidência ocorra

com maior frequência. Neste contexto, o estudo do comportamento de um pesquisador em relação à recorrência da referência aos seus próprios trabalhos e de outros aponta sua história intelectual e seu domínio científico (WHITE, 2001). A reincidência da referência a um autor, em publicações distintas de um pesquisador, é designada por White (2001) de *recitação*²⁸.

O conceito de *recitação* desdobra-se em duas perspectivas: *sincrônica* - quando um pesquisador cita um autor mais de uma vez em um único trabalho; *diacrônica* - quando o pesquisador cita o mesmo autor em obras publicadas em momentos diferentes. A contagem das *recitações* diacrônicas pode ser utilizada como evidência da importância dos autores citados ao longo da carreira dos pesquisadores, visto que estes tipicamente citam muitos autores uma vez e a “*recitação*” distingue, assim, os citados (WHITE, 2001).

A análise de *recitação* diacrônica, ou em termos mais precisos, da reincidência das referências na lista de referências de um pesquisador, revela os autores que norteiam o seu pensamento científico, evidenciando suas características sociais, intelectuais e epistemológicas, como os autores que são seus pontos de referência recorrentes, seja como modelo intelectual ou como oponente (WHITE, 2001; VANZ; CAREGNATO, 2003).

Nesse sentido, destaca-se a necessidade de se conhecer o domínio analisado para interpretar a presença, assim como as ausências, dos nomes nas listas de referência de um pesquisador (MCCAIN, 1990; WHITE, 2001).

Partindo da premissa que, embora as citações e recitações sejam afetadas pelos laços sociais entre pesquisadores (relação entre o citante e o citado), seu elemento propulsor universal decorre da reconhecida relevância científica em relação aos trabalhos citados, a contagem das recitações dadas por pesquisador a um autor em obras distintas publicadas, durante sua trajetória científica, pode ser utilizada como evidência da importância deste autor ao longo da sua atividade científica (WHITE, 2001).

Os pesquisadores habitualmente referenciam diversos autores uma única vez, em função da especificidade do objeto e objetivo tratados em uma pesquisa em particular. Neste cenário, a recitação de um autor na obra de um pesquisador indica a sua importância para a construção do lastro científico do pesquisador.

Assim, é altamente provável que, durante a construção da obra de um pesquisador, sua seleção dos autores constantes seja realizada de forma tão particular e característica como uma impressão digital. Neste cenário, o padrão de recorrência a determinadas referências e, por consequência a seus autores, forma parte significativa da identidade científica de um pesquisador, definida como o conjunto de autores que este recita (WHITE, 2001).

Neste contexto, White (2001) considera que, embora um pesquisador escolha de forma consciente os autores e trabalhos que referencia em cada pesquisa que desenvolve, e, nesse sentido, em cada publicação sua identidade é planejada, é pouco provável que perceba o padrão de citação que se forma,

e adquire consistência, em sua obra ao longo da sua vida científica.

Todavia, apesar do reconhecimento da importância da análise de citação (GARFIELD, 1979; EGGHE; ROUSSEAU, 1990; HJØRLAND, 2002; GLÄNZEL, 2003; MOED, 2010), para a análise e avaliação do comportamento científico em diferentes campos do conhecimento, a literatura científica apresenta críticas relacionadas à dificuldade de se desconhecer os motivos que levam um autor a citar os trabalhos de outros autores, o que leva a uma brecha entre os motivos reais pelos quais o autor citou e aqueles considerados o porquê o autor citou, com os reais podendo ser meritórios e pertinentes ou levianos (MACIAS-CHAPULA, 1998).

Entre as razões consideradas pertinentes, segundo Glänzel (2003), Garfield e Weinstock identificaram 15 motivos meritórios, a saber: 1. Prestar homenagem aos pioneiros; 2. Dar crédito a trabalhos relacionados à pesquisa; 3. Identificar metodologia e equipamentos utilizados; 4. Identificar as leituras fundamentais; 5. Corrigir o próprio trabalho; 6. Corrigir o trabalho de outros; 7. Criticar trabalhos anteriores; 8. Fundamentar afirmação; 9. Alertar sobre a publicação de trabalhos futuros; 10. Dar visibilidade a trabalhos mal disseminados, mal indexados ou não citados; 11. Autenticar dados e classes de fatos; 12. Identificar publicações originais nas quais uma ideia ou conceito foi discutido; 13. Identificar publicações originais ou outro trabalho que descreve um conceito ou termo epônimo; 14. Rejeitar trabalho ou ideias de outros; 15. Disputar reivindicação de prioridade do conhecimento.

Em contrapartida, Thorne (1977), citado por Egghe e Rousseau (1990), apresentou uma lista em que revela uma outra face das possíveis razões associadas às citações realizadas por alguns autores, dissociadas do mérito científico. Entre elas, citações para: publicações em série ou múltiplas (“fatiadas”), bajular figuras eminentes, atender preferências editoriais, beneficiar colegas, atender expectativas da comunidade científica quanto ao reconhecimento de autores eminentes. Esta última associa-se ao Princípio da Vantagem Cumulativa²⁹, enunciado por Price (1976b).

Considera-se, todavia, que, embora os motivos elencados por Thorne sejam realidade na ciência, considerando que a atividade científica é social, o possível viés gerado pelas citações não meritórias e/ou pertinentes, não é suficiente para invalidar a tendência geral do volume de citações recebidas pelos autores em função do mérito científico dos seus conhecimentos produzidos. Esta percepção vai ao encontro da premissa, anteriormente mencionada, presente em White (2001) de que o elemento propulsor universal decorre da reconhecida relevância científica em relação aos trabalhos citados.

Além disso, considera-se que caso a quantidade de citação recebida por um autor seja decorrente principalmente de motivos estratégicos, como aqueles mencionados por Thorne (1977), esta terá uma distribuição de pequena amplitude na comunidade científica, ou seja, muitas citações advindas de um pequeno número de pesquisadores. Esta situação pode ser ilustrada, de forma simplificada, na Figura 8, em que o autor X recebe

várias citações, todavia advindas de um único pesquisador, evidenciando a falta do amplo reconhecimento e impacto da sua obra na comunidade científica.

Figura 8 - Simulação de diferentes citações relativas a uma certa quantidade de citações recebidas por um autor, elucidada a partir da visualização da sua origem (amplitude do número de distintos autores citantes).



Fonte: Ajiferuke e Wolfram (2010, tradução nossa).

Complementarmente à caracterização científica de um pesquisador obtida a partir do seu uso recorrente de referências, que define sua identidade (de citação), a imagem de um pesquisador, estabelecida em função das citações recebidas, ou seja, a partir da análise de quem o cita e com quem é citado simultaneamente (CRONIN; SHAW, 2002), oferece outra perspectiva dos seus atributos científicos.

Neste contexto, assumindo-se que o pesquisador tenha sido citado, sua imagem (de citação) junto à comunidade é definida pelo conjunto de todos os autores com quem este foi citado simultaneamente nas listas

de referência da literatura científica. Desse modo, a imagem de citação de um autor é determinada pela comunidade científica, ou seja, pelos pesquisadores ativos (autores citantes ou citers, em idioma inglês), exceto na medida em que ele contribui para sua imagem pelas autocitações em suas próprias listas de referência (WHITE, 2001).

Em decorrência da dinâmica da construção da ciência, tanto a identidade como a imagem de citação de um pesquisador é suscetível a mudanças ao longo do tempo, a primeira decorrente das possíveis mudanças em seus interesses e cerne científicos, em sua trajetória científica, e a segunda, em função de ser dependente da percepção da comunidade científica, com suas publicações incorporadas por outros autores em novas listas de referência, podendo estas estarem ainda em novos contextos.

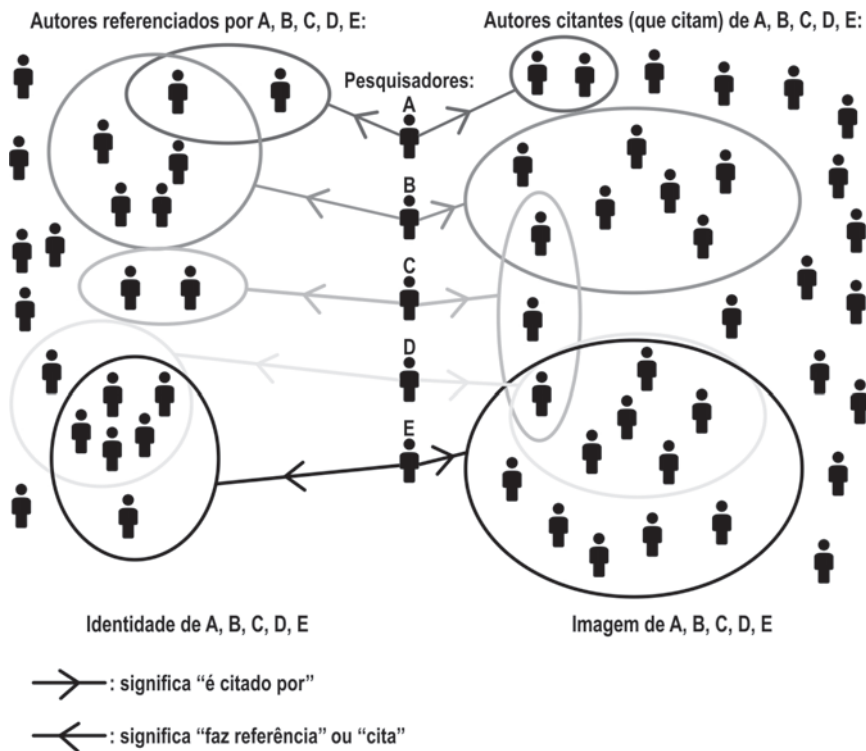
À medida que os anos passam, a identidade e a imagem de um pesquisador tendem a se diferenciarem. Se este não se cita em um artigo, os autores aos quais faz referência serão parte de sua identidade, mas não de sua imagem, a menos que outros pesquisadores o citem em conjunto com aqueles que definem sua identidade. Por outro lado, sua imagem inclui os autores com quem foi citado simultaneamente nas publicações de outros, independente de o pesquisador tê-los referenciado em sua obra e fazerem, assim, parte de sua identidade (WHITE, 2001).

Nesse cenário, a autocitação³⁰ constitui um elo entre os conjuntos da identidade e da imagem e afeta diretamente a imagem de um autor, na medida em que este insere uma referência própria em sua lista de

referências e aparece, assim, citado com os outros autores que compõem sua imagem. Este fato aumenta o total de citações do autor e é considerado por críticos à autocitação como uma forma de incrementar sua contagem de citação e, assim, fortalecer sua posição na comunidade científica.

A Figura 9 apresenta uma ilustração das relações de referência e citação que definem a identidade e a imagem dos pesquisadores em um campo científico em um determinado momento, em que suas identidades são definidas pelos círculos à esquerda destes, nos quais encontram-se os autores referenciados de forma recorrente. Por outro lado, suas imagens são estabelecidas pelos círculos a suas direitas, em função das citações simultâneas realizadas pelos citantes em seus trabalhos científicos, em que ilustra-se que a imagem de B é formada pela citação conjunta com C por um único citante e a imagem de E é formada pela citação conjunta com C, advinda de um autor citante, e com D, advinda de seis autores citantes. Desse modo, a imagem de citação de um autor está associada à noção de cocitação de autores, detalhada na seção 3.2.2 deste capítulo.

Figura 9 - Ilustração das relações de citação que definem a identidade e a imagem dos pesquisadores em um campo científico.



Fonte: elaborado pela autora (2019).

Desse modo, a partir da ilustração presente na Figura 9, compreende-se que em um campo científico podem ser observados diversos domínios científicos, constituídos por grupos de pesquisadores que compartilham referentes teóricos e/ou metodológicos e em função das associações de conhecimento e significados reconhecidos pelos autores ao citar em suas pesquisas, de forma conjunta, pesquisas (e autores) anteriores que subsidiaram a construção do conhecimento destas.

Nesse cenário, a identificação e a visualização das conexões cognitivas e sociais estabelecidas pelas citações, seja na relação

citante→citado (referência), na relação citado↔citado ou citante↔citante, ao longo do tempo, contribui para uma melhor e mais ampla compreensão da estrutura intelectual, suas correntes teóricas e metodológicas, definidas pela comunidade científica. Desse modo, as referências adotadas pelos autores constituem uma forma de marca d'água do paradigma dominante em um domínio científico (CRONIN; SHAW, 2002).

3.2 Análises Relacionais de Citação

Partindo do pressuposto que a citação pode ser tomada como um indicador objetivo da interlocução e/ou proximidade teórico-conceitual ou metodológica entre autores, seja em uma perspectiva intencional e consciente do autor (relação citante - citado), seja em uma perspectiva independente e alheia à intenção original dos autores (citado-citado ou citante-citante), sua análise evidencia as relações vigentes e que definem um domínio científico.

Nesse contexto, a análise de citação univariada não oferece resposta a questões concernentes às relações entre os elementos citados: quais as relações estabelecidas entre os documentos, entre os autores, ou entre os periódicos citados? Entre outras perguntas possíveis relativas às interações entre os elementos citados.

Desse modo, salienta-se que a análise univariada não possibilita conhecer as relações estruturais de conectividade teórico-metodológica

de um domínio, a proximidade, vizinhança, associação e interlocução estabelecida entre os documentos, pesquisadores, periódicos, entre outros, como reconhecidos pela comunidade científica. Para este fim, são necessários dados de outra natureza, relacionais, que levam em conta em sua análise a observação simultânea de duas referências, ou seja, são necessárias medidas que observem a distância ou proximidade entre dois documentos, autores, periódicos, entre outros.

Embora também apresente as referências (citações) presentes na literatura científica como objetos de análise, as análises relacionais de citação baseiam-se em pares de dados ao examinar as referências de uma literatura e tem por finalidade identificar e evidenciar a coocorrência de um par de referências na literatura científica. Para tal, buscam analisar a frequência com que pares de documentos, de autores, de periódicos, de instituições ou de país, aparecem citados juntos na literatura, a fim de conhecer e expressar as particularidades da estrutura de domínios científicos, possibilitando a visualização³¹ das interlocuções, similaridades e conexões construídas por suas comunidades.

3.2.1 Acoplamento bibliográfico de autores

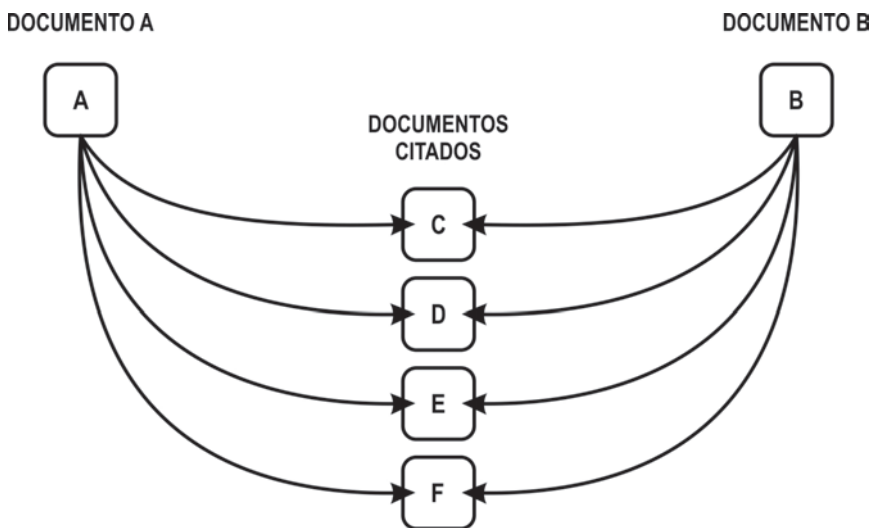
Em 1963, Michael M. Kessler apresentou um método para agrupar artigos científicos e técnicos, baseado em unidades de acoplamento bibliográfico, em que uma unidade de acoplamento entre dois artigos é

definida como um item de referência usado pelos dois artigos (KESSLER, 1963). Se há um item de referência em comum entre dois artigos, eles são ditos estarem bibliograficamente acoplados (EGGHE; ROUSSEAU, 2002).

Desse modo, o acoplamento bibliográfico (AB) entre dois artigos ocorre quando estes referenciam pelo menos uma publicação em comum (Figura 10). Nesse contexto, o acoplamento bibliográfico estabelece uma conexão entre dois artigos ao utilizarem as mesmas referências.

A força de acoplamento bibliográfica entre dois artigos não se altera com o tempo, uma vez que a frequência do acoplamento bibliográfico é fixa, não podendo facilmente contribuir para o estudo das mudanças nos domínios científicos ao longo do tempo (SMALL, 1973) e por isso, é considerada uma análise retrospectiva.

Figura 10 - Documentos A e B acoplados bibliograficamente, em função dos quatro documentos C, D, E, F citados, em comum.



Os documentos A e B estão acoplados porque ambos citam os documentos C, D, E e F

Fonte: Mattos e Dias (2010, adaptada de Garfield, 2001).

Segundo Kessler (1965), o conjunto de referências utilizado pelos autores em seu artigo evidenciam o ambiente intelectual no qual trabalham e, se dois artigos apresentam bibliografias similares, há uma relação implícita entre eles.

Desse modo, no método de Acoplamento Bibliográfico, parte-se da hipótese de que se dois artigos fazem referência a uma mesma fonte, eles apresentam proximidade teórica e/ou metodológica (KESSLER, 1965). Nesse contexto, a intensidade do acoplamento de dois artigos depende da quantidade de referências que eles têm em comum e quanto maior o número de referências em comum, maior será a força de conexão entre eles (EGGHE; ROUSSEAU, 2002; ZHAO; STROTMANN, 2008).

Assim, o Acoplamento Bibliográfico mensura a proximidade entre dois artigos comparando suas referências e quanto maior o número de referências que compartilham, maior a similaridade entre eles, que pode ser temática, teórico, metodológica ou outra particularidade compartilhada (LUCAS; GARCIA-ZORITA; SANZ-CASADO, 2013).

Nesse contexto, a análise de Acoplamento Bibliográfico possibilita o estudo do desenvolvimento das linhas de pesquisa, permitindo identificar os núcleos de pesquisa, os pesquisadores e os artigos mais importantes em um domínio científico (CARVALHO, 1975). Glänzel (2003) considera este o método baseado na análise do conjunto de referências mais importante para a visualização de um domínio.

Destaca-se, todavia, que o método de Acoplamento Bibliográfico

(AB), como um indicador de similaridade entre artigos, foi pouco aplicado para a análise da estrutura e rede de conhecimento em um domínio científico, desde a sua introdução na década de 1960, ao contrário da análise de cocitação que, nas duas últimas décadas, se mostra hegemônica como procedimento para a visualização das redes de conhecimento e sua evolução ao longo do tempo (ZHAO; STROTMANN, 2008).

No entanto, a partir de meados da década de 2000, observa-se um crescimento nos estudos que analisam as redes de conhecimento geradas com base no método AB, como Jarneving (2005, 2007), com acoplamento por documentos, e Boyack, Börner e Klavans (2009) com acoplamento de disciplinas, por meio de periódicos. Meireles, Céndon e Almeida (2014) apresentam um método baseado na associação entre redes neurais artificiais e acoplamento bibliográfico de documentos para gerar grupos de documentos, a fim de contribuir para a identificação de tendências de pesquisa em domínios.

Zhao e Strotmann (2008), em estudo que estende o método de acoplamento bibliográfico para a conexão entre autores, argumentam que, embora a análise de cocitação de autores seja frequentemente aplicada para mapear a estrutura intelectual de um domínio científico, o que de fato é mapeado nesses estudos é a estrutura das influências intelectuais em um domínio como percebida por seus autores ativos (ZHAO; STROTMANN, 2008).

Nesse contexto, Zhao e Strotmann (2008) introduzem a análise

de acoplamento bibliográfico de autores (ABA), como uma ampliação do método proposto por Kessler (1963), a fim de mapear as atividades científicas dos autores e contribuir para a obtenção de uma visualização mais real e aprofundada das estruturas intelectuais de um domínio científico e ampliar a compreensão das análises de redes de citação baseadas nos autores.

No método de acoplamento bibliográfico de autores (ABA), considera-se que quanto mais referências em comum, dois autores têm em suas obras, mais semelhantes são suas investigações (QIU; DONG; YU, 2014).

Segundo Zhao e Strotmann (2008), a análise de acoplamento bibliográfico de autor evita o problema apontado por Small (1973) em relação ao método de Kessler (1963), uma vez que a frequência do acoplamento bibliográfico entre dois autores (isto é, entre suas obras) evolui ao longo do tempo, quando pelo menos um deles continua a publicar. Consideram, ainda, que este método, ao focar nos autores da fonte dos trabalhos (citantes) em lugar dos autores de referências citadas, pode fornecer uma visão mais completa da estrutura, características e desenvolvimento das redes de conhecimento de conhecimento científico.

Segundo os propositores do método, há dois aspectos principais que afetam os resultados da análise ABA: o método de escolha do conjunto de autores para representar o domínio científico estudado e a definição da forma de cômputo da frequência do acoplamento bibliográfico entre os autores.

Em relação à definição do método de cômputo da frequência do acoplamento entre os autores, há duas possibilidades: frequência de acoplamento entre os dois autores citantes ser tomada como o número total de documentos compartilhados em suas listas de referências (Figura 11); ou de forma mais simples, tratar a obra completa de um autor citado como uma única publicação e calcular a frequência do acoplamento entre os dois autores citantes como o número total de autores compartilhados nas listas de referências das suas publicações (Figura 12).

Figura 11 - Ilustração do Acoplamento bibliográfico entre pesquisadores, a partir da frequência de compartilhamento de referências.

Referências (doc) citadas nas obras de Autor 1 e Autor 2										
Pesquisadores Acoplados	Doc1	Doc2	Doc3	Doc4	Doc5	Doc6	Doc7	Doc8	Doc9	Doc10
Autor 1		X	X		X		X		X	X
Autor 2	X	X		X	X	X		X		

Fonte: Grácio (2016).

Na Figura 11, os Pesquisadores 1 e 2 são acoplados pelas referências à Doc2 e Doc5. Desse modo, a frequência de acoplamento bibliográfico entre estes pesquisadores é igual a 2.

Na Figura 12, os pesquisadores 1 e 2 são acoplados pela citação aos autores Autor C, Autor E e Autor F. Desse modo, a frequência de acoplamento bibliográfico entre os pesquisadores 1 e 2 é igual a 3.

Figura 12 - Ilustração do Acoplamento bibliográfico entre pesquisadores, a partir da frequência de compartilhamento de autores citados.

Autores citados nas obras de Autor 1 e Autor 2							
Pesquisadores Acoplados	Autor A	Autor B	Autor C	Autor D	Autor E	Autor F	Autor G
Autor 1	X		X	X	X	X	
Autor 2		X	X		X	X	X

Fonte: Grácio (2016).

Zhao e Strotmann (2008, 2014) utilizam em suas análises o Acoplamento bibliográfico entre autores, a partir da frequência de compartilhamento dos autores citados, como ilustrado na Figura 12.

Segundo Hjørland (2013), entender o acoplamento bibliográfico entre autores significa entender o grau de sobreposição da identidade de citação desses autores. Segundo ainda esse autor, essa sobreposição pode ser parcialmente determinada por diferenças entre os diversos domínios científicos, uma vez que em alguns deles, os autores têm grande liberdade na escolha dos problemas de pesquisa, métodos de investigação e também no que considera literatura relevante. Por outro lado, em outros domínios, os pesquisadores estão bastante restringidos pelas normas e convenções desenvolvidas coletivamente. Desse modo, as identidades de citação apresentam maior variabilidade em alguns domínios do que em outros.

A partir da proposta de Zhao e Strotmann (2008), observam-se pesquisas que têm se dedicado a desenvolver o método ABA, em termos

teóricos e aplicados. Entre elas, destacam-se: Rousseau (2010) que revisa as noções de acoplamento bibliográfico e cocitação e suas generalizações, assim como as diferentes formas de suas aplicações, como método para evidenciar a estrutura intelectual de um domínio científico; e Ma (2012) que apresenta três métodos para o cálculo da força relativa de acoplamento bibliográfico entre os autores.

Atualmente, alguns pesquisadores têm se dedicado a estender o método de acoplamento de autores a outros tipos de relação de compartilhamento entre os autores, como Cabanac (2011) que mediu o acoplamento entre autores, por meio da similaridade da participação concomitante em conferências científicas (QIU; DONG; YU, 2014).

3.2.2 Análise de cocitação de autores

Henry Small propôs, em 1973, uma forma de analisar a ligação entre dois documentos, baseada no estudo da frequência³² com que eles documentos são citados juntos. Denominou essa forma de concorrência de dois documentos na literatura de cocitação.

Com esse procedimento, Small buscou analisar, identificar e descrever a estrutura e a conectividade de uma área do conhecimento científico (BAYER; SMART; McLAUGHIN, 1990), via documentos publicados.

Ao contrário do acoplamento bibliográfico que liga/assemelha documentos, por meio das suas citações em comum, a cocitação identifica

a ligação/semelhança de dois documentos citados, via suas frequências de ocorrência conjunta em uma lista de referências dos autores citantes.

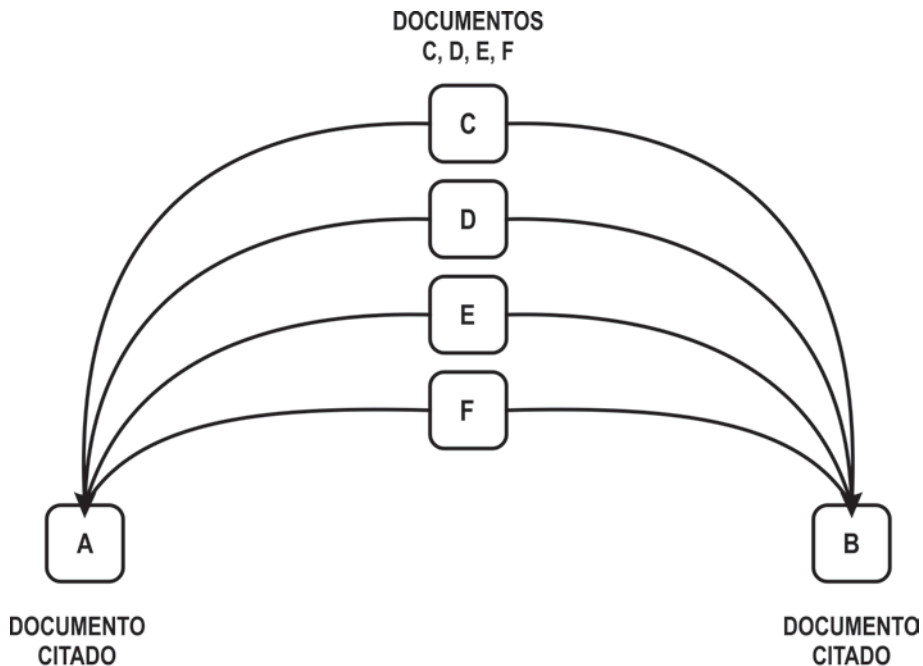
Neste trabalho pioneiro, Small define cocitação como a citação conjunta de dois artigos em uma literatura posterior (Figura 13) e afirma que para dois documentos ser fortemente cocitados, um grande número de autores deve citar os dois trabalhos simultaneamente. Baseado nessa característica da cocitação, o autor destaca que a cocitação é uma relação estabelecida pelos autores citantes (SMALL, 1973).

Segundo ainda o autor, ao se medir a força de cocitação entre dois documentos, se evidencia o grau de associação entre pares de documentos, segundo a compreensão da comunidade de autores citantes, isto é, conforme são reconhecidos pelos citantes (SMALL, 1973). Assim, a proximidade e a interlocução de dois documentos não são determinadas pelos autores dos documentos, mas definidas pela comunidade científica que se apropria do seu conteúdo e estabelece conexões durante o processo de geração de novos conhecimentos.

Desse modo, a intensidade da cocitação entre dois artigos é determinada pelo número de publicações em que ambos os artigos são citados juntos. No momento da publicação, os dois artigos podem parecer não estarem ligados. Suas ligações podem aparecer (e crescer com o tempo), quando estes artigos começam a ser citados conjuntamente na literatura científica. Assim, a força da cocitação é determinada pela reação dos pesquisadores em relação aos artigos publicados (MARSHAKOVA,

1981), ou seja, evidencia a estrutura de conhecimento de uma determinada área, segundo a compreensão da comunidade citante.

Figura 13 - Documentos A e B cocitados, em função dos documentos citantes C, D, E, F



Fonte: Mattos e Dias (2010, adaptada de Garfield, 2001).

A premissa fundamental da análise de cocitação (AC) é que, quando dois ou mais documentos ou autores são citados juntos, em uma pesquisa posterior, há uma proximidade temática, conceitual e/ou metodológica entre os citados, na visão do autor citante (SMIRAGLIA, 2011b) e desse modo, quanto maior o número de vezes que dois documentos são citados juntos, mais provável que eles sejam relacionados

em conteúdo (BELLARDO, 1980).

Segundo Small (2004), os cientistas definem as estruturas de seus campos científicos ao incluírem em suas listas de referências o que consideram relevante na literatura e esta compreensão da importância é consensual na comunidade. Afirma que a estrutura da ciência é, desse modo, gerada por padrões de correconhecimento, uma vez que ao cocitar dois documentos, os autores citantes estão correconhecendo os citados e criando uma associação de significados.

Alinhado a este pensamento, Hjørland (2002) considera que a Análise de Cocitação mostra detalhes e conexões reais entre pares de documentos. Segundo o autor, essas ligações representam o reconhecimento explícito da dependência entre artigos, pesquisadores, campos, abordagens, entre outros.

Desse modo, quanto maior a frequência de cocitação, mais próxima a relação entre os citados, sejam documentos, autores, periódicos, países ou temas. A análise de cocitação evidencia a similaridade, complementaridade, sobreposição de ideias entre os autores citados, mas também pode indicar a contraposição destas, no entendimento dos autores citantes.

Nesse contexto, segundo Hjørland (2013), entender os padrões de cocitação significa entender a história do reconhecimento e impacto acadêmico das publicações, uma vez que uma publicação pode ser potencialmente relevante para pesquisas futuras e pode, assim, potencialmente ser citada.

Como proposta por Small (1973), a análise de cocitação ao se sustentar na identificação de artigos basilares e/ou de impacto na comunidade científica e suas ligações, propicia uma visão microscópica da rede de associações, influência e impacto da produção do conhecimento em uma comunidade científica, com autores de maior visibilidade podendo ser representados por vários documentos na rede de cocitações.

Nos estudos bibliométricos, conforme propostos por Small (1973), a AC utiliza a análise de redes baseada na coocorrência de artigos científicos nas listas de referências de uma literatura científica para evidenciar e traçar as influências e a evolução das ideias, paradigmas e colégios invisíveis na ciência (BAYER; SMART; MCLAUGHIN, 1990). A partir dos padrões de cocitação de uma literatura pode-se identificar e evidenciar a associação entre conceitos-chave, estabelecidos pela comunidade citante.

Destaca-se, assim, que a AC se embasa em documentos específicos como unidade de análise e na observação de quais pares de documentos são citados juntos na literatura e pode ser usada para estabelecer agrupamentos ou “core” de uma literatura científica.

Apesar da relevância dos trabalhos individuais para mudanças nos paradigmas científicos, o conjunto de publicações de um pesquisador representa melhor a sua influência e seu lugar na estrutura da ciência (BAYER; SMART; MCLAUGHLIN, 1990).

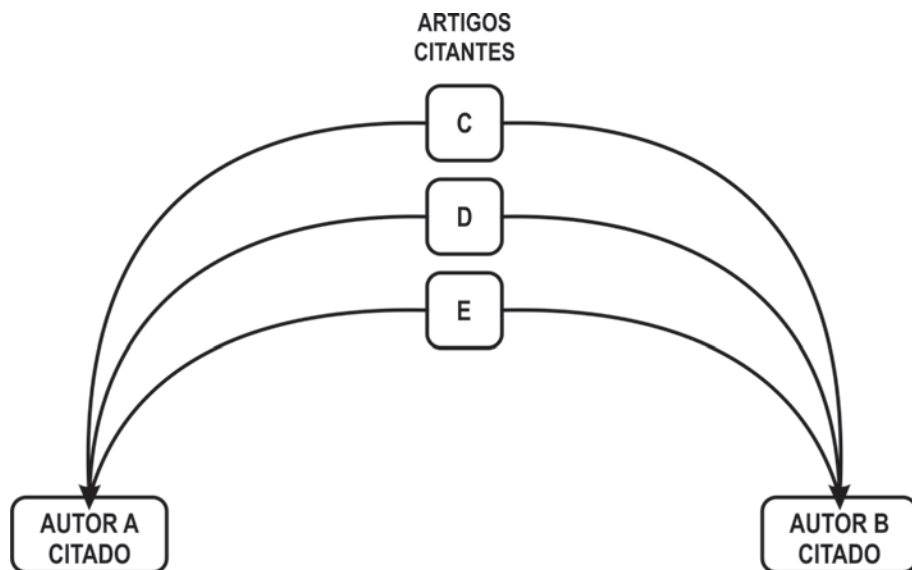
Sustentados na ideia de que o conjunto da obra de um pesquisador representa mais adequadamente a sua influência e seu lugar na estrutura

da ciência que a análise da contribuição individual dos seus artigos, para a evolução teórico-conceitual da sua área do conhecimento, em 1981, White e Griffith propuseram a análise de cocitação baseada no conjunto de contribuições do pesquisador a fim de se visualizar a estrutura intelectual da ciência.

Desse modo, a análise de cocitação de autores (ACA), proposta por White (1981) e White e Griffith (1981), baseia-se na frequência com que um autor está ligado a outro autor, identificada pela citação conjunta de ambos na literatura de uma comunidade científica, sem especificar quais trabalhos dos dois autores foram citados juntos nesta literatura (AHLGREN; JARNEVING; ROUSSEAU, 2003), conforme ilustrado na Figura 14.

Por esta abordagem, assume-se que quanto mais frequentemente dois cientistas são citados juntos e mais semelhantes são seus padrões de cocitação com outros autores, mais estreito o relacionamento entre eles (BAYER; SMART; MCLAUGHLIN, 1990).

Figura 14 - Autores A e B cocitados, em função dos artigos citantes C, D e E.



Os autores A e B são cocitados porque ambos foram referenciados nos artigos C, D e E.

Fonte: Grácio (2016).

Os autores citantes, no desenvolvimento e comunicação de suas pesquisas, estabelecem relações de conteúdo entre os autores citados e criam uma associação de significados desses conteúdos trabalhados, ao fazerem intensa referência simultânea aos autores citados. Assim, a ACA propicia a visualização da estrutural intelectual de uma comunidade científica, evidenciando as associações consolidadas entre os autores citados, na forma como são compreendidas pelos autores citantes dessa comunidade.

Destaca-se, que a proximidade de dois autores detectada por sua alta frequência de cocitação pode ser oriunda de várias situações, entre elas: compartilhamento de referencial teórico e/ou metodológico,

similaridade e/ou complementaridade de conteúdos e oposição teórico e/ou metodológica, entre outras. Todavia, quaisquer que sejam as razões para suas citações conjuntas, na percepção dos autores citantes, há uma proximidade de assunto entre os cocitados, seja por similaridade ou contraposição de ideias, por complementaridade teórico-metodológica ou compartilhamento metodológico. Ainda, quanto maior a frequência de cocitação, mais consolidado está o correconhecimento estabelecido pelos citantes.

Considera-se, assim, que a principal função da análise de cocitação de autores é identificar os autores influentes em uma área do conhecimento e mostrar suas inter-relações, a partir das citações registradas (WHITE; McCAIN, 1998).

McCain (1990) destaca que a ACA tem como objetivo analisar a estrutura intelectual de um determinado domínio, mostrando sua estrutura social e cognitiva. Segundo a autora, a ACA parte do pressuposto que os pesquisadores, ao citarem fontes similares e próximas, abordam problemas de pesquisa semelhantes na comunidade científica a que pertencem. Nesse sentido, a ACA pode ser definida como análise de um conjunto de autores organizados estruturalmente em rede (social e cognitiva) de um determinado domínio.

Ao apresentar metodologicamente seis passos para o desenvolvimento de uma análise de cocitação de autores McCain (1990) destaca, no sexto passo, a relevância da consulta a especialistas do domínio

em estudo, a fim de trazer elementos epistemológicos, teóricos, históricos e sociais para a interpretação e validação do contexto em que os autores foram cocitados (OLIVEIRA; GRÁCIO, 2013). Nesse contexto, reforça-se o paradigma social da ACA, partilhado com a Análise de Domínio, proposta por Hjørland e Albrechtsen (1995) e Hjørland (2002), dado o caráter social, histórico e epistemológico dos motivos para citar, ou não, um autor.

Smiraglia (2007, 2009, 2011a) destaca o importante papel da ACA para a construção de mapas multidimensionais relativos aos parâmetros dos domínios, contribuindo para a visualização do conjunto de temas neles tratados e das suas fronteiras de pesquisa, permitindo a visualização efetiva da intenção e extensão dos domínios. Ainda, a visualização obtida por esse método permite-se observar a evolução do conhecimento e da partilha de informação em um domínio, ou entre domínios, assim como identificar seus paradigmas teóricos nucleares e as mudanças neles ocorridas, por tratar com as percepções da comunidade (citante), contribuindo para se identificar, evidenciar e visualizar como essa comunidade reconhece as similaridades entre os pesquisadores (citados) no domínio (SMIRAGLIA, 2007, 2009, 2011a).

Alinhado a esse entendimento, Gmür (2003) afirma que a frequência de cocitação entre dois autores citados determina como a estrutura de conhecimento de um domínio é percebida por seus pesquisadores.

Ampliando para outras unidades de análise, Miguel, Moya-Anegón e Herreno-Solana (2008) afirmam que a análise de cocitação, seja de

documentos, autores, jornais, especialidades ou áreas de conhecimento, produz representações válidas da estrutura intelectual de um domínio científico.

A importância da análise de cocitação para a visualização e compreensão das estruturas subjacentes de uma comunidade científica, suscitou estudos teóricos e metodológicos relativos à questão da forma mais adequada do indicador utilizado em sua análise: frequência absoluta ou índice relativo de cocitação. Esses estudos têm mostrado resultados algumas vezes discordantes, seja em defesa da utilização da frequência absoluta, seja em defesa de uma das distintas formas de índices relativos (em especial, Cosseno de Salton, Índice de Jaccard e correlação linear), gerando uma significativa literatura na área da Ciência da Informação.

Alguns autores se posicionaram favoravelmente ao uso dos índices relativos (AHLGREN; JARNEVING; ROUSSEAU, 2003; ECK; WALTMAN, 2008; LEYDESDORFF, 2008). Small e Sweeney (1985) observam que os índices relativos de cocitação proporcionam análises que agrupam os referentes teóricos de forma mais representativa da estrutura intelectual de um domínio científico, ao eliminar a dependência do volume de literatura referenciada, que varia entre as diversas áreas e domínios.

Leydesdorff e Vaughan (2006) destacam que as frequências absolutas de cocitação contribuem de forma significativa para a construção das redes de cocitação e avaliação da conexão da rede e proximidade científica entre os autores cocitados. Por outro lado, segundo esses

autores, os índices relativos de cocitação, associados aos procedimentos multivariados de agrupamento dos autores cocitados, são importantes por mostrar, de forma consistente, aspectos subjacentes da estrutura intelectual de um domínio científico, não identificáveis nas análises de cocitação somente com frequências absolutas.

Luukkonen et al. (1993) destacam a relevância de usar os índices absolutos e relativos nas análises relacionais entre autores, uma vez que cada um deles evidencia um aspecto diferente da estrutura de um domínio. Os índices absolutos expõem características relativas à centralidade dos autores e periferia nas redes, ao passo que os índices relativos explicitam a intensidade das relações entre os pares de autores.

No Brasil, alguns pesquisadores também têm tratado das questões relativas à adequação e aspectos representados pelos índices absolutos e relativos de cocitação, para a visualização da estrutura de um domínio científico. Destacam-se os estudos de Mattos e Dias (2010) e Grácio e Oliveira (2013, 2014, 2015b). Os primeiros apresentam pesquisa teórico-metodológica relativa ao debate existente na literatura científica da área acerca da obtenção das frequências de cocitação, construção e transformação destas em matrizes de proximidade e a escolha da medida de proximidade (absoluta ou relativa).

Em um estudo também de natureza teórico-metodológica Grácio e Oliveira (2013) realizam uma análise comparativa da contribuição dos indicadores de proximidade absolutos e relativos - Cosseno de Salton e

Índice de Jaccard -, para a análise de cocitação de autores, em que concluem que os índices relativos diferenciaram a intensidade de proximidade entre os autores não discriminada por frequências absolutas, em função da presença de cada autor no domínio estudado.

Em artigo subsequente, as autoras analisam a contribuição dos indicadores de proximidade absoluto e normalizado pelo Cosseno de Salton, para a análise de cocitação de autores, e apresentam a integração dos resultados relativos às proximidades reveladas em uma representação visual (rede) que reúne as informações advindas das duas medidas, visando contribuir para a melhor compreensão de um domínio do conhecimento científico (GRÁCIO; OLIVEIRA, 2014).

Em continuidade, Grácio e Oliveira (2015) analisam e comparam os resultados obtidos para a ACA, utilizando duas medidas relativas de proximidade entre autores - Coeficiente de Correlação de Pearson (r), considerada medida global de proximidade, e Cosseno de Salton (CS), medida local de proximidade - em que constatam que tendências semelhantes ocorrem para baixa proximidade local e baixa proximidade global.

3.2.3 Estudos comparativos entre ACA e ABA

A partir de um levantamento³³ realizado na base Scopus, com o termo de busca bibliographic coupling, Grácio (2016) recuperou um

total de 120 artigos publicados, em especial, por autores dos Estados Unidos (25), da China (12), de Taiwan (12), Alemanha (11) e Espanha (8) e socializados, mormente, nos periódicos: *Scientometrics* (34), *Journal of Informetrics* (9) e *Journal of the American Society for Information Science and Technology* (8).

Nesse levantamento, observou também que, embora o primeiro artigo sobre Acoplamento Bibliográfico date de 1963, até o início da década de 1990, somente 10% desse conjunto de artigos tinha sido publicado. No período entre 1993 e 2008 (16 anos), publicaram-se outros 28% desse conjunto de artigos. Assim, os estudos sobre acoplamento bibliográfico têm sido mais frequentes nos últimos 7 anos (2009 a 2015), uma vez que a maioria (62%) dos artigos recuperados, datam deste período.

Ainda, o levantamento realizado por Grácio (2016), na mesma base de dados, com os termos de busca *Co-citation analysis* ou *Cocitation analysis*, recuperou um total de 451 artigos, evidenciando que este método têm sido mais usado para se avaliar um domínio científico, em relação ao Acoplamento Bibliográfico, e corroborando observação de Glänzel (2003) e Boyack e Klavans (2010).

Para este conjunto de publicações, embora mais volumoso que o recuperado para a temática Acoplamento Bibliográfico, observaram-se tendências temporais similares, assim como de autoria e principais canais de comunicação. Em relação à autoria, a temática tem sido desenvolvida, em especial, por autores dos Estados Unidos (125), China (65), Canadá

(37), Reino Unido (33), Espanha (30), Alemanha (29) e Taiwan (24) e socializados, principalmente, nos periódicos: *Scientometrics* (91), *Journal of the American Society for Information Science and Technology* (45), *Information Processing and Management* (11) e *Journal of Informetrics* (10).

Em termos temporais, a temática também tem seu maior volume de publicações em anos mais recentes: 12% dos artigos publicados até o início da década de 1990; 33% publicado no período de 1993 a 2008; e 55% dos artigos foram publicados nos últimos 7 anos (2009-2015).

Embora os métodos de acoplamento bibliográfico e de cocitação sejam, atualmente, usados com frequência nas pesquisas relativas às frentes de pesquisa e às estruturas sociais e cognitivas de um domínio científico, a maioria emprega apenas um dos dois métodos e raramente os integram em suas investigações (HUANG; CHANG, 2015). Grácio e Oliveira (2013, 2014) citam diversos estudos de análise de cocitação, em âmbito internacional e nacional, e Huang e Chang (2015) mencionam aplicações do acoplamento bibliográfico, em âmbito internacional.

Recentemente, pesquisadores da área observaram a necessidade de estudos que analisem, de forma comparativa, esses dois métodos relacionais de análise de citação. Na última década, observam-se estudos comparativos entre o acoplamento bibliográfico e análise de cocitação, a fim de saber se estes são indicadores adequados e eficientes de similaridade ou proximidade temática. Entre esses estudos, destacam-se Jarneving (2005), Zhao e Strotmann (2008), Boyack e Klavans (2010),

Lu e Wolfram (2012) e Qiu, Dong e Yu (2014).

A análise de Jarneving (2005), comparando artigos bibliograficamente acoplado com artigos cocitados, conclui que a frente de pesquisa é retratada de formas distintas, de acordo com o método empregado.

Em estudo analisando a contribuição da Análise de Cocitação de Autores e do Acoplamento Bibliográfico de Autores para a visualização da estrutura do domínio científico da Ciência da Informação, Zhao e Strotmann (2008) concluem que os dois métodos se complementam e que, quando combinados, fornecem uma visualização mais abrangente e real da estrutura intelectual do domínio, permitindo observar a sua trajetória evolutiva, uma vez que o método ABA fornece uma imagem da estrutura em termos da atividade de pesquisa em curso e a ACA fornece uma imagem das influências intelectuais sobre as atividades de pesquisa.

Entre os resultados obtidos por Boyack e Klavans (2010), em estudo comparativo sobre a precisão de quatro métodos distintos para mapeamento do domínio da Biomedicina, entre eles o Acoplamento Bibliográfico e a Análise de Cocitação, conclui que cada método pode ser considerado uma forma de representar a frente de pesquisa no domínio estudado e aponta que, entre os métodos relacionais de citação, o acoplamento bibliográfico apresenta resultados ligeiramente superiores que a análise de cocitação para a visualização da estrutura científica do domínio estudado.

No estudo comparativo entre cinco medidas relacionais em nível micro (autores), aplicado à Ciência da Informação, entre elas, análise de cocitação e acoplamento bibliográfico, Lu e Wolfram (2012) concluem que cada método apresenta perspectivas e propriedades diferentes do domínio estudado.

Em estudo para examinar a contribuição de cinco tipos de redes de cocorrência de autores, entre elas acoplamento bibliográfico e análise de cocitação de autores, para a análise da comunicação científica e da estrutura intelectual do domínio da Ciência da Informação, Qiu, Dong e Yu (2014) concluíram que o Acoplamento mostra vantagem significativa ao revelar estrutura científica com mais alta precisão e exatidão, além de identificar mais subestruturas que os outros tipos de rede e que a rede de cocitação de autores tem capacidade moderada para revelar a estrutura de um domínio científico.

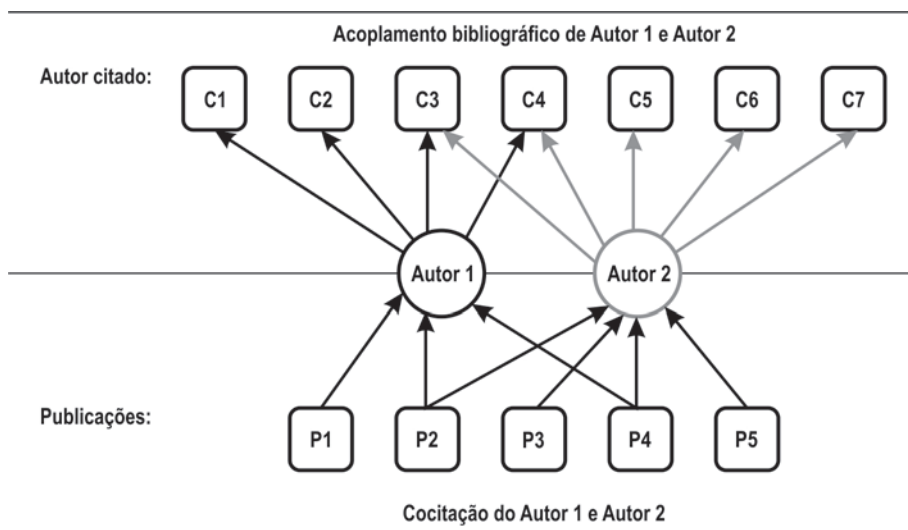
De forma sintética e ilustrativa, a Figura 15 exhibe, simultaneamente, as relações observadas pelo Acoplamento Bibliográfico de Autores (ABA) e pela Análise de Cocitação de Autores (ACA) para a caracterização da interlocução entre autores e da estrutura subjacente a um domínio científico, identificando e evidenciando as conexões sobre as quais se constrói e socializa o conhecimento.

Alguns estudiosos apontam que ao conectar artigos que cita(m) a(s) mesma(s) publicação(s), o acoplamento bibliográfico constitui uma análise com característica retrospectiva (sincrônica), ao passo que a cocitação,

ao conectar os artigos quando aparecem citados concomitantemente na literatura do domínio científico, configura uma análise com característica prospectiva, ou diacrônica (MARSHAKOVA, 1981; GARFIELD, 2001; LUCAS; GARCIA-ZORITA, 2014).

Em decorrência do tempo necessário para o reconhecimento dos autores em um domínio científico, o acoplamento bibliográfico pode oferecer visualizações mais precisas nos estágios iniciais de desenvolvimento de um domínio, quando comparado à cocitação. Ainda, a análise de cocitação restringe a análise aos artigos mais frequentemente utilizados, ao passo que o acoplamento bibliográfico se estende a praticamente todas as publicações referenciadas (GLÄNZEL, 2003).

Figura 15 - Ilustração comparativa do procedimento de análise entre Acoplamento Bibliográfico de Autores e Análise de Cocitação de Autores.



Acoplamento bibliográfico: Autor 1 e Autor 2 estão bibliograficamente acoplados pelos autores citados C3 e C4.
Cocitação de autores: Autor 1 e Autor 2 são cocitados pelas publicações P2 e P4.

Os dois tipos de análise relacional de citações têm sido utilizados para diversas questões consignadas à Ciência da Informação, como para a compreensão da comunicação científica, a frente de pesquisa e a estrutura intelectual de um domínio científico.

As análises relacionais de citação em nível micro (coocorrência de autor) têm prevalecido na análise da comunicação e das estruturas intelectuais de um domínio científico, uma vez que, nos outros níveis de agregação (revistas, instituições, países, entre outros), as redes são decorrentes das relações acadêmicas dos autores, como as relações de colaboração ou de citação realizadas pelos autores (QIU; DONG; YU, 2014).

Além disso, ao entendermos a literatura científica como a expressão mais objetiva da representação de um domínio científico, como resultado da atividade social de pesquisa, o estudo dessa literatura consiste uma abordagem para Análise de Domínio. Desse modo, as análises unidimensionais e relacionais de citação, partilham do paradigma social da Análise de Domínio, defendido por Hjørland (2002, 2017), ao basear-se na frequência de coocorrência de referências e citações a autores em um domínio (representado por sua literatura científica), uma vez que o ato de citar consiste de um processo social e as razões, positivas ou negativas, relevantes ou irrelevantes, pelas quais se cita, ou não se cita, um autor, configuram um aspecto social de um domínio e sua comunicação científica.