

Análise de redes sociais para visualização do comportamento científico

Maria Cláudia Cabrini Grácio
Ely Francina Tannuri de Oliveira

Como citar: GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Análise de redes sociais para visualização do comportamento científico. *In:* GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T.; ROSAS, F. S. (org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020. p. 164-191. DOI: <https://doi.org/10.36311/2020.978-65-86546-91-0.p164-191>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

Capítulo 7

ANÁLISE DE REDES SOCIAIS PARA VISUALIZAÇÃO DO COMPORTAMENTO CIENTÍFICO

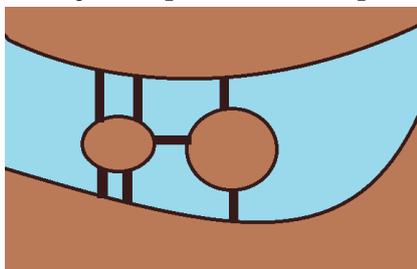
Maria Cláudia Cabrini Grácio
Ely Francina Tannuri de Oliveira

1 INTRODUÇÃO

A Análise de Redes Sociais (ARS) encontra seu fundamento matemático e de análise estrutural especialmente na Teoria dos Grafos, que tem como seu precursor o matemático Leonhard P. Euler (OTTE; ROUSSEAU, 2002).

Conhecido como problema de Euler, o primeiro registro sobre um estudo relacionado com o que contemporaneamente se denomina Teoria dos Grafos data do século XVIII. Ao visitar em 1736 a cidade de Königsberg, conhecida à época por ser moradia de diversos intelectuais, Euler tomou conhecimento de um problema que, embora aparentemente simples, não tinha sido ainda resolvido: no rio que cortava a cidade, havia duas ilhas que eram ligadas entre si por uma ponte e estas se ligavam às margens por mais seis pontes (Figura 1). O problema consistia em encontrar o percurso para um passeio que partindo de uma das margens e, atravessando uma única vez cada uma das sete pontes, retornasse à margem de partida. Euler provou que, para que esse passeio fosse possível, cada massa de terra deveria se ligar à outra por um número par de pontes.

Figura 1- Ilustração do problema das pontes de Euler.



Fonte: adaptado de Boaventura e Jurkiewicz (2009)

Ao longo do século XX, enquanto crescia o interesse dos matemáticos pelos estudos dos grafos, os conceitos de relação, rede e estrutura surgiam, quase independentemente, em várias disciplinas das Ciências Sociais e Comportamentais. Os pioneiros da análise de redes sociais vieram da Sociologia, da Psicologia Social e da Antropologia.

Em 1930, Jacob L. Moreno, psiquiatra e educador considerado o fundador da *sociometria*, explorou métodos para a análise da estrutura da escolha das amizades. Buscou identificar os meios pelos quais as relações nos grupos de pessoas serviam simultaneamente como limitação e oportunidade para suas ações.

Nesse estudo, ao pedir para os estudantes identificarem seus amigos, Moreno e sua equipe trabalharam com dados relacionais (contatos, conexões, amizades e ligações nos grupos). Esse tipo de dado é essencial para a análise da estrutura da ação social, ao relacionar um indivíduo a outro, não podendo ser reduzido às suas propriedades individuais. Assim, relações não são propriedades dos indivíduos, mas dos sistemas de indivíduos, as quais conectam pares de indivíduos dentro de sistemas relacionais mais amplos. Desse modo, é considerado que o método adequado para os estudos baseados em dados relacionais é a análise de redes (SCOTT, 2000).

Moreno denominou *sociograma* sua invenção destinada a representar as propriedades formais das configurações sociais: uma figura na qual os indivíduos (ou, unidades sociais) eram representados

por pontos em espaços bidimensionais e as relações entre os pares de indivíduos eram representadas por linhas ligando os pontos correspondentes, o que possibilitava visualizar e medir a estrutura de um grupo. Para Moreno, as configurações sociais tinham estruturas definidas e discerníveis e mapeá-las em sociogramas permitiria a visualização dos canais através dos quais a informação fluía de um indivíduo (unidade social) para outro e através dos quais um podia influenciar outro. Assim, a construção de sociogramas poderia identificar líderes e indivíduos isolados, descobrir assimetrias e reciprocidade e mapear as cadeias de conexão.

Essa invenção marca o início da sociometria, definida como a medida das relações interpessoais em pequenos grupos e precursora da análise de redes sociais, em que o termo ‘rede social’ designa um conjunto de entidades sociais (atores) e suas ligações (WASSERMAN; FAUST, 1994). Muitos atribuem o primeiro uso do termo “análise de redes sociais” ao antropólogo Barnes, em 1954.

Na década de 1950, os estudiosos descobriram que podiam usar matrizes para registrar e representar os dados de uma rede social. Nesse cenário, segundo Otte e Rousseau (2002), a *Análise de Redes Sociais* (ARS) sofreu muitas influências, principalmente da Matemática e da Ciência da Computação, para a investigação da estrutura social.

Assim, nessa época, ocorre a incorporação dos três principais fundamentos matemáticos dos métodos de rede: teoria dos grafos, teoria estatística e probabilística e modelos algébricos. A teoria dos grafos proporcionou tanto uma representação apropriada de uma rede social como um conjunto de conceitos para o estudo das propriedades formais das redes sociais. A teoria estatística tornou-se importante à medida que os pesquisadores começaram a estudar reciprocidade, equilíbrio e transitividade, entre outras propriedades. Os modelos algébricos têm sido usados para o estudo das redes multirrelacionais.

Atualmente, a ARS é considerada um procedimento interdisciplinar, desenvolvido sob diversas influências, além da Matemática e Ciência da Computação, da Sociologia, Psicologia

Social, Antropologia, entre outros campos. Seu uso vem crescendo significativamente nos últimos 20 anos em função do aumento da quantidade de dados disponíveis para análise, do desenvolvimento tecnológico, ampliando as áreas de conhecimento que a utilizam para representar a informação relacional presente em suas análises. Esta tendência tem sido observada em artigos científicos, diferentes bases de dados e programas de pesquisa. A partir de pesquisas em bases de dados, Otte e Rousseau (2002) e Borgatti e Foster (2003) verificaram que o total de artigos sobre o tema teve um crescimento exponencial no período de 1974 a 2000.

Na Ciência da Informação, a ARS vem ganhando espaço, especialmente, pelo fato de oferecer aporte para a análise de diferentes objetos de pesquisa, juntamente com modelos matemáticos, os quais têm tido uso ainda incipiente na área (MATHEUS; SILVA, 2009).

Sob a perspectiva da Ciência da Informação (CI), os estudos das ligações dentro das organizações por meio da ARS são capazes de identificar e analisar os fluxos de informação entre os atores (unidades sociais). Assim, nos estudos do campo da CI que têm como objeto de investigação os fluxos de informação e a geração de conhecimento, tanto no âmbito empresarial, como social ou científico, entre outras organizações, pode-se contar com uma ampla literatura que utiliza a metodologia de análise de rede sociais, sendo necessário apenas o ajuste para o objeto de pesquisa do campo da CI. No Brasil, o artigo Marteleto (2001) é considerado o primeiro trabalho tratando de ARS no âmbito da Ciência da Informação (VANZ, 2009).

Particularmente, no campo da análise da comunicação científica, beneficiando-se da flexibilidade do conceito de ator (ou agente), a *Análise de Redes Sociais* tem se mostrado uma relevante ferramenta adicional para os estudos bibliométricos e informétricos (WASSERMAN; FAUST, 1994). Nesse escopo, vem se consolidando como uma metodologia recorrente para os estudos destinados às análises de colaboração científica, coautoria, citação, acoplamento bibliométrico, cocitação, coocorrências de palavras, entre outras formas de interação

social na atividade científica, a fim de melhor compreender e visualizar o comportamento de uma comunidade ou campo científico (OTTE; ROSSEAU, 2002).

2 TEORIA DAS REDES SOCIAIS: CONCEITOS FUNDAMENTAIS

O ambiente social pode ser analisado a partir do ponto de vista dos padrões ou regularidades nos relacionamentos entre entidades que nele interagem. Nessa perspectiva, os elos (ou conexões) estabelecidos entre as unidades (entidades) sociais são componentes fundamentais para a compreensão de um fenômeno social.

Em uma rede social, as unidades sociais podem ser pessoas, departamentos dentro de uma organização, instituições, agências de serviço público em uma cidade, estados, países, entre outros, cujas ligações são analisadas (WASSERMAN; FAUST, 1994).

Neste cenário, o conjunto de estudos que se ocupa da investigação empírica de um fenômeno social a partir da identificação dos relacionamentos entre as entidades sociais e a análise dos padrões e as implicações dessas ligações é denominado Análise de Redes Sociais (WASSERMAN; FAUST, 1994). Assim, em ARS, as conexões (ou ligações) entre as unidades (entidades) sociais são o foco primário para analisar e conhecer um grupo e os atributos das entidades que compõem o grupo são secundários, auxiliando a compreensão das características das relações estabelecidas. Nesse contexto, salienta-se que o relacionamento entre um par de entidades sociais é uma propriedade do par e não uma característica inerente das entidades, ou seja, a unidade básica em que as ligações relacionais são medidas são pares de entidades, não a propriedade individual de uma ou outra entidade envolvida na relação.

A Análise de Redes Sociais permite retratar, mensurar e descrever as relações em um conjunto de entidades sociais a fim de caracterizar e analisar a estrutura vigente. Salienta-se, todavia, que embora em uma ARS a ênfase principal seja dada às relações

presentes entre os unidades sociais que compõem a rede, para um amplo entendimento de um fenômeno social, ela não pode prescindir da compreensão das características individuais dos seus componentes (WASSERMAN; FAUST, 1994; OTTE; ROUSSEAU, 2002).

Destaca-se que o estudo da dinâmica de redes deve considerar que as unidades sociais são entidades que evoluem ao longo do tempo. Suas propriedades e atributos mudam, assim como a maneira como interagem, levando a uma reconfiguração global da estrutura da rede quando analisada através do tempo. É essa perspectiva que permite entendermos as redes sociais como sistemas dinâmicos, não considerando apenas suas propriedades topológicas/estruturais, mas também suas propriedades dinâmicas. A dinâmica da rede surge como possibilidade de pesquisa quando observamos o processo que ocorre quando as conexões e entidades surgem ou desaparecem da rede, indicando as alterações de sua estrutura ao longo do tempo (NEWMAN; BARABASI; WATTS, 2006).

Desse modo, ainda que os estudos de redes sociais tenham se desenvolvido a partir de uma visão estruturalista, com grande repercussão na ciência e influência em pesquisas de diversas áreas (WASSERMAN; FAUST, 1994), eles têm caminhado para uma perspectiva de sistemas dinâmicos, por meio da visualização da evolução das tendências utilizando várias redes construídas em janelas temporais consecutivas.

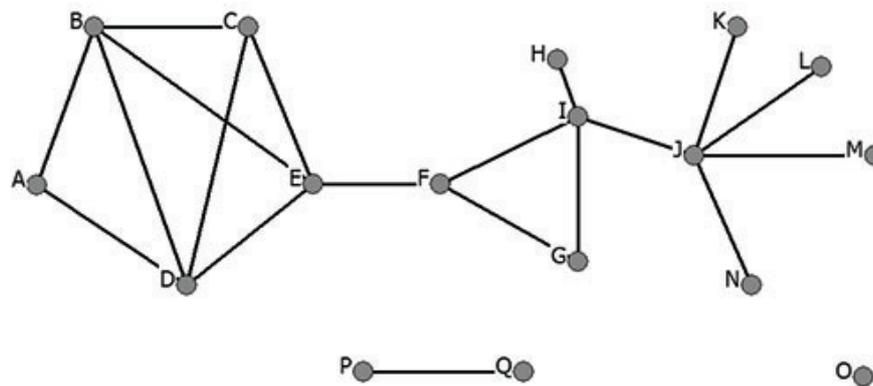
Entre os princípios relevantes da ARS, destaca-se o fato de as unidades sociais e suas ações serem vistos como interdependentes mais que autônomas e independentes e de as conexões entre elas serem canais para transferência ou “fluxo” de recursos, sejam eles materiais ou não materiais (WASSERMAN; FAUST, 1994).

Wasserman e Faust (1994) definem os seguintes conceitos basilares envolvidos nos estudos de Análise de Rede Social: ator, laço ou ligação relacional, díade, tríade, subgrupo, grupo, relação e rede social.

Um *ator* é uma unidade discreta, individual ou coletiva que

corresponde a uma entidade social. Assim, o conceito de ator é flexível. Em uma análise bibliométrica, pode se referir a pesquisadores (nível micro de agregação), assim como a departamentos, grupos de pesquisa, periódicos (nível meso) ou a países, regiões, campos do conhecimento (nível macro), entre outros. A rede social desenhada na Figura 2¹ é composta por 17 atores rotulados pelas letras A a Q.

Figura 2- Ilustração rede social hipotética



Fonte: elaboração própria.

Uma *ligação*, também denominada *laço relacional*, é o elo que estabelece um vínculo entre pares de atores, constituindo o canal para o fluxo de recursos materiais ou não materiais. Entre os tipos de ligação (laços) analisados nos estudos bibliométricos, destacam-se os seguintes: coautoria, citação, cocitação, acoplamento bibliográfico,

¹ A rede da Figura 2 foi gerada por meio do software Ucinet. Destaca-se, todavia, que há diversos softwares disponíveis para gerar uma rede social e calcular seus indicadores; entre eles, mencionam-se: Ucinet, Netdraw, Pajek, Cite Space, VosViewer e Gephi. Com exceção do Ucinet, todos são softwares livre. Para uma apresentação mais detalhada sobre os softwares disponíveis para a geração de redes e cálculo dos seus indicadores, há diversas publicações; entre elas, citam-se: Alejandro e Norman (2005), Pinto et al. (2009), Batagelj e Mrvar (2011) e Van Eck e Waltman (2016).

coocorrência de palavras, entre outras. Na rede ilustrada na Figura 2, observam-se 19 ligações, as quais, em estudos bibliométricos, podem representar coautoria, cocitação ou acoplamento bibliográfico entre os 17 atores ilustrados, entre outras possibilidades. Destaca-se a ausência de laços conectando o ator O com qualquer outro ator da rede, sendo ele, portanto, denominado *ator isolado*.

Uma *díade* consiste em um par de atores e os (possíveis) laços entre eles. Na Figura 2, há 136 díades, resultante do cálculo $(17 \times 16) / 2$, correspondentes do pareamento de cada um dos 17 atores com todos os outros 16 atores da rede. Porém, deste total de 136 díades, somente 19 têm laço entre os 2 atores que definem a díade. A título de exemplificação, a díade formada pelos atores A e B tem um laço (ligação) entre eles e a díade formada pelos atores A e P não apresenta laço entre os atores. Uma *tríade* é um subconjunto de três atores e os (possíveis) laços entre eles. A título de ilustração, na Figura 2, a tríade composta pelos atores A, B e D, possui laços entre todos eles. Os atores A, P e Q formam outra tríade, todavia somente com laço entre P e Q.

Estendendo os conceitos de díade e tríade, define-se *subgrupo* como qualquer subconjunto de atores e todas as ligações entre eles. Na rede ilustrada na Figura 2, os atores A, B, C, D, E e F e suas respectivas ligações constitui um subgrupo. Embora menos intuitivo, outro subgrupo é formado pelos atores A, B, E, F, G e I, com suas respectivas ligações.

Um *grupo* constitui uma coleção completa finita de atores sobre os quais as ligações são medidas. Uma *relação* é uma coleção de ligações de um tipo específico entre os membros de um grupo.

Uma *rede social* é um conjunto ou conjuntos finitos de atores e a relação ou relações definidas entre eles. A título de exemplo, a Figura 2 é uma rede composta pelo conjunto finito de 17 atores, que definem um único grupo, e uma relação, definida pelo conjunto de 19 ligações estabelecidas entre os 17 atores. Em pesquisas bibliométricas, esse grupo de atores poderia estar representando um conjunto de pesquisadores, de periódicos, de instituições ou de países e a relação

observada poderia estar representando a colaboração científica entre eles ou outra relação, como a de cocitação entre os atores.

Outros conceitos importantes apresentados por Wasserman e Faust (1994) envolvidos na análise de redes sociais são: ator ponte, caminho, componente e clique.

Um *ator ponte* é aquele responsável pela ligação de dois subgrupos de uma rede social. É, assim, um ator que faz parte de um subgrupo que interage com um ator de outro subgrupo. Desse modo, sua importância está associada à intermediação do fluxo da informação que realiza entre dois subgrupos. Na Figura 2, os atores E, F, I e J são exemplos de atores ponte.

Um *caminho* é o conjunto de ligações existentes percorridas para conectar dois atores e um *caminho geodésico* é o caminho mais curto entre dois atores.

Um tipo especial de subgrupo é denominado *componente*, definido como um subgrupo em que há um caminho entre todos os seus atores. Se toda a rede formar um componente, diz-se que está *totalmente conectada*. A Figura 2 é composta por um componente principal, formado por 14 atores (A a N), e outro componente formado por dois atores (P e Q), além de um ator isolado (ator O).

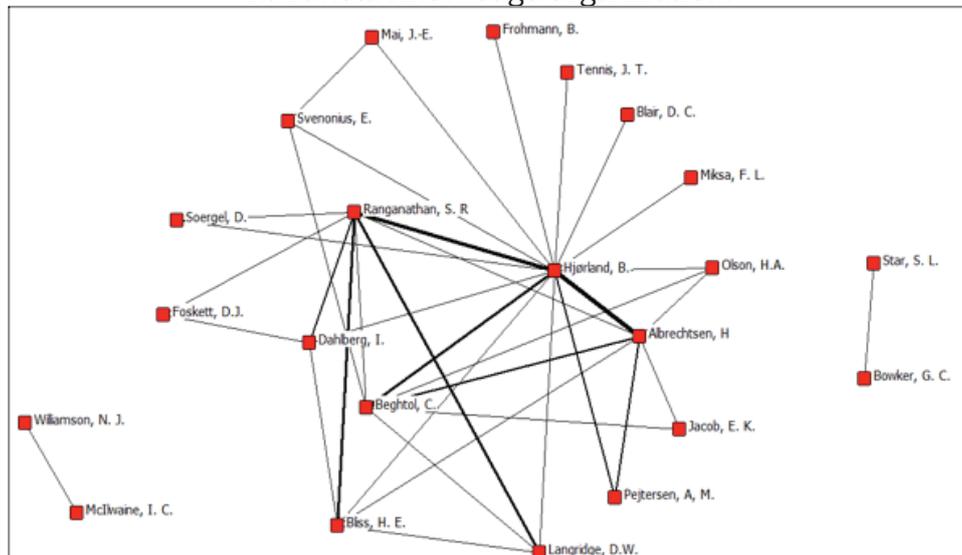
Estendendo esse conceito, define-se uma clique² como um componente no qual todos os atores estão diretamente conectados entre si, isto é, há uma ligação entre quaisquer dois atores que o compõe. Na Figura 2, os atores B, C, D e E formam uma clique, assim como os atores F, G e I, mais do que uma tríade, constituem uma clique.

As redes sociais podem ser classificadas em dois tipos, segundo a quantidade de grupos de atores envolvidos, a saber: *um modo* e *dois modos*.

Nas redes *um modo*, os atores pertencem a apenas um único grupo. Como exemplo, citam-se as redes de coautoria, de cocitação e de coocorrência de palavras-chave, em que os atores formam um

² Esclarece-se que o termo “clique” encontra-se como um substantivo feminino nos textos de Boaventura e Jurkiewicz (2009) e Marteleto (2001) e como substantivo masculino no texto de Matheus e Silva (2009).

Figura 4 - Rede de cocitação de autores na revista Knowledge Organization.



Fonte: Guimarães, Oliveira e Grácio (2012).

Legenda: A espessura das linhas é proporcional à intensidade da cocitação entre os autores cocitados.

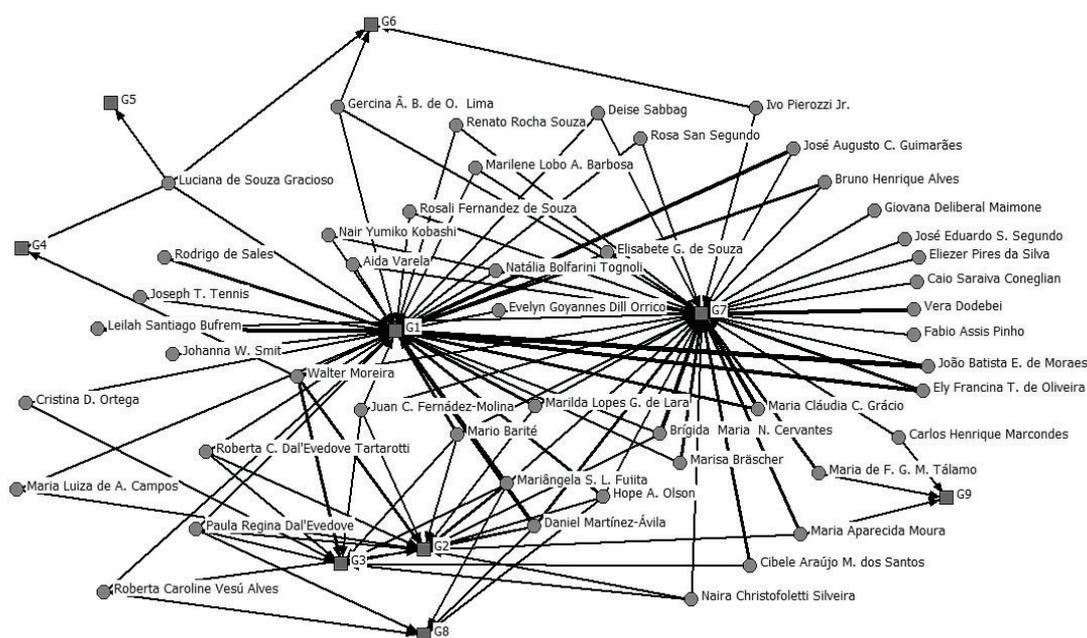
As redes *dois modos* são formadas por dois grupos distintos de atores, cada um deles com atributos particulares. Como exemplo, citam-se as redes de citação, de participação de pesquisadores em eventos, de atuação temática dos pesquisadores, cada uma constituídas por dois grupos de atores, a saber: rede de citação - grupo de atores citantes e grupo de atores citados; rede de participação em eventos científicos - grupo de pesquisadores e grupo de eventos científicos; rede de atuação temática - grupo de pesquisadores e grupo de temáticas trabalhadas. A rede exemplificada na Figura 5 é do tipo *dois modos*.

As redes sociais podem também ser categorizadas em função do tipo de direcionamento da relação envolvida: *direcionada* ou *não direcionada*.

Uma *rede direcionada* (ou *orientada* ou *unidirecional*) é aquela em que a relação não é necessariamente simétrica, envolvendo um ator

transmissor e outro receptor. São representadas por vetores. A rede de citação (ator transmissor: autor citante e ator receptor: autor citado) em um campo ou domínio científico e a rede de adesão dos pesquisadores aos GTs da ANCIB seriam exemplos de redes direcionadas analisadas nos estudos bibliométricos, entre outras.

Figura 5 - Autores mais produtivos e os grupos que constituem os referentes teóricos basilares, segundo Dahlberg (1993).



Fonte: Alves e Oliveira (2016).

Legenda: Os quadrados verdes indicam os diferentes referentes teóricos basilares de Dahlberg e os círculos laranja representam os autores mais produtivos na ISKO Brasil. A espessura das linhas é proporcional à intensidade da fundamentação do autor pelo referente teórico.

Uma rede *não direcionada* (*não orientada ou bidirecional*) é aquela em que a relação entre os atores é recíproca (simétrica). Relações de coautoria, de cocitação e de coocorrência de palavras geram redes bibliométricas não direcionadas. As Figuras 3 e 4 são

compostas por um único grupo de atores (autores) e um único tipo de ligação não direcionada (coautoria e cocitação, respectivamente), por outro lado, a rede presente na Figura 5 é de ligação direcionada.

Além disso, as redes podem ser classificadas em função do tipo de valoração definida na relação estabelecida: *dicotômica* ou *valorada*.

Uma rede é denominada *dicotômica* (ou *não valorada* ou *binária*) quando a ligação vigente entre os atores do grupo não possui mensuração de intensidade, tendo somente duas possibilidades de situação entre eles: ausência de ligação (usualmente rotulada por 0) ou presença da ligação (usualmente rotulada por 1). A rede presente na Figura 2 é um exemplo de rede não valorada, uma vez que os segmentos de reta (ligações) presentes entre os atores são todos de mesma espessura (valor da ligação igual a 1).

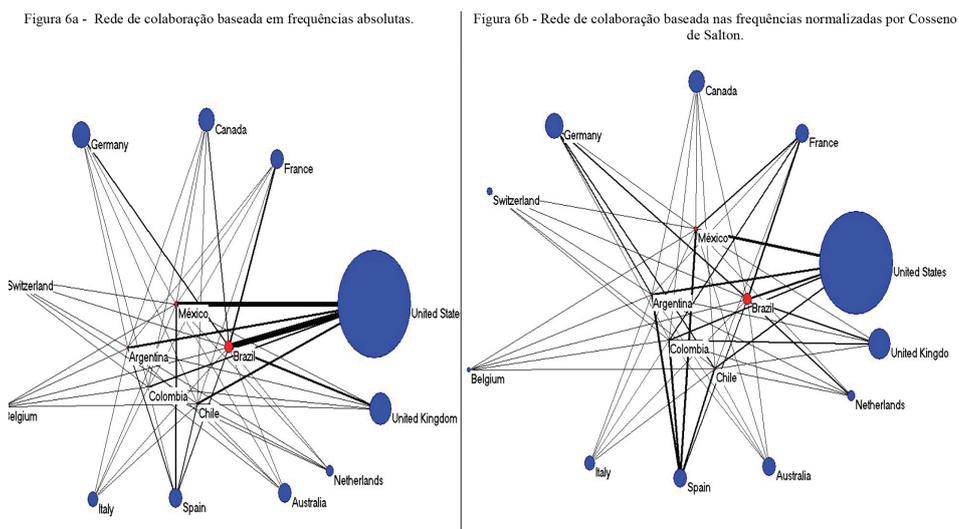
Quando a ligação entre os atores varia em uma escala de mensuração discreta (números naturais) ou contínua (números reais), a rede é denominada *valorada*. Nas redes bibliométricas, uma valoração discreta pode resultar da contagem da frequência de coautorias, de cocitação ou de acoplamento bibliográfico entre autores, assim como da frequência de coocorrência de termos ou de palavras-chave em publicações científicas. Neste contexto, a valoração contínua está, usualmente, associada à normalização³ das frequências absolutas observadas, por meio do Cosseno de Salton, Índice de Jaccard e Coeficiente de Correlação de Pearson. Nas Figuras 3 e 4 são apresentados exemplos de redes com valoração discreta, uma vez que na Figura 3, as espessuras das linhas (ligações) representam as frequências de coautorias entre as 30 instituições que compõem a rede e, na Figura 4, as espessuras das linhas são proporcionais

³ Os indicadores normalizados de ligação padronizam as unidades de medida e revelam aspectos não explicitados nas frequências absolutas de ligação. Matriz com valores normalizados têm sido preferencialmente usadas em análises bibliométricas com procedimentos multivariados, especialmente nas análises de clusters, por revelarem, de modo consistente, aspectos subjacentes à estrutura intelectual de uma área científica, não identificados pelas análises baseadas em frequências absolutas, ao eliminar a dependência do volume da literatura científica (GRÁCIO; OLIVEIRA, 2015).

às frequências de cocitação entre os 22 autores cocitados na rede, resultante da análise dos 57 artigos publicados pelos 16 pesquisadores mais produtivos nos Anais da ISKO Internacional (1993-2011).

A Figura 6 apresenta a rede de coautoria entre os cinco principais países da América Latina produtores de conhecimento científico no campo das Ciências Ambientais, gerada a partir da valoração discreta, resultante do número de artigos em coautoria (Figura 6a), e a partir da valoração contínua, obtida pela normalização da frequência de coautoria por Cosseno de Salton (Figura 6b).

Figura 6 - Rede de colaboração entre os países da América Latina (AL) e seus principais países coautores no campo das Ciências Ambientais (2006-2015).



Fonte: Grácio e Oliveira (2016).

Legenda: A cor vermelha dos círculos indica os países latino-americanos analisados (países ego) e a cor azul indica os principais países coautores dos países latino-americanos. A área dos círculos é proporcional à intensidade da produção científica do país. A espessura das linhas é proporcional à intensidade da coautoria entre os países, expressa em frequência absoluta (Figura 6a) ou normalizada (Figura 6b).

Uma rede pode ainda ser classificada como: *global ou ego*.

Uma rede global, ou sociocêntrica, mapeia os padrões de conexão entre todos os atores que compõem o grupo, podendo ser, ou não, analisados subgrupos no seu interior. Expressa, assim, as ligações existentes na rede como um todo (COSTA; FERREIRA, 2013). Análises de rede baseadas no modelo sociocêntrico são de significativa importância, dado o poder restritivo de uma rede sobre seus membros não ser mediado apenas por suas ligações diretas. Neste tipo de estrutura de rede, o cerne da análise está na concatenação de ligações indiretas, através de uma configuração das ligações com propriedades que existem independentemente entre pares particulares de indivíduos (SCOTT, 2000).

Uma rede egocêntrica, ou local, está focada em um (ou mais) ator (es), denominado(s) ator(es) ego, considerado(s) de referência numa dada estrutura social. Parte-se do conceito sociológico de *star*, da posição de centralidade de certos indivíduos ou instituições que assumem um papel relevante pela influência que exercem sobre os outros participantes da rede (SCOTT, 2000, COSTA; FERREIRA, 2013).

Assim, em uma rede global (ou sociocêntrica), a análise contempla todas as relações possíveis entre os atores do grupo, ao passo que em uma rede egocêntrica, a análise se volta para as ligações que partem de determinado(s) ator(es) ego para os outros atores (alters), uma vez que a atenção está centrada na compreensão do papel de um ponto de referência (ator) específico (SCOTT, 2000, SILVEIRA, 2013).

A rede presente na Figura 3 é do tipo global, uma vez que mapeia todas as ligações (coautorias) entre as 30 instituições com pesquisas em colaboração científica no GT7 da ANCIB no período de 2003 a 2009. As redes presentes nas Figuras 4, 5 e 6 são do tipo egocêntrica, uma vez que diagramam somente as ligações entre os atores de destaque: na Figura 4, somente as citações a partir dos autores mais citados (autores ego) na produção científica dos pesquisadores mais produtivos na ISKO internacional (1993-2011); na Figura 5, os autores mais produtivos (autores ego) na ISKO Brasil; na Figura 6, os principais

países coautores dos cinco países latino-americanos mais produtivos (países ego) em Ciências Ambientais.

3 INDICADORES DE REDES: DENSIDADE E CENTRALIDADE

A teoria dos grafos tem sido útil para a ARS por proporcionar um vocabulário e conceitos que permitem a sistematização e cálculos a partir dos quais os modelos e as propriedades das estruturas sociais podem ser denotados e demonstrados de forma precisa e objetiva.

Um grafo pode ser definido como uma coleção de nós, ou vértices, conectados por um conjunto de segmentos de reta, ou arestas, que, juntos, determinam uma rede (OTTE; ROUSSEAU, 2002). Em linguagem matemática, um grafo (G) é uma estrutura formada por dois conjuntos V (conjunto de nós ou vértices) e A (conjunto de arestas ou segmentos de reta). Em termos formais: $G = (V, A)$.

Quando um grafo é usado para formalizar uma rede social, os vértices representam os atores e as arestas conectando os vértices representam as ligações entre os atores. Desse modo, um grafo é um modelo de uma rede social, que tem sido amplamente utilizado em ARS a fim de representar formalmente as relações sociais e quantificar as propriedades estruturais do grupo social analisado, em que a noção de Grupo em ARS é equivalente ao conjunto V e a noção de relação ao conjunto A da Teoria dos Grafos.

A partir desse modelo, pode-se utilizar as propriedades e indicadores da teoria dos grafos para analisar as redes sociais. A visualização de um grafo que modela uma rede social permite descobrir padrões não detectados por outras metodologias.

Destaca-se que as matrizes são uma forma alternativa de representar os dados de uma rede social. Grafos de redes sociais podem ser representados por matrizes, denominadas *matrizes de adjacência*, quando as linhas e as colunas representam os atores de uma rede *um modo*, e matrizes de incidência, quando as linhas representam os atores de um grupo e as colunas representam atores do outro grupo de

uma rede *dois modos*.

Desse modo, em redes *um modo*, a matriz é quadrada, uma vez que há um único grupo analisado e então os atores que estão nas linhas são os mesmos que estão nas colunas. Em relações não valoradas (dicotômicas), as células dessas matrizes registram a presença, ou ausência, da ligação entre os atores que estão na respectiva linha e coluna. Em relações valoradas, registram o valor da ligação entre esses atores. Além disso, se a relação é não direcionada (recíproca), esta matriz é simétrica.

A matriz presente na Figura 7 apresenta as ligações presentes na Figura 2 (rede *um modo* não valorada), em que o valor 1 representa a ocorrência da ligação e o valor 0 representa a ausência da ligação entre os dois atores que compõem a respectiva célula da matriz. Observa-se que a matriz é quadrada, com 17 linhas e 17 colunas, e simétrica, com o valor presente na célula correspondente à linha *i* e coluna *j* sendo o mesmo presente na célula correspondente à linha *j* e coluna *i*. Optou-se por registrar nas células da diagonal matriz, que corresponderia à ligação de um ator com ele mesmo, o símbolo “-” significando ausência de dado⁴.

⁴ Em redes *um modo* valoradas, antes da conversão de uma matriz com as frequências absolutas de ligação entre os atores para a respectiva matriz com valores normalizados, é necessário decidir acerca dos valores da diagonal, ou seja, os valores correspondentes da ligação do ator consigo mesmo. A literatura bibliométrica propõe algumas alternativas; são elas: utilizar a maior frequência de ligação do ator com os outros da rede; adotar a média da ligação de cada ator como valor da diagonal; utilizar zeros na diagonal; ou considerar a diagonal como um conjunto de valores perdidos (missing). Destaca-se que a última alternativa tem sido adotada mais frequentemente pela comunidade bibliométrica, por ser mais operacional e com menos viés conceitual (GRÁCIO; OLIVEIRA, 2015).

Figura 7 - Matriz com as ligações entre os atores representados na Figura 2

Ator	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
A	-	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
B	1	-	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
C	0	1	-	1	1	0	0	0	0	0	0	0	0	0	0	0	0
D	1	1	1	-	1	0	0	0	0	0	0	0	0	0	0	0	0
E	0	1	1	1	-	1	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	1	-	1	0	1	0	0	0	0	0	0	0	0
G	0	0	0	0	0	1	-	0	1	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0
I	0	0	0	0	0	1	1	1	-	1	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	1	-	1	1	1	1	0	0	0
K	0	0	0	0	0	0	0	0	0	1	-	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	1	0	-	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	1	0	0	-	0	0	0	0
N	0	0	0	0	0	0	0	0	0	1	0	0	0	-	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	1
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-

Fonte: Elaborado pelas autoras

Destaca-se que em redes valoradas, o número 1 registrado nas células da matriz da Figura 7 seria substituído pela intensidade da ligação entre os atores que definem a célula. A título de exemplo, se a rede ilustrada na Figura 2 correspondesse a uma rede de coautoria científica, a matriz da Figura 7 estaria formalizando as ligações existentes, em que o valor 1 corresponderia à presença de coautoria entre eles e o valor 0 a ausência de trabalhos em coautoria entre os

respectivos autores, todavia sem as intensidades de coautoria. Para uma rede valorada de coautoria, seria necessário substituir o número 1 pelas respectivas intensidades de coautoria entre os autores.

A fim de aprofundar a análise da estrutura de uma rede social, utilizam-se indicadores definidos para nós e grafos, entre eles: a densidade, que mede a coesão (da estrutura) do grafo; e medidas que descrevem o papel desempenhado pelos atores individualmente na rede, denominados indicadores de centralidade (dos nós): grau, proximidade e intermediação. Salienta-se que estes indicadores adquirem expressões matemáticas distintas, de acordo com a classificação da rede, i.e., se *um* ou *dois modos*, se valorada ou dicotômica (não valorada), se orientada (unidirecional) ou não orientada (bidirecional), embora tenham conceitualmente o mesmo significado.

As formulações matemáticas seguintes são referentes a redes *um modo*, não orientadas e dicotômicas. Nesse modelo, encontram-se as análises de coautoria ou colaboração científica, de citação, de acoplamento bibliográfico e de coocorrência de termos, para as quais as redes podem ser construídas com suas ligações valoradas, todavia para o cálculo dos seus indicadores de densidade e centralidade, utiliza-se a versão dicotômica da matriz valorada.

Densidade da Rede

A densidade (D) é um indicador que mede a intensidade de coesão (conexão) da rede. Quando há uma ligação direta entre todos os atores da rede, denomina-se esta por *rede completa*. Define-se a densidade de uma rede pelo quociente entre o total de ligações nela existentes e o número máximo possível de ligações (equivalente a ser uma rede completa). Logo, D é calculado por:

$$D = \frac{L}{N \times (N - 1) \div 2} = \frac{2 \times L}{N \times (N - 1)}$$

em que L = número de ligações existentes entre os atores; N = número de atores que formam a rede.

A densidade de uma rede varia entre 0, equivalente a todos os atores estarem isolados, e 1, equivalente a ser uma rede completa. Assim, quanto mais próximo de 1 o valor obtido para D , mais densa é a rede. O valor da densidade pode ser expresso também em porcentagem. A rede ilustrada na Figura 2 tem densidade igual a 0,14 (ou 14%), decorrente do fato de ter $L = 19$ ligações e $N = 17$ atores, constituindo assim uma rede de baixa densidade, considerando que do total de ligações possíveis entre todos os atores, apenas 14% estão presentes. A rede presente na Figura 3 tem densidade igual a 0,048 (ou 4,8%), evidenciando uma baixa relação de interlocução entre as instituições (GRÁCIO; OLIVEIRA, 2011) e a rede presente na Figura 4 tem densidade igual a 0,16 (ou 16%), resultante da presença das 37 ligações entre os 22 autores cocitados na rede, evidenciando que também esta rede tem baixa densidade.

Centralidade de Grau

A centralidade de grau de um ator (nó) é definida como o número de ligações que possui. Permite caracterizar a posição do ator em relação ao grupo na rede. Em termos matemáticos, a centralidade de grau de um ator é definida por:

$$d(i) = \sum_{j=1}^N d_{ij}$$

em que: $l_{ij} = 1$, se existe uma ligação entre os atores i e j , para $1 \leq j \leq N$ e $j \neq i$ e $l_{ij} = 0$, se não há uma ligação entre eles; com N = número de atores na rede.

Em uma rede de coautoria, a centralidade de um autor é equivalente ao número de pesquisadores distintos com os quais o autor publicou trabalhos em coautoria. A título de ilustração, na Figura 2, $d(A) = 2$, uma vez que há ligação do ator A com os atores B e

D, e $d(J) = 5$, advinda das cinco ligações do ator J.

A fim de eliminar a influência do tamanho do grupo na avaliação da centralidade dos atores, permitindo análises comparativas entre grupos de tamanhos distintos, a centralidade de grau de um ator i em uma rede de tamanho N pode ser normalizada dividindo o valor de d por $N-1$. Em termos matemáticos, a centralidade de grau normalizada, denominada $d'(i)$, é obtida por:

$$d'(i) = \frac{d(i)}{N - 1}$$

Esta normalização faz com que o valor da centralidade de grau varie entre 0, equivalente a um ator isolado, e 1, equivalente a um ator conectado com todos os outros atores da rede. Também o valor de centralidade de grau normalizada pode ser expresso em porcentagem.

A Tabela 1 mostra a centralidade de grau dos 17 atores que compõem a rede da Figura 2, calculada por meio do software Ucinet.

Tabela 1 - Centralidade de grau dos 17 atores que compõem a rede da Figura 2

Ator	Centralidade de grau (d)	Centralidade de grau normalizada (d')	Ator	Centralidade de grau (d)	Centralidade de grau normalizada (d')
A	2	13%	J	5	31%
B	4	25%	K	1	06%
C	3	19%	L	1	06%
D	4	25%	M	1	06%
E	4	25%	N	1	06%
F	3	19%	O	-	-
G	2	13%	P	1	06%
H	1	06%	Q	1	06%
I	4	25%			

Fonte: Elaborado pelas autoras

Para o ator A, a centralidade de grau $d'(A) = 2/(17-1) = 0,13$ ou 13%, significando que o ator A está conectado com 13% dos atores da rede e $d'(J) = 5/(17-1) = 0,31$ ou 31%, sendo este o ator com a maior centralidade de grau na rede.

Centralidade de Proximidade

A centralidade de proximidade de um ator i , denominada $c(i)$, é igual à soma da menor distância deste ator a cada um dos outros atores da rede. Formalmente, para um ator i , a centralidade de proximidade, $c(i)$, é definida por:

$$c(i) = \sum_{j=1}^N d_{ij} \text{ para } i \neq j$$

em que d_{ij} é o número de ligações em um caminho geodésico (mais curto) do ator i para o ator j e N = número de atores na rede.

Diferentemente da centralidade de grau, a proximidade é uma medida inversa de centralidade: valores baixos para este indicador significam que o ator está próximo (relacionado) de todos os outros atores da rede por caminhos curtos (pequenas distâncias), ou seja, é um ator central (OTTE; ROUSSEAU, 2002). Desse modo, a fim de tornar este indicador uma medida direta de centralidade, a proximidade normalizada, denominada $c'(i)$, é definida por:

$$c'(i) = \frac{N - 1}{c(i)}$$

Em relação à Figura 2, considerando que a rede é composta por um componente principal (constituído por 14 atores), um segundo componente composto por 2 atores e 1 ator isolado, não é possível calcular a centralidade de proximidade dos seus atores, uma vez que, por definição, para a obtenção da centralidade de proximidade de um ator é necessário que exista um caminho entre este e qualquer outro ator da rede. Para esta rede, o cálculo da centralidade de proximidade

só é possível se restringirmos a análise ao seu componente principal, isto é, calcular a centralidade de proximidade somente dentro do grupo de 14 atores rotulados de A a N.

A Tabela 2 mostra os valores da centralidade de proximidade restrita ao subgrupo de atores que formam o componente principal da Figura 2, obtidos por meio do software Ucinet.

Tabela 2 - Centralidade de proximidade dos 14 atores do componente principal da Figura 2

Ator	Centralidade de proximidade (c)	Centralidade de proximidade normalizada (c')	Ator	Centralidade de proximidade (c)	Centralidade de proximidade normalizada (c')
A	51	0,255	H	39	0,333
B	40	0,325	I	27	0,481
C	41	0,317	J	31	0,419
D	40	0,325	K	43	0,302
E	32	0,406	L	43	0,302
F	28	0,464	M	43	0,302
G	33	0,394	N	43	0,301

Fonte: Elaborado pelas autoras

Dentro do componente principal da Figura 2, o ator com a maior centralidade de proximidade é ator I, com valor igual a $c'(I) = 0,481$, indicando que este é o ator com as menores distâncias em relação a todos os outros atores deste subgrupo. Por outro lado, o ator A é aquele com a menor centralidade de proximidade ($c'(A) = 0,255$) dentro do componente principal analisado, i.e., aquele com as maiores distâncias em relação as demais do subgrupo.

Centralidade de Intermediação

A centralidade de intermediação de um ator é definida como o número de caminhos geodésicos entre quaisquer outros dois atores, que passam por ele. Assim, a centralidade de intermediação mede a capacidade de o ator intermediar o fluxo da informação na rede. Formalmente, a centralidade de intermediação de um ator i , denotada como $b(i)$ é obtida por:

$$b(i) = \sum_{j,k}^N \frac{g_{jik}}{g_{jk}}, \text{ para } i \neq j \neq k$$

onde g_{jk} é o número de caminhos mais curtos do ator j para o ator k ($j, k \neq i$) e g_{jik} é o número de caminhos mais curtos do ator j para o ator k passando pelo ator i e $N = n^{\circ}$ de atores na rede.

A centralidade de intermediação normalizada é obtida por:

$$b'(i) = \frac{2 \times b(i)}{N^2 - 3 \times N + 2}$$

A Tabela 3 mostra a centralidade de intermediação dos 17 atores da rede da Figura 2, calculada por meio do software Ucinet.

Tabela 3 - Centralidade de intermediação dos 17 atores da rede da Figura 2

Ator	Centralidade de intermediação (b)	Centralidade de intermediação normalizada (b')	Ator	Centralidade de intermediação (b)	Centralidade de intermediação normalizada (b')
A	-	-	J	42	35%
B	5,5	4,6%	K	-	-
C	-	-	L	-	-
D	5,5	4,6%	M	-	-
E	36	30%	N	-	-
F	40	33,3%	O	-	-
G	-	-	P	-	-
H	-	-	Q	-	-
I	47	39,2%			

Fonte: Elaborado pelas autoras

O ator I tem centralidade intermediação (d (I)) igual a 47, equivalente a $d'(I) = 39,2\%$, significando que é o ator com o maior poder de intermediação na rede. Por outro lado, a maioria (11) dos atores da rede não faz a intermediação entre quaisquer outros atores da rede, uma vez que seus indicadores de intermediação tem valor igual a zero.

4 CONSIDERAÇÕES FINAIS

Este capítulo apresentou os principais conceitos e procedimentos envolvidos na análise e compreensão de redes sociais. Tratou, ainda, dos indicadores destinados a avaliar a coesão e a centralidade dos atores quanto ao fluxo da informação, úteis para a exata compreensão da estrutura e dinâmica de uma rede.

Quanto aos softwares disponíveis para a construção de uma

rede e o cálculo de indicadores, considera-se que, por fugirem ao escopo deste capítulo, podem ser melhor conhecidos pelos estudiosos e pesquisadores interessados, por meio dos diversos manuais e tutorias disponíveis na literatura técnica da área.

REFERÊNCIAS

- ALEJANDRO, V. A.; NORMAN, A. G. **Manual introdutório à análise de redes sociais**: medidas de centralidade; exemplos práticos com UCINET6.109 e NETDRAW 2.28. 2005. Tradução e adaptação de: AIRES, M. L. L.; LARANJEIRO, J. B.; SILVA, S. C. A. Disponível em: https://www2.unicentro.br/lmqqa/files/2016/05/Manualintrodutorio_ex_ucinet.pdf. Acesso em: 13 jul. 2020.
- ALVES, B. H.; OLIVEIRA, E. F. T. O desenvolvimento do domínio da “Organização do Conhecimento” no contexto da Ciência da Informação a partir da ISKO-Brasil. **Brazilian Journal of Information Studies**, v. 10, n. 1, p. 1-6, 2016.
- BATAGELJ, V.; MRVAR, A. Pajek: **Program for Analysis and Visualization of Large Networks. Reference Manual**. 2011. Disponível em: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/PajekMan.pdf>. Acesso em: 13 jul. 2020.
- BOAVENTURA NETTO, P. O.; JURKIEWICZ, S. **Grafos**: introdução e prática. São Paulo: Editora Blucher, 2009.
- BORGATTI, S. P.; FOSTER, P. C. The Network Paradigm in Organizational Research: A Review and Typology. **Journal of Management**, v. 9, n. 6, p. 991 -1013, 2003.
- COSTA, A.; FERREIRA, C. Redes de coautoria e colaboração institucional nas ciências da saúde: análise evolutiva a partir de star scientists nacionais. **Revista da Faculdade de Letras – Geografia – Universidade do Porto**, v. 2, p. 227-248, 2013.
- DAHLBERG, I. Knowledge organization: its scope and possibilities. **Knowledge Organization**, v.20, n.4, p.211-222, 1993.
- GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Produção e comunicação da informação em CT&I – GT7 da ANCIB: análise bibliométrica no período 2003/2009. **Liinc em revista**, v. 7, n. 1, p. 248 – 263, 2011.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Indicadores de proximidades em análise de cocitação de autores: um estudo comparativo entre coeficiente de Correlação de Pearson e Cosseno de Salton. **Informação & Sociedade**, v. 25, n. 2, p. 105-116, 2015.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Comparing absolute and normalized indicators in scientific collaboration: a study in Environmental Science in Latin America. *In*: INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY INDICATORS (STI 2016), 21., 2016, Valência. **Proceedings SIT 2016: Peripheries, Frontiers and Beyond**. Valência: Universitat Politecnica de Valência, 2016. v. 21. p.1-9.

GUIMARÃES, J. A. C.; OLIVEIRA, E. F. T.; GRACIO, M. C. C. Referentes teóricos em Organização do Conhecimento: uma análise de domínio da revista Knowledge Organization. *In*: INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION CONFERENCE, 12., 2012. **Proceedings [...]**. [S.l.:s.n.], 2012.

MATHEUS, R. F.; SILVA, A. B. O. Fundamentação Básica para Análise de redes Sociais: conceitos, metodologia e modelagem matemática. *In*: POBLACIÓN, D. A.; MUGNAIN, R.; RAMOS, L. M. S. V. C. (org.). **Redes Sociais e Colaborativas em informação científica**. São Paulo: Angellara, 2009. p. 239-287.

MARTELETO, R. M. Análise de Redes Sociais: aplicação nos estudos de transferência da informação. **Ciência da Informação**, v. 30, n. 1, p. 71-81, 2001.

NEWMAN, M.; BARABÁSI, A. -L.; WATTS, D. J. (Eds.). **The Structure and dynamics of networks**. [S.l.]: Princeton University Press, 2006.

OTTE, E.; ROUSSEAU, R. Social network analysis: a powerful strategy, also for the information sciences. **Journal of information Science**. v. 28, n. 6, p. 441-453, 2002.

PINTO, A. L. *et al.* Visualização da informação das redes sociais através de programas de cienciografia. *In*: POBLACIÓN, D. A.; MUGNAIN, R.; RAMOS, L. M. S. V. C. (org.). **Redes sociais e colaborativas em informação científica**. São Paulo: Angellara, 2009. p. 289-312.

SCOTT, J. **Social network analysis**. 2.ed. London: SAGE, 2000.

WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

SILVEIRA, S. M. P. Rede ECOVIDA de Agroecologia: uma inovação

estratégica para o desenvolvimento territorial sustentável na zona costeira catarinense. **Revista Internacional Interdisciplinar INTERthesis**, v. 10, n. 2, p. 181-213, 2013.

VAN ECK, N.; WALTMAN, L. **VOSviewer Manual**. 2016. Disponível em: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.4.pdf. Acesso em: 13 jul. 07 2020.

VANZ, S. A. **As redes de colaboração científica no Brasil (2004-2006)**. 2009. 204 f. Tese (Doutorado em Comunicação e Informação) – Faculdade de Biblioteconomia e Documentação, Universidade Federal do Rio Grande do Sul, Porto Alegre.

WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.