

Proposta de metodologia de mapeamento e avaliação da produção científica da Universidade de São Paulo com foco na estrutura e dinâmica de suas redes de colaboração científica: em busca de novos modelos causais

Dalton Lopes Martins*

Sueli Mara Soares Pinto Ferreira**

Resumo O entendimento das causas e as principais razões que influenciam o modo como os pesquisadores se articulam e constroem suas redes de colaboração científica ainda é uma questão em aberto na pesquisa acadêmica. De fundamental importância para o desenvolvimento de novos indicadores e modos de avaliação da produção científica, o conceito de redes sociais permite operar novos planos de análise, contribuindo com seus aspectos estruturais e dinâmicos ao estudo dos mecanismos e gatilhos causais que levam à constituição dessas redes de colaboração científica. A obtenção de atributos individuais dos pesquisadores, de dados de constituição das redes ao longo do tempo e o modo de desambiguação dos nomes que compõem essas redes de colaboração têm se mostrado os principais desafios de estudos das redes. O objetivo deste artigo é descrever como concebemos uma maneira de estudar as redes de colaboração de uma universidade, com foco específico na Universidade de São Paulo, identificando suas principais estratégias de conectividade e mecanismos causais, além de encontrar as relações entre suas redes e diferentes níveis de produtividade científica de seus participantes. Vale frisar que o artigo apenas descreve as questões da pesquisa e o modo de tratá-las, ficando sua execução para os próximos passos deste trabalho de pesquisa. Para tanto, pretende utilizar como base de análise uma Biblioteca de Produção Científica Institucional em desenvolvimento pelo SiBi/USP, que coleta os artigos publicados por membros da universidade em bases de dados de indexação de revistas nacionais e internacionais, tais como Scielo, Web of Science e BioMed, além da utilização da base de dados institucional para obtenção dos atributos individuais dos pesquisadores participantes dessas redes de colaboração.

Palavras-chave análise de redes sociais, indicadores, cientometria, modelos causais.

Proposal of a methodology for mapping and evaluating the University of São Paulo's scientific production, focusing on collaborative structure and dynamics: in search of new causal models

Abstract The understanding of the causes that influence how researchers articulate and build their scientific collaboration networks is still an open question in academic research. Of fundamental importance for the development of new indicators and methods of evaluation of

* Doutor em Ciências da Informação. FATEC. Endereço: Praça Coronel Fernando Prestes, 30 - Bom Retiro – São Paulo-SP - CEP 01124-060 - Telefone: (11) 3322-2200. E-mail: dmartins@gmail.com

** Doutora em Ciências da Informação. Universidade de São Paulo, Escola da Comunicação e Artes, Departamento de Biblioteconomia e Ciências da Informação. Endereço: Rua da Praça do Relógio, 109 - Bloco K - 4º andar – Cidade Universitária - SP 05508-050. Telefone: (11) 3091- 4195 e 3091-1547. E-mail: sueli.ferreira@gmail.com

scientific literature, the concept of social networking helps operate new levels of analysis, contributing their structural and dynamic aspects to the study of causal mechanisms and triggers that lead to the formation of these networks of scientific collaboration. Obtaining attributes of individual researchers, data on the constitution of networks over time and mode of disambiguation of the names that make up these collaboration networks have been the main challenges in the area of research networks. The purpose of this article is to describe how we designed a way to study a university's collaboration networks, focusing on the University of São Paulo, and identifying their key strategies, connectivity and causal mechanisms, as well as finding links between their networks and different levels of participants' productivity. It should be noted that this article only describes the research questions and how to treat them, leaving their implementation to the next steps of this research. The database used for analysis was the Institutional Scientific Production being developed by Sibi/USP, which collects articles published by members of the university indexed in national and international databases such as Scielo, Web of Science and BioMed, as well as an institutional database to obtain the individual attributes of the researchers participating in these networks.

Keywords social network analysis, indicators, scientometrics, causal model

Introdução

A pesquisa sobre como se estruturam a produção científica de uma universidade e os seus modos de promoção, incentivo e desenvolvimento é um elemento estratégico na concepção de políticas científico-tecnológicas que tenham por objetivo direcionar o modo como a ciência opera, bem como os seus resultados alcançados. Questões sobre como direcionar recursos para determinadas áreas de pesquisa através da construção de editais temáticos, programas de fomento, formação de redes de cooperação e intercâmbio, a distribuição de recursos de pesquisa para instituições, pesquisadores e formas de classificar essas instâncias dependem do modo como os pesquisadores e sua produção são avaliados. Logo, uma das questões centrais é a concepção de metodologias de análise dessa produção científica que forneçam subsídios consistentes de avaliação dos resultados obtidos por uma determinada política, bem como ofertar indicadores de melhorias e mudanças no modo de condução dessas políticas.

A questão se torna ainda mais relevante quando nos propomos analisar a produção científica e as estratégias que viabilizam essa produção de uma universidade como a Universidade de São Paulo (USP). Constituída¹ por 83 unidades e órgãos, mais de 88.000 alunos matriculados e mais de 5.000 docentes em atividade e com 239 programas de pós-graduação, a USP figura como a primeira universidade brasileira no ranking das universidades que possuem mais trabalhos indexados na *Web of Science*, possuindo mais de 53.000 artigos indexados com mais de 416.000 citações (FAPESP, 2012), representando mais que o dobro de artigos da segunda colocada, a Unicamp. Sem dúvida, os números apresentados colocam a USP como uma instituição de destaque no cenário científico brasileiro, sendo ela, portanto, a instituição mais produtiva quando se trata da representação da ciência brasileira no exterior. Poderíamos levantar aqui diversos fatores interessantes, não apenas no contexto da própria Universidade de São Paulo, mas para a ciência brasileira de modo geral, relacionados às condições mais influentes e que mais apoiaram

¹ Dados oriundos do portal USP relativos ao ano de 2011: <http://www5.usp.br/usp-em-numeros/>

a universidade para atingir tais resultados. No entanto, a forma de avaliarmos essas condições, considerando sua produção científica como espaço de análise, é um determinante metodológico que influenciará quais características serão ressaltadas e quais serão consideradas secundárias e de menor influência.

Anunciamos aqui algumas questões de pesquisa em nossa proposta de metodologia de mapeamento em que buscamos identificar potenciais formas metodológicas de resposta. Considerando a produção científica de uma universidade, como mapear suas redes de colaboração entre seus pesquisadores e seus diferentes arranjos organizacionais, tais como grupos de pesquisa, departamentos, centros de pesquisa, fundações, institutos e faculdades? Como são as redes de relação entre os pesquisadores e as unidades organizacionais da Universidade, e o que as condiciona e influencia? Quais são os tipos de redes existentes e de que forma elas se associam com diferentes padrões de produção científica? Que consequências essas redes produzem para a Universidade? Como conseguir o máximo possível de atributos que auxiliem a caracterizar os nós e as conexões estabelecidas entre os atores, sejam eles pessoas ou arranjos organizacionais? Como identificar esses gatilhos causais relacionados ao desempenho na produção científica, bem como às diferentes políticas científicas e tecnológicas que regulam os modos de funcionamento dessa universidade?

Indicadores, redes sociais e seus possíveis modos de olhar

Mapeando a cadeia de indicadores que têm sido classicamente utilizados para avaliação dos esforços e resultados da política científica e tecnológica de forma geral, Mugnaini et al. (2004) apontam quatro níveis de indicadores que se relacionam entre si: indicadores de insumo, indicadores de processo, indicadores de produto e indicadores de impacto.

Os indicadores de insumo estão voltados à avaliação do esforço de fomento, mostrando como e onde os investimentos foram localizados, os grupos de pesquisa existentes e o número de pesquisadores em um determinado contexto. Os indicadores de processo mostram as estratégias de uso desses recursos financeiros e institucionais. Os indicadores de produto referem-se aos resultados mais imediatos da política, apontando o número de artigos publicados e patentes registradas. Por fim, os indicadores de impacto se referem aos efeitos mais a médio prazo e mais perenes dos recursos aplicados pela política de ciência e tecnologia.

Os indicadores acima apresentados permitem uma avaliação dos esforços e resultados gerados por uma política científica, sem deixar evidentes, necessariamente, quais as condições causais que levaram a esses resultados específicos. As principais limitações desses indicadores podem ser apontadas pela falibilidade das premissas teórico-conceituais nas quais são embasados, como descrito por Velho (2001 p. 118):

A principal limitação dos indicadores tradicionais, no entanto, reside na falibilidade das premissas teórico-conceituais que lhes dão sustentação, quais sejam: a linearidade do processo de inovação tecnológica, o status epistemológico especial da ciência e sua neutralidade, a existência de um sistema normativo e de recompensa na ciência que funciona de maneira a garantir que o conhecimento produzido é objetivo, verdadeiro e partilhado entre os praticantes.

De certa forma, esses indicadores anunciam uma ênfase em categorias macrosociológicas, atreladas a essas premissas teórico-conceituais, considerando aquilo que foi colocado no sistema científico, os investimentos feitos, os pesquisadores contratados e os grupos formados, e aquilo que foi produzido em termos de titulações obtidas, artigos e patentes produzidas, bem como possíveis modificações e correlações existentes com outros indicadores sociais. Estamos diante de uma relação que mostra entradas e saídas. Elementos relacionais e processos de articulação social de médio alcance, tais como os padrões de relação nos quais os pesquisadores se inscrevem, suas estratégias de sociabilidade a partir do modo como se mobilizam e são mobilizados, assim como a relação entre esses padrões relacionais e atributos e comportamentos individuais, ficam de fora do alcance daquilo que esses indicadores podem avaliar.

Além disso, estudos recentes apontam importantes limitações em alguns desses indicadores para o tipo de resultado que se propõem apresentar. Campbell (2008) aponta as limitações do fator de impacto, demonstrando como um valor que é atribuído a uma revista científica e é utilizado de forma errônea por muitos pesquisadores e sistemas de avaliação para indicarem resultados de seus próprios artigos, no fundo, pode ser dependente de poucos artigos de alta relevância na revista, não representando um fator que poderia ser generalizado para todos os outros artigos. Lane (2010) sugere que o uso das métricas atuais, tais como índice-h e índice de citação, é limitado, não permitindo capturar e considerar diversas atividades que apoiam e transmitem ideias científicas, tais como orientação, blogs de divulgação, criação de protótipos industriais e mesmo a forma como sua produção científica é acessada, referenciada e utilizada na web. Além disso, Lane aponta para o fato de que um sistema de métricas induz determinado tipo de comportamento dos pesquisadores, levando a modos de funcionamento que podem ser mais ou menos desejáveis conforme nosso entendimento sobre os impactos que um sistema causa no processo de produção científica.

Em busca de entender outros fatores e outros níveis de análise, para além das categorias macrosociológicas que apontamos acima, há diversos estudos recentes que apontam que um dos fatores fundamentais a serem considerados como forma de explicação dos resultados de produtividade científica dos pesquisadores está relacionado ao modo como eles colaboram entre si e terminam por constituir redes de colaboração científica. Eaton *et al.* (1999) sugere uma forte relação entre o número de coautores e a produtividade científica, apontando que quem tem mais coautores tende a publicar mais artigos. Melin (2000), pesquisando a produção científica e os motivos de colaboração de 195 professores universitários, chega a três importantes conclusões: os resultados da pesquisa são difíceis de colocar em um esquema causal, não sendo fácil entender que razões levam a que resultados, apontando que o estudo de como essas relações dependem umas das outras deveria ser pesquisado em futuros trabalhos. No entanto, indica que as principais causas da colaboração científica entre os pesquisadores são pragmáticas e relacionadas a como podem melhorar seus próprios trabalhos em conjunto. Por fim, indica que as colaborações futuras de um pesquisador dependem fortemente de sua rede social de colaboração atual.

Lee e Bozeman (2005) apontam que, apesar da falta de melhores explicações causais diretas, vários elementos ligados à colaboração científica têm sido identificados na literatura como influentes na produtividade científica: divisão do trabalho, habilidades complementares, eficiência no uso do tempo, estímulo intelectual, aprendizado de novas teorias e técnicas entre os colaboradores, companheirismo e parceria na discussão dos resultados da pesquisa, acesso a equipamentos mais modernos, comunicação de novas informações e novas oportunidades de publicação dos resultados do trabalho. Liao (2010), avaliando a diversidade de membros que atuam em colaboração, entendendo por diversidade a quantidade de diferentes parceiros com que

um pesquisador se relaciona em sua rede de relacionamentos e sua intensidade de trocas, conclui que quanto mais um pesquisador for central em sua rede mais citações, mais prêmios e melhor fator de impacto será atribuída a sua produção científica. Abbasi *et al.* (2011) realizando uma pesquisa sobre os efeitos das redes de coautoria no desempenho da produção científica aponta que pesquisadores que estão conectados com mais pesquisadores diferentes possuem melhor desempenho na citação de seus artigos. Além disso, pesquisadores com alta intensidade de conexões com outros pesquisadores (repetidas colaborações) também possuem um melhor desempenho em relação aos que possuem menos conexões. Belgacem e Lamari (2012) encontram importantes evidências da relação entre a qualidade da produção científica e o financiamento à colaboração entre pesquisadores, indicando que quatro níveis de determinação causal deveriam ser levados em consideração nos estudos de colaboração científica, sendo eles os atributos individuais dos pesquisadores, organizacionais, financiamento e atributos disciplinares da área onde se encontram.

Os resultados das pesquisas apresentadas acima denotam quatro importantes características: a limitação das métricas atualmente em uso, o efeito direto entre colaboração e produtividade científica, como um pesquisador se posiciona e se articula em sua rede social e a necessidade de produção de modelos causais que expliquem melhor como esses fatores se relacionam ainda é uma questão em aberto. Logo, dessas características, saltam a colaboração e a rede social como conceitos que poderiam ser melhor trabalhados na busca por melhores indicadores e modelos causais. Um ponto que vale aqui ressaltar e que tomaremos como ponto de partida para o entendimento do que propomos como objetivos de nossa pesquisa é o conceito de redes sociais como um conceito operatório-metodológico (MARTELETO, 2007) que sirva tanto para analisar os modos como se desenvolvem a produção científica quanto as possíveis estratégias de conectividade e eventos causais que nos ajudem a entender melhor os fatores que influenciam essa mesma produção.

Um desdobramento imediato desse uso do conceito de redes sociais é o deslocamento analítico que ele provoca em relação aos indicadores tradicionais de avaliação das políticas científicas e tecnológicas. Esse deslocamento opera uma inclusão dos pesquisadores e suas características ao estudo de sua própria produção. Além disso, esse deslocamento ocorre na maneira de tratar o objeto sociológico de estudo, entendendo que há um reposicionamento em nível intermediário pelo uso do conceito de redes, que não é nem o nível microsociológico do indivíduo e nem o nível macrosociológico da sociedade, sendo possível qualificá-lo como um nível mesosociológico, onde operam as formas sociais resultantes das interações entre os indivíduos (MERCKLÉ, 2011, pg. 14).

Ao buscar entender os processos da produção da ciência e da tecnologia, Bruno Latour (1997) entende que essa produção exige a formação de redes específicas que constituem determinadas características de instituições, profissões e projetos que tenham por objetivo a produção de fatos e artefatos argumentativos e analíticos que se insiram nos debates conceituais então vigentes a respeito de um determinado assunto.

Se a tecnociência pode ser descrita como algo tão poderoso apesar de tão pequeno, tão concentrado e tão diluído, significa que tem as características de uma rede. A palavra rede indica que os recursos estão concentrados em poucos locais – nas laçadas e nos nós – interligados – fios e malhas. Essas conexões transformam os recursos esparsos numa teia que parece se estender por toda parte (LATOUR, 1997, pg. 273).

Ao abrir esse plano de análise da tecnociência, Latour (1997, p. 331) sugere os principais elementos de mapeamento das cadeias de associações que levam a constituição dessas redes: que pontos estão interligados, que dimensões e que força têm essas ligações, quais são os mais legítimos porta-vozes, como todos esses elementos são modificados durante a controvérsia no embate de conversações em torno da produção científica e, por fim, como são feitas as atribuições de causas e efeitos. Uma questão implícita desse plano de análise a partir do uso do conceito de redes sociais é o fato de essa busca por atribuições de causas e efeitos no estudo do mapeamento dessas redes, bem como a identificação de possíveis modelos causais, estar pautada na crença de que as relações sociais não ocorrem de modo simplesmente aleatório, mas ocorrem a partir de racionalidades do comportamento relacional dos atores sociais, sem os quais não poderíamos explicar as estruturas e dinâmicas das redes que constituem (MERCKLÉ, 2011, p. 99).

O uso das redes sociais como conceito operatório-metodológico nos permite sugerir aspectos que fundamentam essas racionalidades a partir da identificação de regularidades em sua estrutura e nos processos dinâmicos que levam a sua constituição. As regularidades de estrutura são determinadas por eventos característicos de conexão entre os nós. As conexões ocorrem de uma determinada forma majoritariamente em detrimento de outras possibilidades pouco exploradas em uma rede em análise. O estudo dessas regularidades nos permite inferir causas, estratégias típicas de conectividade entre os nós, que nos auxiliam a explicar como opera uma rede (MARTINS e FERREIRA, 2012). Já as regularidades dinâmicas estão atreladas aos processos de mudança nas características estruturais da rede. Marques (2010) sugere que essas regularidades podem ser analisadas a partir de alguns princípios que, a rigor, poderiam nos levar a identificarmos gatilhos causais que são responsáveis pela produção de diferenciação e variabilidade entre as redes.

As mudanças nas redes podem seguir padrões mais localizados e conjunturais ou mais estruturais, associados à trajetória e aos tipos de ambientes relacionais a que os indivíduos têm acesso. As mudanças tanto podem criar (ou destruir) oportunidades para relações (o que podemos denominar oportunidades relacionais) como mudar o ambiente em que elas ocorrem, contribuindo para a alteração dos conteúdos dos vínculos. Como consequência, não apenas os tamanhos e as estruturas das redes podem ser alterados, como também são transformados a sociabilidade e os tipos de vínculos (MARQUES, 2010, p. 175).

Vale aqui mencionar uma outra contribuição importante do uso do conceito de redes sociais para a identificação de possíveis causas para as regularidades estruturais e dinâmicas identificadas pelo estudo da análise de redes. O conceito clássico de categorização dos indivíduos a partir de determinados grupos dos quais fazem parte entende que é necessário primeiro categorizar esses indivíduos para depois inferir possíveis comportamentos a partir das categorias em que se encontram. Já a análise de redes sociais propõe uma inversão nesse modo de categorização, entendendo que é a partir da análise da forma específica de relacionamento social que é possível propor categorias que sirvam para apoiar a explicação de como se deram esses relacionamentos, além de facilitarem o entendimento das categorias macrosociais, tais como demográficas e institucionais, pelas quais os atores já são classificados (MERCKLÉ, 2011, p. 15). A análise de redes sociais, entendida desse modo, ganha um papel de relevância ao considerar a regularidade de eventos relacionais como ponto de partida para a proposição de possíveis causas de

fenômenos sociais de interesse relacionados a esses eventos.

É importante ressaltarmos aqui o que entendemos por causalidade entre redes e os atributos sociais dos seus atores. Entendemos que tanto os padrões de relações nos quais os atores estão inseridos, quanto as suas condições de produção científica são construídos ao longo de sua trajetória profissional no campo da pesquisa científica. Marques (2010, p. 19), observando essas relações de causalidade entre os padrões de redes, a sociabilidade e os atributos sociais, afirma que nenhum desses elementos de forma isolada é considerado causa do outro, embora eles se influenciem dinamicamente todo o tempo. Esse entendimento leva a considerarmos que a causalidade entre os padrões de produção científica, os padrões de rede e os atributos dos atores que estudaremos ao longo de nossa pesquisa possuem causalidade múltipla, ou seja: somente colocando-os em análise de forma conjunta teremos condições de inferir gatilhos causais que nos permitirão explicar de outros modos como surgem e se desenvolvem as redes de produção científica da universidade.

Estudos recentes na área da bibliometria e cientometria fornecem importantes indícios do uso de técnicas e indicadores oriundos da análise de redes sociais com foco na colaboração científica como meios de utilizar esse conceito operatório-metodológico das redes para explicar possíveis causas que determinam o desempenho de pesquisadores e instituições em seus modos de produção científica. Parte importante do objetivo desses estudos é identificar novas métricas para avaliação da produção científica que forneçam melhores mecanismos explicativos do que os atualmente em uso.

Bollen *et al.* (2009) realizaram uma análise multivariada de 39 indicadores de avaliação de produção científica, estando entre eles indicadores como fator de impacto, índice-h e indicadores de avaliação de revistas por sua posição de centralidade na rede social de citação em que se encontram. Os resultados mostram que indicadores baseados no cálculo de centralidade de intermediação apresentam melhores avaliações do prestígio de um artigo do que outros, como fator de impacto, por exemplo. Abbasi e Altmann (2010) desenvolveram um modelo baseado na teoria de redes sociais que demonstra como o desempenho na produção científica de pesquisadores é correlacionado com sua rede de coautoria, considerando como medida de desempenho o índice-h. A pesquisa aponta que dois indicadores de análise de redes sociais, grau de centralidade ponderado e eficiência, são positivamente correlacionados com o desempenho acadêmico. Abbasi et al. (2011) analisam a correlação positiva de quatro indicadores de redes sociais com o desempenho acadêmico baseado no índice-g, uma adaptação do índice-h. Os indicadores encontrados são o grau de centralidade, grau de centralidade de intermediação, eficiência e a força de conexão média entre os nós.

Uma condição limitante, com ainda pouca pesquisa na área, apontada por diversos estudos utilizando indicadores oriundos da análise de redes sociais como modo de proposição de causas que orientam a produção científica é a dificuldade de obtenção de atributos dos nós que constituem essas redes (MELIN, 2000; LEE e BOZEMAN, 2005; ALTMANN et al., 2009; ABBASI; ALTMANN, 2010; BELGACEM e LAMARI, 2012). Abbasi et al. (2011) sugerem que o acesso a informações demográficas dos pesquisadores, tais como idade, gênero e nacionalidade, e informações sobre o contexto social no qual se inserem, tais como se são alunos de graduação, professores associados, titulares, entre outros, poderiam contribuir de forma significativa para enriquecer a capacidade explicativa dos modelos baseados na análise de redes sociais.

De forma a produzir uma estrutura de apoio a pesquisa que tenha por objetivo responder às questões acima enunciadas na seção de introdução, entendemos ser necessário responder a quatro questões metodológicas: quais fontes de dados de produção científica e de atributos dos

pesquisadores serão utilizadas, que universo da produção científica elas abrangem, que métodos de análise de redes sociais utilizaremos e como iremos analisar nossos resultados visando responder às questões levantadas anteriormente. A seguir, detalhamos uma proposta de como essas questões poderiam ser abordadas.

Fontes e abrangência dos dados

Há duas questões que devemos tratar em relação às fontes de dados utilizadas. Em primeiro lugar, temos de garantir que tenhamos acesso ao maior número possível de atributos dos pesquisadores. Em segundo lugar, temos de garantir que tenhamos a maior amostra possível da produção científica desses pesquisadores, de modo a estarmos diante de um espaço de análise que represente, se não a totalidade, a maior cobertura dos eventos relacionais que iremos avaliar.

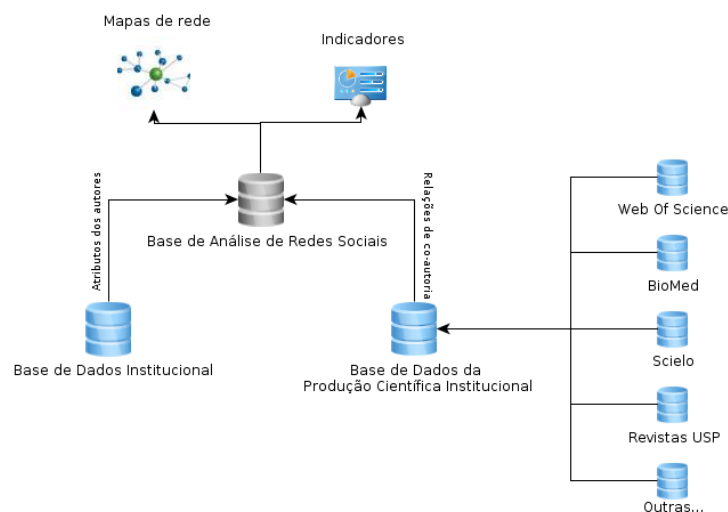
Os atributos dos pesquisadores podem ser obtidos pela base de dados institucional da universidade, criada e mantida pelas suas próprias necessidades administrativas, onde estão armazenados dados como o momento de contratação do pesquisador, em que área organizacional ele atua e por quais já passou ao longo de sua história na instituição, os diferentes cargos e posições ocupadas, sua idade, gênero, entre outros dados que podem se mostrar importantes ao longo da realização deste trabalho.

A produção científica dos pesquisadores já envolve outra ordem de complexidade. Dado que essa produção ocorre de forma distribuída, ou seja, os pesquisadores podem submeter as publicações que produzem para diferentes revistas nacionais e internacionais, não há propriamente um local único de onde poderíamos coletar toda a produção da universidade. O desafio, portanto, é garantir que tenhamos acesso à maior amostra possível dessa produção. Para isso, assim como para outros projetos de seu interesse, o Departamento Técnico do Sistema de Bibliotecas da Universidade de São Paulo (SiBi/USP) tem desenvolvido a Base de Dados da Produção Científica Institucional (BDPI).

A BDPI é constituída pela coleta de metadados e artigos completos da produção científica de pesquisadores da Universidade de São Paulo já indexadas em outras bases de dados de renome, tanto em nível nacional como internacional. Desse modo, a BDPI coleta dados da *Web of Science*, Scielo, BioMed e do próprio portal de revistas da universidade. Entendendo serem essas as bases de dados atualmente mais representativas em termos da qualidade das revistas científicas que ali são indexadas, a BDPI se torna um local privilegiado para coletarmos uma importante e representativa amostra da produção científica da Universidade de São Paulo.

Com base na descrição acima, propomos a arquitetura da informação na Figura 1 como fonte de dados a serem utilizados na realização desta pesquisa.

Figura 1: Arquitetura da informação de apoio a análise de redes sociais de uma universidade.



A agregação dessas bases de dados institucional e de produção científica, no que chamamos aqui de Base de Análise de Redes Sociais, permite selecionarmos apenas os atributos dos pesquisadores que interessarem para o desenvolvimento de nossa pesquisa e relacionarmos esses atributos com as relações de coautoria dos pesquisadores ao longo do desenvolvimento de sua produção científica.

Método de análise de redes sociais

Os indicadores de redes sociais que serão utilizados na realização deste projeto serão os mesmos propostos e utilizados em nossa tese de doutoramento (MARTINS, 2012). O objetivo do uso desses indicadores é produzir uma visão das redes em análise em três níveis: a rede como um todo, os subgrupos ou possíveis comunidades dessa rede e características individuais de cada um de seus atores. Desse modo, temos a nossa disposição diferentes planos de análise, permitindo observarmos e compormos esses diferentes movimentos do que forma e constitui essas redes.

Além disso, nosso objetivo é retratar tanto a estrutura quanto a dinâmica das redes. Para análise estrutural, retomamos aqui a perspectiva de Wasserman e Faust (2009, p. 9) a respeito da análise de redes sociais:

Dada uma coleção de atores, a análise de rede social pode ser utilizada para estudar as variáveis estruturais medidas nos atores do conjunto. A estrutura relacional de um grupo ou um amplo sistema social consiste de padrões de relações entre os conjuntos de atores. O conceito de rede coloca ênfase no fato de que cada indivíduo está conectado a outros indivíduos, cada qual conectado com outros poucos, alguns ou muitos outros. A expressão “rede social” se refere ao conjunto de atores e suas conexões. O analista deve procurar modelar essas

relações para demonstrar a estrutura do grupo. Então, podemos estudar o impacto da estrutura no funcionamento do grupo e/ou a influência da estrutura nos indivíduos dentro do grupo.

Logo, a análise estrutural nos permite identificar as diferentes formas e as regularidades nas formas das redes que podem surgir ao longo do trabalho. Apresentamos no quadro 1, a seguir, os indicadores estruturais que utilizaremos, uma breve síntese de seu significado e a qual nível de análise estão relacionados.

Quadro 1: Indicadores estruturais de análise de redes. Fonte: Martins (2012).

Níveis	Indicadores	Significado
Rede	Densidade	Taxa de conectividade da rede
	Diâmetro	Maior distância entre dois atores numa rede
	Afiliação	Configuração das redes modo-2, caracterizando dois tipos de vértices na rede: atores e organizações
Subgrupos	Centro e periferia	Distinção entre atores de maior e menor grau de centralidade
	Componentes	Subgrupos independentes na rede
	Componentes fortes	Subgrupo fortemente conectado
	Componentes fracos	Subgrupo fracamente conectado
	Cliques	subgrupo completo que contém três ou mais vértices
	m-slices	subgrupo contendo um número mínimo de linhas m e os vértices que são relacionados por essas linhas
	k-core	subgrupo onde cada vértice estabelece relação com um número mínimo k de outros vértices.
Atores	Distância	Número de conexões existentes entre dois atores numa rede
	Grau de centralidade da rede	número de linhas incidentes em um vértice do grafo
	Grau de centralidade por interposição	habilidade de um indivíduo de se conectar aos círculos importantes da rede
	Grau de centralidade por vizinhança	representa a habilidade de um indivíduo monitorar o fluxo de informação e enxergar o que está acontecendo na rede
	Equivalência estrutural	Medida que avalia o quanto dois atores possuem um padrão de conexão semelhante.
	Brokers	Distinção de papéis de intermediação que um ator pode desempenhar, dependendo do contexto da rede.

Já a análise dinâmica recoloca a questão das redes sociais entendendo que os padrões estruturais mudam com o tempo, a partir de determinados tipos de eventos que estimulam ou reduzem o número e a forma como as conexões são estabelecidas, assim como o modo em que os novos nós chegam e os antigos saem da rede. Newman, Barabasi e Watts (2006, p. 7) definem a análise dinâmica a partir dos seguintes princípios:

As conexões que as pessoas fazem afetam a forma da rede, e a forma da rede afeta as conexões que as pessoas fazem. A estrutura da rede social se desenvolve de forma historicamente dependente, de tal maneira que o papel dos participantes e os padrões de comportamento que seguem não podem ser ignorados.

Logo, a análise dinâmica está preocupada em identificar quais são os eventos de conectividade que são responsáveis pelas alterações na estrutura das redes, fazendo com que as mesmas se desenvolvam segundo um determinado padrão de comportamento promovido por seus participantes. Apresentamos no quadro 2, a seguir, os indicadores dinâmicos que utilizaremos, uma breve síntese de seu significado e a qual nível de análise estão relacionados.

Quadro 2: Indicadores dinâmicos de análise de redes. Fonte: Martins (2012).

Níveis	Indicadores	Significado
Rede	Distribuição de probabilidade	Como se distribuem graus de conectividade dos nós
	Evolução no grau médio de centralidade	Como evolui a conectividade média dos nós
	Coefficiente de clusterização	Avalia o grau de influência e transitividade na rede
	Evolução na distância média entre os nós	Como evolui a distância entre os nós
	Evolução do padrão estrutural	Probabilidade de como a estrutura da rede evolui

Como forma de entendermos os impactos que os padrões estruturais e dinâmicos de redes causam na produção científica dos pesquisadores da universidade, utilizaremos três indicadores de produtividade científica que serão considerados como as variáveis independentes de nossa análise. A **quantidade de artigos produzidos** nas bases de dados que constituem a BDPI, a **quantidade de citações recebidas** e o **índice-h** de um autor em determinados extratos de tempo a serem definidos ao longo do desenvolvimento da pesquisa.

A análise dos resultados será dividida em três momentos. **Análise da variabilidade estrutural e dinâmica das redes sociais**, em que iremos identificar as principais diferenças estruturais e dinâmicas existentes entre as redes analisadas. Por diferenças estruturais e dinâmicas entendemos seus padrões de formação, tais como, a quantidade de nós e links participantes, as formas de conectividade características (majoritariamente entre 2 nós, entre 3, em formato estrela, etc.), os padrões e as tendências de transitividade entre os nós, o coeficiente de potência das curvas de distribuição de grau de centralidade, o modo de entrada dos autores nas redes, o intervalo de tempo de colaboração e movimentação social das redes, entre outros. Temos como hipótese aqui que encontraremos essa variabilidade estrutural e dinâmica ao analisarmos redes de coautoria que pertencem a diferentes áreas do conhecimento presentes na universidade, logo pautadas por diferentes modos de produção científica e formas de entender e articular sua produção em coautoria. A partir dos diferentes padrões encontrados, proporemos uma tipologia dessa variabilidade de redes, identificando e categorizando as formas estruturais e dinâmicas mais comuns. Vale lembrar que as redes serão geradas e analisadas considerando os diversos atributos individuais de seus autores, permitindo-nos identificar redes por gênero, cargos ocupados, arranjos organizacionais, idade, entre outros.

Análise das similaridades entre os pesquisadores, em que analisaremos a relação entre os indicadores estruturais e dinâmicos de rede utilizados, a produtividade dos pesquisadores a partir de nossas variáveis independentes acima mencionadas e seus atributos individuais. O objetivo é realizar uma análise de similaridade multivariada de componentes principais (HUSSON *et al.*, 2011, p. 2) buscando identificar uma possível tipologia dos indivíduos: quais os mais similares e dissimilares, quais grupos de indivíduos mais homogêneos em suas similaridades, quais as dimensões de variabilidade que opõe extremos e indivíduos intermediários nas relações analisadas. **Análise dos gatilhos causais** em que, tendo já mapeado a variabilidade das redes

e encontrado as relações entre produtividade científica, atributos individuais e os indicadores estruturais e dinâmicos dessas redes, proporemos alguns dos possíveis gatilhos causais que constituem os elementos de ativação e respostas das redes. É importante ressaltar aqui que, conforme os resultados encontrados, selecionaremos uma quantidade de atores para a realização de entrevistas, realizando um levantamento de informações qualitativas que visam complementar indícios e ilustrar situações mais específicas, sobretudo de ordem motivacional, de forma a colaborar na argumentação em torno dos gatilhos causais. Estimamos, em um primeiro momento, que de 1% a 5% (até 50 pessoas), considerando uma amostra dos nós centrais da rede, devem fazer parte dessa fase da pesquisa.

Dessa forma, entendemos que o enunciado do problema de que pretendemos tratar, bem como os aspectos metodológicos de fundamentação de nossa pesquisa estão aqui estabelecidos, gerando as condições necessárias à realização de nossos objetivos.

Proposta de produtos da análise de dados

Os resultados de tal proposta de análise de dados podem ser organizados em cinco produtos a serem desenvolvidos e sistematizados ao longo do trabalho. **Metodologia de análise de redes sociais de produção científica baseada na coautoria de artigos científicos de uma universidade**, em que conforme o desenvolvimento da pesquisa, sistematizaremos as etapas de trabalho, envolvendo desde a coleta, tratamento, síntese e análise e discussão dos dados em suas diversas etapas e sistemas de informação. Entendemos que essas etapas podem ser replicadas, de maneira geral, em outras universidades que tenham por intenção a realização do estudo aqui proposto.

Mapeamento das redes sociais de produção científica baseada na coautoria de artigos científicos da Universidade de São Paulo, sendo um dos produtos imediatos deste trabalho os mapas de redes sociais da Universidade de São Paulo. Esses mapas retratarão as estruturas de suas redes, bem como as dinâmicas de seu desenvolvimento, considerando as mudanças dos mapas ao longo do tempo. Os mapas serão sistematizados como um resultado do trabalho, podendo servir de referências para futuras análises comparativas de outras instituições e mesmo da própria Universidade em extratos futuros de tempo.

Tipologia de variabilidade dinâmica e estrutural de redes sociais de produção científica, em que a análise da variabilidade dinâmica e estrutural será sistematizada de forma a construir uma tipologia dessa variabilidade, com o objetivo de identificar as principais estruturas características que constituem e formam essas redes, bem como identificar os principais eventos de conectividade que levam à constituição dessas formas. O objetivo da produção dessa tipologia é criar um esforço inicial de catalogação desses processos, visando inclusive alimentar a discussão acadêmica na área, bem como criar condições comparativas e analíticas para outros estudos que vierem a complementar e propor outros olhares possíveis para esses mesmos processos.

Sistematização dos gatilhos causais das redes sociais, sabendo que a identificação de gatilhos causais é um dos temas centrais desta pesquisa. Temos aqui por objetivo propor mecanismos explicativos que nos auxiliem a entender quais são os principais fatores que devem ser levados em consideração no modo como os pesquisadores se articulam e se mobilizam socialmente em torno de seu processo de produção científica. De fato, este é um campo ainda emergente de análise e pesquisa acadêmica, como nossa breve revisão bibliográfica buscou demonstrar.

Acreditamos que um resultado importante deste projeto é a sistematização desses gatilhos e seu modo de apresentação, entendendo que estamos considerando aqui o relacionamento entre estrutura e dinâmica de redes sociais com os atributos individuais e resultados da produtividade científica dos pesquisadores.

Banco de dados de métricas institucionais, produtividade e atributos individuais, onde, por fim, um dos resultados deste projeto será a construção de um banco de dados que poderá ser utilizado por outros estudos futuros de interesse da própria Universidade de São Paulo e do Sistema Integrado de Bibliotecas (SiBi), facilitando que outros pesquisadores proponham outros modos de análise, considerando outros planos da própria bibliometria, cientometria, infometria e webometria de forma geral.

Conclusão

A pesquisa na área de estudos métricos em ciência e tecnologia vem se apropriando, sobretudo no Brasil, nos últimos anos, das metodologias de análise e do conceito de redes sociais como um novo conjunto de ferramentas intelectuais que proporcionam novos olhares e modos de analisar em relação aos problemas informacionais. Ainda que de maneira introdutória, as redes sociais vêm se mostrando uma forma flexível e transversal de modelagem de diferentes modos de entender as questões informacionais, permitindo avançar na compreensão das questões sociais que envolvem direta e indiretamente a Ciência da Informação.

O presente artigo propõe algumas novas técnicas de estudo e formas de tratamento de indicadores da ciência e tecnologia em relação com as redes sociais de pesquisadores e instituições que são analisadas por esses mesmos indicadores. Visando abrir novos caminhos de experimentação e avançar nas técnicas de uso da análise de redes sociais, o artigo delinea seu método de pesquisa, propondo uma arquitetura de informação, indicadores e meios de identificar possíveis relações entre esses indicadores, bem como produtos possíveis de serem obtidos ao seguir os passos de análise aqui propostos.

Estabelecidos esses passos conceituais e a teoria que nos permite desenhar tais modos de posicionar nossa pesquisa, entendemos que nossos próximos passos consistem na aplicação do que aqui vem exposto, buscando demonstrar seus primeiros resultados, dificuldades encontradas, bem como adaptações possíveis e novos recursos a serem incorporados na análise de redes sociais, visando contribuir aos avanços dos estudos métricos em ciência e tecnologia na Ciência da Informação brasileira, assim como com outras universidades que tiverem interesse em avaliar sua produção científica a partir dessas referências metodológicas.

Artigo recebido em 15/02/2013 e aprovado em 22/03/2013.

Referências

ABBASI, Alireza, ALTMANN, Jorn. On the correlation between research performance and social network analysis measures applied to research collaboration networks. In: HICSS '11 HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 44., 2011, Hawaii. *Proceedings...* [S.l.: s.n.], 2011. p. 1-10.

_____; HOSSAIN, Liaquat. Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, v. 5, n. 4, p. 594-607, Oct. 2011.

_____; HWANG, Junseok. Evaluating the productivity of researchers and their communities: the rp-index and the cp-index. *International Journal of Computer Science and Applications*, v. 6, n. 2, p. 104-118, 2009.

BELGACEM, Ines; LAMARI, Moktar. Scientists' collaboration in the social sciences field: investigating the determinants of scholarly collaboration in the canadian context 2001-2008. *International Journal of Information Studies*, v. 4, n. 2, p. 83-94, Apr. 2012.

BOLLEN, Johan et al. A principal component analysis of 39 scientific impact measures. *PLOS One*, v. 4, n. 6, June 2009.

CAMPBELL, Philip. Escape from the impact factor. *Ethics in Science and Environmental Politics*, v. 8, p. 5-7, 2008.

EATON, J. P. et al. Structural analysis of co-author relationships and author productivity in selected outlets for consumer behavior research. *Journal of Consumer Psychology*, v. 8, n. 1, p. 39- 59, 1999.

FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO - FAPESP. Trabalhos indexados das universidades brasileiras com maior volume de produção. *Revista da FAPESP*, n. 194, abr. 2012.

HUSSON, François, LÊ, Sébastien. *Exploratory multivariate analysis by example using r*. [S.l.]: CRC Press, 2010. 240 p.

LANE, Jane. Let's make science metrics more scientific. *Nature*, n. 464, p. 488-489, 25 Mar. 2010.

LATOUR, Bruno. *Ciência em ação: como seguir cientistas e engenheiros sociedade afora*. Campinas: Ed. Unesp, 1998. 438 p.

LEE, Sooho; BOZEMAN, Barry. The impact of research collaboration on scientific productivity. *Social Studies of Science*, v. 35, n. 673, 2005.

LIAO, Chien Hsiang. How to improve research quality?: examining the impacts of collaboration intensity and member diversity in collaboration networks. *Scientometrics*, v. 86, n. 3, p. 747-761, Mar. 2011.

MARQUES, Eduardo. *Redes sociais, segregação e pobreza*. Campinas: Editora Unesp, 2010. 216 p.

MARTELETO, Regina Maria. Informação, rede e redes sociais- fundamentos e transversalidades. *Informação & Informação*, v. 12, n. esp., 2007.

MARTINS, Dalton Lopes. *Análise de redes sociais de colaboração científica no ambiente de uma federação de bibliotecas digitais*. Tese (Doutorado em Ciência da Informação)- Universidade de São Paulo, 2012.

_____; FERREIRA, Sueli Mara P. Estruturas de hierarquia e eventos geradores em redes de colaboração científica: estudo exploratório entre relações de coautoria e em bancas de defesas de teses e dissertações. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 3., 2012, Rio Grande do Sul. *Anais...* Cd-rom.

MELIN, Göran. Pragmatism and self-organization research collaboration on the individual level. *Research Policy*, n. 29, p. 31-40, 2000.

MERCKLÉ, Pierre. *Sociologie des réseaux sociaux*. [S.l.: s.n.], 2011.

MUGNAINI, Rogério; JANNUZZI, Paulo de Martino; QUONIAN, Luc. Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. *Ciência da Informação*, v. 33, n. 2, p. 123-131, maio/ago. 2004.

NEWMAN, Mark E. J.; BARABASI, Alfred L.; WATTS, Duncan. *The structure and dynamics of networks*. Princeton: Princeton University Press, 2006.

_____. *Networks: an introduction*. Oxford. 2010.

VELHO, Léa Maria Strini. Estratégias para um sistema de indicadores de C&T no Brasil. *Parcerias Estratégicas*, n. 13, p. 109-121, dez. 2001.

WASSERMAN, Stanley; FAUST, Katherine. *Social network analysis: methods and applications*. Cambridge: Cambridge University Press, 1994. 857 p.