

REDES DE CO-AUTORIA DOS PROFESSORES DA CIÊNCIA DA INFORMAÇÃO: UM RETRATO DA COLABORAÇÃO CIENTÍFICA DESSA DISCIPLINA NO BRASIL

PROFESSORS OF INFORMATION SCIENCE CO-AUTHORSHIP NETWORK: A PICTURE OF SCIENTIFIC COLLABORATION IN THIS FIELD IN BRAZIL

Antonio Braz de Oliveira e Silva

Analista do IBGE, doutorando em Ciência da Informação pela Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG) e membro do Netic - Núcleo de Estudos em Tecnologias para Informação e Conhecimento (www.netic.com.br). E-mail: abraz@netic.com.br.

Fernando Silva Parreiras

Doutorando em Ciência da Computação pela *Universität Koblenz*, Alemanha, mestre em Ciência da Informação (ECI/UFMG) e membro do Netic (www.netic.com.br). E-mail: fparreiras@netic.com.br.

Renato Fabiano Matheus

Analista do Banco Central, doutorando em Ciência da Informação (ECI/UFMG), mestre em Ciência da Informação (ECI/UFMG) e membro do Netic (www.netic.com.br). E-mail: renato.fabiano@bcb.gov.br; www.rfmatheus.com.br.

Wladimir Cardoso Brandão

Mestre em Ciência da Informação (ECI/UFMG) e membro do Netic (www.netic.com.br). E-mail: wbrandao@netic.com.br

Resumo: Este artigo apresenta a rede de colaboração científica entre professores dos programas de Pós-Graduação na área de Ciência da Informação (CI) em funcionamento no Brasil. Tal colaboração é identificada a partir da co-autoria de artigos em revistas e eventos científicos. A fonte de informação adotada foi a base de dados do projeto 'RedeCI' (www.redeci.netic.com.br) sendo a metodologia empregada a análise de rede social (ARS). Além da rede individual dos professores, foram também consideradas, as suas universidades de origem e suas linhas de pesquisa, a fim de identificar a colaboração entre instituições. Os resultados mostraram a existência de uma rede pouco conectada, com muitos componentes independentes. O artigo representa mais um esforço do projeto 'RedeCI' em direção a uma maior compreensão da área de CI no Brasil, por meio da identificação das redes de colaboração e dos colégios invisíveis nela existentes.

Palavras-chave: redes de co-autoria, análise de redes sociais, ciência da informação, bibliometria, cienciométrica.

Abstract: This paper presents the scientific collaboration network among professors of the Pos-graduation Centers in Information Science (IS) in Brazil. Such collaboration is identified from the co-authorship in papers from journals and scientific events. The main information source is the database of the 'RedeCI' project (in English: "Brazilian Information Science Network", www.redeci.netic.com.br). Such database comprises papers published in IS journals or presented in the IS Congress in Brazil, with some additional attributes for the professors (Pos graduation Center and research area). The structure of coauthorship networks was investigated at the level of professors by using all the papers existing in the 'RedeCI' project combining statistical analysis with social network analysis (SNA). The attributes were used to complement the informations obtained by the ARS and identifying the institutional linkages behind the coauthorship network. The results show a poorly connected network, with many partitions. The authors stress the benefits brought out by the use of SNA and the relevance of database like 'RedeCI' projects.

Keywords: *co-authorship network, social network analysis, information science, bibliometrics, scientometrics.*

1 INTRODUÇÃO

A preocupação em compreender e explicar as mudanças e o desenvolvimento no pensamento científico é o elemento central nos estudos das instituições científicas (CRANE, 1972), sendo a produção do conhecimento científico uma questão essencial para a Ciência da Informação (CI) desde a sua origem (BUSH, 1945).

A partir de tais considerações e de acordo com Crane (1972), a questão sociológica que se coloca é saber “se as comunidades científicas e as variações nos padrões de comunicação entre os cientistas afetam, de fato, o desenvolvimento do conhecimento” (CRANE, 1972, p.11). Em caso afirmativo, seria adequado, segundo a autora, examinar as diferentes culturas dessas comunidades, para se definir com a maior clareza possível as interações entre os componentes cognitivos e sociais da ciência. Nesse sentido, a primeira tarefa consiste “na definição de meios que permitam identificar essas comunidades” (CRANE, 1972, p.12), uma vez que os cientistas de cada área apresentam diferentes tipos de relações sociais entre si, tornando necessário o uso e a combinação de vários indicadores na análise de sua estrutura social. Sobre os possíveis indicadores, Crane (1972, p. 41) ainda afirma que “[...] os mais importantes indicadores da organização social em uma área de pesquisa [são]: a discussão informal das pesquisas, a publicação em colaboração, o relacionamento entre os professores e a influência dos colegas na definição dos problemas e técnicas de pesquisa”. Além disso, ainda se deve levar em consideração que as áreas de pesquisa não são “inteiramente face a face [...] e seus membros são dispersos geograficamente” (CRANE, 1972, p.42). Como se pode notar, as diferentes formas de compartilhamento de conhecimentos entre os pesquisadores envolvem a troca de conhecimento tácito pelo processo de socialização, seja ele presencial ou não. Dessa forma, podem existir colégios invisíveis, nos quais os laços entre as pessoas não obedecem às relações formais de hierarquia, vínculo funcional, linhas de financiamento, ou, mesmo, nacionalidades.

Considerando as dificuldades e a complexidade do cenário delineado, alguns dos indicadores mencionados são mais fáceis de identificar e analisar do que outros. Tais indicadores podem ser usados para identificar o contorno geral da comunidade, ou, ainda, para ajudar a explicitar os colégios de colaboração científica existentes. Um indicador com tal característica, muito usado na identificação de comunidades científicas, é o estudo da co-autoria em trabalhos científicos. Trata-se de uma área de aplicação de técnicas bibliométricas - que em muitos casos podem ser combinadas com outras técnicas - dentre as quais se destaca a análise de redes sociais (ARS).

A ARS permite que se analisem redes de contatos entre pessoas (WASSERMAN; FAUST, 1999), inclusive na forma de co-autoria. As características da rede de co-autoria fornecem informações para a compreensão da organização social da área e sua produção científica, na forma de trabalhos em colaboração. Conforme o pressuposto apresentado por Silva *et al.* (2005), tendo como referência apenas a rede dos professores do Programa de Pós-Graduação em Ciência da Informação (PPGCI / UFMG), uma rede densa de artigos em co-autoria indicaria a execução de um programa de pesquisa vigoroso, com os resultados divulgados representando a colaboração dos professores.

O objetivo deste artigo é apresentar um retrato da Ciência da Informação, tendo por base a rede de colaboração na forma de co-autoria de artigos em revistas e eventos científicos da área, entre professores dos programas de Pós-Graduação em funcionamento no Brasil. O artigo segue outros trabalhos que também buscam colaborar para o entendimento da produção científica na área de CI no Brasil e promover a disseminação dos métodos de ARS (SILVA *et al.*, 2005; MATHEUS, SILVA, 2006; PARREIRAS *et al.*, 2006).

Este trabalho é organizado da seguinte forma: em seguida a essa introdução, na Seção 2 é apresentada uma visão geral da análise de redes sociais (ARS), bem como os conceitos básicos e medidas usadas no decorrer do artigo; na Seção 3 são apresentadas as relações entre a ARS e a bibliometria; a Seção 4 apresenta a base de dados do projeto 'RedeCI' e algumas estatísticas sobre os autores e artigos e, em seguida, a metodologia usada para o levantamento da rede de co-autoria dos professores dos programas de pós-graduação em CI; a Seção 5 analisa os principais resultados; e a Seção 6 aponta para o baixo nível de cooperação entre os professores de diferentes instituições medido pela co-publicação de artigos. Destaca também a importância da ARS como um instrumento valioso de análise para a CI, desde que se disponha de bases de dados para estudos quantitativos. São apresentadas, também, algumas questões lançadas a partir dos resultados, os quais deverão ser respondidos em trabalhos futuros.

2 VISÃO GERAL DA ARS – CONCEITOS BÁSICOS

O uso da ARS vem crescendo significativamente nos últimos 20 anos. Tal crescimento vem ocorrendo em função do aumento da quantidade de dados disponíveis para análise, do desenvolvimento nas áreas de informática e processamento de dados – com o conseqüente aumento do poder computacional à disposição dos pesquisadores – e da ampliação dos assuntos de interesse e das áreas de conhecimento que dela se utilizam. Alguns pesquisadores demonstraram essa tendência a partir de pesquisas realizadas em base de dados de artigos científicos e programas de pesquisa (OTTE, ROUSSEAU, 2002; BORGATTI, FOSTER, 2003).

A fim de se utilizar a ARS como instrumento de pesquisa, a definição de alguns conceitos essenciais se faz necessária, começando pelo próprio conceito de rede social (do inglês *social network*): “[...] *um ou mais conjuntos finitos de atores [e eventos] e todas as relações definidas entre eles*” (WASSERMAN, FAUST, 1999, p. 29). Dessa forma, para se entender adequadamente a definição de rede social deve-se entender os conceitos de ator, atributo, laço relacional, grafo, e algumas medidas da rede.¹

Um *ator*, em ARS, é uma unidade discreta que pode corresponder a diferentes entidades, como uma pessoa, um conjunto discreto de pessoas agregados em uma unidade social coletiva, subgrupos, organizações e outras coletividades. O conceito de ator é flexível, permitindo diferentes níveis de agregação, o que possibilita sua adequação a diferentes problemas de pesquisa. Os *atributos* (*attributes*) de um ator são suas características individuais como nome, sexo e idade. Embora a ARS focalize, principalmente, os laços relacionais, os atributos também podem ser analisados, inclusive com o auxílio de outros métodos estatísticos que estão além da ARS. O conjunto de atributos de uma rede é denominado *composição* da rede social. O *laço relacional* (*relational tie*) é responsável por estabelecer a ligação entre pares de atores. Uma *relação* em uma rede (*relation*) define todo o conjunto de laços que respeitam o mesmo critério de relacionamento, dado um conjunto de atores. As relações têm duas propriedades importantes: direcionamento e valoração. No primeiro caso, as relações podem ser direcionais, caso no qual têm um ator como transmissor e outro como receptor, ou não-direcionais, caso no qual a relação é recíproca. No caso da valoração, as relações podem ser dicotômicas, o que implica na presença (valor 1) ou ausência (valor 0) de uma relação, ou valoradas, com valores discretos ou contínuos. A *distância geodésica* $d(n_i, n_j)$, entre um par de atores é o número de laços que existem no caminho mais curto entre eles.

Uma rede (ou grafo) não direcionada $G(N, L)$ (*undirected graph*) consiste de dois conjuntos de informações: um conjunto de *atores* (ou nós), representado por $N = \{n_1, n_2, \dots, n_g\}$ e um ou mais conjunto(s) de *laços* representados por $L = \{l_1, l_2, \dots, l_L\}$ entre

pares de atores. O número total de atores existente em uma rede social é representado por g e o número total de laços por L . O laço (não direcionado) responsável por ligar os atores n_i e n_j é representado por $l_k = (n_i, n_j)$. Em uma rede social (não direcionada), o grau (*nodal degree*), denotado por $d(n_i)$, é o número de laços incidentes em um ator, ou ainda, de forma equivalente, o número de atores adjacentes a ele. O grau de um ator pode variar de 0 , caso no qual o ator é isolado, até $g - 1$, caso em que o ator está em contato com todos os demais atores da rede.

Torna-se importante, também, destacar o significado de algumas medidas de importância de um ator em uma rede, como o grau de centralidade (*degree centrality*), o grau de proximidade (*closeness centrality*) e o grau de intermediação (*betweenness centrality*). Na ARS, essas características estruturais mostram a posição de cada ator na estrutura. O grau de centralidade para um ator é dada por $C_D(n_i) = d(n_i)$; ou seja, é simplesmente o grau do ator.

O grau de proximidade de um ator mede o quanto o ator está próximo de todos os demais atores da rede. Para calcular a centralidade de proximidade se soma a distância geodésica do ator em relação a todos os demais atores da rede, e depois inverte-se, uma vez que quanto maior a distância menor a proximidade: $C_C(n_i) = [\sum_{j=1}^g d(n_i, n_j)]^{-1}$. O grau de intermediação,

por seu turno, analisa o quanto um ator está no caminho geodésico entre outros atores na rede. Seja g_{jk} o número de caminhos geodésicos (mais curtos e de mesmo tamanho) que ligam os atores j e k , e $g_{jk}(n_i)$ o número de caminhos geodésicos; no total de g_{jk} , que passa pelo ator

n_i , o índice de centralidade de intermediação é dado por $C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}$ ii.

A densidade de uma rede é dada pelo número de laços existentes em uma rede em relação ao máximo possível: $D = 2L / (g * (g - 1))$ (WASSERMAN; FAUST, 1999, p. 101). Para muitos tipos de análise, ela representa a proximidade dos membros do grupo e, nesse caso, pode-se supor que uma 'sub-rede' que tenha maior densidade e seja mais coesa (WASSERMAN; FAUST, 1999, p. 270-273) favoreça o fluxo de determinados tipos de informações entre seus membros. Numa rede, se existe um caminho entre dois atores, tais atores são ditos *alcançáveis* entre si (*reachable*). Se todos os atores de uma rede são alcançáveis entre si, então a rede é *conectada* e os dados dos atores e seus relacionamentos são representados em um grafo composto por um único componente. Se a rede não é inteiramente conectada, ela é composta por dois ou mais *componentes*, com as mesmas propriedades, sendo que um componente de um grafo é um sub-grafo no qual os atores são alcançáveis internamente, mas não a partir dos demais sub-grafos. Em outras palavras, cada componente é um subconjunto dos atores e laços que formam uma rede social (WASSERMAN; FAUST, 1999, p. 109).

3 ARS, BIBLIOMETRIA, CO-AUTORIA E COOPERAÇÃO

A bibliometria é descrita como um campo de pesquisa verdadeiramente interdisciplinar. Sua metodologia inclui elementos da matemática, das ciências sociais e naturais, da engenharia, e das ciências da vida (GLÄNZEL, 2003). Uma das sub-áreas de estudo da bibliometria se relaciona com o estudo das co-autorias como indicadores da produção científica.

Apesar do crescimento acentuado das publicações com múltiplos autores, frutos da crescente cooperação entre instituições e da ampliação das pesquisas multidisciplinares, deve-se levar em conta que existem vários fatores associados à indicação de um colaborador que transcendem a colaboração *per se*. Glänzel (2003), baseando-se nos estudos de Beaver (2001),

enumera 18 motivos pelos quais os cientistas cooperam entre si, sendo que alguns estão associados à montagem de redes ou colégios invisíveis, associação com pesquisadores de maior renome ou à obtenção de prestígio. Embora não entrando na descrição da metodologia de ARS, várias análises são feitas a partir das redes de co-publicação e co-autoria, inclusive entre países. Na descrição de ferramentas bibliométricas feitas por Glänzel (2003), observa-se que várias delas combinam as técnicas de bibliometria tradicional com a ARS (GLÄNZEL,2003).

Na mesma linha, Katz e Martin (1997) consideram que a co-autoria é apenas um indicador parcial da colaboração entre pesquisadores. Inicialmente, os autores analisam o significado de colaboração em pesquisa, mostrando que se trata de um termo que não tem uma definição única de uso geral. Eles apontam as dificuldades em se medir a colaboração, mesmo com entrevistas e questionários, em função da complexidade da natureza da interação humana. O artigo analisa, ainda, a premissa de que a colaboração é positiva para os resultados das pesquisas, concluindo que pouco foi feito para explicar como e por que ocorre a colaboração.

Com relação ao aumento do número de autores por publicação, fazendo-se uma combinação do que vários pesquisadores na área detectaram, são destacadas 10 razões para que isso se dê. O aumento da participação de vários pesquisadores de diferentes formações nas pesquisas evidencia a sua característica multidisciplinar e explica o aumento da colaboração, mas outros fatores também são relevantes tais como a área de conhecimento de origem, os contatos informais, a distância física e a existência de colégios invisíveis (KATZ, MARTIN,1997).

A própria distinção entre colaboração e co-autoria poderia ser estudada com auxílio da ARS. Outros trabalhos que evidenciam a importância da ARS como ferramenta relevante nos estudos de comunidades de prática e na análise bibliométrica também podem ser citados. Aplicando a ARS em dois departamentos da mesma área (Biologia), mas de duas diferentes universidades suecas, são analisadas as redes de co-autoria e de citações diretas e indiretas, buscando identificar a integração dos diferentes grupos de pesquisa e a posição dos pesquisadores na rede, segundo a sua produtividade (MAHLCK, PERSSON, 2000). Em outro trabalho, tomando por base as redes de co-autoria de 385 autores de diferentes nacionalidades, Kretschmer (2004) usa a ARS e as demais informações bibliométricas (produção e produtividade) para definir os atributos dos autores e analisar a posição, na rede, daqueles com as mesmas características ou atributos, numa tentativa de analisar a relação entre os laços sociais e a produção científica, conforme a preocupação de Crane (1972), já citada.

Também na linha de análise de redes e estrutura social dos colaboradores, e buscando a compreensão sobre a dinâmica e o crescimento dos grupos de pesquisadores e o seu impacto sobre as redes de cooperação, Yoshikane e Kageura (2004) usam a ARS para estudar a estrutura de cooperação entre pesquisadores japoneses de 4 áreas: 2 da área de engenharia (engenharia elétrica e processamento de dados) e 2 na área de Química (bioquímica e polímeros). Ainda no que diz respeito às redes de colaboração entre cientistas, Newman (2001) constrói redes de co-autoria, para o período de 1995 a 1999, a partir de grandes bases de dados americanas (variando de 13 mil artigos e 12 mil autores na área de Ciência da Computação, a 2 milhões e 1,5 milhões, respectivamente, na área pesquisa biomédica). Os resultados evidenciam duas características sociais relevantes nessas diferentes comunidades científicas: as redes são do tipo “mundo pequeno” (*small-world*), ou seja, dois cientistas escolhidos aleatoriamente estão separados por um pequeno número de passos; e a presença de *clusters* ou seja, a maioria se conecta entre si através de autores intermediários e 2 cientistas têm 30% ou mais de probabilidade de colaborarem entre si se ambos já colaboraram com um terceiro cientista. (NEWMAN, 2001).

Em outro estudo sobre as redes de colaboração entre autores que publicaram na revista ‘Scientometrics’ entre 1978 e 2004, Hou, Kretschmer e Liu (2006) trabalharam uma base contendo 1630 autores e 1927 artigos, sendo 875 em co-autoria. O número médio de artigos por autor é de 1,18 e de autores por artigo é de 1,73. Os autores fizeram um corte, extraindo da base um subconjunto para os autores com três ou mais publicações, resultando em 234 autores, sendo que destes, 163 publicaram em co-autoria. Na análise dos resultados foi usada a metodologia de ARS, com a identificação dos *clusters* e a estimativa das medidas de centralidade. Adicionalmente, foram introduzidas outras características dos *clusters*, como a linha de pesquisa dos artigos (obtida a partir da frequência de termos nos seus títulos); e os resultados indicaram a presença de um grande componente, que agrega cerca de 59% dos autores, e uma baixa densidade da rede, ou seja, uma rede fracamente conectada.

No Brasil, aplicando-se a metodologia de ARS, foi feito um estudo sobre a rede de co-autoria englobando os professores do PPGCI/UFMG, com o objetivo de identificar: i) se ela é densa, com muitos artigos publicados nessa condição, envolvendo a maior parte dos professores; ii) se existe colaboração entre os professores das diferentes linhas de pesquisa (SILVA *et al.* 2005). Os resultados estariam manifestados nas publicações totais e, especialmente, nas conjuntas. A base de dados foi construída a partir da base do CNPq – Lattes (CNPQ, 2005) e foi a origem do projeto “RedeCI”. Os resultados mostraram que a rede de colaboração é bastante esparsa e muito da produção em colaboração se dá entre um número reduzido de colaboradores preferenciais. As metodologias utilizadas nos dois últimos trabalhos acima mencionados são próximas entre si e daquela utilizada no presente artigo.

4 METODOLOGIA

Este trabalho analisa a produção (em co-autoria) de artigos científicos por professores dos programas de pós-graduação em Ciência da Informação no Brasil. Os artigos analisados são oriundos de revistas e eventos da área, previamente cadastrados na base de dados da RedeCI (PARREIRAS *et al.*, 2006), enquanto a vinculação do professores aos programas de Pós-Graduação e suas linhas de pesquisa foi identificada a partir de informações disponíveis no sítio da ANCIB (ANCIB, 2006)ⁱⁱⁱ, e verificada, quando possível, no sítio Web da respectiva instituição. Dessa forma, além da rede individual dos professores, ao se adicionar na análise a instituição de origem e a linha de pesquisa, foi possível, também, elaborar uma primeira interpretação da colaboração entre instituições.

Procurou-se, então, responder às seguintes perguntas: 1) existe uma rede de colaboração, na forma de publicação em co-autoria, entre os professores de diferentes programas de pós-graduação em CI no Brasil? 2) quais as características desta rede? Ou seja: 2.1) a rede é densa, com elevado número de artigos publicados em co-autoria? 2.2) existe colaboração entre os professores de diferentes instituições e linhas de pesquisa?

4.1 DESCRIÇÃO DA BASE DE DADOS DA ‘REDECI’

A base de dados utilizada foi a do projeto ‘RedeCI’ (2006), versão de julho de 2006. Os elementos dessa base de dados são os autores e artigos publicados nas principais revistas da área: Revista Ciência da Informação, Datagrama zero, Encontros Bibli, Informação e Sociedade, Informática Pública, Perspectivas em Ciência da Informação, Revista Digital de Biblioteconomia e Ciência da Informação e Transinformação, ou apresentados nos principais eventos da área: CINFORM e ENANCIB. Os resultados se encontram à disposição no sítio Web^{iv} do projeto.

Na versão aqui usada, a base conta com 1604 autores e 1767 artigos, sendo 1056 artigos (60%) com um único autor e 612 em co-autoria (40%). A maioria dos autores da base publica apenas uma vez ou participa de um único evento. Esses autores, denominados transientes, representaram 60% do total; tomando-se apenas os periódicos, somente 18,6%

dos autores publicaram em duas ou mais revistas da área. Dos 1604 autores, 116 (7,2%) são professores distribuídos em 9 instituições distintas. Eles participaram, individualmente ou em co-autoria, de 612 (34,6%) dos 1767 artigos da base. Do total de professores, 83 escreveram em parceria com outros professores ou autores, sendo que 54 escreveram artigos uns com os outros - 144 artigos no total - e 73 escreveram artigos com outros autores não-professores - 412 artigos no total^v. Do restante, 12 (10,4% dos professores) escreveram isoladamente e outros 21 não publicaram artigos.

A análise da produtividade científica e da distribuição dos artigos e autores foi feita com base na Lei de Lotka^{vi}, um dos pioneiros da bibliometria (GLÄNZEL, 2003). Os parâmetros de concentração da produção foram estimados e testados com base no aplicativo desenvolvido por Rousseau e Rousseau (2000) e os resultados, para a 'RedeCI', mostram uma participação de cerca de 70% de autores com apenas uma única contribuição, participação maior que a prevista pela formulação básica da Lei^{vii}. Os valores obtidos chamam a atenção para a possibilidade de que grande parte das contribuições seja de autores que concluíram um trabalho de pesquisa referente à elaboração da sua dissertação ou tese e que não permanecem na área de CI.

Algumas das constatações acima são consistentes com outros estudos bibliométricos da área. Em estudo anterior, levando-se em consideração uma única publicação – Ciência da Informação no período de 1990 a 1999 –, foram feitas análises tomando-se por base os atributos dos artigos, tais como o número de autores, os temas, o volume e o fascículo no qual foram publicados (MUELLER, PECEGUEIRO, 2001). Com relação à autoria, aproximadamente 78% dos 248 artigos apresentavam um único autor e, daqueles com múltipla autoria (54 artigos), a maior parte apresentava apenas dois autores. Dessa forma, o número total de autores pesquisados chegou a 270, sendo que 83% do total (225) apresentaram um único artigo nas revistas estudadas. Com o objetivo de analisar a produção científica dos docentes dos cursos de pós-graduação em CI registrados até 1999, Población e Noronha (2002) realizaram um estudo nessa linha. Inicialmente, com base nos atributos dos docentes (área e ano da titulação, linhas de pesquisa, instituição de origem, dentre outros) elaboraram o quadro sobre a evolução da formação dos professores vinculados ao programa. Em seguida, associaram os professores segundo as linhas de pesquisa identificadas e instituições de origem, assim como a sua produção (1108 documentos no período de 1990 a 1999). O total foi dividido em literatura branca (59,8%) e literatura cinzenta (40,2%). A maior parte dos documentos (aproximadamente 74%) tinha um único autor e 15% tinham dois autores.

A utilização da base de dados da 'RedeCI' permite a obtenção de resultados consistentes com outros estudos sobre a área de CI e permite avançar na compreensão das redes de colaboração existentes na área, através do uso da metodologia de ARS. Entretanto, na conclusão de seu trabalho, Población e Noronha, (2002) destacaram uma tendência de ampliação do número de trabalhos com dois ou mais autores, como reflexo da constituição de núcleos de pesquisa multidisciplinares, da ampliação da cooperação entre instituições e da maior valorização dos trabalhos em parceria por parte das agências patrocinadoras de pesquisa. O presente artigo testa tal afirmação em um dos indicadores da colaboração: as publicações em co-autoria.

4.2 METODOLOGIA DE SELEÇÃO DOS DADOS

Com base nos sítios dos programas de pós-graduação em Ciência da Informação e no sítio da ANCIB (2006), foram selecionados os professores/pesquisadores que serão usados na presente pesquisa. Implicitamente supõe-se que eles, por trabalharem na área em instituições de ensino e pesquisa, são os elementos estruturantes da comunidade científica dessa própria área. Além do nome dos professores, foram adicionadas à base de dados as suas instituições

de origem e as suas linhas de pesquisa. Como limitação, essa forma de coleta das informações retrata apenas a situação atual dos programas e, como a base de dados se refere a um período de aproximadamente 12 anos (de 1994 a 2005), a existência de professores que se retiraram do programa antes de 2006 não é considerada. Da mesma forma, não são consideradas as mudanças de vínculo institucional para os que mudaram de instituição.

Foram selecionados 116 professores de 9 instituições, sendo que 54 professores apresentaram trabalhos de co-autoria entre si nas publicações e eventos considerados (TABELA 1). A lista encontra-se no Anexo 1. Em termos de linhas de pesquisa, as 21 linhas descritas nos sítios mencionados foram agregadas em quatro: epistemologia, ensino e pesquisa em CI (com 7 professores), gestão da informação e do conhecimento (44), informação, cultura e sociedade (31) e organização e uso da informação (34).

TABELA 1
Número de professores por instituição de ensino - Brasil.

Instituição	Total	Com co-autoria	%
PUCCAMP	8	4	50
UFBA	12	4	33
UFF/IBICT	18	14	78
UFMG	20	11	55
UFPB	7	0	0
UFSC	11	3	27
UNB	13	7	54
UNESP	11	5	45
USP	16	6	38
Total	116	54	47

Fonte: os autores.

Na base dos 54 professores foi aplicada a metodologia de ARS para identificação de grupos (*clusters*) de autores e cálculo das medidas de centralidade. Foram geradas duas matrizes quadradas (os mesmos autores aparecem nas linhas e nas colunas, isto é, os valores i e j são iguais) e simétrica em relação à diagonal principal, isto é, os laços aqui estudados são recíprocos (A é co-autor de um artigo com B e vice-versa). Não existem os 'auto-laços', isto é, os autores não têm laços consigo mesmo. A primeira matriz é denominada 'Matriz dicotômica ou binária', pois indica apenas a existência de laços (no caso, de co-autorias) e a segunda é denominada 'Matriz valorada', pois informa a quantidade de relacionamentos de um determinado tipo, no caso, o número de artigos publicados por cada par de pesquisadores.

O nome dos professores foi substituído pela denominação da instituição, seguida por uma numeração seqüencial, de forma a facilitar a visualização das instituições às quais os professores estão vinculados^{viii}. A rede social obtida a partir da primeira matriz pode ser visualizada na Figura 1.

5 ANÁLISE DOS DADOS

O grupo de professores que apresentaram trabalhos em co-autoria não forma uma rede densa e interligada. A análise da rede indica que ela possui uma baixa densidade, com apenas 3,3% dos laços possíveis presentes. Essa baixa conexão se reflete na fragmentação em diversos componentes, conforme se observa na Figura 1. Foram encontrados 15 componentes, sendo cinco deles compostos por mais de três professores. Aprofundando-se a análise para esses componentes, observa-se que três são compostos por professores de uma mesma instituição: i) USP – com os professores dessa instituição identificados pelos números 11, 12 e 13, todos da linha de pesquisa 'organização e uso da informação'; ii) UFSC – com os professores 05 (linha de pesquisa 'informação, cultura e sociedade'), 09 e 11 (ambos da linha

‘epistemologia, ensino e pesquisa em CI’); e iii) UNESP - como os professores 02, 04 (ambos da linha ‘gestão da informação e do conhecimento’) e 06, 10 e 11 (todos da linha ‘organização e uso da informação’), sendo este último central, nesse componente.

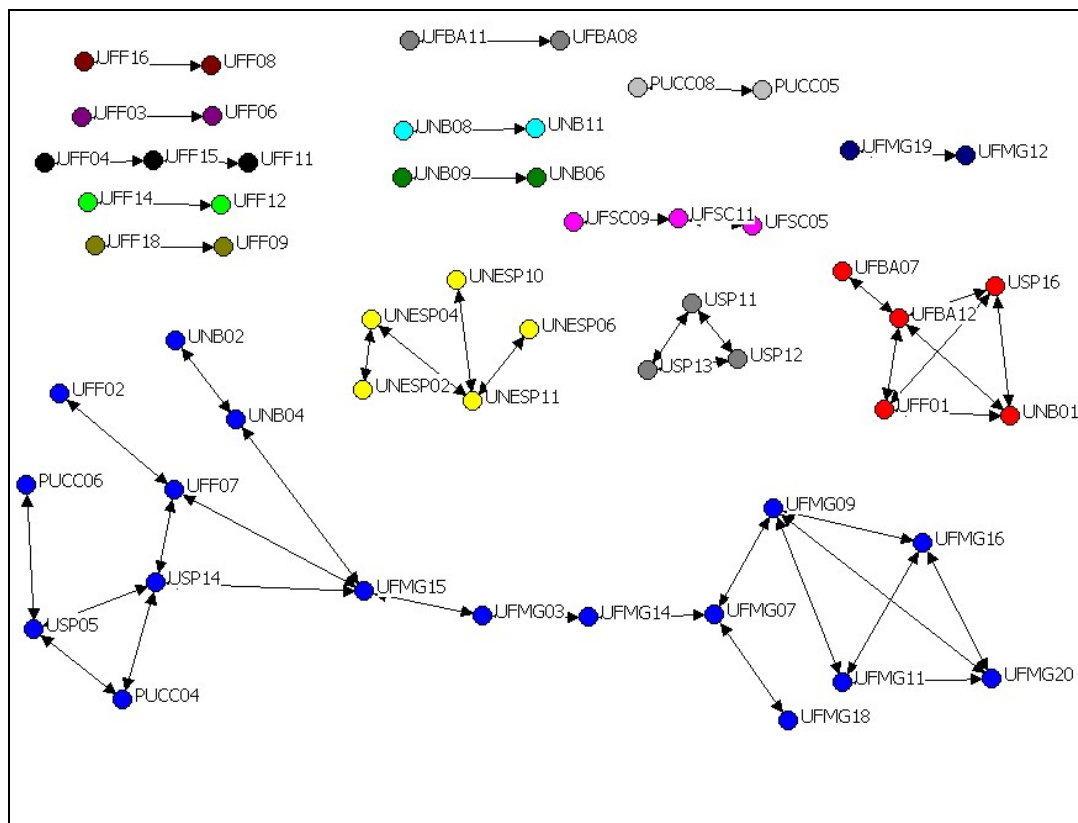


Figura 1 – Rede de co-autoria dos professores dos programas de pós-graduação em CI (Lista com nomes no Anexo 1).

Fonte: os autores, com uso do UCINET (BORGATTI; EVERETT; FREEMAN, 2002).

Os dois componentes que envolvem duas ou mais instituições, conforme a Figura 1, são:

- com 5 professores: UFBA07 e UFBA12 (ambos da linha ‘informação, cultura e sociedade’), UFF01 (‘gestão da informação e do conhecimento’), UNB01 (‘organização e uso da informação’) e USP16 (‘informação, cultura e sociedade’);
- com 17 professores: UFF02, UFF07, UFMG07, UFMG09, UFMG18 e UNB02 (‘gestão da informação e do conhecimento’), UFMG11 e UFMG20 (‘informação, cultura e sociedade’), PUCC04, PUCC06, UFMG03, UFMG14, UFMG15, UFMG16, UNB04, USP05 e USP14 (‘organização e uso da informação’).

Analisando-se mais detidamente esse último componente, observa-se que a ligação entre o subconjunto de professores da UFMG e os demais participantes é dada por um único professor – o UFMG15. No outro subconjunto os professores da UNB estariam isolados se não fosse sua relação com esse professor, que também é fundamental para a ligação com as demais instituições. Além dele, destacam-se no subconjunto da UFMG o professor UFMG07 e no outro subconjunto os professores UFF07 e USP14. Essas características ficam evidentes quando se calcula as medidas de centralidade da rede (QUADRO 1).

Para o aprofundamento da análise da ligação entre os professores de instituições distintas, foi feita uma análise, com base na Plataforma Lattes (CNPq, 2006) das relações adicionais que poderiam existir entre os professores, destacando-se os grupos de pesquisa e a relação orientando-orientador. Chegou-se aos seguintes resultados: i) no componente com 5

professores, todos mantêm relações entre si (com exceção de UFBA07) graças a uma única publicação. O laço restante representa a ligação entre dois professores de uma mesma instituição; ii) no segundo grupo, a ligação entre os professores UFMG15, UFF07 e USP14 é observada em apenas um artigo, da mesma forma que entre os professores UFMG15 e UNB04. Não foi observada entre eles nenhuma relação de orientação ou participação em grupos de pesquisa. Dessa forma, a posição central do professor UFMG15 é resultado da publicação de 2 artigos e sua ligação com o grupo composto pelos professores da UFMG; iii) já a ligação entre os professores USP14, USP05 e PUC04 é mais rica, contendo maior número de relações. Existem na base três artigos com essa estrutura de autores. Todos participam de pelo mesmo um grupo de pesquisa em comum e o professor USP14 orientou o professor USP05 na obtenção de seu título. Este último também participa de outro grupo de pesquisa com o professor PUC06 e o laço apresentado mostra a co-autoria de um artigo^{ix}.

QUADRO 1

Posição dos professores segundo as medidas de centralidade da rede
(somente os 10 primeiros, ordenados pelo grau de intermediação (*Betweenness centrality*))

Professores	Grau de centralidade (degree centrality)	Grau de proximidade ^x (closeness centrality)	Grau de intermediação (Betweenness centrality)
UFMG15	1	1	1
UFMG03	15	2	2
UFMG14	16	3	3
UFMG07	5	5	4
USP14	3	4	5
UFMG09	2	8	6
UFF07	6	6	7
UNB04	17	7	8
USP05	7	9	9
UNESP11	8	19	10

Fonte: os autores, com uso do UCINET (BORGATTI; EVERETT; FREEMAN, 2002)

6 CONCLUSÕES

Os resultados obtidos com a ARS, aplicada sobre as redes de co-autoria de professores dos programas de pós-graduação em Ciência da Informação no Brasil, indicam uma baixa cooperação entre eles, tendo-se em vista que as revistas e eventos selecionados são os mais importantes para se acompanhar a produção científica nessa área de conhecimento no Brasil.

Mesmo considerando que se trata de uma primeira exploração e que existe um longo trabalho a ser realizado na análise da base de dados de co-autoria com uso da metodologia de ARS, os resultados confirmam outras indicações sobre a fraca conexão entre os professores em termos de colaboração na publicação de trabalhos, como a apresentada por Silva *et al.* (2005) para o programa da UFMG. Resultado similar é confirmado por Muller e Pecegueiro (2001) e por Población e Noronha, (2002), que demonstraram a presença significativa de artigos individuais na área ou em revistas da área. No último estudo, bem como no presente, foram considerados apenas os docentes das instituições de ensino no Brasil.

A primeira representação obtida com a rede de co-autoria não reforça a idéia de que existam ‘colégios invisíveis’ na área, uma vez que a colaboração em co-autoria encontrada foi principalmente intra-institucional. Obviamente, devem ser empreendidos outros estudos no futuro, acrescentando outras informações sobre as redes de cooperação inter-institucionais, como exemplificado pela composição dos grupos de pesquisa da CAPES, para a obtenção de um retrato aprofundado da área de CI no Brasil. Além disso, algumas investigações diretas e metodologias adicionais, como entrevistas com os pesquisadores, podem ser necessárias para

se atingir o objetivo de estudar de forma mais ampla os chamados colégios invisíveis e a colaboração entre instituições. Um resultado interessante, e que merece ser mais amplamente analisado, foi a existência de uma maior coesão social – como medida pelas co-autorias e reforçada por outros laços entre pesquisadores, como a co-participação em grupos de pesquisa e orientação de pesquisa de dissertação e tese - entre os professores de programas distintos, mas com proximidade geográfica (USP e PUC Campinas). Esse tipo de questão é relevante para os pesquisadores da CI e da sociologia da ciência.

Os resultados são apenas uma parte da exploração da base de dados construída no projeto 'RedeCI'. Buscou-se destacar a colaboração dos professores, materializada nas co-autorias em artigos. A mesma base de dados permite outros tipos de análise, como, por exemplo, sobre as publicações mais utilizadas, a produtividade por linha de pesquisa e, mesmo sobre a rede completa dos professores, além do subconjunto aqui analisado.

Os autores desse artigo gostariam de reforçar a tese de que a ARS e os resultados estatísticos não eliminam a necessidade de uma análise qualitativa aprofundada. No entanto, deve-se enfatizar que o uso da ARS permite uma abordagem da colaboração entre os professores, destacando fatos que não seriam detectáveis com o uso de métodos estatísticos tradicionais. Entender os mecanismos sociais por trás das conexões e como eles facilitam os fluxos de informações e de conhecimento entre os pesquisadores são problemas de pesquisa da CI. No presente artigo, foi descrito e analisado um dos aspectos da colaboração entre pesquisadores, mas sabe-se que, para a compreensão mais ampla das suas características fundamentais, há um longo caminho de pesquisa pela frente.

7 AGRADECIMENTO

Antonio Braz de Oliveira e Silva e Renato Matheus realizaram o presente trabalho com o apoio dos Programas de Pós-Graduação do IBGE e BACEN, respectivamente.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANCIB. Associação Nacional de Pesquisa em Ciência da Informação. Disponível em <<http://www.ancib.org.br/>>. Acesso em: 15 julho de 2006.
- BEAVER, D. D. Reflections on scientific collaboration (and its study): past, present, and future. **Scientometrics**, [s.l.], v. 52, n. 3, p.365-377, jan. 2001.
- BORGATTI, S. P., EVERETT, M. G., FREEMAN, L. C. 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.
- BORGATTI, S. P., FOSTER, P. C. The network paradigm in organizational research: a review and typology. **Journal of Management**, Miami, v.29, n.6, p.991-1013, dez. 2003.
- BUSH, V. As we may think. **Atlantic Monthly**. v.176, n.1, p.110-116, jul. 1945.
- CNPq. Conselho Nacional de Desenvolvimento Científico e Tecnológico. Plataforma Lattes. Brasília, 2005. Disponível em: <<http://lattes.cnpq.br/pl/>>. Acesso em: 08 ago. 2006.
- CRANE, D. **Invisible Colleges: diffusion of knowledge in scientific communities**. Chicago and London: The University of Chicago Press, 1972.
- GLÄNZEL, W. **Bibliometrics as a research field: a course on theory and application of bibliometric indicators**. Courses Handout, 2003.
- HOU, Haiyan, KRETSCHMER, Hildrun, LIU, Zeyuan. The Structure of Scientific Collaboration Networks in Scientometrics. In: International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting, 2006, Nancy (France). Disponível em: <<http://eprints.rclis.org/archive/00006101/01/446D95.pdf>>. Acesso em: 28 abr. 2006.
- KATZ, J. S., MARTIN, B. R. What is Research Collaboration? **Research Policy**, v.26, p.1-18, 1997.

KRETSCHMER, H. Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the Web. **Scientometrics**, [on line], v. 60, n. 3, p. 409-420, jan. 2004.

MATHEUS, Renato Fabiano, SILVA, Antonio Braz de O. Análise de redes sociais como um método para a Ciência da Informação. **DataGramZero: Revista de Ciência da Informação**, [s.l.], v. 7, n. 2, p.1-15, abr. 2006. artigo 3. Disponível em: <http://www.dgz.org.br/abr06/F_I_art.htm>. Acesso em 17 jul. 2006.

MAHLCK, P., PERSSON, O. Socio-bibliometric mapping of intra-departmental networks. **Scientometrics**, [on line], v. 49, n. 1, p. 81-91, 2000.

MUELLER, S. P. M., PECEGUEIRO, Claudia Maria. O periódico Ciência da Informação na década de 90: um retrato da área refletido em seus artigos. **Ciência da Informação**, Brasília, v. 30, n. 2, p.47-63, maio/ago. 2001.

NEWMAN, M. E. J. From the Cover: The structure of scientific collaboration networks. **Proceedings of The National Academy of Sciences of USA (PNAS)**, [on line], v. 98, p. 404-409, jan. 2001. Disponível em: <<http://www.pnas.org/cgi/reprint/98/2/404>>. Acesso em: 01 mar. 2005.

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

PARREIRAS, Fernando Silva *et al.* RedeCI: Quem é quem na Ciência da Informação no Brasil. **Perspectivas em Ciência da Informação**, v. 11, n. 3, 2006. (no prelo).

POBLACIÓN, D. A., NORONHA, D. P. Produção das literaturas “branca” e “cinzenta” pelos docentes/doutores dos programas de pós-graduação em ciência da informação no Brasil. **Ciência da Informação**, Brasília, v. 31, n. 2, p.98-106, mai. 2002.

RedeCI. Disponível em <<http://www.redeci.netic.com.br>>. Acesso em 08 ago. 2006.

ROUSSEAU, B., ROUSSEAU, R. LOTKA: A program to fit a power law distribution to observed frequency data. **Cybermetrics**, [s.l.], v. 4, n. 2, p.1-13, 2000.. Disponível em: <<http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html>>. Aplicativo disponível em <<http://www.cindoc.csic.es/cybermetrics/articles/lotka102.exe>>. em: 25 out. 2005.

SILVA, Antonio Braz de Oliveira e; MATHEUS, Renato Fabiano; PARREIRAS, Fernando Silva; PARREIRAS, Tatiane Aparecida Silva. Estudo da Rede de Co-Autoria e da Interdisciplinaridade na Produção Científica Através de Métodos de Análise de Redes Sociais: Avaliação do Caso do PPGCI/UFMG. In: ENANCIB, 6, Florianópolis, 2005.

WASSERMAN, S., FAUST, K. **Social Network Analysis: methods and applications**. Cambridge: Cambridge University Press, 1999. 857p.

YOSHIKANE, F., KAGEURA, K. Comparative analysis of coauthorship networks of different domains: The growth and change of networks. **Scientometrics**, [on line], v. 60, n. 3, p. 435-446, jan. 2004.

ⁱ Para um estudo aprofundado do tema sugere-se a leitura dos principais manuais da área (ver, por exemplo, WASSERMAN, FAUST, 1999).

ⁱⁱ Todas essas medidas podem ser normalizadas, o que permite a comparação entre atores de redes diferentes.

ⁱⁱⁱ As informações foram, posteriormente, confirmadas nos sítios das respectivas instituições.

^{iv} <http://www.redeci.netic.com.br>

^v Deve-se mencionar que professores que se afastaram de suas instituições, por aposentadoria ou outro motivo qualquer e não constam do seu sítio *Web*, não são tratados como tais.

^{vi} Considera que a lei de Lotka pode ser descrita como uma lei de potência da forma C/k^β e, como β tende para 2, C seria igual a $6/(\pi)^2$ ou, aproximadamente, 0,61. Assim, se uma bibliografia pode ser descrita por essa lei de potência, aproximadamente 61% dos autores teriam contribuído com apenas uma publicação.

^{vii} A adequação das estimativas dos parâmetros foi feita por meio do teste de Kolmogorov-Smirnov., feito pelo mesmo aplicativo que calcula os parâmetros da função (ROUSSEAU, ROUSSEAU, 2000) indicou que os resultados acima eram, estatisticamente, significativos.

^{viii} A lista dos professores, com seus atributos, encontra-se no Anexo 1. A lista completa de professores pode ser obtida no sítio da 'RedeCi' (<http://redec.netic.com.br/>)

^{ix} Outros resultados sobre a rede ('*Bonacich's Power*' e '*Reach Centrality*') confirmam e reforçam as conclusões iniciais aqui apresentadas (rede pouco conectada e fragmentada) e não serão apresentadas nesse artigo, em função da exigüidade de espaço.

^x Tecnicamente, por se uma rede fragmentada, essa medida não poderia se calculada. No entanto, o UCINET o faz, com o uso de um artifício (BORGATTI; EVERETT; FREEMAN, 2002)

ANEXO 1

Lista de professores dos Centros de Pós-Graduação em Ciência da Informação com co-autoria com outros professores

Instituição	Denominação usada na ARS	Professor	Linha de Pesquisa
UFF/IBICT	UFF01	aldo de albuquerque barreto	gestão da informação e do conhecimento
UFMG	UFMG11	ana maria rezende cabral	informação, cultura e sociedade
UNB	UNB01	antonio lisboa carvalho de miranda	organização e uso da informação
UFMG	UFMG14	beatriz valadares condon	organização e uso da informação
UFF/IBICT	UFF11	carlos henrique marcondes	gestão da informação e do conhecimento
USP	USP11	daisy pires noronha	organização e uso da informação
USP	USP12	dinah aguiar poblacion	organização e uso da informação
USP	USP13	edmir perrotti	informação, cultura e sociedade
UFSC	UFSC05	edna lucia da silva	gestão da informação e do conhecimento
UFMG	UFMG15	eduardo jose wense dias	organização e uso da informação
UNB	UNB06	emir jose suaiden	gestão da informação e do conhecimento
UFF/IBICT	UFF12	geraldo moreira prado	epistemologia, ensino e pesquisa em ci, etc
UFMG	UFMG16	gercina angela borem de oliveira lima	organização e uso da informação
UFBA	UFBA07	helena pereira da silva bemfica	informação, cultura e sociedade
UFF/IBICT	UFF14	isa maria freire	gestão da informação e do conhecimento
UNB	UNB08	jaime robredo	organização e uso da informação
UNESP	UNESP06	joao batista ernesto de Moraes	organização e uso da informação
USP	USP14	johanna w smit	organização e uso da informação
UFMG	UFMG18	jorge tadeu de ramos neves	gestão da informação e do conhecimento
UFF/IBICT	UFF15	jose maria jardim	informação, cultura e sociedade
USP	USP16	jose teixeira coelho netto	informação, cultura e sociedade
UFBA	UFBA08	katia carvalho	informação, cultura e sociedade
UNB	UNB09	kira tarapanoff	gestão da informação e do conhecimento
UFF/IBICT	UFF16	lena vania ribeiro pinheiro	epistemologia, ensino e pesquisa em ci, etc
UFMG	UFMG19	lidia alvarenga	organização e uso da informação
UFMG	UFMG20	lúgia maria moreira dumont	informação, cultura e sociedade
UFF/IBICT	UFF18	liz-rejane legey	informação, cultura e sociedade
UFSC	UFSC09	magda chagas pereira	epistemologia, ensino e pesquisa em ci, etc
UFMG	UFMG03	maria aparecida moura	organização e uso da informação
PUCCAMP	PUCC04	maria de fatima goncalves moreira talamo	organização e uso da informação
UNESP	UNESP10	maria isabel asperti nardi	organização e uso da informação
UFF/IBICT	UFF02	maria luiza de almeida campos	gestão da informação e do conhecimento
UFF/IBICT	UFF03	maria nelida gonzalez de gomez	epistemologia, ensino e pesquisa em ci, etc
UFF/IBICT	UFF04	maria odila fonseca	gestão da informação e do conhecimento
UNESP	UNESP11	mariangela spotti lopes fujita	organização e uso da informação
UFMG	UFMG07	marta araujo tavares ferreira	gestão da informação e do conhecimento
UNESP	UNESP02	miriam celi pimentel porto foresti	gestão da informação e do conhecimento
UFSC	UFSC11	miriam vieira da cunha	epistemologia, ensino e pesquisa em ci, etc
UFMG	UFMG09	monica erichsen nassif borges	gestão da informação e do conhecimento
UNB	UNB11	murilo bastos da cunha	gestão da informação e do conhecimento
USP	USP05	nair yumiko kobashi	organização e uso da informação
UFBA	UFBA11	nanci elizabeth e gomes oddone	informação, cultura e sociedade
PUCCAMP	PUCC05	orandi mina falsarella	gestão da informação e do conhecimento
UFBA	UFBA12	othon jambeiro	informação, cultura e sociedade
PUCCAMP	PUCC06	raimundo nonato macedo dos santos	organização e uso da informação
UFMG	UFMG12	renato rocha souza	organização e uso da informação
UFF/IBICT	UFF06	rosa ines de novaes cordeiro	gestão da informação e do conhecimento
UFF/IBICT	UFF07	rosali fernandez de souza	gestão da informação e do conhecimento
UFF/IBICT	UFF08	sandra lucia rebel gomes	gestão da informação e do conhecimento
UFF/IBICT	UFF09	sarita albagli	informação, cultura e sociedade
UNESP	UNESP04	silvana aparecida borsetti gregorio vidotti	gestão da informação e do conhecimento
UNB	UNB02	sofia galvao baptista	gestão da informação e do conhecimento
UNB	UNB04	suzana pinheiro machado mueller	organização e uso da informação
PUCCAMP	PUCC08	vera silvia marao beraquet	gestão da informação e do conhecimento

Fonte: Projeto "RedeCi" (www.redeci.netic.com.br)