

# Análise de redes sociais como metodologia de apoio para a discussão da interdisciplinaridade na ciência da informação

## Antonio Braz de Oliveira e Silva\*

Analista do IBGE, doutorando em ciência da informação, Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG), e membro do Núcleo de Estudos em Tecnologias para Informação e Conhecimento (Netic) ([www.netic.com.br](http://www.netic.com.br))  
E-mail: [abr@netic.com.br](mailto:abr@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](http://www.netic.com.br))  
E-mail: [renato.fabiano@bcb.gov.br](mailto:renato.fabiano@bcb.gov.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](http://www.netic.com.br)).  
E-mail: [fparreiras@netic.com.br](mailto:fparreiras@netic.com.br)

## Tatiane A. Silva Parreiras

Bacharela em ciência da informação – PUC-MG, assistente de pesquisa do Netic ([www.netic.com.br](http://www.netic.com.br))  
E-mail: [tatianeparreiras@netic.com.br](mailto:tatianeparreiras@netic.com.br)

---

## Resumo

*Este artigo apresenta a Análise de Redes Sociais (ARS) como um método a ser aplicado em estudos na ciência da informação (CI), inclusive para a compreensão da estruturação da pesquisa nessa área do conhecimento. Esta é, normalmente, apresentada como uma área do conhecimento interdisciplinar, e as diferentes linhas de pesquisa existentes nos Programas de Pós-Graduação, no Brasil, recebem influências de diferentes áreas do conhecimento. A aplicação da ARS no estudo da rede de co-autoria dos professores do PPGCI/UFMG permite tanto a apresentação do método, quanto a obtenção de resultados empíricos para alimentar a discussão sobre a CI. As redes de colaboração entre os professores refletem, em tese, a participação em programas de pesquisa da área. A colaboração entre professores das diferentes linhas se relaciona com a integração das diferentes interdisciplinaridades que elas apresentam. O artigo expõe, inicialmente, uma discussão sobre a interdisciplinaridade na CI; em seguida dá uma visão geral da ARS e apresenta estudos da área, destacando-se a análise de redes de co-autoria. Finalmente, descreve-se a metodologia da pesquisa e os principais resultados obtidos da ARS. A conclusão destaca os resultados e reforça a importância da ARS como método para a CI, o qual pode ser explorado com sucesso, inclusive em pesquisas sobre redes de co-autoria e interdisciplinaridade.*

## Palavras-chave

*Redes sociais. Ciência da informação. Metodologia. Interdisciplinaridade. Redes de co-autoria.*

## Social network analysis as a method to support the debates about the information science and its interdisciplinary nature

### Abstract

*This paper discusses the Social Network Analysis (SNA) as a method to be broadly applied in Information Science (IS) research. Information science is normally considered as an interdisciplinary field. In this regard, research lines that comprise the research work in post-graduate programs in Brazil are influenced by different fields of knowledge. The co-authorship analysis of professors networks in the post-graduate program in information science at the University of Minas Gerais (PPGCI/UFMG) carried out by using SNA had a twofold output. The first one was to present the method itself. The second, to obtain empirical results that can be used to implement the discussion about the discipline. Collaborative networks between professors from different research lines is related to the integration of interdisciplinary characteristics found in the disciplines themselves. The article briefly presents the interdisciplinary nature of IS and gives an overview of the theoretical basis of SNA. Moreover, it presents studies about subject-matters related to IS, particularly the co-authorship network analysis. Finally, the methodological approach of this research and the main results are presented. In conclusion, the authors stress the benefits brought about by the use of SNA as a methodological tool. It can actually be fruitfully explored to support researches in both co-authorship networks and debates about Information Science and its interdisciplinary nature.*

### Keywords

*Social networks. Information science. Methodology. Interdisciplinarity. Co-authorship networks.*

---

\* Antonio Braz e Renato Matheus realizaram o presente trabalho com o apoio dos Programas de Pós-graduação do IBGE e Bacen, respectivamente.

## INTRODUÇÃO

A ciência da informação é apresentada, de uma maneira geral, como uma área do conhecimento relativamente recente e interdisciplinar, o que impede, em muitas situações, que se delimite, claramente, o seu campo de atuação. As discussões sobre seu programa de pesquisa e metodologia são colocadas em termos do seu relacionamento com outras disciplinas, e a simples identificação dos assuntos de interesse da CI oferece dificuldades decorrentes dessa característica.

Os programas de pós-graduação (*stricto sensu*) da área, que, no Brasil, representam as instituições de pesquisa, são estruturados em linhas de pesquisa, a maioria comum ao conjunto das instituições. Elas organizam suas linhas de pesquisa em três vertentes básicas: i) tratamento e uso da informação; ii) informação, cultura e sociedade; iii) gestão da informação e do conhecimento. As exceções são o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)/ Universidade Federal Fluminense (UFF), que têm uma linha de pesquisa em *teoria, epistemologia, interdisciplinaridade em CI*, e a Universidade Federal de Santa Catarina (UFSC), com uma linha na área de *ensino e formação dos profissionais da informação*. Essas linhas contemplam os aspectos práticos e teóricos desejáveis nos programas de pesquisa. No entanto, as diferenças encontradas nos termos que descrevem o conteúdo de cada uma delas são significativas, podendo-se inferir que cada uma tem uma característica interdisciplinar distinta, em termos de influência de outras áreas de conhecimento.

Em termos da convivência das linhas de pesquisa em um mesmo programa, seria possível esperar que as diferentes interdisciplinaridades permitissem diferentes abordagens para o mesmo problema de pesquisa e que as diferentes abordagens teóricas fossem se aproximando, na construção do campo da ciência da informação. Essa conjectura pode ser abordada, do ponto de vista de sua verificação, de vários ângulos e com diferentes instrumentais. Na verdade, não se pode analisar um campo de conhecimento tomando-se apenas um deles.

O objetivo do artigo é propor a metodologia de análise de redes sociais (ARS ou SNA, da expressão em inglês *Social Network Analysis*) como uma ferramenta para a análise da produção científica. Ela permite, por exemplo, a identificação de 'colégios invisíveis' e a observação de alguns aspectos da interdisciplinaridade decorrentes da colaboração de pesquisadores de áreas distintas. A metodologia de ARS aplicada às redes de colaboração

permite que a interdisciplinaridade das áreas de conhecimento possa ter um dos seus múltiplos aspectos capturados e analisados. Entender a formação destas redes, por exemplo, com relação à produção de pesquisas ou artigos em parceria, é uma das formas de se analisar a estruturação de um campo do conhecimento, conforme a literatura da área (CRANE, 1972). Sua importância é, ainda, maior para a ciência da informação, pois a metodologia adquire duas grandes funções: serve para a análise da sua própria produção científica, da mesma forma que para qualquer área do conhecimento e, ao mesmo tempo, constitui uma ferramenta complementar àquelas já empregadas nas análises bibliométricas. Para justificar a proposição, foi feito um estudo de caso, tendo como alvo o PPGCI/UFMG. O uso desse programa como referência para a aplicação da metodologia escolhida deve-se à sua importância para a CI no Brasil. Dessa forma, o estudo descrito nesse artigo se propõe a analisar a rede de co-autoria entre os professores do PPGCI/UFMG, procurando identificar: i) se ela é densa, como muitos artigos publicados nessa condição, envolvendo a maior parte dos professores; ii) se existe colaboração entre os professores das diferentes linhas de pesquisa. Esses resultados servem de base para se analisarem as relações de colaboração entre estas linhas que, conforme se depreende da descrição de seu conteúdo, possuem diferentes interdisciplinaridades.

Em outras palavras, utiliza-se a metodologia de análise de redes sociais para estudar a rede de co-autoria dos professores de um programa representativo na área – o PPGCI/UFMG –, tentando-se associar a colaboração entre os professores da mesma área ao vigor do programa de pesquisa e, entre aqueles de diferentes áreas, à intenção interdisciplinar das mesmas. Os resultados estariam manifestados nas publicações totais e, especialmente, nas conjuntas. Essa abordagem é uma das possíveis, uma vez que os pesquisadores podem ter (e, normalmente, têm) variadas formas de relações com membros de sua comunidade de pesquisa (outros professores do programa, alunos, pesquisadores de outras instituições ou de outras áreas e outros) e a compreensão final da organização social de uma área só poderia ser compreendida com a adição dessas outras formas de relacionamento (CRANE, 1972). No entanto, cada abordagem fornece pistas para as demais e permite que se testem algumas suposições sobre o comportamento social dos pesquisadores.

Este artigo está organizado em mais quatro seções, além dessa Introdução. Na seção 2, 'A Ciência da Informação (CI) e a Interdisciplinaridade', analisam-se a característica interdisciplinar da área, amplamente

mencionada na literatura, e as conseqüentes dificuldades de se organizar a pesquisa dentro da CI. Associa, também, as diferentes influências à estruturação das linhas de pesquisa que moldam os programas no Brasil, para apresentar a conjectura de que a colaboração entre as diferentes linhas seria um esforço para se delimitar o campo da CI. A seção 3, 'ARS e as redes de co-autoria', discute a metodologia de ARS e sua aplicação em estudos de co-autoria em diferentes países e comunidades de pesquisa. Apresenta, ainda, a notação matemática básica para se compreender a ARS. A seção 4, 'Estudo da rede de co-autoria entre os professores do PPGCI/UFMG', apresenta o Programa e a justificativa de sua escolha para a presente pesquisa, assim como a metodologia empregada no levantamento, crítica e tratamento dos dados, a análise de dados, as principais estatísticas descritivas, uma avaliação da distribuição da produção entre os professores, com base na distribuição de Lotka e no índice de Gini, e a ARS propriamente dita, apoiada em tabelas e grafos. Finalmente, na seção 5, 'Conclusão', avaliam-se as suposições e conjecturas ante os resultados e avança-se a proposição de estudos futuros.

## A CIÊNCIA DA INFORMAÇÃO E A INTERDISCIPLINARIDADE

A ciência da informação é, de uma maneira geral, apresentada como uma área do conhecimento relativamente recente e interdisciplinar, o que impede, em muitas situações, que se delimite o seu campo de atuação. As discussões do seu programa de pesquisa e metodologia são colocadas em termos do seu relacionamento com outras disciplinas, e a simples identificação dos assuntos de interesse da CI oferece dificuldades decorrentes dessa característica (TARGINO, 1995; CARDOSO, 1996; GÓMEZ, 2001; GOMES, 2001; LE COADIC, 1997; MACHLUP; MANSFIELD, 1983), uma vez que a área traz a influência de vários campos do conhecimento, tais como a biblioteconomia, a sociologia, a administração, a ciência da computação e a comunicação.

A discussão sobre a natureza da CI parece ser tão antiga quanto a história desse campo de conhecimento. De acordo com Vakkari (1991), durante os anos 70 a discussão tinha um viés teórico e conceitual, mas, nos anos 90, parece ter retornado e se ampliado, incluindo a relação entre biblioteconomia e ciência da informação e, sobretudo, a identificação e a definição dos conceitos centrais da disciplina, tais como conhecimento, informação e necessidade de informação.

A institucionalização social da CI pode ser entendida tomando-se emprestado o conceito de campo desenvolvido pelo sociólogo Pierre Bourdieu. Segundo ele, a atuação dos indivíduos se processa em contextos sociais específicos, que tocam em determinados aspectos da vida social. O campo é definido como o *locus* das lutas no qual os agentes buscam manter ou alterar a distribuição do capital (específico ao campo, isto é, relevante para o processo em questão), assim como a sua posição dentro dele. Portanto, os campos podem ser analisados independentemente das características que apresentam os atores individuais que dele fazem parte (BOURDIEU, 2000).

A partir de determinado ponto, qualquer área começa a se definir como um espaço relativamente autônomo, com suas próprias leis de funcionamento, seus teóricos, suas revistas etc., chegando-se a ponto em que já não se pode compreender o que é produzido em um campo sem se conhecer a sua história. Dessa forma, os campos se comportam como os demais organismos sociais, nascendo, desenvolvendo-se e, em algum momento, desaparecendo. Em termos de sistemas biológicos, os indivíduos com interesse no campo não lutam internamente só para alterar sua situação de poder, mas também para conservar e ampliar os espaços do campo *vis-à-vis* os demais.

Compreender a sua institucionalização cognitiva, isto é, o consenso e a clareza na formulação de teorias, os critérios para a relevância dos problemas e a definição e a aceitação das soluções encontradas, assim como dos métodos usados, significaria delimitar os objetos de pesquisa do campo, ou seja, especificar quais são os principais objetivos e as áreas promissoras para pesquisa. Dessa forma, esse aspecto da institucionalização está intimamente relacionado com a definição do *status* científico do campo estudado. A compreensão desse *status* passa por diferentes e complexas discussões e abordagens, indo da discussão básica do que é ciência, passando pelas origens históricas do campo, sem esquecer a discussão filosófica que analisa as diferenças entre as ciências naturais e as ciências sociais.

Para se chegar à compreensão desse *status*, no caso da ciência da informação, seria interessante observar a abordagem, do ponto de vista da história das ciências, feita por Thomas Kuhn. Em palestra intitulada O Problema com a Filosofia Histórica da Ciência, proferida no Departamento de História da Ciência da Universidade de Harvard, na série de conferências denominada *Rothschild Lectures*, ele faz uma análise do

papel dos filósofos da ciência na mudança da imagem que se tem da ciência, destacando a necessidade de, do ponto de vista histórico, entender-se o processo pelo qual as diferenças entre as crenças dos diversos grupos de pesquisadores vão convergir para um consenso. Do ponto de vista histórico, observou-se o surgimento de novas especialidades ao longo da história, uma analogia com a especialização observada, nos seres vivos, pela biologia. Ao longo dos últimos séculos, especialmente na área das ciências naturais, as especializações se transformaram em ciência, criando um campo próprio, com revistas etc. – um processo que adquire a forma de uma árvore (uma árvore genealógica, por assim dizer) (KUHN, 1992).

Uma vez que se cria um ramo novo nessa árvore, ele deve conter sua própria delimitação e o seu próprio objeto de interesse. A ciência da informação parece estar no caminho inverso, ou seja, com o enxerto de dois outros ramos (biblioteconomia e cibernética), pretende-se criar um novo, mais robusto, que se aproveite dos paradigmas dos anteriores e abra, ainda, novos espaços de investigação. Assim, aparentemente, a CI encontra-se diante de um paradoxo: se o caminho da transformação de especializações em ciências é a necessidade de se analisarem os novos problemas surgidos nas disciplinas já existentes, a partir de uma nova visão e de perspectivas de novas direções de desenvolvimento, criando, assim, um novo campo de conhecimento e uma nova ciência, por que se valer permanentemente dos métodos de outras disciplinas? Para discutir essa questão, Gómez (2000) utiliza, como ponto de partida, o conceito de programa de pesquisa, introduzido por Imre Lakatos, historiador da ciência, na década de 60. Um programa de pesquisa científica é um aglomerado de teorias conectadas que derivam de um núcleo central, composto pelas crenças comuns que unem os seguidores do programa. Segundo ela:

[...] a pesquisa em ciência da informação apresentaria um problema particular que podemos identificar de modo quase imediato: Se existe grande diversidade na definição das heurísticas afirmativas, as que definem as estratégias metodológicas de construção do objeto e que permitem a estabilização acumulativa do domínio, maior é a dificuldade para estabelecer as heurísticas negativas, as que definem o que não poderia ser considerado objeto do conhecimento da ciência da informação, condição diferencial que facilita e propicia as relações de reconhecimento e complementaridade com outras disciplinas. E isto acontece na ciência da informação por um lado, pela referência intrínseca de seu objeto a todos os outros

modos de produção de saberes, gerando constantemente novas treliças interdiscursivas e, por outro lado, pela natureza estratificada e poli-epistemológica dos fenômenos ou processos de informação. *Desde suas primeiras manifestações, apresentava-se, assim, a ciência da informação, como conjunto de saberes agregados por questões antes que por teorias* (GÓMEZ, 2000). (Grifo dos autores).

A classificação da produção científica desta área de conhecimento é um tema capaz de gerar diferentes taxonomias (ODDONE; GOMES, 2003). Apesar de tal dificuldade, Gomes (2003), a partir de uma síntese de diversos estudos produzidos no Brasil, dentre os quais Oliveira (1998; 1999) e Mueller e Pecegueiro (2001), concluiu que os assuntos mais pesquisados pela CI, no Brasil, estudam os seguintes temas: i) usuários, transferência e uso da informação e da biblioteca; ii) processamento e recuperação da informação (entrada, tratamento, armazenamento, recuperação e disseminação da informação) (GOMES, 2003, p. 17). Para simplificação, neste texto o primeiro tema é referenciado por meio da expressão usos da informação – ou USOS – e o segundo tema por meio da expressão recuperação da informação – ou RI. Nas pesquisas em CI, o primeiro tema – USOS – está ligado mais fortemente aos indivíduos, e o segundo – RI – está tradicionalmente ligado à recuperação da informação registrada, mais especificamente aos sistemas de informação.

Ao analisar a pesquisa em CI, Dias (2002) a apresenta como uma área do conhecimento que abarca uma série de especialidades ou subáreas, algumas consensualmente aceitas (biblioteconomia, por exemplo) e atribui a origem da utilização da expressão CI, nesse sentido genérico, ao fato de assim ser utilizada na tabela de áreas do conhecimento do Conselho Nacional do Desenvolvimento Científico e Tecnológico (CNPq). Nela, a CI se subdividiria em três grandes subáreas: teoria da informação, biblioteconomia e arquivologia. Segundo o autor, a teoria da informação se caracterizaria como a especialidade do campo que lida com a informação especializada. Ele tenta delimitar o campo a partir da definição das disciplinas básicas, a saber: 1) *organização da informação* – visa à organização das informações / documentos que vão dar sustentação a sistemas de informação e de recuperação da informação (SIRIs); 2) *Busca em Sistemas de Informação* – estudo das funções de busca e da formulação das estratégias de busca. No caso da informação especializada, trata-se da interação entre o conhecimento do usuário das complexas fontes de informação e de seus respectivos instrumentos de acesso.

Dessa forma, ele tenta definir a heurística negativa do programa de ensino em CI, afirmando “que os conteúdos que estejam fora das disciplinas básicas não podem jamais constituir o principal em um programa [de ensino], mas ter apenas um caráter complementar” (DIAS, 2002).

Pode-se, então, aceitar que as áreas de pesquisa e as disciplinas obrigatórias possam ser, de fato, agrupadas em duas grandes áreas: USOS e RI. Resta o problema de associá-las às linhas de pesquisa existentes nos programas de pós-graduação da área. Em discussão acerca dos desafios da CI, um dos autores deste artigo (MATHEUS, 2005) argumenta que os programas de pesquisa para a área deveriam agregar, simultaneamente, três abordagens complementares: i) pesquisas práticas e empíricas; ii) pesquisas teóricas, que devem interagir com as pesquisas empíricas por meio de temas definidos pelos programas de pesquisa; iii) pesquisas associadas à filosofia da

informação, que têm maior liberdade em relação a temas para debater filosófica e epistemologicamente a CI.

Com base no referencial exposto, pode-se tentar associar as linhas de pesquisa com essa proposta e com as duas grandes áreas. Os cursos de pós-graduação (*stricto sensu*) organizam suas linhas de pesquisa em três vertentes básicas: i) tratamento e uso da informação; ii) informação, cultura e sociedade; iii) gestão da informação e do conhecimento. As exceções são o Ibiect/UFF, que têm uma linha de pesquisa em *teoria, epistemologia, interdisciplinaridade em CI*, e a UFSC, com uma linha na área de *ensino e formação dos profissionais da informação*.

Tomando-se as três linhas comuns, pode-se dizer que elas contemplam os aspectos práticos e teóricos desejáveis nos programas de pesquisa e que podem ser analisadas segundo as duas grandes áreas que estruturam a produção na CI, conforme pode ser observado no quadro 1.

#### QUADRO 1

Linhas de pesquisa da ciência da informação e sua relação com as áreas e as características dos programas de pesquisa

Programa de pesquisa	Nome	Conteúdo	Usuários, transferência e uso da informação e da biblioteca (USOS)	Processamento e recuperação da informação (RI)
Pesquisas práticas e empíricas; pesquisas teóricas	Tratamento e uso da informação	Organização do conhecimento, representação da informação, indexação, índice, classificação do conhecimento, comunicação científica, colégios invisíveis, análise de citações.	Usuários de sistema de informação e especialistas em informação, com ênfase em bibliotecas e conhecimento científico.	Classificação, indexação e recuperação de informação; automação.
Pesquisas práticas e empíricas; pesquisas teóricas	Informação, cultura e sociedade	Informação, sociedade contemporânea, redes sociais, informação, estado e sociedade civil; informação, espaço e práticas sociais; informação, cultura e tecnologia; informação e educação práticas informacionais e o desenvolvimento da cidadania; análise dos impactos das tecnologias da informação e da comunicação na sociedade em geral.	Cidadãos em geral e grupos sociais específicos.	Sistemas públicos de informação, em qualquer meio; governo eletrônico; redes de informação em comunidades ou grupos.
Pesquisas práticas e empíricas; pesquisas teóricas	Gestão da informação e do conhecimento	Acesso, disseminação e uso da informação em organizações, fontes e serviços de informação para negócios, gestão da informação e do conhecimento tecnológicos, aprendizagem organizacional, empreendedorismo, gestão estratégica, inteligência empresarial, monitoração ambiental, tecnologias da informação para a gestão, a informação no processo decisório das organizações.	Usuários de sistema de informação e especialistas em informação, com ênfase em organizações.	Classificação, indexação e recuperação de informação; automação. Sistemas de gestão e apoio à tomada de decisão.

Fonte: Ancib.

Conforme a tabela, pode-se dizer que as linhas de pesquisa oferecem as características básicas de subáreas de produção do conhecimento para serem tomadas como unidades de análise sobre a ciência da informação. Além disso, pelas diferenças encontradas nos termos que descrevem o conteúdo, pode-se entender que cada uma delas tem uma característica interdisciplinar distinta, em termos de influência de outras áreas de conhecimento. No caso de 'Tratamento e uso da informação', as influências seriam, principalmente, das áreas de biblioteconomia, arquivologia, lingüística e ciência da computação. Na linha 'Informação, cultura e sociedade', as maiores influências seriam das ciências humanas, em especial a sociologia e a ciência política (não esquecendo que essa influência se estende, inclusive, aos métodos de pesquisa). Finalmente, na linha de 'Gestão da informação e do conhecimento', estariam relacionadas as ciências sociais aplicadas, especialmente administração e economia, além da ciência da computação.

Em termos da convivência das linhas de pesquisa em um mesmo programa, seria esperado que as diferentes interdisciplinaridades permitissem diferentes abordagens para o mesmo problema de pesquisa e que as diferentes abordagens teóricas fossem se aproximando na construção do campo da ciência da informação. Essa conjectura pode ser abordada, do ponto de vista de sua verificação, de vários ângulos e com diferentes instrumentais. Na verdade, não se pode analisar um campo de conhecimento tomando-se apenas um desses ângulos, conforme já havia sido mencionado (ver, a esse respeito, Crane, 1972).

### ARS E AS REDES DE CO-AUTORIA

A análise de redes sociais (ARS ou SNA, da expressão em inglês *Social Network Analysis*\*) é uma abordagem oriunda da sociologia, da psicologia social e da antropologia (FREEMAN, 1996; WASSERMAN; FAUST, 1999). A análise de redes sociais interessa a pesquisadores de vários campos do conhecimento que, na tentativa de compreender o seu impacto sobre a vida social, deram origem a diversas metodologias de análise que têm como base as relações entre os indivíduos, em uma estrutura em forma de redes. As redes são sistemas compostos por 'nós' e conexões entre eles, que, nas ciências sociais, são representados por sujeitos sociais

---

\* Os termos originais em inglês são utilizados para que se evitem ambigüidades, uma vez que existem traduções diferentes para o mesmo termo em português.

(indivíduos, grupos, organizações etc.) conectados por algum tipo de relação. De forma genérica, pode-se estudar o sistema visando apenas a entender como ele se comporta e como as conexões influenciam esse comportamento, com aplicações na área de saúde pública (e.g., estudos epidemiológicos), de tecnologia da informação (e.g., a mesma idéia para os vírus de computador), da sociologia (e.g., os movimentos sociais), da economia (e.g., mercados e economias de rede) e da matemática aplicada (e.g., otimização de algoritmos) (WATTS, 1999).

O uso da ARS vem crescendo significativamente nos últimos 20 anos, conforme demonstrado por pesquisadores na área, a partir de pesquisas em base de dados de artigos científicos e em programas de pesquisa (OTTE; ROUSSEAU, 2002; BORGATTI; FOSTER, 2003). 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 utilizam a ARS e a publicação de inúmeros manuais sobre o tema.

A ARS não é uma proposta nova para a CI, embora não se possa afirmar que seja, como metodologia, uma ferramenta amplamente utilizada na área. A experiência internacional, muito mais ampla que a brasileira, mostra que alguns dos principais autores da área de ARS publicaram em revistas científicas da CI. Em sua pesquisa, Otte e Rousseau (2002) fazem uma rápida revisão de estudos na área, tomando como referência o início dos anos 70 e relacionando-os, principalmente, com redes de informação, redes de co-autoria, de pesquisadores e de citações. Depois, combinando a base de dados Library and Information Science Abstracts (Lisa) e da lista dos principais autores sobre o tema de ARS, pesquisaram sua relevância na área de CI. Dos 47 autores mais prolíficos (com seis ou mais artigos), 12 haviam escrito artigos presentes nessa base de dados (independentemente de serem os primeiros autores), mas não são pesquisadores da CI. Já no Brasil, ainda são poucos os estudos que utilizam essa metodologia, e as referências na área de CI são reduzidas, sendo que o trabalho de Regina Marteleto parece ser pioneiro (MARTELETO, 2001). Pesquisando-se a base de periódicos PERI da ECI/UFMG, com cerca de 7 mil artigos indexados nas áreas de biblioteconomia e ciência da informação, obtiveram-se 16 artigos para a pesquisa 'redes sociais' e quatro para 'análise de redes sociais'.

Mesmo para uma pesquisa rápida, os resultados indicam o baixo uso dessa ferramenta na CI\*.

Em contrapartida, muitas pesquisas que utilizam a ARS e que poderiam ser apontadas como tendo uma interseção com os programas na área de CI podem ser encontradas em publicações de outras áreas. Por exemplo, as ligações estudadas através da ARS dentro das organizações e das empresas, para identificar e analisar os fluxos de informação entre os atores e seus impactos na geração de conhecimento, ou as relações de autoridade e seu papel no fluxo de informações podem ser encontradas em revistas da área de administração e sociologia. Outros exemplos de estudos utilizando a ARS que podem, também, ser considerados como pertinentes à CI são a comunicações entre atores (nós) para a obtenção de informações vantajosas (GRANOVETTER, 1973; BURT, 2000; BURT, 1995; BORGATTI; CROOS, 2003); o envio de mensagens eletrônicas entre pessoas, como feito no trabalho sobre troca de mensagens eletrônicas (FREEMAN's EIES – *Electronic Information Exchange System* – study *apud* WASSERMAN; FAUST, 1999, p. 62) e as relações de autoridade formal ou de aconselhamento técnico em uma organização (KRACKHARDT; HANSON, 1993, MOLINA, 2000; KRACKHARDT, 1987); a análise de redes de empresas em *clusters* ou sistemas produtivos locais (SAXENIAN, 1996; MÁCIAS, 2002); o estudo de redes de pequenas e médias empresas (ROCHA, 2003), empreendedorismo e redes familiares (LIN *et al.*, 2001); e as redes entre grandes empresas e seus fornecedores e os fluxos de conhecimento embutidos nessas relações (CARLEIAL, 2001).

Há, entretanto, uma área na qual a ferramenta de ARS é mais utilizada, combinando seus resultados com outros métodos quantitativos já empregados na CI. Beneficiando-se da flexibilidade do conceito de ator, a análise de rede tem sido empregada como uma ferramenta adicional para os estudos nas áreas de bibliometria e infometria. Como até mesmo artigos científicos podem ser tratados como atores, neste caso as citações fazendo a ligação entre os atores (WASSERMAN; FAUST, 1999, p. 51), os conceitos de redes permitem a identificação de

grupos de pesquisadores e comunidades de prática, lideranças e autores principais, assim como a interpretação social das redes de citações bibliométricas (MAHLCK; PERSSON, 2000).

Uma rede de co-autoria é uma rede na qual os nós são os professores / pesquisadores, e há conexão entre eles sempre que partilham a autoria de um artigo. A visualização da rede, na forma de grafos, é considerada, pelos autores da área, mais intuitiva do que a visualização na forma de matrizes, embora os dados coletados sejam, normalmente, apresentados dessa forma.

Na nomenclatura de grafos usada na ARS, trata-se de uma rede não-direcional, isto é, as ligações entre nós independem de sua origem. Assim:

A rede de co-autoria vai ser representada pelo grafo  $G(N,L)$ ;

O conjunto  $N$  de pesquisadores serão os nós,  $N = \{n_1, n_2, \dots, n_g\}$ ; sendo  $g$  o número de nós;

As relações de co-autoria pelas *linhas*,  $L = \{l_1, l_2, \dots, l_l\}$ , ou seja, as relações de co-autoria são dadas por pares de pesquisadores conectados por uma linha;

$\exists l_k = (n_i, n_j)$  se os pesquisadores  $n_i$  e  $n_j$  possuem uma relação de co-autoria, para  $i \neq j$ ;

$\neg \exists l_k = (n_i, n_j)$ , para  $i = j$ ;

$l_k = l_q = (n_i, n_j) = (n_j, n_i)$ , uma vez que as relações são do tipo não-direcionais.

Em termos matriciais, a matriz  $X$ , denominada matriz de adjacência, é a matriz social ou sociomatrix (do inglês *sociomatrix*), na qual, nas linhas (i) e nas colunas (j), estão dispostos os autores da rede de co-autoria estudada. Trata-se de uma matriz quadrada, uma vez que os mesmos autores aparecem nas linhas e nas colunas sociais (os valores i e j são iguais), e simétrica em relação à diagonal principal, isto é, os laços aqui estudados são recíprocos. Não existem os autolaços, isto é, os autores não têm laços consigo mesmos\*. Assim:

Seja  $X$  a matriz de co-autoria, tal que:

$x_{ij} = 1$ , se existe uma relação de co-autoria; para  $i \neq j$ ;

\* Nesse caso, seria um *loop*, isto é um autolaço ou autoligação. Eles não são possíveis em muitas modelagens (por exemplo, não é possível estabelecer um laço de amizade consigo próprio).

\* A base de periódicos PERI contém artigos de periódicos e trabalhos publicados em anais de eventos técnico-científicos, refletindo a literatura nacional nas áreas de biblioteconomia, ciência da informação, arquivística e outras interdisciplinares. A pesquisa, realizada em 10/03/2005 e com o uso do *Microis*, obteve 16 artigos para a pesquisa 'redes\*sociais' e quatro artigos para 'análise\*redes\*sociais' (pesquisa *booleana* com o operador '\*', que equivale à interseção ou 'and') para cerca de 7 mil artigos indexados.

$x_{ij} = 0$ , se não existe uma relação de co-autoria; para  $i \neq j$ ;

$x_{ij} = 0$ , para  $i = j$ , ou seja,  $x_{ij} = 0^*$ ;

$x_{ij} = x_{ji}$ , uma vez que as relações são do tipo não direcionais.

Uma segunda matriz pode, também, informar o número de artigos publicados em co-autoria, em vez apenas da existência de sua relação. Dessa forma:

$x_{ij} = x_{ji} = q$ , sendo  $q$  o número de artigos publicados em co-autoria;

$x_{ij} = x_{ji} = 0$ , se não existe uma relação de co-autoria.

A primeira matriz é denominada matriz 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. As redes de co-autoria (A é co-autor de um artigo com B) têm sido estudadas, da mesma forma que as redes de citações (B cita o autor C) e co-citações (B e C são co-citados por D), para se entender como funcionam as comunidades científicas, de forma a se analisar a construção dos colégios invisíveis e das comunidades de prática. A metodologia de análise de redes permite analisar e representar graficamente essas comunidades e testar algumas hipóteses sobre o comportamento de colaboração dos pesquisadores, agregando, ainda, informações sobre os seus atributos, tais como área de pesquisa e de formação, tempo de formação, departamento, instituição a que pertencem, sexo etc. As pesquisas fornecem aos formuladores de políticas informações sobre a topografia e a morfologia das diferentes áreas de conhecimento, sobre a existência ou não de subgrupos fechados de pesquisa, sobre as relações entre pesquisadores de várias universidades e, mesmo, países. Entender como as conexões se desenvolvem, o que facilita e o que emperra os fluxos de informações e conhecimentos são problemas de pesquisa da CI. Os autores estudados tomam o cuidado de salientar que o uso da ARS e de métodos estatísticos não eliminam a necessidade de uma análise qualitativa aprofundada.

Nessa linha, tomando por base as redes de co-autoria de grande porte (inclusive internacional, em uma base de 385 autores), Kretschmer (2004) usa a ARS e as demais

\* Os valores da diagonal principal são nulos. Se os valores numéricos nas células da diagonal principal forem não nulos, eles indicam a presença de um autolaço. Como foi visto, os autolaços são denominados *loops* na nomenclatura de grafos.

informações bibliométricas (produção e produtividade), para definir os atributos dos autores e analisar a posição na rede daqueles com os mesmos atributos. Primeiramente, ele analisa a estratificação social dos cientistas, a partir da hipótese de que cientistas com os mesmos atributos têm maior frequência de citações entre si (*birds of a feather flock together\**), isto é, de que a rede não seria uniforme, mas formada por agrupamentos (*clusters*). Usando-se o conceito de distância geodésica\*\*, o autor testa, basicamente, três hipóteses: i) se existe uma conexão entre a estrutura dos *clusters* e a produtividade dos cientistas; ii) se existe uma relação entre a distância geodésica e a produtividade; iii) se a estratificação social é maior quando são menores as distâncias geodésicas. O autor chama a atenção para a importância dos resultados desse tipo de pesquisa como insumo para a política científica. Suas preocupações apontam para novas linhas de pesquisa, sobre a comparação das redes ao longo do tempo e sobre a possibilidade de se identificarem novas áreas do conhecimento quando elas ainda são embrionárias, assim como as redes de colaboradores e pesquisadores existentes na Web. Também na linha de análise de redes e estrutura social dos colaboradores, White *et al.* (2004) apontam para a necessidade de mais pesquisas na área, já que a resposta à pergunta do título do artigo, se as citações refletem a estrutura social (*Does citation reflect social structure?*), não pôde ser respondida de forma categórica. Eles reforçam a necessidade de se compreenderem os aspectos sociológicos e de poder subjacentes às redes de pesquisadores.

Sobre a dinâmica dos grupos de pesquisadores, Yoshikane e Kageura (2004) usam a ARS para estudar a estrutura de cooperação entre pesquisadores japoneses de quatro áreas: duas de engenharia (elétrica e processamento de dados) e duas na área de química (bioquímica e polímeros). Eles constatam o crescimento relativo do número de artigos produzidos em co-autoria, assim como no número de autores por artigo, fruto da crescente interdisciplinaridade das áreas, com impactos significativos sobre as redes de cooperação. Concluíram que, embora seja um fenômeno geral, as características da pesquisa em cada área (questões operacionais, tradição

\* Literalmente, pássaros da mesma plumagem voam juntos. Ditado inglês para mostrar que pessoas com as mesmas características ou interesses tendem a passar muito tempo juntas. Originalmente, era usado de forma negativa, para tratar de características não aprovadas por quem usa o ditado.

\*\* A distância geodésica,  $d(n_i, n_j)$ , entre um par de nós, é o número de laços ou ligações que existe no caminho mais curto entre eles, sendo que, caso não exista tal caminho, a distância pode ser considerada indefinida ou infinita.

de trabalho em grandes ou pequenos grupos etc.) ainda determinam a estrutura de rede. Recomendam, como resultado adicional, que a dinâmica das redes deve ser acompanhada pelos responsáveis pela política científica.

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, tais como MEDLINE (pesquisa biomédica, com, aproximadamente, 2,2 milhões de artigos e 1,5 milhões de autores), *Los Alamos e e-Print Archive* (física teórica, com cerca de 98 mil artigos e 53 mil autores), SPIRES (física experimental de altas energias, 67 e 57 mil, autores e artigos, respectivamente) e NCSTRL (ciência da computação, 13 mil e 12 mil, respectivamente). Os resultados evidenciam as redes do tipo mundo pequeno (*small-world*); ou seja, dois cientistas escolhidos aleatoriamente estão separados por um pequeno número de passos (a distância geodésica é curta, cerca de cinco ou seis passos), a presença de *clusters* (em todas as bases existe um componente ou subconjunto com cerca de 50% a 80% dos autores, dependendo da área, ou seja, a maioria se conecta entre si por meio de autores intermediários e dois cientistas têm 30% ou mais de probabilidade de colaborarem entre si, se ambos já colaboraram como um terceiro cientista, isto é, formam uma tríade). Constatam-se, ainda, diferenças nas diversas áreas pesquisadas, como as que existem entre o tamanho das redes de colaboração: são muito maiores na área de física experimental de altas energias; como o fato de a área biomédica mostrar menor tendência à formação de tríades e, portanto, de *clusters* (NEWMAN, 2001). Uma conclusão importante, especialmente para informar aos formuladores de política científica, é que a ciência funciona bem quando a comunidade de pesquisadores é densamente conectada.

Aplicando a metodologia de redes para estudar a formação de comunidades em torno de cientistas e pesquisadores, Molina, Muñoz e Domenech (2002) estudaram a estrutura de co-autorias a partir de três pesquisadores selecionados: um matemático, um médico oncologista e um biólogo molecular. Os autores assinalam que as redes de co-autoria representam uma abordagem para estudar as comunidades científicas e seus colégios invisíveis e devem ser acopladas a outras (por exemplo, o número de teses orientadas, participação em congressos, participação em grupos de pesquisa etc.), mas possuem grandes vantagens, tanto no acesso às bases de dados, quanto pela facilidade no seu manuseio e na visualização dos resultados. Foi usada a metodologia de amostragem

do tipo bola de neve\*, isto é, a partir dos três autores iniciais (estágio zero), foram selecionados outros pesquisadores com os quais eles possuem uma relação de co-autoria (primeiro estágio), em seguida repete-se o processo para esses pesquisadores (segundo estágio) e assim sucessivamente até que não haja novas indicações. As redes resultantes, denominadas na literatura redes do tipo egocêntrica (cada um dos três pesquisadores originais é o *ego* de sua rede, e os demais são os seus *alteres* (WASSERMAN; FAUST, 1999, p. 35), permitem uma aproximação da rede de influências de cada pesquisador, além de se identificarem características próprias de cada área de pesquisa. Foram usadas várias medidas de centralidade e medidas de poder, conforme a metodologia de ARS\*\*. A justificativa para a utilização de artigos (e não livros ou outras referências) é que, nas áreas pesquisadas, os estudos de cientometria mostram que as referências a artigos representam 80% do total das referências (MOLINA; MUÑOZ; DOMENECH, 2002, p.8).

O conceito de mapas sociobibliométricos (MAHLCK; PERSSON, 2000) reforça a complementaridade das metodologias de ARS e bibliometria. Aplicando a ARS em dois departamentos da mesma área (biologia), mas de duas diferentes universidades suecas, eles analisam as redes de co-autoria e de citações. Os autores testam basicamente duas hipóteses: a primeira sobre a posição dos pesquisadores mais produtivos na rede e na hierarquia dos departamentos; a segunda sobre a integração de diferentes grupos de pesquisa. A base de dados sobre as co-autorias foi obtida eletronicamente\*\*\* e continha todos os artigos produzidos pelos pesquisadores. Os dados sobre os atributos dos pesquisadores (sexo, ano de formação, sua posição no departamento, trocas de departamento etc.) foram coletados em pesquisas diretas,

\* A amostragem via bola de neve (do inglês *snowball sampling*) é uma técnica realista e adaptativa, na qual os atores que irão fazer parte da pesquisa são indicados diretamente pelos próprios pesquisados. De acordo com tal técnica, pergunta-se a um predeterminado grupo de atores ('zona de primeira ordem' (Goodman, 1961, *apud* WASSERMAN; FAUST, 1999, p. 34) ou 'primeiro estágio' (ROTHENBERG, 1995, p. 105)) com quem ele tem laços, resposta que serve como indicação do próximo grupo de atores na rede a ser pesquisado ("segundo estágio", ou "zona de segunda ordem"). A pesquisa prossegue até que não sejam indicados novos atores. Tal técnica de amostragem parece ter suas origens no trabalho de Goodman (1961), citada por Rothenberg, (1995, p. 104). No estudo citado, em vez de se consultarem diretamente os autores, utiliza-se uma base de dados com as informações de co-autoria.

\*\* Para mais detalhes sobre essas medidas, ver Wasserman e Faust.

\*\*\* CD-ROM da *Science Citation Index*, para o período 1986-1996.

assim como os dados sobre a estrutura dos grupos de pesquisa. Sobre a produtividade dos autores, eles salientam que muitos artigos apresentam vários autores e que há uma tendência de aumento no número de autores, especialmente nas áreas de ciências da vida e ciências naturais (*life and natural science*). No entanto, observa-se a presença de poucos autores com a grande produção respondendo pela maior parte da produção do departamento\*. Finalmente, ele observa que diferentes grupos de pesquisa tendem a se organizar em *clusters*, isto é, com a maior parte da produção tendo origem entre pesquisadores do mesmo *cluster*, com pouca colaboração entre os diferentes grupos de pesquisa. Nesse caso, embora não explicado de forma explícita no texto, os grupos de pesquisa seriam concorrentes, mas não interdisciplinares.

Resumidamente, podem ser destacadas as seguintes contribuições: i) observa-se o crescimento no número de autores por artigo; ii) essa tendência é mais estudada nas ciências naturais e da vida; iii) mesmo em áreas com características muito distintas entre si, em termos de funcionamento de seus programas de pesquisa, das redes de colaboração e da organização operacional de seus laboratórios, essa tendência se verifica, representando a crescente necessidade de colaboração de pesquisadores com diferentes especializações; iv) esse fato estaria relacionado com a crescente interdisciplinaridade das pesquisas; v) observa-se uma lei de potência nas redes, isto é, poucos pesquisadores muito produtivos em meio ao conjunto geral de muitos autores com poucas publicações; vi) há uma tendência de organização das redes em torno desses pesquisadores / professores mais produtivos; vii) as organizações internas das instituições de pesquisa (no caso aqui mencionado, os departamentos) podem criar grupos de pesquisa que funcionam como *clusters*, isolados entre si, embora possam ser, todos eles, interdisciplinares.

## ESTUDO DA REDE DE CO-AUTORIA ENTRE OS PROFESSORES DO PPGCI/UFMG

Conforme mencionado, a ciência da informação é um campo do conhecimento relativamente jovem, e caracterizado, pela maior parte dos autores da área, como interdisciplinar. Essas características se manifestam, também, na estruturação dos cursos de pós-graduação e

em suas linhas de pesquisa. As influências de outras áreas do conhecimento (biblioteconomia e arquivologia, ciência da computação, administração, sociologia, comunicação, para relacionar apenas as mais citadas), na estruturação de tais cursos, devem também manifestar-se na alocação dos professores. Assim, o pressuposto é o de que estes tentem se alocar nas linhas de pesquisa para as quais possuem vantagem em termos de capital humano e intelectual.

O objetivo desta seção é analisar a rede de co-autoria entre 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. Esses resultados permitirão embasar, conforme apresentado pela literatura analisada, as discussões sobre a interdisciplinaridade do programa estudado.

O primeiro objetivo está associado ao pressuposto de que uma rede densa de artigos em co-autoria indica a execução de um programa de pesquisa vigoroso, com resultados obtidos a partir da colaboração dos professores. O segundo objetivo, por sua vez, relaciona-se com o pressuposto de que a colaboração entre os professores de diferentes linhas indicaria uma relação próxima entre as diversas disciplinas que influenciam a área, portanto, a consolidação de pesquisas interdisciplinares.

Dessa forma, a consolidação de um programa de pesquisa da CI a partir das suas características interdisciplinares poderia ser observada na crescente colaboração entre os professores das diversas linhas de pesquisa – a busca de um conhecimento comum, conceitos, definições e metodologias. Não se deve perder de vista que há uma tendência contraditória entre a colaboração entre os professores das diferentes linhas (formação de redes) e a formação de *clusters* fechados em cada uma delas. De qualquer forma, se for verificada a mesma distribuição já observada em outras áreas e países, de que há uma tendência de organização das redes em torno de poucos professores mais produtivos, ela pode ser observada apenas para o conjunto das redes, se a colaboração for ampla e entre as linhas, ou dentro de cada linha, se houver tendência para a formação de *clusters*.

## Justificativa da escolha do PPGCI/UFMG

O PPGCI/UFMG existe desde 1976 (mestrado), com o nível de doutorado implantado em 1997. Desde então,

---

\* Essa estrutura pode ser representada por uma lei de potência. A distribuição de *Lotka*, muito usada em estudos bibliométricos, evidencia, justamente, a existência de poucos autores com muitas publicações e de vários autores com poucas.

até dezembro de 2004, foram defendidas 189 dissertações de mestrado e 17 teses de doutorado. Sua avaliação na Capes é 5, maior entre os programas, posição dividida com a Universidade de Brasília (UnB). Atualmente conta com 19 professores, todos com titulação de doutor, maior contingente de docentes entre os programas existentes no Brasil\*. Dessa forma, pode ser considerado representativo na área da CI, no Brasil. O objetivo de se estudar o período de 1997 até o final de 2004 leva em conta a implantação do doutorado, que deve ser considerado um marco para os professores da casa.

Atualmente, o PPGCI/UFMG se divide em três linhas de pesquisa que serão usadas como referência neste estudo: informação, cultura e sociedade (ICS); organização e uso da informação (OUI); gestão da informação e do conhecimento (GIC).

### Metodologia de coleta e tratamento dos dados

A análise da rede de co-autoria entre os professores do PPGCI/UFMG, com artigos em publicações e em congressos, apresentados a partir de 1997, tem como base de dados o CNPq – Lattes. Foram seguidos os seguintes procedimentos:

1. inicialmente, obteve-se a lista de professores do Programa a partir do seu sítio na Internet;
2. a lista foi confirmada com a Secretaria do Programa, acrescentando-se novo professor, chegando-se, com isso, a 19 pesquisadores;
3. para atender ao período de estudo desejado, foram acrescentados os professores que se aposentaram no período considerado, mas que estavam em atividade durante pelo menos um dos anos em estudo. Essa informação foi obtida na área de recursos humanos da ECI/UFMG, e chegou-se a uma lista de 23 professores;
4. com a lista completa de professores, foram feitas consultas ao CNPq – Lattes, para a obtenção das

informações sobre as publicações (referência, co-autores, palavras-chave, grande área, área e subárea e as denominadas informações adicionais);

5. para os quatro professores que não tinham informação no CNPq – Lattes, apenas um apresentava as relações de co-autoria que se desejava estudar\*. Dessa forma, aqueles que não possuem informações e sobre os quais não se descobriu a publicação de artigos ou a participação em congressos foram eliminados da lista\*\*. As consultas foram feitas nos meses de março e abril de 2005;

6. além das informações mencionadas, foram acrescentadas aquelas referentes à formação dos professores (graduação, mestrado e doutorado), à área da sua tese e da sua dissertação, ao ano de entrada no Programa e a sua linha de pesquisa;

7. foram adotados alguns procedimentos de crítica: em primeiro lugar, houve a correção da grafia dos nomes e sua unificação para permitir a construção das relações de co-autoria; em segundo lugar, como as relações de co-autoria são do tipo não-direcionais, o mesmo artigo deveria estar mencionado por todos os seus autores. Caso isso não ocorresse, seria feita a inclusão na base de dados montada para a pesquisa, em termos da notação matricial adotada:

Seja  $X$  a matriz de co-autoria original, a matriz de co-autoria final foi construída como  $\chi = X \cup X'$ , no qual o símbolo' representa a transposição da matriz. Dessa forma, garante-se que  $x_{ij} = x_{ji}$

8. os artigos de cada professor / autor foram classificados da seguinte forma:

- a. individuais;
- b. em co-autoria com outros professores do PPGCI/UFMG;
- c. em co-autoria com outros professores da ECI/UFMG não listados acima (listagem obtida no sítio da ECI/UFMG);

\* No sítio da ECI/UFMG (<http://www.eci.ufmg.br>, acesso em 10/05/2005), estão listados 18 professores, assim como no sítio da Ancib – Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (<http://www.ancib.org.br/>). No entanto, a lista fornecida pela Secretaria do Programa informa 19 professores, com a inclusão da professora Gercina Ângela Borém de Oliveira Lima. O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) / Universidade Federal Fluminense (UFF) conta com 18 pesquisadores (<http://www.ibict.br/secao.php?cat=Pós-Graduação%20em%20CI/Pesquisadores>, acesso em 10/05/2005). O PPGCI/UFMG representa pouco mais de 16% dos pesquisadores alocados aos programas de pós-graduação.

\* Foi incluída a professora Isis Paim.

\*\* Essa decisão tem implicações para algumas medidas que podem ser realizadas com base na ARS, como, por exemplo, a densidade das redes. No entanto, a análise qualitativa do objeto de estudo indicou, como mais razoável, a opção adotada. Houve, de fato, uma mudança no comportamento dos professores em função das necessidades de publicação, e aqueles que se aposentaram no período e não possuem o registro no CNPq – Lattes não conviveram com essa alteração.

d. com alunos que foram orientados em suas pesquisas de mestrado e doutorado (com base na lista de dissertações e teses defendidas, obtida junto à Secretaria do Programa);

e. com outros colaboradores.

Em função do reduzido número de artigos classificados em (c), na análise os valores foram agregados ao total de outros colaboradores, categoria (e).

A escolha dessa base de dados está relacionada com a crescente preocupação dos professores de manter atualizadas suas produções científicas, uma vez que são avaliados por isso, tanto em sua vida acadêmica, quanto na concessão de financiamentos de pesquisa. Ainda assim, não se eliminam, totalmente, os riscos de omissão (por exemplo, todos os autores de um artigo não atualizarem suas informações) ou os erros de informação (referência ao co-autor por outro sobrenome ou a existência de homônimos).

### Análise dos resultados

Inicialmente, foram feitas as análises estatísticas mais tradicionais, para, em seguida, efetuar-se a análise da rede de co-autorias. A análise vai obedecer à seguinte estrutura: comparação do número de publicações nos períodos, antes e após 1997; produção por autor e sua distribuição, segundo a classificação adotada; participação de cada autor no total da produção do Programa; número médio de autores por artigo, número total de colaboradores.

Serão analisadas, para complementar as informações anteriores, a ligação entre a formação dos professores e a sua linha de pesquisa, assim como a classificação dos artigos publicados segundo as áreas e sua ligação com as linhas de pesquisa do programa.

Embora deva ser tomada com cautela a comparação entre o período que vai até 1996, inclusive, e o período de 1997 a 2004, uma vez que muitos professores não faziam ainda parte do Programa, o que se observa é um crescimento

TABELA 1  
Número de artigos, em revistas científicas e congressos, dos professores do PPGCI, por período

Linha de Pesquisa	Individuais	Com professores do PPGCI	Com alunos/ orientandos	Com outros	Total
<b>Até 1996</b>					
Informação, Cultura e Sociedade (ICS)	29	8	2	27	66
Organização e Uso da Informação (OUI)	26	2	1	7	36
Gestão da Informação e do Conhecimento (GIC)	24	8	1	16	49
<b>Total</b>	<b>79</b>	<b>18</b>	<b>4</b>	<b>50</b>	<b>151</b>
<b>De 1997 até 2004</b>					
Informação, Cultura e Sociedade (ICS)	53	18	21	35	127
Organização e Uso da Informação (OUI)	51	15	14	27	107
Gestão da Informação e do Conhecimento (GIC)	23	42	84	34	183
<b>Total</b>	<b>127</b>	<b>75</b>	<b>119</b>	<b>96</b>	<b>417</b>

Fonte: CNPq – Lattes.

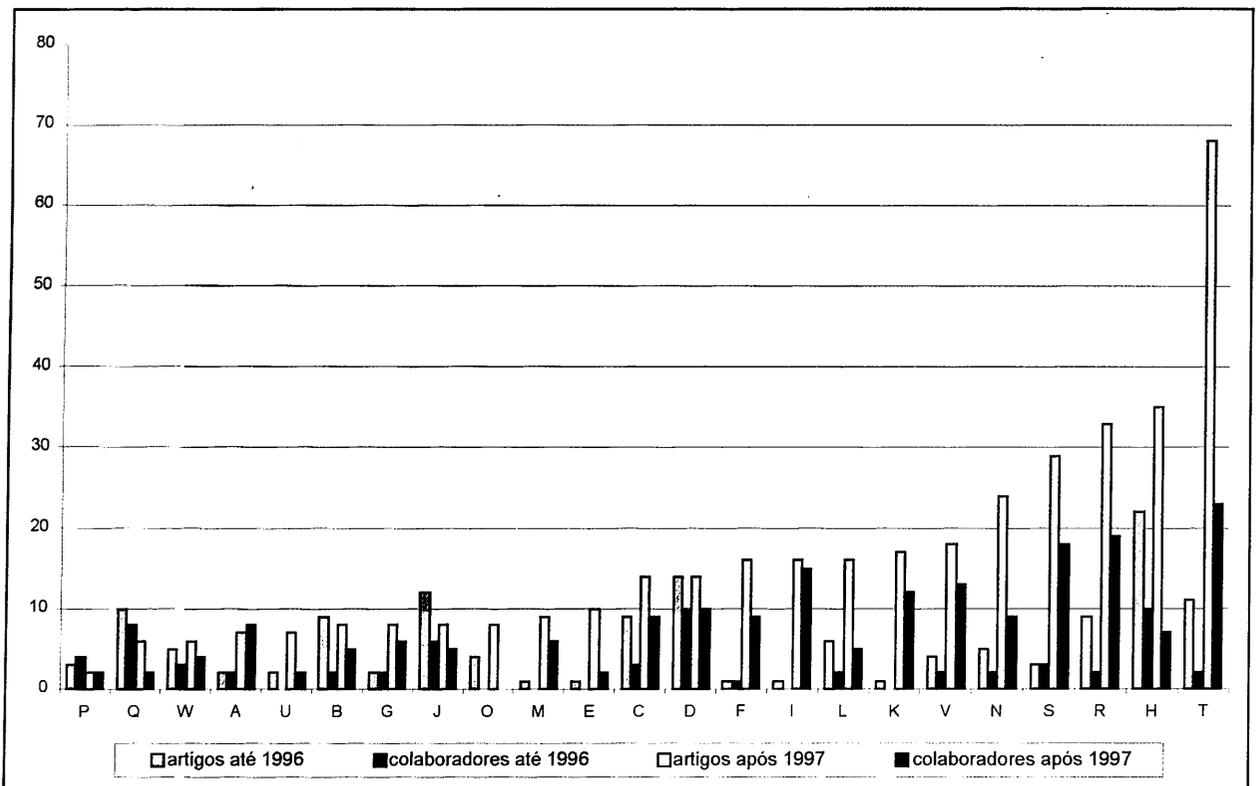
significativo no número de artigos publicados. Até 1997, haviam sido publicados 137 artigos, sendo que 57,6% (79 artigos) deles eram individuais e apenas 11,1% (18) em co-autoria com outros professores do Programa. No período seguinte, o total saltou para 379 artigos, sendo aproximadamente 33,5% individuais (127) e 19,8% (75) em co-autoria com outros professores do Programa. A maior mudança, entretanto, foi no número de publicações em colaboração com alunos do programa (orientandos): o total saltou de 4 para 116, representando, respectivamente, 2,7 e 28% do total de publicações em cada período\* (tabela 1).

\* O total apresentado na tabela supera o total de artigos, uma vez que o seu objetivo é destacar a cooperação na forma de co-autoria. Dessa forma, se um artigo foi escrito por dois professores do programa e um aluno (três autores), ele vai aparecer tanto na lista de artigos com outros professores, quanto na lista de alunos do programa. Dessa forma, o total de artigos para o período após 1997 é 379 e, considerando-se as duplas contagens, de 418. A diferença de 39 é distribuída uniformemente, e as duas distribuições têm uma correlação (Pearson) de 0,9939. Com isso, as análises sobre as distribuições dos artigos entre as áreas não são significativamente afetadas.

O aumento na produção se deveu à contribuição da maioria dos professores, pois somente para quatro deles se observou queda ou manutenção do número de artigos. O número de colaboradores também aumentou significativamente, em taxa maior que a do total de artigos. O número total de artigos cresceu 2,5 vezes, o número de artigos com professores do programa cresceu pouco mais de quatro vezes, enquanto aqueles com alunos orientandos foi multiplicado por 30. Embora haja outros fatores que impulsionaram o aumento na produção individual, há forte correlação entre o número de colaboradores e o número de artigos, de 0,8268 para o período até 1996 e de 0,7942 para o período de 1997 até 2004. O gráfico 1 mostra esses resultados de forma bastante clara (O gráfico foi ordenado a partir da produção observada no segundo período, enquanto as letras apresentadas no eixo das abscissas (horizontal) representam os professores).

Enquanto no período inicial, a maior parte da produção por professor se apresentasse na forma de artigos individuais, no período final considerado, há uma distribuição diferenciada entre as diferentes modalidades. Até 1996, 17 dos 23 professores tinham 50% ou mais de sua produção representada por artigos individuais, somente oito apresentavam artigos em co-autoria com outros professores do Programa e apenas três com alunos/orientandos (quatro não haviam escrito artigos individuais). Após 1997, 13 docentes apresentavam a maior parte de sua produção na forma de artigos individuais. Em compensação, somente seis não apresentavam artigos em co-autoria com os colegas docentes do Programa, sete não produziram em colaboração com os seus orinetandos do PPGCI e apenas um não produziu nenhum artigo individual.

GRÁFICO 1  
Produção por professor e por período



A-Alcenir Soares dos Reis; B-Ana Maria de Rezende Cabral; C-Ana Maria Pereira Cardoso; D-Bernadete Santos Campello; E-Helena Maria Tarchi Crivellari; F-Ligia Maria Moreira Dumont; G-Maria Eugenia Albino Andrade; H-Regina Maria Marteleto; I-Beatriz Valadares Cendón; J-Eduardo José Wense Dias; K-Gercina Ângela Borém de Oliveira Lima; L-Lidia Alvarenga; M-Madalena Martins Lopes Naves; N-Maria Aparecida Moura; O-Marlene de Oliveira; P-Vilma Moreira dos Santos; Q-Isis Paim; R-Jorge Tadeu de Ramos Neves; S-Marcello Peixoto Bax; T-Marta Araujo Tavares Ferreira; U-Marta Pinheiro Aun; V-Monica Erichsen Nassif Borges; W-Ricardo Rodrigues Barbosa.

Fonte: CNPq – Lattes.

Analisando-se mais detidamente a produção do segundo período, aquele definido como central para o objetivo do texto, podem-se obter mais informações sobre a produção dos professores.

Em termos da produção e da produtividade dos autores, a literatura consultada evidencia a grande possibilidade de existência de um grupo de professores / pesquisadores muito produtivos em torno do qual se organizam as redes. No caso estudado, o universo é bastante reduzido, e o professor é chamado a publicar artigos e a participar de congressos por conta da avaliação ao qual é submetido e para concorrer ao financiamento de projetos de pesquisa, dentre outras coisas. Seria de se esperar uma produção mais bem distribuída entre os seus membros.

Para avaliar essa distribuição, foram realizadas duas análises. A primeira com base na distribuição de Lotka\*, a segunda, utilizando-se Índice de Gini\*\*.

A distribuição de Lotka pode ser descrita como uma lei de potência do tipo:

$$f(k) = C/k^\beta, k = 1, 2, \dots$$

Os parâmetros  $C$  e  $\beta$  não são independentes e se relacionam (ROUSSEAU; ROUSSEAU, 2000)\*\*\*. Reescrevendo-se a equação como uma distribuição estatística (de forma que a soma de todos os  $n$  seja igual a 1), de tal forma que:

\* Alfred Lotka (1880 – 1949), químico, ecologista, demógrafo e matemático, focou conhecido na área de bibliometria em um artigo pouco representativo de sua obra. Ele mostrou que o número de autores com  $n$  publicações em uma bibliografia pode ser descrito como uma lei de potência da forma  $C/k^\beta$ . Ele demonstrou que  $\beta$  tende para 2 e, nesse caso,  $C$  seria igual a  $6/(\pi)^2$ , ou seja, aproximadamente 0,61. Assim, se uma bibliografia pode ser descrita por essa lei de potência, 61% dos autores teriam contribuído com apenas uma publicação. Ver <http://users.pandora.be/ronald.rousseau/html/lotka.html>.

\*\* Corrado Gini (1884 – 1965), estatístico, demógrafo e sociólogo italiano, desenvolveu o coeficiente, que recebeu o seu nome, para mensurar a desigualdade de renda em uma sociedade. No entanto, pode ser utilizado para medir qualquer tipo de distribuição desigual. Ver [http://en.wikipedia.org/wiki/Gini\\_coefficient](http://en.wikipedia.org/wiki/Gini_coefficient).

\*\*\* Há autores que criticam essa formulação, pois se trataria de um caso particular das leis de potência, nas quais, em geral, os parâmetros são independentes. Um dos comentaristas do artigo mencionado e do software de cálculo da distribuição de Lotka que ele apresenta, Eric Archambault (<http://cybermetrics.cindoc.csic.es/pruebas/v411c1.htm>), considera útil o seu desenvolvimento, mas mostra outras maneiras de se calcular regressões de distribuições hiperbólicas (lei de potência), utilizando-se a planilha Excel. Os resultados para o trabalho aqui apresentado são os mesmos, isto é, as curvas estimadas para as distribuições estudadas são, praticamente, idênticas. Dessa forma, como o aplicativo desenvolvido é de fácil acesso e utilização e calcula o teste *Kolmogorov-Smirnov*, seus resultados foram aqueles utilizados nessa pesquisa.

$$\sum_{k=1}^{\infty} C/k^\beta = 1$$

pode-se estimar mais facilmente os seus parâmetros. Para se testar a adequação das estimativas dos parâmetros em uma dada distribuição, utiliza-se o teste de *Kolmogorov-Smirnov*, calculado com base nos desvios absolutos entre os valores das funções de distribuição observados e as distribuições teóricas (ROUSSEAU, B.; ROUSSEAU, R. LOTKA, 2000).

O coeficiente de Gini é um número entre 0 (perfeita igualdade) e 1 (perfeita desigualdade), enquanto o índice de Gini é o coeficiente expresso em porcentagem, sendo igual ao coeficiente multiplicado por 100. No presente trabalho, se todos os professores contribuíssem com a mesma quantidade de artigos, o coeficiente calculado seria igual a 0. Se todos os professores não publicassem, exceto um deles, o coeficiente calculado seria igual a 1. O coeficiente pode ser estimado a partir da seguinte fórmula de cálculo (Fórmula de *Brown*):

$$G = \left| 1 - \sum_{k=0}^{k=n-1} ((X_{k+1} - X_k)(Y_{k+1} + Y_k)) \right|$$

Na qual:

$G$  = coeficiente de Gini;

$X$  = proporção acumulada da população;

$Y$  = proporção acumulada da variável estudada (no caso, publicações);

$k$  = faixas da população e da variável estudada,  $k = 1, 2, \dots$

Os resultados estimados encontram-se na tabela 2. O que se pode observar é que as duas informações são consistentes. O teste aplicado à distribuição indica que, para o total de publicações\* e para as publicações individuais, não há evidências de que a distribuição dos artigos siga a lei de potência. Para esses itens, o coeficiente de Gini apresentou valores abaixo de 0,40. No entanto, para o que se pretende analisar – as redes de co-autoria –, os resultados confirmam uma concentração dos trabalhos em torno de poucos pesquisadores mais

\* Foi usado o total de artigos, exclusive a dupla contagem. No entanto, os resultados da distribuição seriam os mesmo se fossem usados os resultados com a dupla contagem. Ver nota 14.

produtivos, para os três tipos de colaboração analisados (com outros professores do Programa, com alunos / orientandos do Programa e os demais). Deve-se, entretanto, levar em conta que há forte concentração da produção conjunta em torno de poucos professores e que as redes de colaboração entre professores é relativamente esparsa. Apenas dois professores elaboraram 14 artigos em conjunto, ou seja, respondem por 37% do total de artigos. Para uma análise mais aprofundada das relações de co-autoria entre os professores do Programa e sua representação visual, utiliza-se a metodologia de análise de redes sociais (ARS).

Antes de se passar à ARS, é necessário associar a localização do professor do programa nas linhas de pesquisa, assim como sua produção. A alocação do professor em determinada linha pode levar em conta critérios não acadêmicos, como, por exemplo, a existência de vagas. Pode, ainda, não ser dinâmico o suficiente para acompanhar as novas preocupações de certo pesquisador. Para se determinar a relação entre a alocação do professor e sua produção, foi feita uma associação entre as palavras-chave de cada artigo com as linhas de pesquisa. Foram criadas, ainda, duas novas áreas adicionais para acomodar a produção: a relacionada com a 'epistemologia, ensino e pesquisa em ciência da informação, estudos e análises sobre o profissional da informação' e o agrupamento denominado 'outros'. Conforme se observa na tabela 3, a seguir, há forte associação entre a localização do professor e sua produção. A parte 'A' da tabela mostra que a maior parte dos artigos está classificada na mesma área do seu autor. Se forem desconsiderados os artigos classificados nas linhas 'D' e 'E', os resultados subiriam para 70%, 79% e 85% para as linhas ICS, OUI e GIC, respectivamente.

Mudando-se o enfoque da análise para os artigos produzidos, observa-se, na parte 'B' da tabela 3, que a maior parte deles é produzida nas próprias linhas. Assim, 88,5% dos artigos da linha ICS são produzidos por professores que pertencem a esta mesma linha, valores que chegam a 61,7%, na linha OUI, e 86,2%, na linha GIC. O restante representaria a contribuição interlinhas de pesquisa. Deve-se levar em conta que oito professores

TABELA 2

Parâmetros da distribuição de Lotka e coeficiente de Gini, por tipo de colaboração nas publicações, período 1997-2004

Descrição	Total	Tipos de publicação				
		Co-autoria com outros professores do PPGCI	Artigos Individuais	Colaboração com alunos / orientandos	Colaboração com outros pesquisadores	Todos, exceto os artigos individuais
Distribuição de Lotka						
C	0,2680	0,4936	0,3775	0,4038	0,4031	0,3129
Beta	1,3192	1,6041	1,4909	1,5371	1,5359	1,3860
r Crítico 1% (teste mogorov-Smirnov)	0,3475	0,3953	0,3475	0,4075	0,3953	0,3475
D (Valor estimado pelo teste)	0,5133	0,3808	0,3755	0,2163	0,2854	0,2963
Resultado do teste	Falhou	OK	Falhou	OK	OK	OK
Coeficiente de Gini						
G	0,3976	0,5728	0,3696	0,7103	0,5879	0,5334

produzem majoritariamente fora de suas linhas. Se forem desconsiderados os artigos classificados nas linhas 'D' e 'E', restariam quatro professores produzindo nessas condições. A análise da produção em co-autoria com a metodologia de ARS traz mais elementos para a compreensão da organização do Programa.

### A análise da rede de co-autoria com base na ARS

A ARS é feita a partir dos dados das duas matrizes mencionadas na seção 3: a 'Matriz binária', com informações que indicam apenas a existência de laços de co-autorias e a 'Matriz valorada' com as informações sobre o número de colaborações para o período de 1997 a 2004. A rede de co-autoria estudada contém artigos com dois ou mais autores. Assim, as informações analisadas referem-se ao número de colaborações, e não de artigos (um artigo com a participação dos professores A, B e C vai ser representado na forma de três colaborações: A e B; A e C; B e C).

Primeiramente, será analisada a estrutura da rede obtida e, em seguida, serão feitas análises a partir de algumas medidas de centralidade.

A representação da colaboração entre os professores pela Matriz Binária mostra reduzido número de colaboradores. O total de 23 professores (oito nas linhas de pesquisa ICS e OUI e sete na linha de pesquisa GIC)

TABELA 3

Distribuição da produção dos professores, segundo as linhas de pesquisa e a classificação dos artigos, período 1997-2004 (em percentual)

Linhas de pesquisa dos professores		Classificação dos Artigos					Total
		ICS	OUI	GIC	D	E	
		Informação, Cultura e Sociedade	Organização e Uso da Informação	Gestão da Informação e do Conhecimento	Epistemologia, Ensino e Pesquisa em CI etc.	Outras	
<b>Parte A - Distribuição dos artigos segundo as linhas de pesquisa dos professores</b>							
ICS	Informação, Cultura e Sociedade	48,2	13,4	7,1	10,7	20,5	100,0
OUI	Organização e Uso da Informação	1,0	50,0	12,0	22,0	15,0	100,0
GIC	Gestão da Informação e do Conhecimento	3,6	9,6	74,9	9,0	3,0	100,0
<b>Parte B - Distribuição dos artigos segundo a classificação dos artigos</b>							
ICS	Informação, Cultura e Sociedade	88,5	18,5	5,5	24,5	53,5	29,6
OUI	Organização e Uso da Informação	1,6	61,7	8,3	44,9	34,9	26,4
GIC	Gestão da Informação e do Conhecimento	9,8	19,8	86,2	30,6	11,6	44,1
<b>Total</b>		100,0	100,0	100,0	100,0	100,0	100,0

proporcionaria um potencial de 253 pares de co-autoria, mas apenas 28 estão presentes (densidade de 11%). A baixa densidade interna às três linhas de pesquisa, no entanto, é maior que aquela observada entre as linhas, indicando que a colaboração interna é maior que a externa (tabela 4).

A tabela 5 mostra a comparação da distribuição dos laços com a produção de artigos, isto é, entre os laços existentes e sua produtividade. Observa-se que, exceto para a linha de pesquisa ICS, a produção interna (intra) é mais relevante que a encontrada em colaboração com outras áreas (inter). Os resultados se apresentam da mesma forma, tanto para os laços de colaboração, quanto para a quantidade de artigos produzidos.

Para o conjunto da rede de co-autoria, observa-se, então, baixo nível de cooperação, tanto dentro quanto fora das linhas de pesquisa. Há maior interação interna à linha, quando se observam a produção absoluta e os resultados relativos, mas o número total de laços externos indica que as barreiras à colaboração entre as linhas não são significativamente diferentes das existentes dentro de uma mesma linha. Ou seja, as diferenças entre os laços internos e externos não seriam significativamente

TABELA 4  
Densidade da matriz de co-autoria, período 1997-2004 (em percentual em relação ao máximo)

Linhas de Pesquisa	(ICS)	(OUI)	(GIC)
Informação, Cultura e Sociedade (ICS)	14,3	6,3	8,9
Organização e Uso da Informação (OUI)	6,3	17,9	10,7
Gestão da Informação e do Conhecimento (GIC)	8,9	10,7	19,1

TABELA 5  
Distribuição dos laços e da produção de artigos, período 1997-2004 (em percentual em relação ao total)

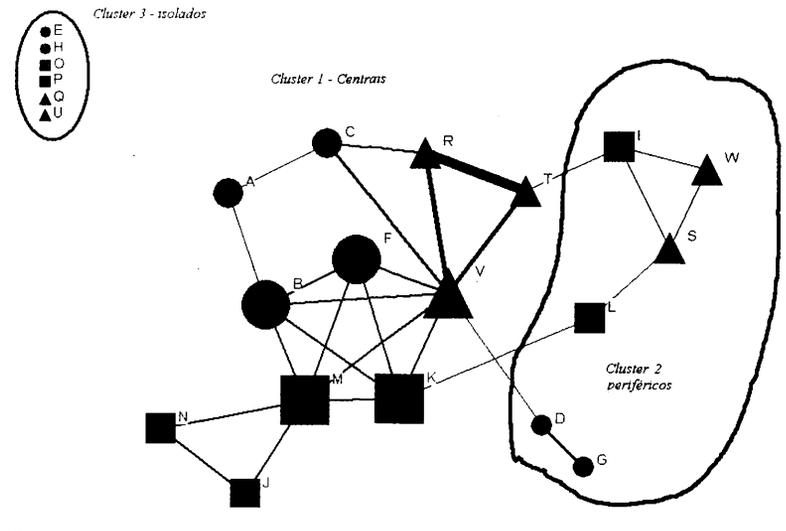
Linhas de Pesquisa	Laços		Produção	
	Internos	Externos	Interna	Externa
Informação, Cultura e Sociedade (ICS)	47,1	52,9	43,8	56,3
Organização e Uso da Informação (OUI)	50,0	50,0	52,9	47,1
Gestão da Informação e do Conhecimento (GIC)	42,1	57,9	71,2	28,8

diferentes para indicar a existência de comportamentos de colaboração distintos intra e interlinhas de pesquisa. O elemento-chave desse ponto da análise é a baixa densidade das redes de colaboração.

Analisar-se-ão, agora, as relações entre os professores, buscando-se identificar aqueles com maior participação na rede. A representação da rede em forma de grafos facilita a visualização dos múltiplos laços entre os professores e a compreensão das medidas usadas na ARS. A figura 1 representa o grafo da rede de co-autorias entre os professores do PPGCI/UFMG, construído com as seguintes características: as letras são os professores, a espessura das linhas representa a quantidade de artigos produzidos em colaboração; as formas dos nós representam as linhas de pesquisa (o círculo, a linha ICS; o quadrado, a linha OUI; o triângulo, a linha GIC).

Observa-se que os professores E, H, O, P, Q e U não estão conectados aos demais (isolados). Dos restantes, alguns estão conectados apenas com um, outros com dois professores. Denomina-se grau nodal (*nodal degree*), denotado por  $d(n_i)$ , o número de linhas incidentes em um nó, ou, ainda, de forma equivalente, o número de nós adjacentes a ele (WASSERMAN; FAUST, 1999, p. 100). O grau de um nó pode variar de 0, caso em que o nó é isolado, até  $g - 1$ , caso no qual o nó está em contato com todos os demais nós do grafo; ou seja,  $0 \leq d(n_i) \leq g$ . Já a *distância geodésica*,  $d(n_i, n_j)$  entre um par de nós é o número de laços que existe no caminho mais curto entre eles, sendo que, caso não exista tal caminho, a distância pode ser considerada indefinida ou infinita. O grau representa a quantidade de colaboradores, e a distância geodésica indica as distâncias entre eles, de forma que se pode identificar a importância das ligações indiretas entre os professores. A tabela 6 mostra essas medidas para o conjunto de professores.

FIGURA 1  
Redes de co-autoria entre os professores do PPGCI/UFMG, período 1997-2004



Fonte: Ucinet.

TABELA 6  
Grau nodal e média e desvio padrão das distâncias geodésicas, período 1997-2004

Linha de pesquisa	Professor	Distância Geodésica		Grau nodal
		Média	Desvio Padrão	
GIC	V	1,5	0,8	8
OUI	M	1,8	1,0	6
ICS	B	1,9	1,0	5
OUI	K	1,8	0,8	5
ICS	F	1,9	0,9	4
ICS	C	2,2	1,1	3
OUI	I	2,5	1,2	3
GIC	R	2,1	0,9	3
GIC	S	2,8	1,4	3
GIC	T	2,0	0,8	3
ICS	A	2,6	1,2	2
ICS	D	2,3	1,0	2
OUI	J	2,6	1,2	2
OUI	L	2,3	1,0	2
OUI	N	2,6	1,2	2
GIC	W	3,2	1,5	2
ICS	G	3,2	1,2	1
ICS	E	0,0	0,0	0
ICS	H	0,0	0,0	0
OUI	O	0,0	0,0	0
OUI	P	0,0	0,0	0
GIC	Q	0,0	0,0	0
GIC	U	0,0	0,0	0

Fonte: Ucinet.

Os professores B, F, K, M e V, individualmente, ocupam a mesma posição na rede, de acordo com os seus graus nodais (maior ou igual a três e conectados entre si). Os demais professores conectados aos demais têm a mesma posição, se tomados os graus nodais superiores a dois e conectados entre si, à exceção dos professores D e G, sendo este último isolado dos demais, a menos de uma única ligação.

Em uma análise centro-periferia\*, levando-se em conta apenas a matriz binária, os resultados mostram a constituição de um grupo central com os professores B, F, K, M e V, ficando os demais na periferia. A densidade do centro é 100%, ou seja, todos estão conectados entre si. A densidade das demais partições – centro-periferia e periferia-periferia – são de 8,9% e 6,5%, respectivamente, de modo significativo mais baixas, indicando poucos relacionamentos. Quando se utiliza a matriz valorada, os resultados se modificam, com o centro representado pelos professores R, T e V, em função da quantidade de artigos produzidos entre eles. Observa-se que o professor V aparece em ambas as análises.

Prosseguindo-se na análise, foram estimadas outras medidas sobre a posição dos professores na rede. A centralidade de grau para um ator é dada por  $C_D(n_i) = d(n_i)$ , ou seja, é simplesmente o grau do nó. Tal medida pode ser normalizada, a fim de ter um valor entre 0 e 1 (ou 0 e 100%, se os resultados forem multiplicados por 100) e para permitir a comparação entre atores de redes diferentes, dividindo-se o grau do nó pelo grau máximo que qualquer nó daquela rede pode ter, ou seja, o número de nós no grafo menos 1 (o próprio nó), chegando-se a  $C_D(n_i) = \frac{d(n_i)}{g-1}$ , sendo  $0 \leq C_D(n_i) \leq 1$ . No caso da matriz valorada, o grau corresponde à soma dos valores dos laços.

A centralidade de intermediação, por seu turno, analisa o quanto um nó está no caminho geodésico entre outros nós. Existe, portanto, uma assunção implícita de que os caminhos mais curtos são os únicos a serem avaliados por tal índice de centralidade e, adicionalmente, de que qualquer caminho mais curto tem a mesma probabilidade de ser escolhido em relação a outros de mesmo valor. Considerando tais premissas, Freeman (1977), por Wasserman e Faust (1999, p. 190), define  $g_{jk}$  como o número de caminhos geodésicos (mais curtos e de mesmo tamanho) que ligam os nós j e k, e  $g_p(n_i)$  como o número de tais caminhos, no total de  $g_{jk}$ , que passa pelo nó  $n_i$ , propondo, assim, um índice de centralidade de intermediação

$$C_B(n_i) = \sum_{j \neq k} \frac{g_{jk}(n_i)}{g_{jk}}$$

que mede, para um nó, a soma de probabilidades de o mesmo estar no caminho geodésico entre todos os demais nós do grafo.

Utilizando-se as medidas mencionadas, pode-se determinar a centralidade de cada professor na rede de co-autoria, com as maiores medidas indicando a importância do ator na rede, seja em função de suas ligações, seja em função de sua produção (tabela 7).

TABELA 7  
Medidas de centralidade, ordenada pela coluna (a) período 1997-2004

Linha de pesquisa	Professor	Grau de centralidade de Freeman			
		Grau Nodal Normalizado (a)	Grau Nodal – valorado (b)	Intermediação (c)	Centralidade (d)
GIC	V	0,3636	21	22,00	0,642
OUI	M	0,2727	12	28,00	0,400
ICS	B	0,2273	9	9,67	0,387
OUI	K	0,2273	9	22,08	0,369
ICS	F	0,1818	8	2,00	0,368
ICS	C	0,1364	6	4,33	0,170
OUI	I	0,1364	3	13,92	0,079
GIC	R	0,1364	18	2,75	0,156
GIC	S	0,1364	3	6,75	0,221
GIC	T	0,1364	16	22,92	0,176
ICS	A	0,0909	2	0,50	0,099
ICS	D	0,0909	4	15,00	0,024
OUI	J	0,0909	4	0,00	0,098
OUI	L	0,0909	2	13,08	0,074
OUI	N	0,0909	4	0,00	0,098
GIC	W	0,0909	2	0,00	0,029
ICS	G	0,0455	3	0,00	0,099
ICS	E	0,0000	0	0,00	0,000
ICS	H	0,0000	0	0,00	0,000
OUI	O	0,0000	0	0,00	0,000
OUI	P	0,0000	0	0,00	0,000
GIC	Q	0,0000	0	0,00	0,000
GIC	U	0,0000	0	0,00	0,000
	Média	0,111	5,478	8,652	0,1133
	desvio padrão	0,095	5,985	13,879	0,100

Fonte: Ucinet.

Obs.: os maiores valores significam posição mais destacada.

Observa-se que os professores em posição de destaque B, F, K, M e V já haviam sido mencionados nas análises anteriores. Para se testar a manutenção da ordenação em cada medida, isto é, se a posição de cada professor se altera em função da medida calculada, foram calculados os coeficientes de ordem de Spearman. Ele é definido como:

\* Utilizou-se o algoritmo de correlação já definido do Ucinet.

$$R = 1 - 6 \frac{\sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)}, \text{ onde:}$$

é a primeira ordenação;

é a segunda ordenação;

é o número de elementos.

Esse coeficiente varia entre -1 (total inversão da ordem) a 1 (manutenção exata da mesma ordenação) e diz respeito apenas à ordem que cada professor ocupa em cada indicador calculado, e o objetivo é ver se eles são consistentes com relação à posição de cada professor. Os resultados, na tabela 8, indicam que a ordem se mantém em todas as medidas de centralidade estimadas, ou seja, não alteram a interpretação da posição relativa de cada professor no conjunto.

Concluindo essa seção, pode-se dizer que a importância do professor, conforme mensurada pela rede de co-autoria, deve-se tanto à sua produção, quanto ao número de colaboradores. Destacam-se, assim, os professores B, F, K, M e V, por sua posição na rede, distribuídos pelas três linhas de pesquisa (duas da linha de informação, cultura e sociedade; segunda da linha organização e uso da informação; e uma da gestão da informação e do conhecimento) e os professores R e T, pela sua imensa produção conjunta e seu relacionamento comum com o professor V. Este professor, por reunir as duas condições mencionadas, é a figura mais central da rede. Para finalizar a ARS no PPGCI/ UFMG, em termos do conjunto, observa-se a criação de 3 *clusters*: o *cluster* 1, com os professores denominados centrais na rede de co-autoria; o *cluster* 2, com os professores mais periféricos; o *cluster* três, com os isolados.

## CONCLUSÃO

O objetivo do artigo foi justificar a proposição de se utilizar a metodologia de análise de redes sociais (ARS) como ferramenta para subsidiar as discussões de interesse da ciência da informação (CI). Ela permite tanto que se olhe para dentro – a análise da sua própria produção –, quanto para fora – a análise da produção científica nas demais área de conhecimento, constituindo-se, neste caso, em uma metodologia complementar para os estudos bibliométricos. Como ferramenta para indicar as redes de colaboração de pesquisadores, ela permite um olhar sobre a interdisciplinaridade de um campo de

TABELA 8  
Coeficientes de ordem de Spearman

	Grau nodal normalizado	Grau nodal valorado	Intermediação	Centralidade
Grau Nodal Normalizado	1			
Grau Nodal – valorado	0,8962	1		
Intermediação	0,8686	0,8360	1	
Centralidade	0,9407	0,9476	0,8409	1

conhecimento, ao facilitar a visualização e análise dessas redes, destacando a colaboração de professores de diferentes áreas. Para tal, foi feita a sua aplicação em um programa de pós-graduação da ciência da informação. Como pano de fundo da abordagem escolhida, foi feita uma discussão sobre a interdisciplinaridade da CI e sobre como a organização dos programas de pesquisa da área reflete as distintas influências de outras disciplinas.

Dessa forma, foi exposto que a organização dos diferentes programas apresenta, em comum, três linhas de pesquisa, que, por sua vez, apresentam ligações com outras áreas do conhecimento de forma distinta entre si. Ou seja, se a ciência da informação é, normalmente, apresentada como uma área de conhecimento das ciências sociais com características marcadamente interdisciplinares, cada uma das linhas possui interdisciplinaridade distinta. Assim, dentro dos programas, essa organização deveria contribuir para atrair professores com diferentes formações que, mediante a realização de pesquisas e colaboração entre eles, pudessem avançar na construção do campo da ciência da informação.

Essa colaboração se dá de várias formas, e não existe a possibilidade de um único ângulo de observação abarcar toda a complexidade de uma área de conhecimento. Uma aproximação possível é a análise de co-autorias entre os professores de um mesmo programa. A publicação em co-autoria deveria representar a preocupação dos professores com os resultados das suas pesquisas e com discussões envolvendo a delimitação e ampliação do campo da ciência da informação. Tomando-se como referência o PPGCI da UFMG, foi feita a análise da produção acadêmica dos professores, na forma de artigos para revistas e congressos. Mesmo sendo um ambiente fechado e com características bastante particulares, como, por exemplo, a necessidade de publicação de forma

permanente, observaram-se os resultados apontados na bibliografia consultada.

De fato, existe um grupo reduzido de professores que respondem pela maior parte da produção; a colaboração se dá, mais intensamente, entre professores de uma mesma linha de pesquisa, embora haja colaboração entre uma parte deles situados em diferentes linhas. No entanto, para a preocupação sobre o desenvolvimento do campo, os resultados são pouco animadores. A rede de colaboração é bastante esparsa e muito da produção se dá entre colaboradores preferenciais. O crescimento observado da produção nos últimos anos está, em grande medida, associado à ampliação da pós-graduação, como demonstra o crescimento da produção em conjunto com alunos do próprio Programa.

Por fim, deve-se destacar que o uso da ARS permitiu uma análise mais aprofundada da colaboração entre os professores, destacando fatos que não seriam detectáveis com o uso de métodos estatísticos tradicionais, também usados nesse trabalho. Assim é uma metodologia poderosa, em especial se usada em conjunto com outras ferramentas da área de bibliometria.

Os resultados mostraram que, se a colaboração entre autores de diferentes áreas de pesquisa significa um esforço para a construção da identidade cognitiva do campo da CI, um longo caminho ainda precisa ser percorrido. Estes resultados, por si só, não permitem uma conclusão definitiva, mas lançam algumas idéias sobre esse ponto, sugerindo novos trabalhos empíricos. Eles representam apenas uma parte da exploração possível da base de dados construída, na qual se buscou destacar a colaboração dos professores materializada nas co-autorias em artigos. Trabalhos futuros utilizando a mesma base de dados permitirão, por exemplo, analisar as publicações mais utilizadas como veículo de divulgação e a associação da formação de cada professor (graduação, mestrado, doutorado e linhas de pesquisa) com sua linha de atuação, dentre outras possibilidades. Já o uso da ARS pode ser ampliado para, por exemplo, análise de citações nos artigos pesquisados.

Como resultado acessório, a pesquisa mostrou as dificuldades de se obterem informações necessárias para estudos quantitativos na área de CI, no Brasil. Não se encontram disponíveis, em uma só base de dados, informações como as utilizadas neste trabalho, dificultando a aplicação dessa metodologia ou de outras normalmente empregadas em estudos bibliométricos. Aqui, no país, não existe uma base ampla e completa o

suficiente para estudos sobre o assunto, análise de citações e co-citações, para fazer sentido nos itens relevantes\*.

Finalmente, fica a alerta, entre descrever um aspecto de uma área de produção de conhecimento, mesmo que relevante para a análise de um fenômeno, e a compreensão das suas características fundamentais, há um longo caminho de pesquisa. As redes de co-autoria existentes entre os professores de um programa informam que a colaboração interdisciplinar não está muito aproximada, mas essa afirmação, para ser conclusiva, precisaria ser complementada por outras pesquisas quantitativas, nos moldes da proposta, mas com outras variáveis, e qualitativas. Como uma das vantagens da metodologia de pesquisa proposta, destaca-se a possibilidade de ser replicada em outros temas complementares, em outros programas e para o conjunto da ciência da informação.

\* Com o presente artigo, as dificuldades encontradas com relação a bases de dados estimularam a montagem de um projeto mais amplo de análise da produção científica na área da ciência da informação no Brasil, denominado RedeCI (<http://redeci.netic.com.br/>).

## AGRADECIMENTOS

Os autores agradecem aos avaliadores anônimos pelas sugestões que permitiram que esse artigo fosse substancialmente melhorado. As falhas remanescentes são de nossa inteira responsabilidade.

Artigo submetido em 20/10/2005 e aceito em 10/07/2006.

## REFERÊNCIAS

- BORGATTI, S. P.; CROOS, R. A relational view of information seeking and learning in social networks. *Management Science*, Evanston, v. 49, n. 4, p. 432-445, Apr. 2003.
- \_\_\_\_\_; FOSTER, P. C. The network paradigm in organizational research: a review and typology. *Journal of Management*, Athens, v. 29, n. 6, p. 991-1013, 2003.
- \_\_\_\_\_; EVERETT, M. G.; FREEMAN, Linton C. UCINFIN: for Windows: software for social network analysis: user's guide. [S.l.]: Analytic Technologies, [200-?].
- BOURDIEU, Pierre. *O campo econômico*. Campinas: Editora Papirus, 2000.
- \_\_\_\_\_. Introdução a uma sociologia reflexiva. In: \_\_\_\_\_. *O poder simbólico*. Rio de Janeiro: Bertrand Brasil, 1998.

BURT, Ronald S. *Structural holes: the social structure of competition*. London: Harvard University Press, 1995.

\_\_\_\_\_.; HOGARTH, Robin M.; MICHAUD Claude. The social capital of french and american managers. *Forthcoming in Organization Science*, v. 11, n. 2, p. 123-147, Mar./Abril, 2000.

CARDOSO, Ana Maria Pereira. Pós modernidade e informação: conceitos complementares?. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 63-79, jan./jun. 1996.

CARLEIAL, Liana. *Redes industriais de subcontratação: um enfoque de sistema nacional de inovação, um estudo das indústrias eletrônica, metalmeccânica e de confecções da Região Metropolitana de Curitiba*. São Paulo: Hucitec, 2001.

CASANUEVA ROCHA, Cristóbal. Relaciones estratégicas entre pymes: contraste de hipótesis empresariales mediante ARS. *Revista Hispana para el análisis de redes sociais*, v. 4, n. 4, jun. 2003. Disponível em: <<http://revista-redes.rediris.es>>. Acesso em: 15 jul. 2004.

CRANE, Diana. *Invisible Colleges: diffusion of knowledge in scientific communities*. London: University of Chicago Press, 1972.

DIAS, Eduardo Wense. Ensino e pesquisa em ciência da informação. *DataGramaZero: Revista de Ciência da Informação*, v. 3, n. 5, out. 2002. Disponível em: <<http://www.dgz.org.br/>>. Acesso em: 10 out. 2003.

FREEMAN, Linton C. Some antecedents of social network analysis. *CONNECTIONS*, v. 19, n. 1, p. 39-42, 1996.

GARCÍA MACÍAS, Alejandro. Redes sociales y "clusters" empresariales. *Revista Hispana para el análisis de redes sociais*, v. 1, n. 6, enero 2002. Disponível em: <[http://revista-redes.rediris.es/html-vol1/vol1\\_6.htm](http://revista-redes.rediris.es/html-vol1/vol1_6.htm)>. Acesso em: 15 jul. 2004.

GOMES, Henriette Ferreira. Interdisciplinaridade e Ciência da Informação: de característica a critério delineador de seu núcleo principal. *DataGramaZero: Revista de Ciência da Informação*, v. 2, n. 4, ago. 2001. Disponível em: <<http://www.dgz.org.br/>>. Acesso em: 10 out. 2003.

GOMES, Maria Yêda Falcão Soares de Filgueiras. A produção científica em biblioteconomia e ciência da informação no Brasil: tendências temáticas e metodológicas. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003. *Anais...* Belo Horizonte: UFMG, 2003. CD-ROM.

GONZALEZ DE GOMEZ, Maria Nélide. Para uma reflexão epistemológica acerca da ciência da informação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 6, n. 1, p.5-18, jan./jun. 2001.

\_\_\_\_\_. Metodologia de pesquisa no campo da ciência da informação. *DataGramaZero: Revista de Ciência da Informação*, v. 1, n. 6, dez. 2000. Disponível em: <[http://www.dgz.org.br/dez00/F\\_1\\_aut.htm](http://www.dgz.org.br/dez00/F_1_aut.htm)>. Acesso em: 10 out. 2003.

GRANOVETTER, Mark S. The strength of weak ties. *American Journal of Sociology*, Chicago, v. 78, n. 6, May. 1973.

KRACKHARDT, David. Cognitive social structures. *Social Networks*, v. 9, p. 109-134, 1987.

\_\_\_\_\_.; HANSON, Jeffrey R. Informal networks: the company behind the chart. *Harvard Business Review*, v. 71, n. 4, p.104-111, July/Aug., 1993.

KRETSCHMER, Hildrun. Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the Web. *Scientometrics*, v. 60, n. 3, p. 409-420, Jan. 2004. Disponível em: <<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3eb2dfb4aa89cd6796b917f7991&referrer=parent&backto=issue,13,2;journal,12,53;linkingpublicationresults,1:101080,1>>. Acesso em: 01 abr. 2005.

KUHN, Thomas S. *O problema com a filosofia histórica da ciência= The trouble with the historical philosophy of science*. Cambridge Mass: Harvard University Press. Disponível em: <<http://www.fafich.ufmg.br/~scientia/>>. Acesso em: 15 jan. 2004.

LE COADIC, Yves-François. *A ciência da informação*. Brasília: Briquet de Lemos/Livros, 1996. 119 p.

LIN, Nan; FU, Yang-Chih; HSUNG, Ray-May. The position generator: measurement techniques for investigations of social capital. In: LIN, Nan; COOK, Karen; BURT, Ronald (Ed.). *Social capital: theory and research*. New York: Aldine de Gruyter, 2001.

MACHLUP, Fritz; MANSFIELD, Una (Ed.). *The study of information: interdisciplinary messages*. [S.l.]: John Wiley & Sons, 1983.

MAHLCK, Paula; PERSSON, Olle. Socio-bibliometric mapping of intra-departmental networks. *Scientometrics*, v. 49, n. 1, p. 81-91, 2000. Disponível em: <<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3eb2dfb4aa89cd6796b917f7991&referrer=parent&backto=issue,5,9;journal,47,53;linkingpublicationresults,1:101080,1>>. Acesso em: 01 abr. 2005.

MARTELETO, Regina Maria. Análise de redes sociais: aplicação nos estudos de transferência da informação. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 71-81, 2001.

MATHEUS, Renato Fabiano. *Desafios para a ciência da informação: enfrentando dificuldades paradigmáticas, dilemas e paradoxos através de programas de pesquisa interdisciplinares*. 2005. Disponível em <<http://dici.ibict.br/view/year/2005.html>>. Acesso em: 11 ago. 2005.

MOLINA, José Luis. El organigrama informal en las organizaciones: una aproximación desde el análisis de redes sociales. *Revista Catalana de Sociologia*, p. 65-86, nov. 2000.

\_\_\_\_\_. The informal organizational chart in organizations: an approach from the social network analysis. *Connections*, v. 24, n. 1, p. 78-91.

\_\_\_\_\_.; MUÑOZ, Juan Manuel; DOMENECH, Miquel. Redes de publicaciones científicas: un análisis de la estructura de coautorías. *Revista Hispana para el análisis de redes sociais*, v. 1, n. 3, enero 2002. Disponível em: <<http://revista-redes.rediris.es/>>. Acesso em: 10 jan. 2005.

MUELLER, Suzana Pinheiro Machado; PECEGUEIRO, Cláudia Maria Pinho de Abreu. 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*, v. 98, p. 404-409, Jan. 2001.

Disponível em: <<http://www.pnas.org/cgi/reprint/98/2/404>>. Acesso em: 01 mar. 2005.

ODDONE, Nanci; GOMES, Maria Yêda Falcão Soares de Filgueiras. Uma nova taxonomia para a ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. *Anais...* Belo Horizonte: UFMG, 2003.

OLIVEIRA, Marlene de. *A investigação científica na ciência da informação: análise da pesquisa financiada pelo CNPq*. 1998. 221 p. Tese (Doutorado em Ciência da Informação)– Faculdade de Estudos Sociais Aplicados, Universidade de Brasília, Brasília, 1998.

\_\_\_\_\_. Características das dissertações produzidas no curso de mestrado em Ciência da Informação da UFPB. *Informação & Sociedade: Estudos*, João Pessoa, v. 9, n. 2, p. 465-488, 1999.

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

ROUSSEAU, B.; ROUSSEAU, R. Lotka: a program to fit a power law distribution to observed frequency data. *Cybermetrics*, Madrid, v. 4, n. 4, 2000. Disponível em: <<http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html>>. Acesso em: 15 mar. 2005.

SAXENIAN, Annalee. *Regional Advantage: culture and competition in Silicon Valley and route 128*. Massachusetts: Harvard University Press, 1996.

TARGINO, Maria das Graças. A interdisciplinaridade da ciência da informação como área de pesquisa. *Revista Informação & Sociedade*, v. 5, n. 1, 1995. Disponível em: <<http://www.informacaoesociedade.ufpb.br/519501.pdf>>. Acesso em: 28 jun. 2004.

VAKKARI, Pertti. Open the horizon of expectations. In: THE INTERNATIONAL CONFERENCE HELD FOR THE CELEBRATION OF 20TH ANNIVERSARY OF THE DEPARTMENT OF INFORMATION STUDIES, 20., 1991, Tampere. *Proceedings...* Tampere: Taylor Graham, 1991. p. 26-28.

YOSHIKANE, Fuyuki; KAGEURA, Kyo. Comparative analysis of coauthorship networks of different domains: The growth and change of networks. *Scientometrics*, v. 60, n. 3, p. 435-446, Jan. 2004. Disponível em: <<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3eb2dfb4aa89ed6796b917f7991&referrer=parent&backto=issue,15,22;journal,12,53;linkingpublicationresults,1:101080,1>>. Acesso em: 01 abr. 2005.

WASSERMAN, Stanley; FAUST, Katherine. Social network analysis: methods and applications. In: STRUCTURAL analysis in social the social sciences series. Cambridge: Cambridge University Press, 1994. 857 p. v. 8.

WATTS, Duncan J. *Small worlds: the dynamics of networks between order and randomness*. New Jersey: Princeton University Press, 1999.