

A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na Plataforma Lattes

Renato Balancieri

Doutorando do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento - UFSC. E-mail: renato@stela.ufsc.br

Alessandro Botelho Bovo

Doutorando do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento - UFSC. E-mail: abbovo@stela.ufsc.br

Vinícius Medina Kern

Professor do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento - UFSC e professor da Univali São José. E-mail: kern@stela.ufsc.br

Roberto Carlos dos Santos Pacheco

Professor do Departamento de Informática e Estatística e do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento - UFSC. E-mail: pacheco@stela.ufsc.br

Ricardo Miranda Barcia

Professor do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento - UFSC. E-mail: rbarcia@egc.ufsc.br

INTRODUÇÃO

Os esforços colaborativos envolvendo pesquisadores de diferentes países foram detectados já no século XIX (BEAVER; ROSEN, 1978). A colaboração científica pode ser um empreendimento cooperativo que envolve metas comuns, esforço coordenado e resultados ou produtos (trabalhos científicos) com responsabilidade e mérito compartilhados. Assim, a colaboração científica oferece uma fonte de apoio para melhorar o resultado e maximizar o potencial da produção científica (WEISZ; ROCCO, 1996).

Em uma ótica social, importante ganho resultante da formação de redes científicas quando comparada a pesquisadores isolados é a ampliação do repertório de abordagens e ferramentas que advém do intercâmbio de informações e da fertilização cruzada que se verifica quando grupos distintos juntam esforços no sentido de determinada meta (WEISZ; ROCCO, 1996).

As novas TICs fizeram surgir uma gama de novas possibilidades para a análise de redes sociais e, conseqüentemente, de redes de colaboração em ciência, tecnologia e inovação (CT&I). O advento da Internet é, sem dúvida, o fato mais significativo. A formação de comunidades virtuais na internet é promovida com suporte tecnológico, valendo-se da criação de padrões, principalmente os baseados na tecnologia XML, e da disponibilidade de serviços, como aqueles oferecidos usando a tecnologia de *web services*.

Com o avanço das TICs e das possibilidades de efetivação de intercâmbio, o processo de estudo, indução e formação de redes nas mais diversas atividades humanas tem sido revisitado. Barabási (2003), por exemplo, trata da identificação de *hubs* (vértices centrais), enquanto Watts (2003) aborda o postulado dos seis graus máximos de separação entre quaisquer dois elementos de uma rede.

Particularmente em ciência e em tecnologia, novas questões se colocam a partir da disponibilidade de fontes e sistemas de informação em CT&I. Essas podem ser resumidas como as abordagens de estudo de redes científicas podem ser revisitadas e combinadas à disponibilidade de informação sobre as atividades em CT&I.

Resumo

As redes de pesquisa impulsionam a criação do conhecimento e o processo de inovação resultantes do intercâmbio de informações e, sobretudo, da junção de competências de grupos que unem esforços na busca de metas comuns. Este artigo apresenta breve histórico dos estudos relativos às redes de colaboração científica, sua evolução cronológica e as principais abordagens de estudo. Discute-se particularmente como as análises de redes de pesquisa podem ser revisitadas à luz das possibilidades recentes surgidas com as novas Tecnologias de Informação e de Comunicação (TICs). Para tal, apresentam-se exemplos de sistemas de conhecimento no âmbito da Plataforma Lattes: Egressos, Colaboradores e Redes-GP. Esses sistemas permitem executar, com grandes volumes de dados, análises de redes por meio de algoritmos descritos na literatura, bem como criar novas formas de análise possibilitadas pelas TICs.

Palavras-chave

Colaboração científica; Análise de redes sociais; Redes de pesquisa; Tecnologias da informação e da comunicação.

An analysis of scientific collaboration networks under the new technologies of information and communication: a study in Lattes Platform

Abstract

Scientific networks have fostered knowledge creation and the innovation process by means of information interchange and mainly by the combination of competences of different groups searching for common goals. In this paper we outline the chronology and review the main approaches to study collaborative scientific networks. More specifically the article addresses the possibilities brought by the new Information and Communication Technologies (ICTs) to scientific network research by discussing examples of knowledge systems from the Lattes Platform: Egressos (Alumni), Colaboradores (Collaborators), and Redes-GP (Research Group Networks). Those systems allow for the realization, with wide amounts of data, of social network analysis, both using algorithms from the literature and creating new approaches brought about by the ICTs.

Keywords

Scientific collaboration; Social network analysis; Research networks; Information and communication technologies

Passados 30 anos, o campo de estudo de redes sociais como as TICs avançou no sentido de alcançar efetividade e variedade de sistemas de análise de relacionamentos.

O presente artigo apresenta o histórico de redes de colaboração científica, os principais resultados com relação às definições necessárias para o desenvolvimento de sistemas de conhecimento em redes de colaboração científica e discute exemplos na Plataforma Lattes.

ANÁLISE DE REDES DE COLABORAÇÃO CIENTÍFICA: UM BREVE HISTÓRICO

Nesta seção, apresenta-se breve histórico das áreas de redes de colaboração científica. A apresentação está separada por períodos, incluindo-se cronologia, autores e comentário sobre os estudos realizados em cada fase.

A figura 1 apresenta uma cronologia do estudo de redes de colaboração científica.

PERÍODO ATÉ A DÉCADA DE 1960

Smith (1958) foi um dos primeiros autores a observar crescimento na incidência de artigos em co-autoria e a sugerir que tais artigos pudessem ser usados como uma medida aproximada da colaboração entre grupos de pesquisadores. Ele notou também que os resultados finais

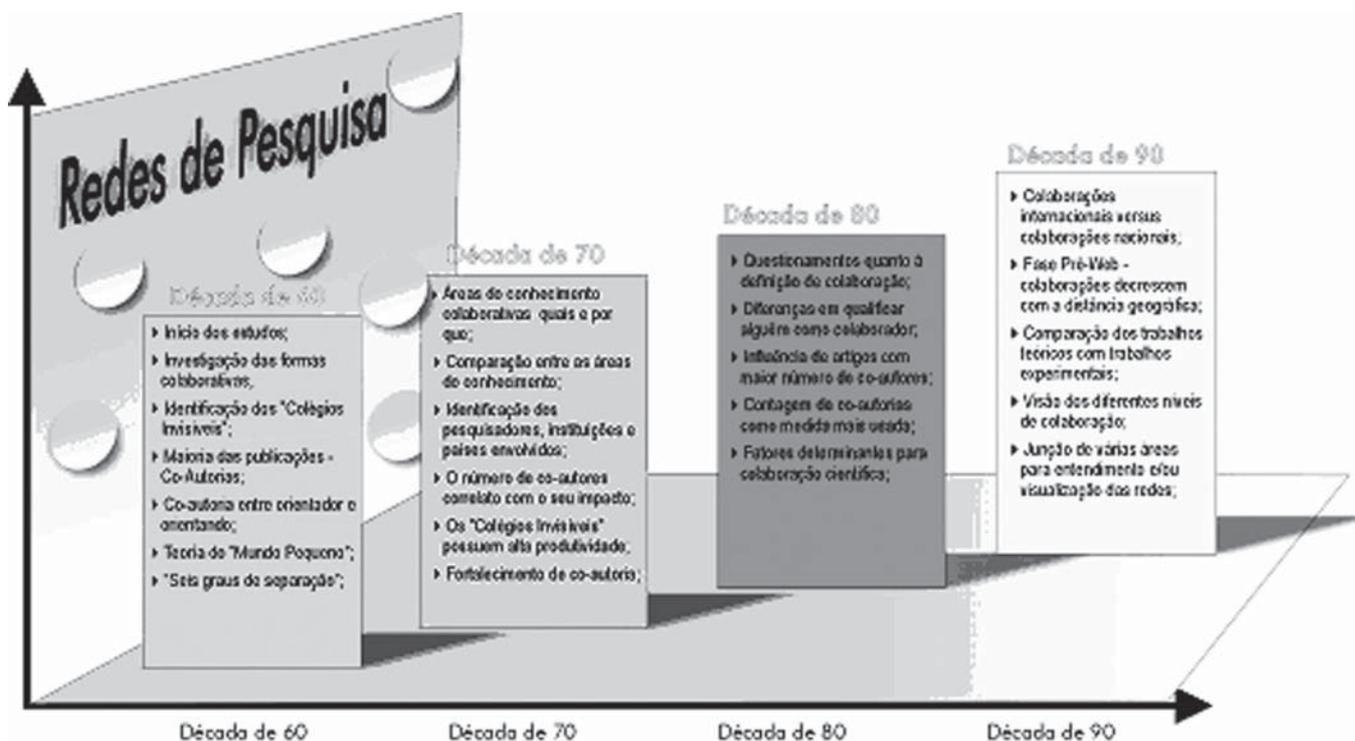
de um projeto científico são freqüentemente expressos por meio da publicação documentada pelos pesquisadores envolvidos, independentemente da natureza da cooperação científica – seja ela interpessoal, interinstitucional ou internacional.

As observações de Smith (1958) foram testadas empiricamente por Price (1963), que forneceu evidências do aumento de autorias múltiplas na ciência. De acordo com Price, podendo ser encontrada nas mais diferentes formas, a colaboração científica se dava, freqüentemente, no âmbito dos chamados “colégios invisíveis”. Esses colégios constituíam-se em comunidades informais de pesquisadores que se comunicavam, trocavam informações e experiências e também publicavam formalmente seus resultados no campo do conhecimento científico.

Hagstrom (1965) constatou que alguns professores não incluíam seus orientandos na sua relação de colaboradores. Durante entrevistas, Hagstrom perguntava aos professores que tinham escrito artigos se o trabalho foi realizado em colaboração com outros, e número considerável de docentes respondia “não”, embora a maioria ou a totalidade de seus artigos tivesse sido feita em co-autoria com estudantes.

Price e Beaver (1966) desenvolveram estudos mostrando que os pesquisadores mantêm encontros em congressos,

FIGURA 1
Cronologia de redes de colaboração científica (BALANCIERI, 2004)



conferências, reuniões sobre suas especialidades, visitam-se por meio de intercâmbios institucionais ou realizam trabalhos em colaboração. Esse tipo de organização transcende os limites do departamento, da instituição de um país e abrange cientistas de todos os lugares do mundo onde houver atividade científica relevante na área ou na especialidade em questão. Os autores concluíram que a maior parte das colaborações em pesquisa inicia-se com relações informais, várias delas estabelecidas durante o período de treinamento do pesquisador.

Milgram (1967) apresenta o “problema do mundo pequeno” (*small world*). O modelo do “mundo pequeno” afirma que cada ator em uma rede (independentemente do seu tamanho e da densidade) pode encontrar outro ator com seis passos em média – também conhecido como “seis graus de separação”.

Portanto, o período até a década de 1960 é marcado pelo início dos estudos na área de colaboração científica. Por meio de estudos teóricos e testes empíricos, iniciou-se a investigação das formas com que se davam os relacionamentos de colaboração, identificando que os mais freqüentes aconteciam no âmbito dos “colégios invisíveis”. Nesse período, observou-se também que essas publicações, na maioria das vezes, ocorrem na forma de co-autoria e que o início da colaboração começa nas relações entre orientador e orientando. Esse período é marcado também pela teoria do “mundo pequeno”. Dentro da colaboração científica, esta teoria teve grande importância, pois se começou a pensar sobre as distâncias entre um pesquisador e outro, formadas pelas redes de pesquisadores com artigos escritos em co-autoria.

DÉCADA DE 1970

Em 1970, Norman Storer afirma que parte significativa do que se sabia sobre colaboração científica fora investigado por meio de técnicas bibliométricas. Entre várias outras revelações, tais estudos forneceram evidência para algo que se supunha intuitivamente: o grau de cooperação varia significativamente entre as várias áreas do conhecimento em função de suas características cognitivas e organizacionais. Constatou-se, por exemplo, que as ciências básicas e as ciências naturais apresentam maior índice de cooperação do que as ciências aplicadas e as ciências sociais (STORER, 1970).

Em 1971, Medows e O'Connor afirmam que, quaisquer que sejam as razões que levam à colaboração ou as unidades que as executam, assume-se freqüentemente que seus resultados finais serão publicados de forma a identificar os pesquisadores envolvidos, assim como suas instituições e seus países de origem. Com base nessa premissa, define-

se a cooperação científica como sendo o conjunto de trabalhos cooperativos desenvolvidos entre dois ou mais pesquisadores e identificados por meio de artigos co-assinados (MEDOWS; O'CONNOR, 1971).

Em 1972, Diana Crane mostra que os “colégios invisíveis” são caracterizados por sua alta produtividade, por compartilhar prioridades de pesquisa, por treinar estudantes, por produzir e monitorar o conhecimento em seu campo (CRANE, 1972).

Ainda em 1972, Janice Lodahl e Gerald Gordon, em suas pesquisas, mostram que o fato de as ciências básicas e naturais apresentarem índice maior de cooperação do que as ciências aplicadas e sociais se deve ao caráter universal das ciências básicas e ao grau de maturidade e consenso paradigmático das ciências naturais. Estes se contrapõem ao caráter localizado e contingencial das ciências aplicadas, que geram resultados mais facilmente apropriáveis, e à falta de consenso paradigmático das ciências sociais, que dificulta o processo de negociação entre os pesquisadores e a tomada de decisão (LODAHL; GORDON, 1972).

No mesmo ano, um estudo de Nudelman e Landers sugere que o total de crédito dado por uma comunidade científica para todos os autores de um artigo de autoria conjunta é em média maior do que o crédito alocado para autor de artigo de autoria única. Ou seja, o número de co-autores parece ser fortemente correlacionado com o impacto de um artigo (NUDELMAN; LANDERS, 1972).

Em 1978, Beaver e Rosen afirmam que o primeiro artigo científico escrito em colaboração, entre diferentes pesquisadores, apareceu em 1678 (BEAVER; ROSEN, 1978).

Em 1979, Beaver e Rosen afirmaram que, dada a variedade de colaborações que podem se estabelecer, não é de se estranhar que as razões que levam os cientistas a colaborar entre si sejam também das mais diferentes naturezas. A mais óbvia é, certamente, a necessidade da contribuição especializada de outrem para alcançar os objetivos da pesquisa, o que também inclui a necessidade de trabalhar próximo de outrem para desenvolver novas habilidades e conhecimento tácito, como é tipicamente o caso das relações entre mestre e aprendiz (BEAVER; ROSEN, 1979).

No mesmo ano, Frame e Carpenter reafirmam aquilo que Storer, em 1970, já havia concluído: as ciências básicas e naturais cooperam mais do que as ciências aplicadas e sociais (FRAME; CARPENTER, 1979).

Portanto, na década de 1970, as pesquisas desenvolvidas estiveram direcionadas à variabilidade das redes entre

áreas do conhecimento e à dinâmica dessas relações. Algumas dessas pesquisas só foram possíveis mediante o uso de técnicas bibliométricas. Nesse período, fortaleceu-se a idéia de cooperação científica por co-autoria.

DÉCADA DE 1980

Em 1980, William Goffman e Kenneth Warren mostram que as pesquisas desenvolvidas por grandes grupos tendem a ter mais influência no meio científico (GOFFMAN; WARREN, 1980).

Em 1982, Stefaniak afirmou que a relação direta entre colaboração e co-autoria tem sido questionada com base em evidências de que diferentes países, áreas do conhecimento, instituições, grupos de pesquisa e indivíduos têm tradições e critérios significativamente diferentes para qualificar alguém como co-autor (STEFANIAK, 1982).

Em 1983, Subramanyam mostrou que as interações entre os pesquisadores também variam em intensidade, podendo ser desde substantivas até muito tênues. Algumas vezes um pesquisador pode ser considerado colaborador e até mesmo aparecer como co-autor, simplesmente por fornecer material sobre o qual ele tem controle ou por realizar alguma operação de rotina. Em outros casos, pesquisadores de diferentes organizações podem colaborar mediante o uso coletivo de dados ou idéias, ou da execução, de maneira separada, de diferentes partes de um projeto, integrando posteriormente os resultados e a análise. Têm sido até mesmo registradas ocasiões em que uma sugestão brilhante feita por um pesquisador no curso de uma discussão informal acabou tendo maior impacto na direção e nos resultados de projetos de pesquisa do que semanas de atividades intensas de colaboração entre cientistas no laboratório. No mesmo estudo, Subramanyam mostrou que a contagem de artigos realizados em co-autoria tem sido a medida mais comumente usada para detectar a ocorrência, a abrangência e os participantes das colaborações científicas (SUBRAMANYAM, 1983).

Pesquisas como a de Lawani (1986), Pravdic e Oluic-Vukovic (1986) mostraram que o desejo dos pesquisadores em aumentar sua visibilidade e, conseqüentemente, seu reconhecimento pelos pares também tem sido apontado como um fator que estimula a colaboração científica.

Portanto, a década de 1980 é marcada pelos questionamentos relativos à definição de colaboração, dada a variabilidade dos critérios de definição do que é um colaborador. Além disso, estudou-se a colaboração sob a ótica do impacto dos trabalhos científicos, comprovadamente maior quando estes são resultantes de um coletivo de pesquisadores.

PERÍODO A PARTIR DE 1990

Em 1990, Francis Narin e Edith Whitlow encontraram evidências de que artigos de co-autoria internacional são citados duas vezes mais do que um artigo de um único país (NARIN; WHITLOW, 1990).

No ano seguinte, Narin *et alii* apontam que, em temas de pesquisa em que a União Européia resolveu investir recursos para pesquisa cooperativa – comportamento de materiais radioativos no ambiente, manufatura de ferro e aço, agricultura e ciência de alimentos e outros –, houve um aumento em artigos co-autorados por pesquisadores de diferentes países significativamente maior do que em outras áreas (NARIN *et alii*, 1991).

Em 1992, Leclerc *et alii* apresentam um exemplo do uso da contagem de co-autoria de artigos como medida de colaboração científica (LECLERC *et alii*, 1992).

No mesmo ano, Kodama afirma que um fator preponderante que tem motivado a colaboração científica é o importante aumento dos campos interdisciplinares. Ele deixa claro que alguns dos mais significativos avanços científicos surgiram da integração ou fusão dos campos anteriormente separados (KODAMA, 1992).

Ainda neste ano, Luukonen *et alii* mostram que os fatores que levam os pesquisadores a colaborar são agrupados em cognitivos econômicos e sociais. Tais fatores têm importância relativa variada para explicar as diferenças nas taxas de colaboração das áreas do conhecimento e dos diversos países (LUUKONEN *et alii*, 1992).

Em 1993, no período *pré-web*, Katz mostra que as colaborações decrescem exponencialmente com a distância geográfica que separa os pesquisadores institucionais. Assim a proximidade regional parece encorajar a colaboração, já que os pesquisadores tendem a gerar mais comunicação informal (KATZ, 1993).

Em 1994, Katz demonstra que colaborações são também motivadas pela necessidade de compartilhar o uso de equipamentos cada vez mais caros e complexos e pelos novos padrões de financiamento adotados pelas agências (KATZ, 1994).

Em 1995, Colonomos mostra que houve investimento acadêmico significativo em estudos de redes a partir do campo das relações internacionais, influenciando a história recente das ciências políticas. A origem da reflexão se dá logo ao fim da II Guerra Mundial e tem progresso com o fim da Guerra Fria, quando há uma redefinição dos atores nas relações internacionais. Novos caminhos de

pesquisa surgem, nascidos mediante a reflexão de ordem e desordem nos sistemas políticos ligados à elaboração da noção de globalização (COLONOMOS, 1995).

Em 1996, Weisz e Roco constataam número crescente de projetos de pesquisa e ensino que haviam sido concebidos e executados na forma de cooperação em que diferentes instituições assumiam tarefas distintas. Assim, uma rede pode ser, então, definida como uma organização de coesão tênue, consistindo de diferentes indivíduos ou grupos ligados entre si por vínculos de naturezas diversas. Essas redes são tipicamente centros “não-físicos” que contam com meios de comunicação avançados a fim de promover a interação de participantes com qualificações complementares. O grau de participação de cada unidade participante é flexível (WEISZ; ROCCO, 1996).

Katz e Martin (1997) mostram que, em geral, os trabalhos teóricos produzem artigos com poucos co-autores em comparação a trabalhos experimentais, ou seja, experimentalistas tendem a colaborar mais do que teóricos. Ainda neste mesmo trabalho, os autores apresentam diferentes níveis de colaboração, podendo a colaboração ocorrer entre e dentro de diferentes níveis.

Os prefixos “inter” e “intra”, respectivamente, têm sido adotados para distinguir essas categorias. Os diferentes níveis de colaboração (indivíduos, grupos, departamentos, instituições e setores, nas mais diferentes combinações dessas unidades, dentro de uma mesma nação ou envolvendo nações diferentes), nas formas inter e intra, estão sumarizados no quadro 1.

QUADRO 1

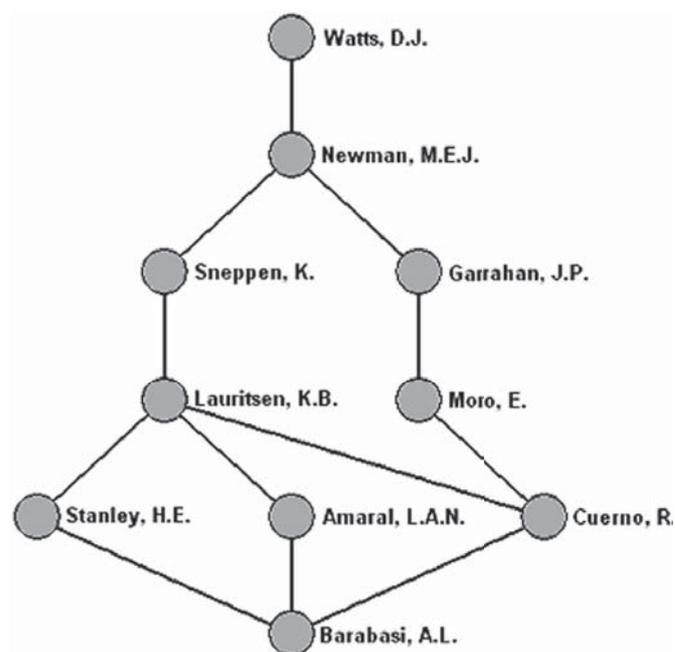
Diferentes níveis de colaboração e distinção entre as formas inter e intra (KATZ; MARTIN, 1997)

NÍVEL	Intra	Inter
Individual	-	Entre indivíduos
Grupo	Entre indivíduos do mesmo grupo de pesquisa	Entre grupos (por exemplo, no mesmo departamento)
Departamento	Entre indivíduos ou grupos no mesmo departamento	Entre departamentos (na mesma instituição)
Instituição	Entre indivíduos ou entre deptos na mesma instituição	Entre instituições
Setor	Entre instituições no mesmo setor	Entre instituições em diferentes setores
Nação	Entre instituições no mesmo país	Entre instituições em diferentes países

Em 2000, Newman define rede sob o aspecto social como um conjunto de pessoas ou grupos que possuem conexões de algum tipo com um ou com todos os outros integrantes da rede. Na linguagem das análises de redes sociais, as pessoas ou os grupos são chamados de “atores”, e as conexões, de “ligações”. Ambos podem ser definidos em diferentes caminhos, dependendo da questão de interesse. Um ator pode ser uma única pessoa, um grupo ou uma empresa. Uma ligação pode ser uma amizade entre duas pessoas, uma colaboração ou um membro comum entre dois grupos, ou ainda um relacionamento de negócios entre duas empresas (NEWMAN, 2000). Na figura 2 exemplifica-se uma rede que mostra os atores como pesquisadores e as ligações como co-autoria em artigo científico.

FIGURA 2

Exemplo de uma rede de colaboração científica entre Duncan Wattz e Albert-László Barabási (NEWMAN, 2000)



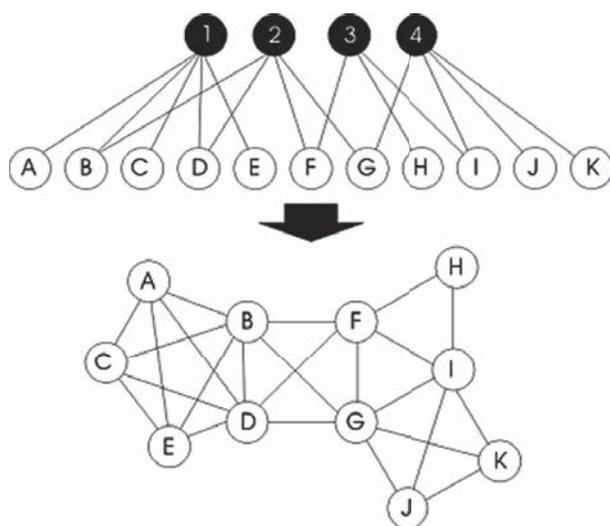
A abordagem de Newman amplia o conjunto de possibilidades para a análise de redes, pois acrescenta-lhes propriedades estatísticas, como número de artigos escritos por autor, número de autores por artigo, número de colaboradores dos cientistas da rede, a distância entre a rede de um pesquisador e a outra rede. Também permite produzir uma variedade de medidas de conexidade dentro da rede, tais como proximidade e intermediação.

Nessa forma de conceber redes como configurações de relacionamentos, existe alto grau de transitividade,

o que significa que há maior probabilidade de duas pessoas estarem relacionadas, se elas tiverem um ou mais relacionamentos em comum, ou seja, a probabilidade de pesquisadores colaborarem aumenta com o número de outros colaboradores em comum (NEWMAN, 2001b). Outra conclusão de Newman é que a probabilidade de um pesquisador, em particular, adquirir novos colaboradores aumenta com o número de pesquisadores com que ele colaborou no passado (NEWMAN, 2001a).

Outra possibilidade dessa análise é a aplicação da Teoria de Grafos para mensurar as propriedades das redes. Um exemplo pode ser visto na figura 3 a seguir, onde se ilustra uma rede de co-autoria. No topo, é apresentado um grafo bipartido representando pesquisadores (A a K) e artigos científicos (1 a 4), com linhas representando autoria. Embaixo, é exibida a projeção da mesma rede apenas com os cientistas.

FIGURA 3
Exemplo de Rede de Co-autoria (NEWMAN; STROGATZ; WATTS, 2001)



Um exemplo de análise utilizando Teoria dos Grafos é o cálculo da distância entre os atores da rede. A idéia de distância é o grau de afastamento de um vértice (ator) em relação aos outros vértices do grafo. Pode ser calculada por algoritmos de caminho mais curto (BOAVENTURA NETTO, 2001). No grafo da figura 3, a distância mais curta entre o pesquisador A e o pesquisador K é de dois colaboradores (A-B-G-K ou A-D-G-K).

Outro exemplo de análise é a determinação do(s) ponto(s) central(is). Estuda-se a localização do ator em relação à rede total, identificando indivíduos considerados “importantes” para a conexão da rede. Podem ser calculados por meio da

matriz de caminhos mínimos (BOAVENTURA NETTO, 2001). No exemplo da figura 3, em termos de número de conexões, os vértices (atores) B, D e G são os mais conectados, com seis atores cada. Logo, são os atores centrais.

Ainda nesse contexto, uma descrição mais detalhada das possibilidades de aplicação da Teoria dos Grafos (e os respectivos métodos de resolução) na análise redes de colaboração é apresentada por Bovo (2004).

Portanto, a partir de 1990, as pesquisas confirmam a hipótese do fator de impacto de trabalho coletivo e buscam analisar outros fatores que influenciam redes de colaboração científica. Entre esses, estão a distância geográfica e a natureza do trabalho científico. O resultado mais relevante desse período é a combinação da área de redes sociais com as investigações em redes científicas, trazendo uma taxonomia para a área de redes de colaboração. Outra característica presente nesse período é a junção de várias áreas de análise para o entendimento e/ou a visualização das redes de relacionamentos formadas pela colaboração científica (e.g., Link Analysis, Teoria dos Grafos, ciência da informação e redes sociais).

PRINCIPAIS RESULTADOS DA PESQUISA EM REDES SOCIAIS SOB A ÓTICA DAS TICs

Para que as TICs possam contribuir nas atividades de análise, inspeção e fomento a redes, é fundamental que se parta de uma definição de referência. As quatro décadas de estudos sobre redes sociais permitiram a formação dessa definição abrangente, que teve gênese em estudos exclusivos à participação de autores em artigos publicados (e.g., MEDOWS; O'CONNOR, 1971) e culminou na concepção de redes como uma coesão de indivíduos ou grupos baseada em diferentes formas e origens de vinculação, desde cognitivas e sociais a econômicas (e.g., WEISZ; ROCO, 1996; NEWMAN, 2000).

As TICs podem ser base para o desenvolvimento de sistemas em diversas áreas de interesse para redes sociais. Para tal, há um conjunto de perguntas cujas respostas permitem aos engenheiros de *software* e de conhecimento conceber sistemas de apoio à decisão e de descoberta de conhecimento em redes sociais. No quadro 2, está parte dessas perguntas com exemplos de pesquisas que procuram respondê-las.

As respostas às questões no quadro 2 estabelecem um marco sobre o qual as TICs podem trazer diversas contribuições à área de redes sociais. Nas próximas seções, apresentam-se exemplos das possibilidades que as TICs trazem a essa área,

QUADRO 2

Principais conclusões sobre os estudos de redes sociais, sob a ótica das possibilidades das TICs

Autores	Questões de interesse às TICs
MEDOWS; O'CONNOR, 1971	O que é Cooperação científica? Caracteriza-se por trabalhos cooperativos identificados por artigos co-assinados
SUBRAMANYAM, 1983	O que caracteriza uma co-autoria em um artigo? Há uma variedade de formas de colaborar.
LUUKONEN et al., 1992	Por que pesquisadores colaboram? Fatores cognitivos, econômicos e sociais, com diferenças por áreas do conhecimento.
KATZ, 1994	O que influencia a formação de redes? Exemplo: padrões de financiamento das agências e necessidade de compartilhar equipamentos também influenciam a formação de redes.
WEISZ; ROCO, 1996	O que é uma rede? É concebida como uma coesão tênue com diferentes indivíduos ou grupos conectados por vínculos de diversas naturezas.
KATZ; MARTIN, 1997	Como se comportam os vértices da rede? O grau de cooperação varia com a natureza do trabalho (experimentalistas cooperam mais que teóricos)
NEWMAN, 2000	O que é uma rede? Rede definida como conjunto de pessoas ou grupos com conexões originadas por diferentes formas de relacionamento social.

tomando-se por base os recentes avanços que a pesquisa em engenharia de *software* e em engenharia do conhecimento trouxeram à Plataforma Lattes (Disponível em: <<http://lattes.cnpq.br>>) – conjunto de sistemas de informação, bases de dados e portais Web voltados para a gestão de CT&I – do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

PERSPECTIVAS DIANTE DAS NOVAS TICs: ANÁLISE DE REDES NA PLATAFORMA LATTES

As perspectivas que as novas TICs trazem à análise de redes sociais têm origem nos seguintes fatos: (a) a consolidação das pesquisas em redes sociais que permitem as conclusões do quadro 2; (b) os avanços nas TICs que permitiram maior produtividade no desenvolvimento de sistemas de conhecimento; (c) a disponibilidade de uma ampla variedade de portais de informação resultantes das novas tecnologias e plataformas de governo eletrônico.

Nesse cenário, as novas plataformas de governo eletrônico, segundo Holmes (2001) citado por Bermejo (2004), podem dispor de uma gama de sistemas de conhecimento especialmente projetados para análise e indução de redes de pesquisa.

Um dos exemplos recentes é o conjunto de sistemas de conhecimento da Plataforma Lattes, uma arquitetura de informações em CT&I desenvolvida para o CNPq gerir suas atividades de fomento e para integrar em um mesmo ambiente os diversos atores ligados ao sistema nacional de inovação do país. Além de viabilizar a interoperabilidade dos sistemas de informação das agências federais, a Plataforma Lattes tem racionalizado o processo de gestão de CT&I.

A estrutura arquitetônica da Plataforma Lattes é composta de níveis conceituais (camadas) levados à prática por meio de instrumentos e métodos que compreendem desde o arquivo de dados sistematizados nas unidades de análise até a extração de conhecimento referente à informação nacional sobre CT&I (REDE SCIENTI, 2003).

As informações de currículos, grupos de pesquisa e projetos em CT&I realizados no país são públicas e permitem a extração de novos conhecimentos sobre C&T (e.g., ROMÃO *et alii*, 2000).

No âmbito das redes de cooperação, interessam particularmente as camadas da arquitetura Lattes em que se encontram sistemas, algoritmos e metodologias dedicados à gestão e geração de conhecimento. Nesse contexto, podem-se identificar dois campos de estudos: (a) proposição de algoritmos e sistemas no âmbito do projeto Lattes; (b) propostas que revelam novos conhecimentos sobre a atividade científica e tecnológica no país a partir da análise de fontes de informação fornecidas pela Plataforma (e.g., ROMÃO, 2002) ou propostas de novos modelos em avaliação de C&T, com testes baseados nas informações das bases Lattes (e.g., NIEDERAUER, 2002; GONÇALVES, 2000).

Nos últimos dois anos, também são buscados, nas bases Lattes, subsídios à aplicação de instrumentos de análise de redes de colaboração técnico-científica, permitindo tanto a replicação de estudos encontrados na literatura (e.g., análise de redes de co-autoria – BALANCIERI, 2004), novas pesquisas sobre essas redes, especialmente quando se consideram as metodologias recentes de análise de redes sociais (e.g., BARABÁSI, 2003).

No âmbito da Plataforma Lattes, três sistemas de conhecimento tratam da análise de redes: Lattes Egressos, Lattes Colaboradores e Lattes Redes-GP. Esses sistemas são descritos a seguir, juntamente com o Diretório Lattes de Teses e Dissertações, que tem recursos para induzir a formação de redes.

LATTES EGRESSOS

Um dos objetivos da Plataforma Lattes foi o atendimento à demanda de informação de todos os atores de um sistema nacional de informação em CT&I. Entre eles, estão os cursos de formação acadêmica e, mais especificamente, seus gestores. Uma das informações essenciais à tomada de decisão sobre seus cursos é, também, uma das mais difíceis de se obter: o perfil dos egressos do curso. Onde estão? Em que estão trabalhando? O que já realizaram em termos de C&T? Qual é, enfim, o perfil dos alunos já formados por um curso? São perguntas para as quais as respostas estão geralmente distantes dos gestores dos cursos, em virtude do fato de os egressos estarem além do limite de atuação do curso presente.

O Sistema Lattes Egressos, segundo Bovo (2004), foi concebido para ajudar os gestores dos cursos a obter informações sobre os egressos de seu curso que continuam ligados ao sistema nacional de C&T. Baseado na técnica de Link Analysis, o sistema permite a apresentação de análises sobre o perfil de egressos de cursos de uma instituição. As informações são apresentadas de forma gráfica no site do Lattes Egressos (Disponível em: <<http://lattes.cnpq.br/lattesegressos>>).

Para utilizá-lo, os interessados escolhem a instituição e o curso que desejam analisar e o critério a partir do qual gostariam de estudar o perfil dos egressos (figura 4).

Com a apresentação em grafo espacial, há diversas possibilidades de análise sobre egressos, entre elas: atuação profissional, área de atuação, endereço profissional, formação acadêmica, formação complementar, identificação e produção C&T. Um exemplo é o perfil de atuação atual dos egressos, segundo a área do conhecimento em que os egressos do curso atuam no momento (figura 5). Nessa figura, o vértice central representa a universidade de origem

FIGURA 4
Site do Lattes Egressos

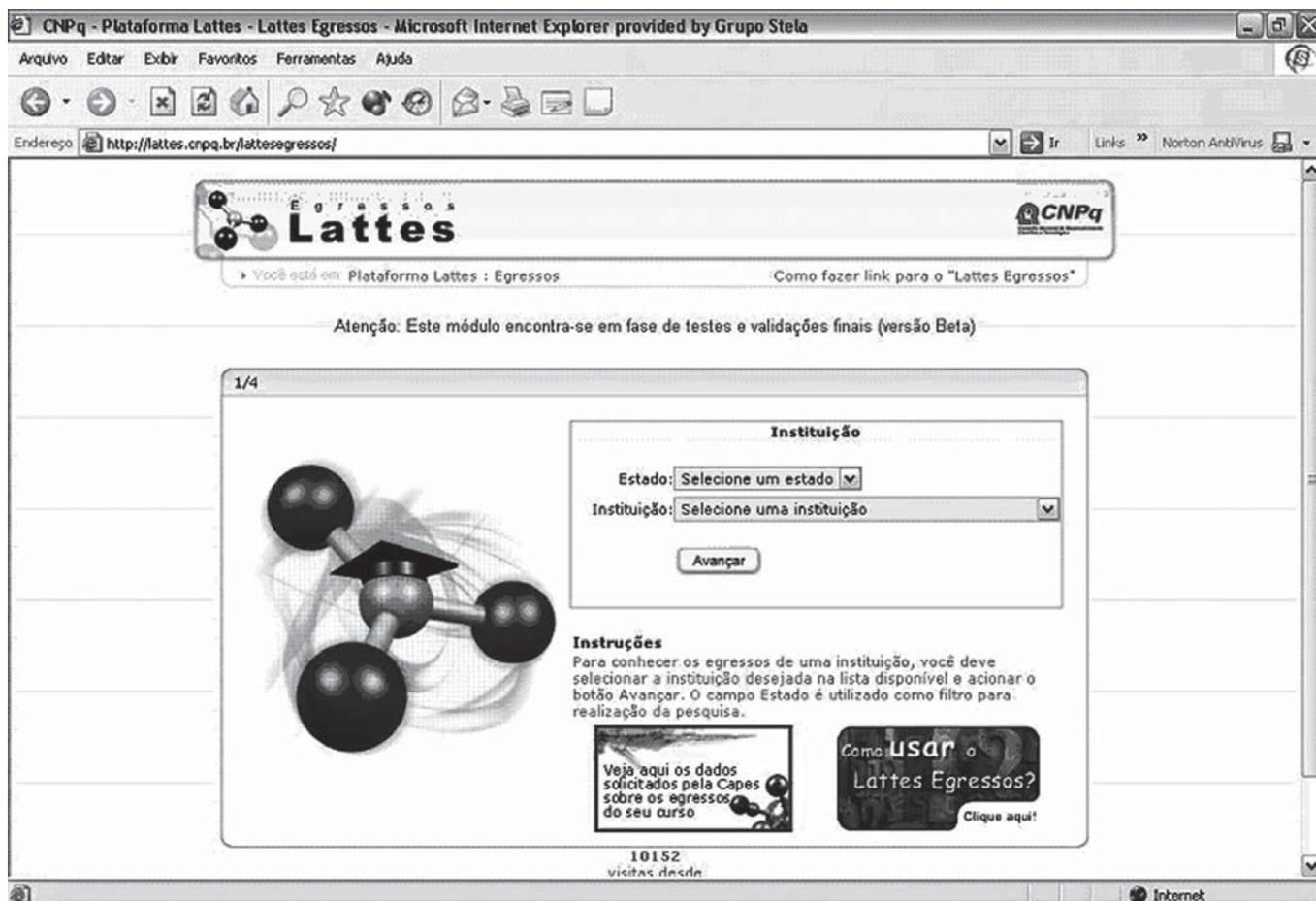
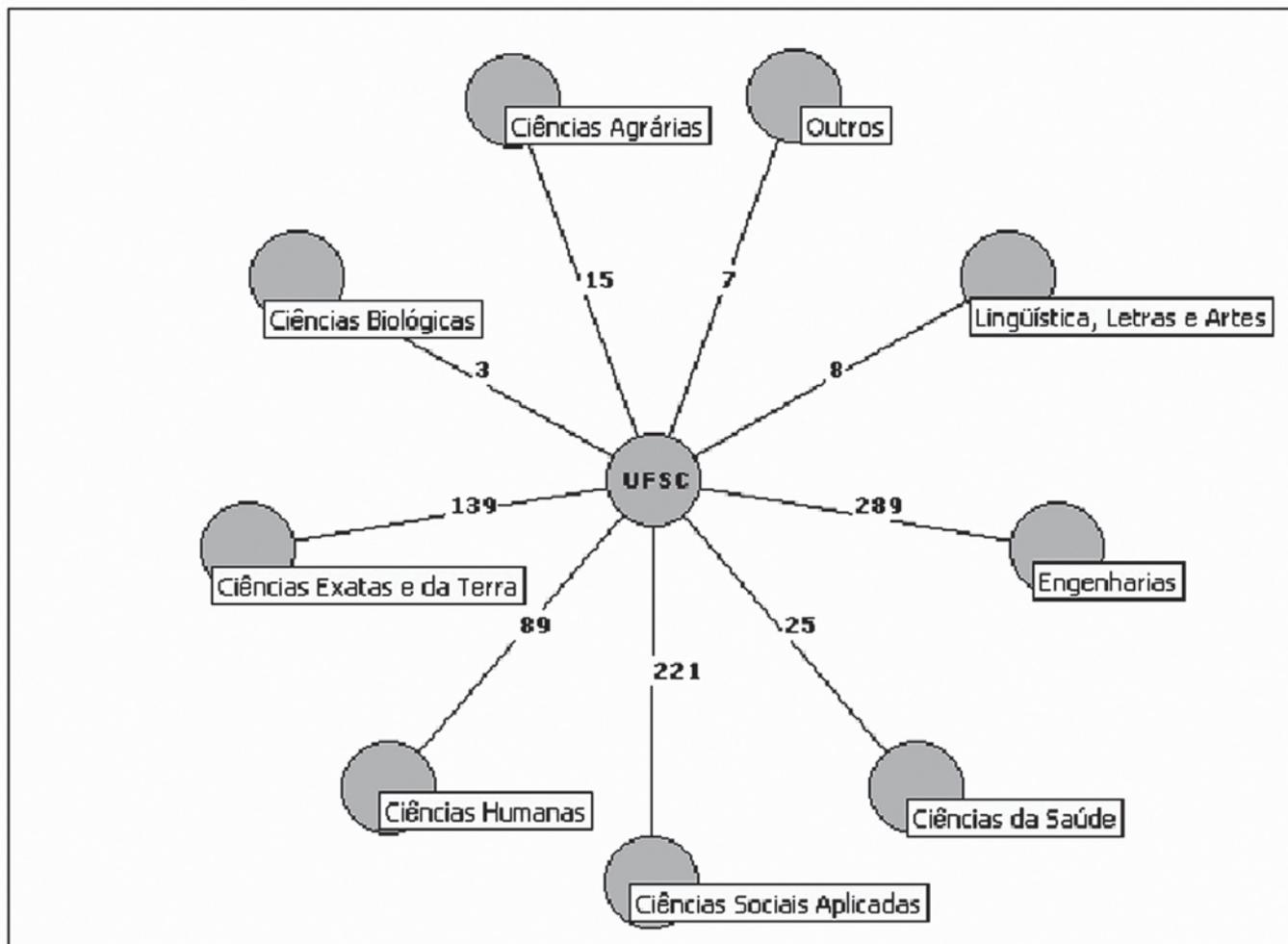


FIGURA 5

Visualização de um grafo no Lattes Egressos: grandes áreas do conhecimento em que atuam os formados do curso no momento



Egressos da instituição UFSC, da área de Engenharia de Produção, nível doutorado, distribuídos por grande área de atuação.

dos egressos, e os demais vértices representam as grandes áreas de conhecimento em que esses egressos estão atuando. Os valores representados nas ligações entre os vértices correspondem ao número de egressos em cada grande área. Clicando-se nesses valores, é possível verificar a lista dos egressos, sua titulação máxima e respectivo *e-mail*.

Um comparativo com cursos de outras instituições da mesma área do conhecimento permite ver se o curso em análise está formando profissionais de caráter multidisciplinar ou se está se concentrando em seu domínio de conhecimento.

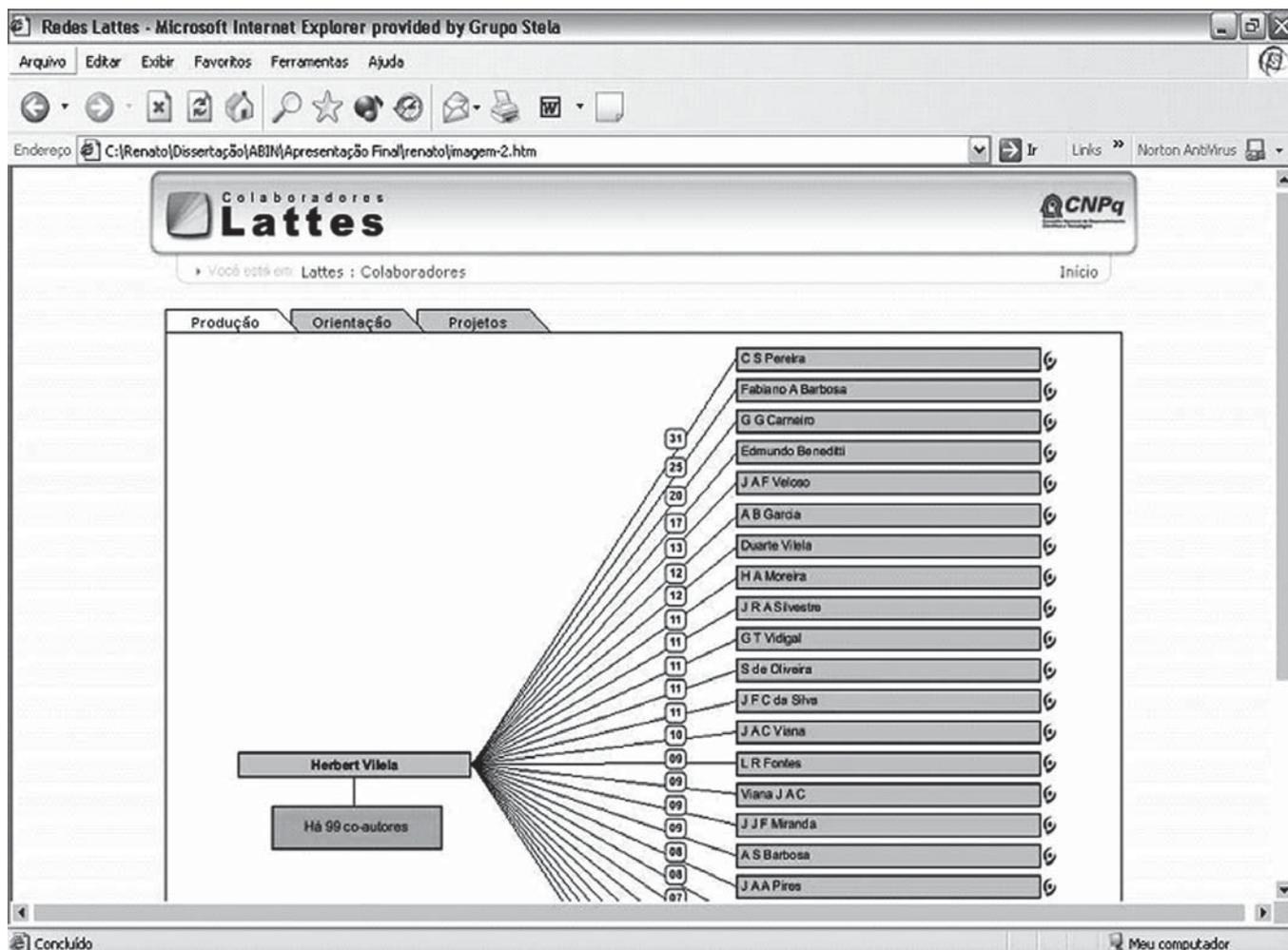
LATTES COLABORADORES

Centrado na unidade de informação curricular, o *Sistema Lattes Colaboradores* permite a visualização espacial

de diversas relações dedutíveis a partir dos registros curriculares.

O sistema permite apresentar sub-redes de colaborações de pesquisadores formadas mediante o relacionamento de co-autoria em produção C&T, orientação e participação em projetos. Essas redes de co-autoria são representadas por um grafo valorado em que os vértices são os pesquisadores e as ligações entre os vértices são os totais de co-autorias entre cada par de co-autores (figura 6). Nessa figura, são apresentados os colaboradores (vértices mais à direita) de um determinado pesquisador (vértice mais à esquerda). Os valores representados nas ligações entre os vértices correspondem à quantidade de produção científica realizada em co-autoria. Para visualizar essas produções, basta clicar sobre esses valores.

FIGURA 6
Site do Lattes Colaboradores



Com o uso de Link Analysis, é possível estudar, de forma dinâmica e configurável, todas as relações passíveis de serem analisadas na forma de grafos. O objetivo é permitir que se possa compreender o significado das interconexões entre os co-autores ao visualizar as ligações e, especialmente, que se possa descobrir relações que não se compreendiam ou cuja explicação é mais bem apresentada por meio de grafos, ou ainda validar as informações contidas no currículo Lattes (e.g., inclusão de co-autores faltantes no currículo, a partir da observação do autor de que sua sub-rede não reflete a realidade de suas interações sociais em pesquisa).

LATTES REDES-GP

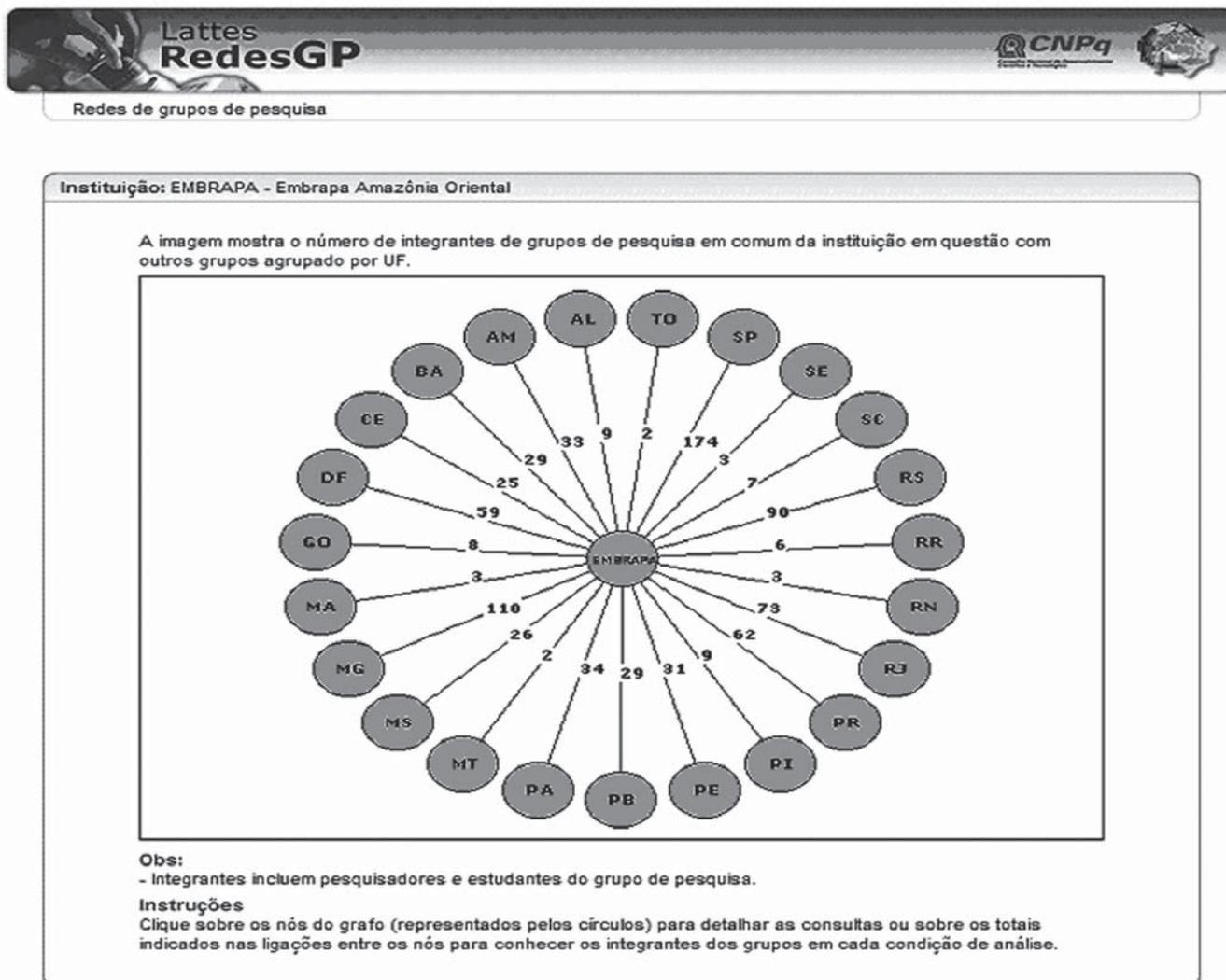
Como mencionado anteriormente, entre os sistemas de conhecimento da Plataforma Lattes estão aqueles que têm por objetivo aprofundar a análise sobre as formas de colaboração científica em um sistema nacional de

inovação. Um dos projetos que se relacionam com essa questão denomina-se "Lattes Redes-GP". O acrônimo "GP" indica grupos de pesquisa, e o objetivo desse projeto é a construção de sistemas de análise da atividade científica e tecnológica organizada na forma de redes de grupos de pesquisa. Utiliza como base as informações do Diretório de Grupos de Pesquisa e tem fundamentos na área de cienciometria e em algoritmos de extração de conhecimento.

Um dos focos de pesquisa deste projeto é o estudo de redes de pesquisa formadas pela participação de pesquisadores/estudantes em mais de um grupo. Dessa forma, é possível verificar o grau de colaboração entre os grupos de pesquisa por áreas de conhecimento e/ou instituições mediante a participação de seus integrantes. A figura 7, por exemplo, mostra o número de integrantes (valor da aresta) em comum da Embrapa (vértice central) com outros grupos de pesquisa agrupados por UF (demais vértices).

FIGURA 7

Visualização de um grafo no Lattes RedesGP: número de integrantes de grupos de pesquisa em comum da instituição relacionado com outros grupos organizados por UF



Os diferentes grafos formados pelas análises do sistema Lattes Redes-GP podem permitir a aplicação de algoritmos usuais da literatura de redes sociais. Um exemplo é o cálculo da “distância” entre um pesquisador e outro nas redes de pesquisa (e.g., Número de Erdős – NEWMAN, 2000). Outro exemplo é a busca por *hubs* (vértices centrais) em redes de relacionamentos, ou seja, pesquisadores a partir dos quais se pode chegar a todos os demais integrantes da rede em análise em poucas ligações (e.g., BALANCIERI, 2004).

O acrônimo “GP” também sugere que as mesmas análises e sistemas de conhecimento possam ser desenvolvidos para outras unidades de informação da Plataforma Lattes, como “Redes-CV”, “Redes-PP” ou “Redes-PC” para currículos, projetos de pesquisa e periódicos científicos, respectivamente.

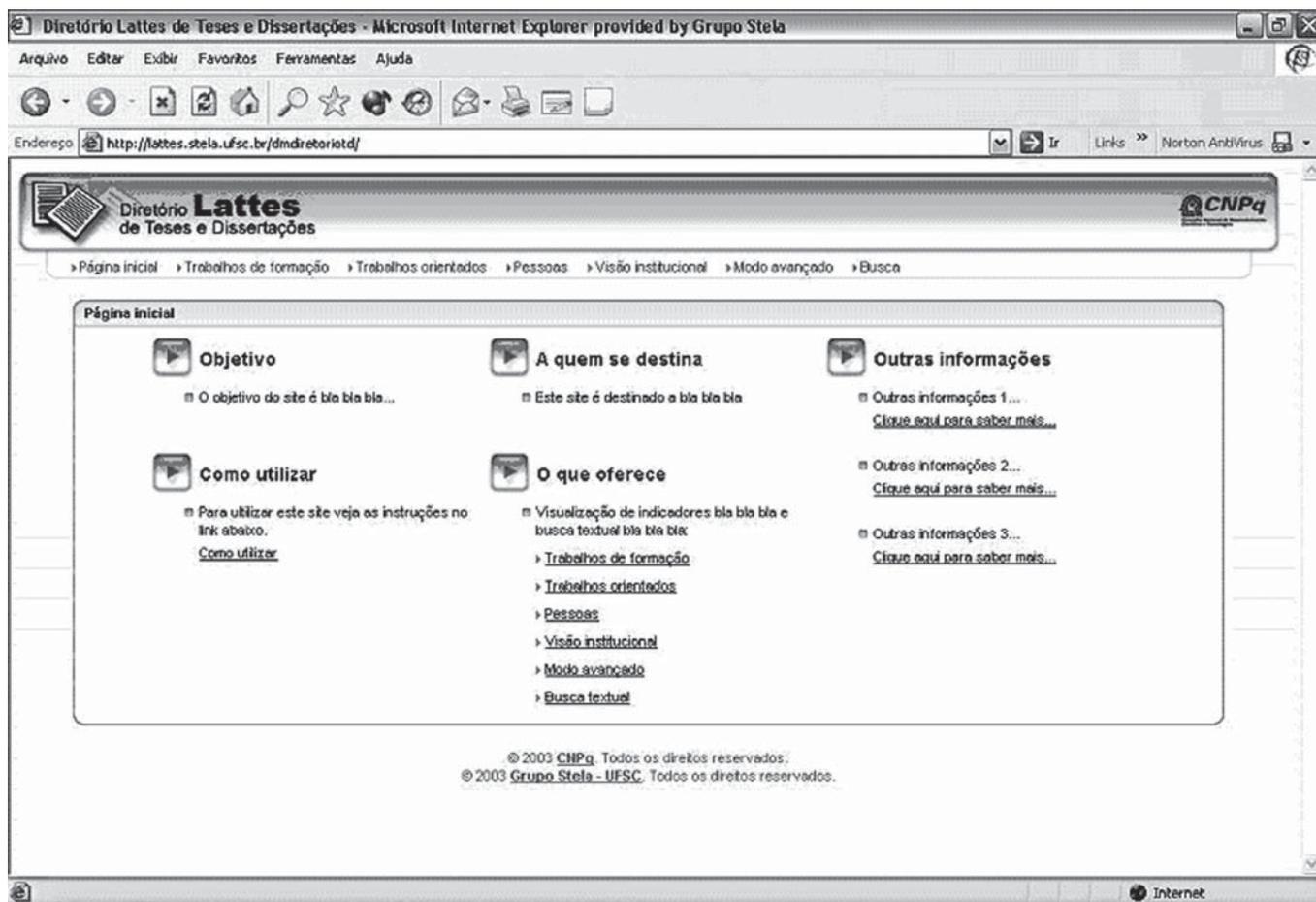
DIRETÓRIO LATTES DE TESES E DISSERTAÇÕES

Analisar e apresentar relacionamentos existentes entre pesquisadores, grupos ou projetos não são as únicas possibilidades que as novas TICs trazem à análise de redes sociais. A indução de redes sociais também é um dos principais recursos viabilizados por ambientes de cooperação e por comunidades de prática em temas específicos.

É nesse contexto que se elaborou o projeto Diretório de Teses e Dissertações (figura 8). O projeto explora as diversas possibilidades de interatividade que se relevam a partir das informações de trabalhos de conclusões nos níveis de graduação, pós-graduação *lato* e *stricto sensu*. Esse diretório traz indicadores e informações à totalidade de produtores de conhecimento (estudantes, docentes

FIGURA 8

Site do Diretório Lattes de Teses e Dissertações



e pesquisadores), gestores de C&T e a interessados em produtos de cursos de formação no país. Buscas sobre participações em bancas, teses e atividades de orientação podem tanto revelar cooperações não declaradas em publicações conjuntas, quanto estabelecer fontes de contatos entre interessados. O passo seguinte a esse tipo de sistema é a formação de comunidades de prática, nas quais os recursos são especificamente formatados para a cooperação dos pares (RHEINGOLD, 1994; PREECE, 2000).

CONSIDERAÇÕES FINAIS

Há duas áreas de desenvolvimento que adquirem significado especial quando o objetivo é impulsionar a cooperação técnico-científica: redes sociais e TICs.

As quatro décadas de estudos em redes sociais na pesquisa produziram uma definição-referência para redes de colaboração técnico-científica (i.e., quem são seus vértices, vínculos e quais são as formas de relacionamento social, cognitivo e profissional entre esses vértices).

Em contapartida, as novas TICs estabeleceram plataformas de governo eletrônico que se consolidaram no mapeamento das atividades técnico-científicas dos sistemas nacionais de inovação, produzindo fontes de informação indisponíveis antes do advento da Internet e, mais recentemente, sistemas de conhecimento que multiplicaram os recursos de gestão do conhecimento em CT&I.

Esse cenário abre novas e estimulantes possibilidades à indução, à análise e ao fomento a redes de pesquisa. A aplicação de TICs permite, por exemplo, revelar relacionamentos ocultos, inspecionar características das diferentes formas de cooperação técnico-científica e estabelecer subsídios à tomada de decisão sobre o fomento à cooperação.

Neste artigo, reviram-se os principais desenvolvimentos que permitem a formação do marco referencial para a definição de trabalho em redes científicas e discutiram-se as novas possibilidades das TICs com base em exemplos já existentes de aplicação dessas abordagens a um sistema nacional de informações em C&T (Plataforma Lattes).

É relevante observar que as informações e recursos viabilizados pelas novas TICs ainda não fazem parte dos processos de gestão da C&T. Entretanto, é de se esperar que a aplicação continuada das TICs às redes em C&T e a relevância da atividade coletiva no processo de inovação possam alterar esse quadro. Informações sobre a capacidade de multiplicação de conhecimento, facilidade de cooperação e alcance das relações interpessoais poderão, no futuro, constar entre os indicadores que subsidiam as decisões de agências de fomento quanto ao financiamento e planejamento em C&T.

RECONHECIMENTO

Este trabalho consiste em um dos resultados das atividades realizadas para o CNPq pelo Grupo Stela (UFSC), entre 2002 e 2004, referentes à pesquisa e ao desenvolvimento da Plataforma Lattes.

Artigo recebido em 24/01/2005 e aceito
para publicação em 15/08/2005.

REFERÊNCIAS

- BALANCIERI, Renato. *Análise de redes de pesquisa em uma plataforma de gestão em ciência e tecnologia: uma aplicação à Plataforma Lattes*. 2004. Dissertação (Mestrado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2004. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?5621>>. Acesso em: 5 jan. 2005.
- BARABÁSI, Albert-László. *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*. [S. l.]: A Plume Book, 2003.
- BEAVER, Donald B.; ROSEN, R. Studies in scientific collaboration: part III: professionalization and the natural history of modern scientific co-authorship. *Scientometrics*, v. 1, p. 231-245, 1979.
- _____. Studies in Scientific Collaboration: part I: the professional origins of scientific co-authorship. *Scientometrics*, v. 1, p. 64-84, 1978.
- BERMEJO, Paulo Henrique. *Metodologia para definição de unidades de informação para plataforma de governo eletrônico: uma aplicação à Plataforma Lattes*. 2004. Dissertação (Mestrado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2004. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?5799>>. Acesso em: 5 jan. 2005.
- BOAVENTURA NETTO, Paulo Oswaldo. *Grafos: teoria, modelos e algoritmos*. 2. ed. São Paulo: E. Blücher, 2001.
- BOVO, Alessandro Botelho. *Um método de tradução de fontes de informações em um formato padrão que viabiliza a extração de conhecimento por meio de link analysis e teoria dos grafos*. 2004. Dissertação (Mestrado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2004. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?5798>>. Acesso em: 5 jan. 2005.
- COLONOMOS, Ariel. *Sociologie des réseaux transnationaux; communautés, entreprises et individus: lien social et système international*. Paris: l'Harmattan, 1995.
- CRANE, Diana. *Invisible colleges*. Chicago: University of Chicago Press, 1972.
- FRAME, J. D.; CARPENTER, M. P. Internacional research collaboration. *Social Studies of Science*, v. 9, p. 481-497, 1979.
- GOFFMAN, William; WARREN, Kenneth S. *Scientific information systems and the principle of selectivity*. New York: Praeger, 1980. p. 127.
- GONÇALVES, A. L. *Utilização de técnicas de mineração de dados em bases de C&T: uma análise dos grupos de pesquisa no Brasil*. 2000. Dissertação (Mestrado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2000. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?1136>>. Acesso em: 11 jan. 2005.
- HAGSTROM, Warren O. *The scientific community*. New York: Basic Books, 1965.
- KATZ, J. Sylvan. Geographical proximity and scientific collaboration. *Scientometrics*, v. 31, n. 1, p. 31-43, 1994.
- _____. *Bibliometric assessment of intranational university-university collaboration*. 1993. Tese (Doutorado em Ciências Políticas)- Science Policy Research Unit, University of Sussex, Brighton, UK, 1993.
- KATZ, J. Sylvan; MARTIN, Ben R. What is Research Collaboration?. *Research Policy*, v. 26, p. 1-18, 1997.
- KODAMA, F. Technology fusion and the new R&D. *Harvard Business Review*, p. 70-78, July/Aug. 1992.
- LAWANI, S. M. Some bibliometric correlates of quality in scientific research. *Scientometrics*, v. 9, p. 325-342, 1986.
- LECLERC, M. et al. Scientific co-operation between Canada and European Community. *Science and Public Policy*, v. 19, n. 1, p. 15-24, Feb. 1992.
- LIPNAK, Jéssica; STAMP, Jeffrey. *Networks, redes de conexão: pessoas conectando-se com pessoas*. São Paulo: Aquarela, 1992.
- LODAHL, Janice B.; GORDON, Gerald. The structure of scientific fields and the functioning of university graduate departments. *American Sociological Review*, v. 37, p. 57-72, 1972.
- LUUKONEN, T.; PERSSON, O.; SIVERTSEN, G. Understanding patterns of scientific collaboration. *Science, Technology and Human Values*, v. 17, p. 101-126, 1992.
- MARTINS, Sandra Regina. *Metodologia de geração automática de textos para gestão de conhecimento em plataformas de governo eletrônico: um estudo de caso na Plataforma Lattes*. 2004. Dissertação (Mestrado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2004. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?6074>>. Acesso em: 5 jan. 2005.
- MEDOWS, A. J.; O'CONNOR, J. G. Bibliographic statistics as a guide to growth point in science. *Science Studies*, v. 1, p. 95-99, 1971.
- MILGRAM, Stanley. The small world problem. *Psychology Today*, v. 1, p. 61-67, 1967.
- NARIN, Francis; WHITLOW, Edith S. *Measurement of scientific cooperation and coauthorship in CEC-Related areas of science*. Luxembourg: Office for Official Publications of the European Communities, 1990. Report EUR 12900, v. 1.

A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na Plataforma Lattes

- _____; STEVENS, Kimberly; WHITLOW, Edith S. Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, v. 21, n. 3, p. 313-323, 1991.
- NEWMAN, M. E. J. *Clustering and preferential attachment in growing networks*. Santa Fé: The Santa Fé Institute, 2001a. Paper 01-03-021.
- _____. The structure of scientific collaboration networks. *Proceedings of National Academy Sciences*, p. 404-409, 2001b.
- _____. *Who is the best connected scientist?: a study of scientific coauthorship networks*. Santa Fé: The Santa Fé Institute, 2000. Paper 00-12-064.
- _____; STROGATZ, S. H.; WATTS, D. J. *Random graphs with arbitrary degree distributions and their applications*. Santa Fé: The Santa Fé Institute, 2001. Paper 00-07-042.
- NIEDERAUER, C. A. P. *Ethos: um modelo para medir a produtividade relativa de pesquisadores baseado na análise por envoltória de dados*. 2002. Tese (Doutorado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2002. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?3060>>. Acesso em: 11 jan. 2005.
- NUDELMAN, A. E.; LANDERS, C. E. The failure of 100 divided by 3 to equal 33 1/3. *The American Sociologist*, v. 7, n. 9, 1972.
- PRAVDIC, N.; OLUIC-VUKOVIC, V. Dual approach to multiple authorship in the study of collaborator/scientific output relationship. *Scientometrics*, v. 10, p. 259-280, 1986.
- PREECE, Jenny. *Online communities: designing, usability, supporting sociability*. [S. l.]: Wiley, 2000.
- PRICE, Derek J. de Solla. *Little science, big science*. New Cork: Columbia University Press, 1963.
- _____; BEAVER, Donald B. Collaboration in an invisible college. *American Psychologist*, v. 21, p. 1011-1018, 1966.
- REDE SCIENTI. *Rede internacional de fontes de informação e conhecimento para a gestão de ciência, tecnologia e inovação: documento de referência*. Disponível em: <<http://www.scienti.info/DocReferencia.pdf>>. Acesso em: 22 set. 2003.
- RHEINGOLD, Howard. *La comunidad virtual: una sociedad sin fronteras*. Barcelona: Gedisa Editorial, 1994. (Colección Limites de La Ciencia).
- ROMÃO, W. *Descoberta de conhecimento relevante em bancos de dados sobre ciência e tecnologia*. 2002. Tese (Doutorado em Engenharia de Produção)– Universidade Federal de Santa Catarina, 2002. Disponível em: <<http://teses.eps.ufsc.br/Resumo.asp?3150>>. Acesso em: 12 jan. 2005.
- _____ et al. Planejamento em C&T: uma abordagem para descoberta de conhecimento relevante em banco de dados de grupos de pesquisa. *Revista Tecnológica*, Maringá, v. 9, p. 139-152, 2000.
- SELL, Denilson et al. A framework to improve semantic web services discovery and integration in an e-gov knowledge network. *Lecture Notes in Computer Science*, Heidelberg, v. 3257, 2004.
- SMITH, Michael. The trend toward multiple authorship in psychology. *American Psychologist*, v. 13, p. 596-599, 1958.
- STEFANIAK, B. Individual and multiple authorship of papers in chemistry and physics. *Scientometrics*, v. 4, p. 331-337, 1982.
- STORER, Norman W. The internationality of science and the nationality of scientists. *International Science Journal*, v. 22, p. 87-104, 1970.
- SUBRAMANYAM, K. Bibliometric study of research collaboration: a review. *Journal of Information Science*, v. 6, p. 35-59, 1983.
- WATTS, Duncan. *Six degrees: the science of a connected age*. New York: W. W. Norton & Company, 2003.
- WEISZ, J.; ROCO, M. C. *Redes de pesquisa e educação em engenharia nas américas*. Rio de Janeiro: FINEP, 1996.