

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em
Engenharia de Produção

**ANÁLISE DE REDES DE PESQUISA EM UMA PLATAFORMA DE
GESTÃO EM CIÊNCIA E TECNOLOGIA: UMA APLICAÇÃO À
PLATAFORMA LATTES**

Renato Balancieri

Dissertação apresentada ao
Programa de Pós-Graduação em
Engenharia de Produção da
Universidade Federal de Santa Catarina
como requisito parcial para obtenção
do título de Mestre em
Engenharia de Produção

Florianópolis
2004

Renato Balancieri

**ANÁLISE DE REDES DE PESQUISA EM UMA PLATAFORMA DE
GESTÃO EM CIÊNCIA E TECNOLOGIA: UMA APLICAÇÃO À
PLATAFORMA LATTES**

Esta dissertação foi julgada e aprovada para a obtenção do grau de **Mestre em Engenharia de Produção** no **Programa de Pós-Graduação em Engenharia de Produção** da Universidade Federal de Santa Catarina

Florianópolis, 2 de abril de 2004.

Prof. Edson Pacheco Paladini, Dr.
Coordenador do Programa

BANCA EXAMINADORA

Prof. Roberto C. S. Pacheco, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof. Vinícius Medina Kern, Dr.
Universidade Federal de Santa Catarina

Prof. Aran Bey Tcholakian Morales, Dr.
Universidade Federal de Santa Catarina

*À minha amada esposa e companheira,
Adriana, que alimenta meu coração
com seu amor e carinho a cada dia.
Aos meus pais, João e Zelma, a quem
muito devo pelo seu amor e confiança.*

Agradecimentos

Desejo manifestar minha enorme gratidão a todos que direta ou indiretamente contribuíram para a realização deste trabalho.

Primeiramente a Deus, por ter me dado a vida, a força necessária em todos os momentos, iluminação, saúde e imensa proteção. Por ter me presenteado com uma esposa maravilhosa, uma ótima família e excelentes amigos.

Ao meu orientador, professor Roberto Carlos dos Santos Pacheco, e aos professores Ricardo Miranda Barcia e Carlos Alberto Pittaluga Niederauer, pessoas formidáveis que contribuíram muito para o desenvolvimento do trabalho e para minha formação como pessoa e pesquisador.

Às professoras Maria Madalena Dias, Tânia Calvi Tait e Elisa Hatsue Moriya, da Universidade Estadual de Maringá. Sem o incentivo de vocês, dificilmente teria deixado Maringá rumo a Florianópolis para realizar o mestrado.

Ao Grupo Stela, onde foi possível a realização deste trabalho.

Aos amigos que de alguma maneira participaram deste trabalho e contribuíram para a sua realização, em especial a Alessandro Botelho Bovo, Marlon Candido Guérios, Paulo Henrique de Souza Bermejo, Wagner Watanabe e a todos os integrantes do Grupo de Oração Universitário (GOU), que me apoiaram e me incentivaram com suas orações.

Aos meus pais, João e Zelma, meus sogros, Otávio e Tereza, e demais familiares, pelo incentivo, confiança e exemplos de força e dedicação.

Finalmente, gostaria de agradecer à minha esposa, por quem tenho profunda admiração e que tem estado sempre ao meu lado, cheia de amor, compreensão, carinho, apoio e incentivo. Obrigado, Adriana Beletato dos Santos Balancieri.

A maravilhosa disposição e harmonia do universo só pode ter tido origem segundo o plano de um Ser que tudo sabe e que tudo pode. Isto fica sendo a minha última e mais elevada descoberta.

Isaac Newton

RESUMO

BALANCIERI, Renato. **Análise de Redes de Pesquisa em uma Plataforma de Gestão em Ciência e Tecnologia**: Uma Aplicação à Plataforma Lattes. 2004. 117 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis 2004.

A colaboração científica é um empreendimento que envolve metas comuns, esforço coordenado e resultados ou produtos através dos quais os colaboradores compartilham a responsabilidade e o crédito. Essas redes têm, em sua base, relação com plataformas de gestão de Ciência, Tecnologia e Inovação (CT&I), ou seja, incluem a escolha de uma fonte de informação e de uma possibilidade de análise sobre como redes se estruturam. Assim, o objetivo desta dissertação é combinar a importância da colaboração para CT&I, os estudos disponíveis e a possibilidade de efetivá-los por meio de instrumentos que se valham da existência de uma plataforma de informação na área. No estudo realizado através das redes de relacionamentos pôde-se evidenciar um fenômeno conhecido na literatura como “mundo pequeno”, que tem vértices centrais (hubs), os quais possibilitam alcançar toda a rede em poucas ligações. A técnica empregada foi a de Link Analysis, que permitiu estudar essas redes de relacionamentos. O trabalho apresentou a detecção desses hubs em redes de co-autoria de pesquisa. Inicialmente foram selecionados pesquisadores, representados como vértices, e construídos links entre pesquisadores que possuíam co-autoria. Em seguida, restringiu-se a exibição de links para vértices com certo número mínimo de links e, assim, identificaram-se os hubs. O experimento abordou os 36 pesquisadores de Engenharia de Produção que possuem o maior número de itens de produção em seus currículos Lattes, e a análise revelou que os três hubs pertencem ao mesmo programa de pós-graduação. Esse tipo de resultado permite aumentar o nível de conhecimento sobre as redes de relacionamento.

Palavras-chave: redes de colaboração; redes de pesquisa; plataforma de gestão de CT&I; instrumentos de análise.

ABSTRACT

BALANCIERI, Renato. **Análise de Redes de Pesquisa em uma Plataforma de Gestão em Ciência e Tecnologia**: Uma Aplicação à Plataforma Lattes. 2004. 117 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis 2004.

The scientific collaboration is an enterprise that involves common goals, coordinated efforts and results or products through which the collaborators share the responsibility and the credit. These networks have, in their base, relation with Science, Technology and Innovation management platforms (CT&I). In other words, they include the choice of an information source and of an analysis possibility of the way networks are structured. Thus, the goal of this dissertation is to combine the importance of the collaboration for CT&I, the studies available in the field and the possibility of making them effective by means of instruments that count on the existence of an information platform in the area. In the present study, it is highlighted, through the use of relationship networks, a phenomenon known in the literature as “small world”, which has central vertexes (hubs) that make it possible to cover the whole network by using few links. The technique employed was “Link Analysis”, which allowed the study of these relationship networks. This study showed the detection of these hubs in research co-authorship networks. Firstly, the researchers, represented by vertexes, were selected and links among researchers with co-authorship were built. Secondly, the exhibition of links was restricted to a given (minimum) number of links and then the hubs were identified. The experiment approached the 36 most productive Brazilian researchers in the Production Engineering field, according to their Lattes curriculums. The analysis revealed that all three researchers (hubs) belong to the same graduate program. This result allows enhancing the level of knowledge about relationship networks.

Keywords: collaboration networks; research networks; management platforms CT&I.; analysis instruments.

LISTA DE FIGURAS

Figura 1.1 - Metodologia do trabalho	20
Figura 2.1 - Abrangência dos objetos de estudo da bibliometria, cienciometria e informetria (PACHECO; KERN, 2001).....	40
Figura 2.2 - Rede de Citação.....	41
Figura 2.3 - Grafo Regular: “Mundo Paroquiano” (WATTS; STROGATZ, 1998)	50
Figura 2.4 - Grafo Aleatório: “Mundo Estilhaçado” (WATTS; STROGATZ, 1998).....	50
Figura 2.5 - Modelo do “Mundo Pequeno” (WATTS; STROGATZ, 1998).....	52
Figura 2.6 - Cronologia de Redes de Colaboração Científica	63
Figura 3.1 - Arquitetura Conceitual da Plataforma Lattes (Fontes: STELA, 2002).....	70
Figura 3.2 - Unidades de Análise da Plataforma Lattes (Fonte: STELA, 2002).....	72
Figura 3.3 - Diagrama esquemático da representação de informações sobre currículos segundo a LMPL (Fonte: CNPq, 2003)	74
Figura 3.4 - Detalhamento do diagrama da Figura 3.3, mostrando a organização dos dados gerais de qualquer currículo (Fonte: CNPq, 2003).....	75
Figura 3.5 - Diagrama esquemático da representação de informações sobre grupo de pesquisa segundo a LMPL (Fonte: CNPq, 2003).....	75
Figura 3.6 - Diagrama esquemático da representação de informações sobre projeto de pesquisa segundo a LMPL (Fonte: CNPq, 2003).....	76
Figura 3.7 - Diagrama de Venn usando um conjunto	79
Figura 3.8 - Diagrama de Venn usando dois conjuntos.....	79
Figura 3.9 - Diagrama de Venn usando três conjuntos	80
Figura 3.10 - Diagrama de Venn usando quatro conjuntos em formato de círculo	80
Figura 3.11 - Diagrama de Venn usando quatro conjuntos em formato de elipse	81
Figura 3.12 - Exemplo de aplicação de Link Analysis em CT&I	82
Figura 3.13 - Exemplo de um grafo.....	84
Figura 3.14 - Exemplo de subgrafos	85
Figura 3.15 - Exemplo de Grafo Orientado	85
Figura 3.16 - Exemplo de um grafo com Caminhos	86
Figura 3.17 - Exemplo de um grafo com um ciclo	86
Figura 3.18 - Exemplo de Grafo Conexo (G1) e Grafo Desconexo (G2)	87
Figura 3.19 - Exemplo de Grafo Orientado fortemente conexo	87
Figura 3.20 - Exemplo de grau de entrada 2 (dois) e grau de saída 1 (um)	88
Figura 3.21 - Grafo com dois cliques: {2,3,4} e {4,5,6,7}	88
Figura 3.22 - Grafo Planar com 5 (cinco) regiões.....	89

Figura 4.1 - Interface Principal do Sistema Lattes Redes.....	103
Figura 4.2 - Resultado da busca por área de atuação	104
Figura 4.3 - Visualização de uma Rede de Co-Autoria com todos os relacionamentos...	105
Figura 4.4 - Grafo Aleatório: “Mundo Estilhaçado” (WATTS; STROGATZ, 1998).....	105
Figura 4.5 - Visualização de uma Rede de Co-Autoria com vértices de grau 7.....	106
Figura 4.6 - Visualização de uma Rede de Co-Autoria com vértices de grau 8.....	107
Figura 1a - Visualização de uma Rede de Co-Autoria	115
Figura 1b - Visualização de uma Rede de Co-Autoria com vértices de grau 5.....	115
Figura 1c - Visualização de uma Rede de Co-Autoria com vértices de grau 6	116
Figura 1d - Visualização de uma Rede de Co-Autoria com vértices de grau 7.....	116
Figura 1e - Visualização de uma Rede de Co-Autoria com vértices de grau 8.....	117

LISTA DE TABELAS

Tabela 2.1 - Algumas medidas estatísticas de Redes Sociais (Fonte: UCINET V).....	28
Tabela 2.2 - Diferentes níveis de colaboração e distinção entre as formas inter e intra (KATZ; MARTIN, 1997).....	35
Tabela 2.3 - Publicações do período até a década de 1960	55
Tabela 2.4 - Publicações do período da década de 1970	57
Tabela 2.5 - Publicações do período da década de 1980	58
Tabela 2.6 - Publicações a partir de 1990.....	62
Tabela 3.1 - Algumas questões sobre Link Analysis	83
Tabela 4.1 - Redes de Pesquisa sob o aspecto das unidades de análise.....	91
Tabela 4.2 - Relacionamentos entre as unidades de informação de uma plataforma de gestão de CT&I.....	92
Tabela 4.3 - Possibilidades de análise de redes de pesquisa na Plataforma Lattes.....	96
Tabela 4.4 - Conceitos-chave do exemplo de Redes Institucionais de Grupos de Pesquisa	97
Tabela 4.5 - Índice de co-participações por Grande Área de Conhecimento.....	100
Tabela 4.6 - Total de co-participações entre as Grandes Áreas.....	101
Tabela 4.7 - Percentual de co-participações entre as Grandes Áreas.....	102

LISTA DE SIGLAS

CNPq	- Conselho Nacional de Desenvolvimento Científico e Tecnológico
CT&I	- Ciência, Tecnologia e Inovação
DTD	- Data Type Definition
FINEP	- Financiadora de Estudos e Projetos
FNDCT	- Fundo Nacional de Desenvolvimento Científico e Tecnológico
HTML	- Hypertext Markup Language
LMPL	- Linguagem de Marcação da Plataforma Lattes
OICYT	- Organismos Internacionais de Ciência e Tecnologia
ONCYT	- Organismos Nacionais de Ciência e Tecnologia
OPS	- Organização Pan-americana de Saúde
PADCT	- Programa de Apoio ao Desenvolvimento Científico e Tecnológico
PBDCT	- Plano Básico de Desenvolvimento Científico e Tecnológico
SCienTI	- Rede Internacional de Fontes de Informação e Conhecimento para a Gestão da Ciência, Tecnologia e Inovação
SCM	- Supply Chain Management
SGDB	- Sistemas Gerenciadores de Banco de Dados
XML	- Extensible Markup Language

SUMÁRIO

RESUMO	vi
ABSTRACT	vii
LISTA DE FIGURAS	viii
LISTA DE TABELAS	x
LISTA DE SIGLAS	xi
1 INTRODUÇÃO	14
1.2 Objetivo geral.....	16
1.3 Justificativa	17
1.4 Estrutura do trabalho.....	18
1.5 Metodologia	19
1.5.1 Classificação da pesquisa.....	21
1.5.2 Delimitação do trabalho.....	21
2 REDES DE COLABORAÇÃO E REDES DE PESQUISA	22
2.1 Introdução.....	22
2.2 Redes de colaboração nas Ciências Sociais.....	24
2.2.1 Metodologia de análise de Redes Sociais.....	25
2.3 Redes de pesquisa	30
2.3.1 Fatores que contribuem para a colaboração em Redes de Pesquisa.....	31
2.3.2 Níveis de colaboração.....	34
2.3.3 Taxonomias para o estudo de Redes de Pesquisa	37
2.3.4 Cadeias de Suprimento (<i>Supply Chain</i>) e Redes de Pesquisa	43
2.4 O Mundo Pequeno (<i>Six Degrees Separation</i>)	46
2.4.1 Breve histórico do Mundo Pequeno.....	47
2.4.2 “Mundo Paroquiano” e “Mundo Estilhaçado”	49
2.5 Redes de Colaboração e Redes de Pesquisa: Um breve histórico.....	54
2.5.1 Período até a década de 60	55
2.5.2 Período da década de 70	56
2.5.3 Período da década de 80	57
2.5.4 Período a partir de 1990.....	59
2.6 Considerações Finais.....	63
3 PLATAFORMA PARA GESTÃO DE CIÊNCIA, TECNOLOGIA E INOVAÇÃO	64
3.1 Introdução.....	64
3.2 Plataforma de Gestão de CT&I	66
3.3 Plataforma Lattes.....	67

3.3.1 Breve histórico	68
3.3.2 Arquitetura da Plataforma Lattes	69
3.3.2.1 Unidades de Informação.....	70
3.3.2.1.1 Abertura e Padronização XML.....	72
3.3.2.2 Sistemas e Fontes de Informação	76
3.3.2.3 Portais e Serviços Web	77
3.3.2.4 Sistemas de Conhecimento.....	77
3.4 Considerações Finais.....	89
4 INSTRUMENTOS PARA ANÁLISE DE REDES DE PESQUISA EM CT&I NA	
PLATAFORMA LATTES DO CNPq.....	90
4.1 Introdução.....	90
4.2 Unidades de Análise e Redes de Pesquisa.....	90
4.3 Metodologia de Análise de Redes de Pesquisa em uma Plataforma de Gestão de CT&I	92
4.4 Possibilidades de Análise de Redes de Pesquisa na Plataforma Lattes.....	94
4.4.1 Instrumentos de análise de redes de pesquisa na Plataforma Lattes	96
4.4.1.1 Aplicação do Método Proposto: Redes de Co-Participação em Grupos de Pesquisa por Instituição	97
4.4.1.2 Dados Estatísticos: Redes de Co-Participação em Grupos de Pesquisa por Grande Área do Conhecimento.....	99
4.4.1.3 Sistema Lattes Redes	102
4.5 Conclusão.....	107
5 CONCLUSÕES E RECOMENDAÇÕES	108
5.1 Conclusões	108
5.2 Trabalhos futuros.....	110
APÊNDICE A - APLICAÇÃO PARA EXTRAÇÃO DE CONHECIMENTO EM REDES DE PESQUISA DA BASE DE DADOS DO DIRETÓRIO DOS GRUPOS DE PESQUISA DO BRASIL	111
1.1 Algoritmo para cálculo de conjuntos independentes maximais de um grafo (BOAVENTURA NETTO, 1996; BERGE, 1985).....	111
1.2 A Base de Dados do Diretório dos Grupos de Pesquisa do Brasil.....	112
1.3 Todas as grandes áreas	112
1.4 Grande área “Engenharias”.....	113
1.5 Grande área “Ciências Agrárias”.....	113
APÊNDICE B – LATTES REDES.....	115

1 INTRODUÇÃO

É, com efeito, natural ao homem aspirar ao conhecimento da verdade.

São Tomás de Aquino

A ciência é uma atividade altamente institucionalizada. O conhecimento passa a ter um papel fundamental como insumo produtivo, em um processo que tem suas origens no Renascimento e se acentua fortemente na Revolução Industrial (GÓMEZ, 1987).

Sua gestão, a utilização e principalmente o acesso se tornam cada vez mais atividades econômicas críticas. Nessa ciência tão institucionalizada não existe praticamente lugar para os trabalhos individuais, capazes de darem conta de uma descoberta científica do início ao fim. A ciência é fundamentalmente um trabalho coletivo e cumulativo, em que pesquisadores e grupos de pesquisa analisam os resultados já obtidos por seus pares e têm como objetivo “acrescentar um tijolo a mais em um vasto edifício ou derrubar um pilar e pôr outro” (MARCONDES, 2002).

Na atualidade, mais do que nunca, a sociedade percebe, com nitidez, a força da ciência no desenvolvimento dos povos, tanto em descobertas de amplo consenso, a exemplo do Projeto Genoma Humano, encarregado de decifrar o código genético humano, como em casos polêmicos, tais como a proliferação dos alimentos transgênicos. Na realidade, convive-se diuturnamente com o binômio “ciência” e “tecnologia” em bancos, restaurantes, postos de gasolina, lojas, televisão, semáforos, laboratórios, consultórios médicos, etc. Por outro lado, pensar sobre a relevância da ciência demanda reconhecer a importância da informação científica, do conhecimento científico, da comunidade científica e, por conseguinte, da colaboração científica.

A colaboração científica é um fenômeno tão antigo quanto a própria ciência (BEAVER; ROSEN, 1978). Tal colaboração pode ser encontrada nas mais diferentes formas (PRICE, 1963). De uma maneira genérica, a colaboração científica é um empreendimento cooperativo que envolve metas comuns, esforço coordenado e resultados ou produtos através dos quais os colaboradores compartilham a

responsabilidade e o crédito. Assim, a colaboração científica oferece uma fonte de apoio para melhorar o resultado e maximizar o potencial.

Assim sendo, cada vez mais projetos de pesquisa e ensino têm sido concebidos e executados na forma de colaboração científica, em que diferentes pesquisadores e/ou instituições assumem tarefas distintas com vistas a um dado objetivo, formando as redes de colaboração (WEISZ; ROCO, 1996).

O propósito para a formação de redes colaborativas é combinar, em torno de uma meta comum, capacidades diversas a fim de compartilhar custos e esforços.

Numa ótica social, um importante ganho resultante da formação de redes científicas, quando comparado a pesquisadores isolados, é o aumento na criatividade que advém do intercâmbio de informações e da fertilização cruzada, verificados quando grupos distintos juntam esforços para alcançar uma dada meta. A reunião das mais qualificadas pessoas e instituições para uma determinada tarefa tecnológica ou de pesquisa ou ensino bem como a sinergia resultante da variedade de recursos e competências são as principais vantagens que induzem a formação de redes de pesquisa (WEISZ; ROCO, 1996).

Os estudos de redes de pesquisa têm, em sua base, relação com plataformas de gestão de Ciência, Tecnologia e Inovação (CT&I), ou seja, incluem a escolha de uma fonte de informação e de uma possibilidade de análise sobre como redes se estruturam.

A área de gestão de CT&I consiste em um conjunto de atividades ligadas à produção, disseminação e utilização do conhecimento e de atividades ligadas ao planejamento, à avaliação e ao fomento.

Em todos os níveis de gestão de CT&I deve-se desenvolver os processos de planejamento, monitoramento e avaliação, de forma a assegurar relevância, viabilidade e execução adequada e eficiente de pesquisas.

Na Era da Informação, é de suma importância a utilização de sistemas de informação que forneçam subsídios para a tomada de decisão diária e para o planejamento de ações futuras. Dessa forma, o advento da sociedade da informação traz novas e importantes possibilidades e oportunidades à colaboração em CT&I. Entre as novas oportunidades está a existência de plataformas de governo

eletrônico, em que as diferentes fontes e unidades de informação estão mapeadas e registradas.

Portanto, dado que CT&I ocorrem de forma coletiva, tem-se um novo conjunto de artefatos de informação para mapear as atividades de CT&I e há um novo conjunto de tecnologias de informação e soluções para análise de relacionamentos. Assim, a questão de pesquisa resultante será combinar a importância da colaboração para CT&I, os estudos disponíveis e a possibilidade de efetivá-los por meio de instrumentos que se valham da existência de uma plataforma de informações na área.

1.1 Questões de pesquisa

Esta dissertação tem por finalidade responder às questões de pesquisa a seguir.

- ? Os métodos disponíveis na literatura para análise de redes de pesquisa em CT&I podem ser aplicados em uma plataforma de gestão de CT&I a partir das unidades de informação disponíveis? De que forma?
- ? A disponibilidade de uma plataforma de gestão de CT&I amplia as possibilidades de análise de redes?
- ? É possível que os instrumentos construídos com base na disponibilidade de uma plataforma de informações ajudem no processo de entendimento sobre como a atividade científica e tecnológica se organiza na forma de coletividade?

1.2 Objetivo geral

O objetivo geral desta dissertação consiste em elaborar e desenvolver instrumentos para a análise de redes de pesquisa com base em estudos sobre a colaboração na atividade científica e aplicá-los em uma plataforma de informações para a gestão em CT&I.

1.2.1 Objetivos específicos

Para alcançar o objetivo geral, a dissertação deverá contemplar os objetivos específicos listados a seguir.

- ? Elaborar uma fundamentação teórica na área de redes de colaboração, incluindo as abordagens de redes sociais com ênfase nos estudos em redes de pesquisa.
- ? Estudar a estrutura e metodologia de construção de plataforma de sistemas de informação para gestão de CT&I, com ênfase na proposição da estrutura das fontes de informação, visando estabelecer relacionamentos entre as unidades de análise.
- ? A partir do estudo da área de redes de pesquisa e da arquitetura de informações em CT&I, elaborar um instrumento que permita a análise dos diversos relacionamentos entre atores de um sistema de CT&I.
- ? Desenvolver o sistema de análise de redes em CT&I que permita o estudo de redes de pesquisa, segundo diversas formas de avaliação de relações, disponíveis na literatura ou viabilizadas pela existência de informações sobre atividades científicas e tecnológicas.
- ? Aplicar o sistema de análise de redes de pesquisa à Plataforma Lattes de Gestão em CT&I.

1.3 Justificativa

Dado que CT&I são fatores de desenvolvimento social, têm natureza de promoverem e fundamentarem-se na atividade colaborativa, e visto que o advento da sociedade da informação trouxe novos e excelentes recursos para a análise de relacionamentos, espera-se promover os aspectos relevantes de CT&I por meio de novos instrumentos de análise da atividade científica. Em resumo:

- ? CT&I se faz colaborativamente;
- ? colaborações implicam em relacionamentos;
- ? as novas tecnologias da informação permitem mapeamento sistêmico das relações em CT&I;
- ? conhecer as relações permite: a) melhorar o apoio à tomada de decisão em CT&I; b) instrumentalizar o planejamento em CT&I ao elucidar relações existentes, hiatos de colaboração, etc.; c) fundamentar os novos ambientes de cooperação com relacionamentos já existentes ou induzi-los; d) facilitar a

atividade científica colaborativa em ambientes virtuais, incluindo atores que já colaboram ou que têm afinidades potenciais para colaboração.

Sendo assim, o processo de gestão de CT&I é essencial para consolidar e ampliar a capacitação científica e tecnológica no País (SEMINÁRIO VIRTUAL, 2003). Nesse contexto, acredita-se que este trabalho trará diversos benefícios, podendo destacar a criação de meios para que analistas de informação em CT&I possam estudar as formas de organização da pesquisa brasileira, segundo redes de colaboração científica.

A partir das análises das inter-relações entre os atores do processo de CT&I, os tomadores de decisão estarão aptos a inspecionar, de forma dinâmica, o modo com que esses interagem em cada área de construção do processo científico e tecnológico do País.

A disponibilidade de instrumentos de análise de redes de pesquisa também poderá servir de base para atividades de inspeção de oportunidades, indução de formação de redes, análise de resultados de investimentos e investigação de possibilidades de incentivo ao intercâmbio acadêmico e empresarial, através das Redes de Excelência.

1.4 Estrutura do trabalho

O trabalho possui a estrutura descrita a seguir.

- ? *Capítulo 1: Introdução e contextualização* – contém a descrição do tema e o contexto da pesquisa, a justificativa, os objetivos e a metodologia empregada.
- ? *Capítulo 2: Redes de Colaboração e Redes de Pesquisa* – são apresentados a área de redes de colaboração, suas diferentes abordagens e as áreas de estudo, os fatores e os níveis de colaboração, as redes de colaboração sob o ponto de vista da sociologia, o histórico de redes e, sobretudo, os estudos específicos sobre as redes de pesquisa.
- ? *Capítulo 3: Plataforma para Gestão de Ciência, Tecnologia e Inovação* – é apresentada uma plataforma de gestão de CT&I enfatizando-se os aspectos normalmente presentes nas análises de redes.

- ? *Capítulo 4: Aplicação: Instrumentos para análise de redes de pesquisa em CT&I na Plataforma Lattes do CNPq* – é apresentada a viabilidade de aplicação do método proposto na dissertação para identificar áreas de estudo de cooperação na área de CT&I.
- ? *Capítulo 5: Conclusões e recomendações* – são abordados os resultados da aplicação de instrumentos nas plataformas de CT&I com base nas análises de redes de pesquisa.

1.5 Metodologia

A realização deste trabalho passa pelo levantamento dos estudos referentes às redes de colaboração bem como pelo estudo da arquitetura e descrição das plataformas de gestão em CT&I. Para efetivar os objetivos, o trabalho fundamenta-se em três etapas: (a) estudo da área de redes de colaboração; (b) estudo da arquitetura das plataformas de gestão de CT&I; e (c) análise de redes de pesquisa em uma plataforma de gestão em CT&I. Para a realização do trabalho, propõe-se a metodologia descrita a seguir.

A primeira etapa é constituída por:

- i. estudar e identificar as principais características das redes sociais;
- ii. estudar e identificar as principais características das redes de pesquisa.

A segunda etapa é constituída por:

- i. identificar e estudar os componentes de uma arquitetura de CT&I;
- ii. descrever as unidades de análise pertencentes a uma plataforma de gestão de CT&I e presentes nas análises de redes de pesquisa.

A terceira etapa é constituída por:

- i. elaborar e desenvolver instrumentos de análise de redes de pesquisa em uma plataforma de gestão de CT&I;
- ii. aplicar o sistema de análise de redes de pesquisa à Plataforma Lattes.

Essas etapas, portanto, englobam duas partes complementares de estudo: de um lado o estudo da área de redes de pesquisa e de redes sociais, cujo principal objetivo é conhecer a taxonomia, os métodos de análise, a classificação e a inspeção de redes; de outro a arquitetura de plataforma E-Gov de informações sobre

o sistema nacional de CT&I, em particular quanto à forma com que os subdomínios são mapeados e como as relações entre eles são formalizadas. Essas duas frentes de análise são combinadas como bases de instrumentos de estudos de redes de pesquisa, contemplando tanto abordagens já conhecidas da literatura de redes de pesquisa como novas formas de inspeção viabilizadas por esses estudos.

A Figura 1.1 apresenta uma visão esquemática da metodologia de construção do trabalho.

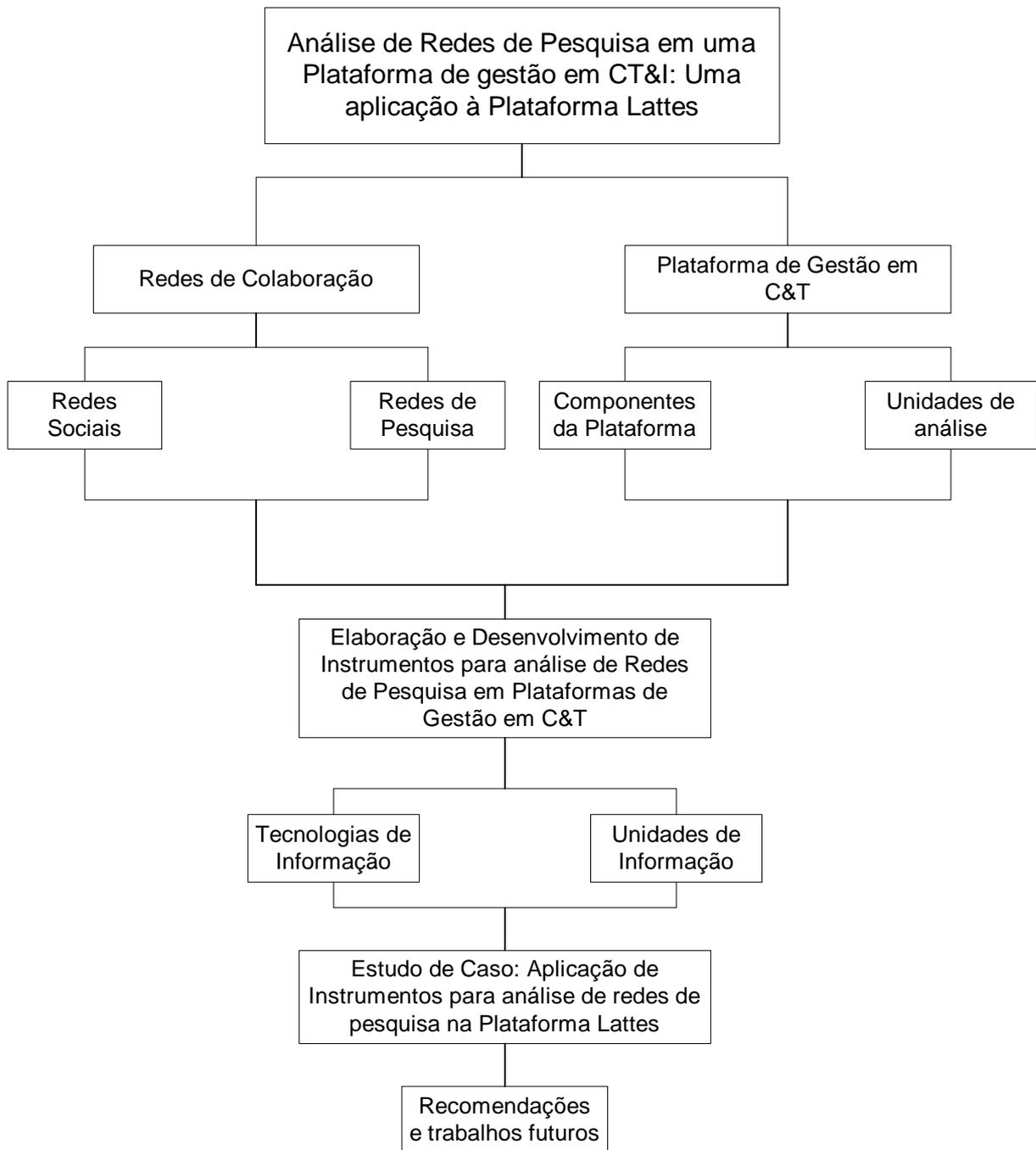


Figura 1.1 - Metodologia do trabalho

1.5.1 Classificação da pesquisa

✍ Natureza

Pesquisa Aplicada (aplicação prática dirigida à análise de redes de pesquisa com base em uma plataforma de gestão em CT&I).

✍ Abordagem do Problema

Pesquisa Qualitativa (não existe a preocupação da análise estatística para a comprovação).

✍ Objetivos

Pesquisa Exploratória (Levantamento bibliográfico).

✍ Procedimentos técnicos da Pesquisa

Pesquisa bibliográfica.

Estudo de caso.

1.5.2 Delimitação do trabalho

Para a análise de Redes de Pesquisa em uma Plataforma de Gestão em CT&I proposta nesta dissertação, foi necessário estabelecer os seguintes elementos delimitadores:

- ? as análises não consideram todas as formas de relacionamentos em CT&I, ou seja, o nível de generalidade dos instrumentos não é pleno, mas está condicionado aos relacionamentos que se conseguiu analisar;
- ? nem todas as áreas de estudo de cooperação foram abrangidas, mesmo em CT&I, isto é, os instrumentos não conseguem replicar todas as diferentes formas de relacionamentos;
- ? o escopo do trabalho não tratou as dinâmicas de formação e evolução das redes colaborativas, de acordo com as diferentes topologias.

2 REDES DE COLABORAÇÃO E REDES DE PESQUISA

De onde viemos? Aonde vamos?
 Viajemos e compreendamos nosso destino (...)
 Na era dos computadores, temos mais idéias e sonhos.
 Agora estamos diante do desafio do amanhã.
 Com o aumento do conhecimento e da tecnologia,
 Nós mudamos nossas vidas e nossos mundos.
 Dos confins do espaço às profundezas do mar.
 Nós construímos uma vasta rede (...)
 Pierre Badin

2.1 Introdução

No presente capítulo aborda-se o tema das Redes de Colaboração, a área de Redes Sociais e Redes de Pesquisa. Um dos objetivos é o desenvolvimento de um levantamento do histórico da área de Redes Sociais, identificando-se os enfoques de análise de cada época e o destaque para os métodos aplicáveis a Plataformas de Informação em CT&I.

Assim sendo, neste capítulo pretende-se tratar a formação das redes de colaborações científicas, suas diferentes abordagens e áreas de estudo. Inicialmente, serão abordadas as redes de colaborações científicas sob o ponto de vista da Sociologia, e em seguida serão apresentadas as redes de pesquisa sob o aspecto das colaborações científicas.

A colaboração científica é um fenômeno tão antigo quanto a própria ciência, e os esforços colaborativos envolvendo pesquisadores de países diferentes foram detectados já no século XIX. O primeiro artigo científico escrito em colaboração entre diferentes pesquisadores surgiu em 1678, segundo relatam Beaver e Rosen (1978).

Tal colaboração pode tomar diferentes formas, desde um aconselhamento e uma discussão geral de idéias até uma participação ativa em um projeto de pesquisa específico. Essas contribuições também podem variar em nível, de substanciais até praticamente vazias. Algumas vezes um pesquisador pode ser considerado colaborador e até mesmo aparecer como co-autor simplesmente por fornecer

material ou realizar alguma operação de rotina. Em outros casos, pesquisadores de diferentes organizações podem colaborar através do uso coletivo de dados ou idéias, de correspondências ou discussões em conferências, de visitas ou da execução, de maneira separada, de diferentes partes de um projeto e da integração posterior dos resultados e da análise.

Dada a variedade de colaborações que podem ser estabelecidas, 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 de contribuição especializada para alcançar os objetivos da pesquisa, o que também inclui a necessidade de se trabalhar próximo a outrem para adquirir novas habilidades e conhecimento tácito, como é tipicamente o caso das relações entre mestre e aprendiz.

Colaborações são também motivadas por uma necessidade de otimização de recursos materiais e financeiros, seja pelas exigências de equipamentos cada vez mais caros e complexos, seja pelos novos padrões ou níveis de financiamento. Tais fatores – que têm sido agrupados em cognitivos, econômicos e sociais – têm importância relativa variada para explicar as diferenças nas taxas de colaboração das áreas do conhecimento e dos diversos países.

Independentemente da natureza de um projeto de cooperação científica – seja ele interpessoal, interinstitucional ou internacional –, seus resultados finais serão freqüentemente expressos por meio da publicação dos resultados produzidos pelos pesquisadores envolvidos.

Dessa forma, as redes informais que os cientistas estabelecem entre si são partes essenciais do conhecimento científico. O produto final – o trabalho publicado – é resultado da série de informações e conhecimentos que são trocados informalmente no âmbito de redes de suporte, as quais são consolidadas ao longo do tempo pelas várias comunidades científicas (MACEDO, 1999).

Os “colégios invisíveis” identificados nos anos 60 por Price constituíram-se em comunidades informais de cientistas, que, comunicando-se e reunindo-se periodicamente para a troca de experiências e informações, passaram a desempenhar um papel fundamental na publicação formal e na disseminação dos avanços no campo do conhecimento científico. Contrastados com os canais mais

formais de comunicação acadêmica, apresentam a particular vantagem da atualidade da informação, da oportunidade de feedbacks e do compartilhamento dos estágios formativos do desenvolvimento de uma idéia e sua transmissão interdisciplinar (GRESHAM, 1994).

2.2 Redes de colaboração nas Ciências Sociais

Redes de colaboração, sob o aspecto social, são um conjunto de pessoas ou grupos que possuem conexões de algum tipo com um ou com todos os outros integrantes da rede (NEWMAN, 2000). Na linguagem das análises de rede social, as pessoas ou os grupos são chamados de “atores”, e as conexões, de “ligações”. Ambos, atores e ligações, 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.

O trabalho em redes de conexões é tão antigo quanto a história da humanidade (BEAVER; ROSEN, 1979), mas, apenas nas últimas décadas, as pessoas passaram a percebê-lo como uma ferramenta organizacional.

O que é novo no trabalho em redes de conexões é a promessa de uma forma global de organização com raízes na participação individual, uma forma que reconhece a independência enquanto apóia a interdependência. O trabalho em redes de conexões pode conduzir a uma perspectiva global baseada na experiência pessoal (LIPNACK; STAMPS, 1992).

Houve grande investimento acadêmico nos estudos de redes de colaboração a partir do campo das relações internacionais, tendo grande significado na história recente das ciências políticas. A origem da reflexão se dá no fim da II Guerra Mundial e tem progresso com o fim da Guerra Fria, quando há redefinição dos atores nas relações internacionais. Novos caminhos de pesquisa surgem, nascidos por meio da reflexão de ordem e desordem nos sistemas políticos, ligados à elaboração da noção de globalização (COLONOMOS, 1995).

2.2.1 Metodologia de análise de Redes Sociais

A análise de redes sociais pode ser considerada como uma metodologia que se aplica ao estudo das relações entre entidades e objetos de qualquer natureza. Originalmente, a análise de redes era aplicada aos sistemas de telecomunicações e computação, circuitos eletromagnéticos, sistemas de engenharia (transportes) e sistemas geográficos (estudos de bacias hidrográficas, por exemplo). Adaptada às relações sociais que constituem os tijolos elementares de toda sociedade humana, a análise de redes se mostrou relevante para a compreensão de problemas complexos, como a integração entre estrutura social (macro) e ação individual (micro) (DEGENNE; FORSÉ, 1999; SCOTT, 2000).

Desde os estudos clássicos de redes sociais até os mais recentes, concorda-se que não existe uma "teoria de redes sociais" e que o conceito pode ser empregado com diversas teorias sociais, necessitando de dados empíricos complementares, além da identificação dos elos e das relações entre indivíduos¹. A análise de redes sociais pode ser aplicada no estudo de diferentes situações e questões sociais.

A análise de redes não constitui um fim em si mesma. Ela é o meio para a realização de uma análise estrutural cujo objetivo dos fenômenos analisados é mostrar que a forma da rede é explicativa. O intuito é demonstrar que a análise de uma díade (interação entre duas pessoas) só tem sentido em relação ao conjunto das outras díades da rede, porque a sua posição estrutural tem necessariamente um efeito sobre sua forma, seu conteúdo e sua função. Portanto, a função de uma relação depende da posição estrutural dos elos, e o mesmo ocorre com o status e o papel de um ator. Uma rede não se reduz a uma simples soma de relações, e a sua forma exerce uma influência sobre cada relação (DEGENNE; FORSE, 1994).

Nos espaços informais, as redes sociais são iniciadas a partir da tomada de consciência de uma comunidade de interesses e/ou de valores entre seus

¹ Os antropólogos Barnes e Both, autores clássicos e pioneiros dos estudos de redes sociais, mostram que o conceito de redes pode ser empregado junto a diversos marcos conceituais. É mais um ferramental de análise do que propriamente um conceito analítico, ou construto teórico reconhecido. Ambos apontam para a necessidade de uma pesquisa de caráter etnográfico, para fazer funcionar o conceito. (BARNES, 1972; BOTH, 1976). Emirbayer (1994) faz uma revisão de estudos de redes sociais, ressaltando seus pressupostos teóricos. Apesar do crescimento dessa área de estudos, o autor afirma que ela não possui bases teóricas e críticas consistentes, sendo fortemente marcada pela empiria. Segundo o autor, a análise de redes sociais "[...] não é uma 'teoria' formal ou unitária que especifica leis, proposições ou correlações distintas, mas antes uma estratégia ampla para investigar a estrutura social" (p. 1.414).

participantes. Entre as motivações mais significativas para o desenvolvimento das redes sociais estão os assuntos que relacionam os níveis de organização social-global, nacional, regional, estadual, local, comunitário. Independentemente das questões que se busca resolver, muitas vezes a participação em redes sociais envolve direitos, responsabilidades e vários níveis de tomada de decisões.

De forma diferente das instituições, as redes sociais não supõem necessariamente um centro hierárquico e uma organização vertical, elas são definidas pela multiplicidade quantitativa e qualitativa dos elos entre os seus diferentes membros, e orientadas por uma lógica associativa. Sua estrutura extensa e horizontal não exclui a existência de relações de poder e de dependência nas associações internas e nas relações com unidades externas (COLONOMOS, 1995).

Estudar a informação através das redes sociais significa considerar as relações de poder que advêm de uma organização não hierárquica e espontânea, e procurar entender até que ponto a dinâmica do conhecimento e da informação interfere nesse processo.

A análise de redes sociais estabelece um novo paradigma na pesquisa sobre a estrutura social. Para estudar como os comportamentos ou as opiniões dos indivíduos dependem das estruturas nas quais eles se inserem, a unidade de análise não está nos atributos individuais (classe, sexo, idade, gênero), mas sim no conjunto de relações que os indivíduos estabelecem através das suas interações com os outros. A estrutura é apreendida concretamente como uma rede de relações e de limitações que pesa sobre as escolhas, as orientações, os comportamentos, as opiniões dos indivíduos. A característica fundamental é lidar com dados relacionais (WASSERMAN; FAUST, 1994; HANNEMAN, 2000), ou seja, dados que expressam relações (conexões ou laços) entre objetos (nós, indivíduos, grupos) diversos. Assim, o foco da análise é deslocado dos atributos individuais (abordagem tradicional nas ciências sociais) para as relações que esses indivíduos estabelecem com outros co-participantes em determinado contexto social.

Pelo motivo de se basear na análise de dados relacionais e não em atributos, a análise de redes sociais restringe a aplicação de diversos modelos estatísticos que supõem a independência quanto à ocorrência e à distribuição dos eventos em uma população. Se o foco da análise são as relações estruturais, fica evidente a interdependência entre os atores e suas interações.

O que era um grande problema de legitimidade perante a comunidade científica há poucos anos, parece começar a mudar ultimamente. Apesar das restrições vinculadas aos modelos relacionais típicos de redes, a análise de redes sociais vem conquistando adeptos nas ciências sociais. Entre os fatores para seu sucesso estão a flexibilidade e a singularidade das redes.

Primeiro, é fácil conceber a existência de redes, bastando, para tanto, representar graficamente as relações entre diferentes objetos e procurar padrões estruturais que as ordenem.

Segundo, as redes têm a singularidade de atuarem como estruturas emergentes, ou seja, próprias de sistemas complexos e dinâmicos – não-linearidade (HANNEMAN, 1997; KLÜVER; SCHMIDT, 1999). Em outras palavras, as redes revelam estruturas sociais (padrões de interação) que podem evoluir de forma não-linear e, portanto, produzir consequências imprevistas sobre determinado contexto. O importante é que essa ação emergente tem implicações nas macro e microdimensões, e a compreensão desse processo crítico nos aproxima muito mais da realidade cotidiana.

Wellman (1999) afirma que a essência da análise de redes sociais está no foco sobre as relações e estruturas sociais – onde quer que possam ser localizadas e quaisquer que sejam os atores envolvidos. A análise de redes sociais não afirma que o mundo é sempre composto de indivíduos normativamente orientados e agregados em grupos fechados ou regiões isoladas. A análise de redes concebe a estrutura social como organizações padrão de membros e seus relacionamentos em uma rede.

Esses padrões de interação, ou recorrências estruturais, podem ser compreendidos como unidades topológicas das redes (KNOKE; KUKLINSKY, 1982; SCOTT, 2000), ou seja, estruturas mais ou menos duráveis que seguem, teoricamente, determinações lógicas intrínsecas do sistema do qual fazem parte.

Essa constatação faz com que alguns teóricos da análise de redes sociais imaginem a possibilidade de constituição de uma “teoria social das redes”, toda ela baseada em modelos topológicos, descritos integralmente pela propriedade dos grafos (KNOKE; KUKLINSKY, 1982; KLÜVER; SCHMIDT, 1999). Significa dizer que, mesmo que não fosse possível isolar o problema da independência estatística, a

generalização (das hipóteses e teorias das redes sociais) seria possível a partir dos efeitos topológicos e métricos descritos pela Teoria dos Grafos.

Independente dessa possibilidade, mais recentemente, uma força-tarefa de analistas de redes tem sido concentrada no desenvolvimento de modelos estatísticos apropriados aos dados relacionais, que, de forma consistente, superam o obstáculo da não-independência².

Finalmente, as medidas estatísticas de redes sociais podem ser divididas em duas categorias: i) descritivas e ii) de análise estrutural.

Medidas descritivas	
Densidade (<i>density</i>)	É a proporção de laços efetivos entre laços possíveis. Uma medida do grau de inserção dos atores na rede.
Centralidade (<i>centrality</i>)	Localização do ator em relação à rede total.
Proximidade (<i>closeness</i>)	Grau de proximidade em relação a outros atores da rede.
Intermediação (<i>betweenness</i>)	Mede o grau de intervenção de um ator em outros atores da rede.
Distância geodésica (<i>distance</i>)	Mede o grau de afastamento de uma localização em relação a outros atores.
Alcance (<i>reachability</i>)	Mede a extensão do contato que um ator tem com outros atores na rede.
Subgrupos (<i>cliques</i>)	Mede o grau de concentração e formação de subgrupos em uma rede.
Medidas estruturais	
Densidade (<i>density</i>)	A densidade da rede mede o grau de coesão e homogeneidade.
Transitividade (<i>transitivity</i>)	Mede o grau de flexibilidade e cooperação de uma rede.
Equivalência estrutural	Mede a posição relativa de um ator na rede.
Equivalência regular	Medida menos estrita que a anterior - mede literalmente o papel social.
Buraco estrutural	Mede o grau de coesão e competição de uma rede.

Tabela 2.1 - Algumas medidas estatísticas de Redes Sociais (Fonte: UCINET V³)

Como afirmam Knoke e Kuklinsky (1982), as medidas estatísticas capturam as propriedades emergentes (não-lineares) dos sistemas sociais que não podem ser avaliadas pela simples agregação de atributos dos membros individuais. Além disso,

² O princípio desses modelos se baseia no “problema do mundo pequeno” (*Small World*), também estudado pela “estatística física”. O modelo do “mundo pequeno” afirma que os atores em uma rede (independente do seu tamanho e densidade) podem encontrar outro ator através de seis passos em média – também conhecidos como “seis graus de separação” (MILGRAM, 1967; DEGENNE; FORSÉ, 1999; WATTS, 1999).

³ O UCINET V (EVERETT; BORGATTI, 2001) é o software de redes mais popular existente no mercado.

tais propriedades emergentes podem afetar significativamente tanto a performance do sistema quanto o comportamento dos membros de uma rede.

Os algoritmos que definem as estatísticas descritivas são mais comuns e de simples aplicação; os algoritmos de análise estrutural são bem mais complexos, pois revelam a estrutura invisível subjacente à rede (medidas de capital social, buracos estruturais e laços simmelianos são dados por esses algoritmos).

A importância desta última categoria é crucial para a compreensão dos processos de formação, manutenção e evolução das redes sociais (BURT, 1992; ENGLE, 1999; DEGENNE; FORSÉ, 1999). Através da identificação das posições e dos papéis desempenhados na estrutura social de uma rede é possível determinar os padrões de relações entre os atores e, conseqüentemente, comparar diferentes processos interativos (estabelecendo as bases para prováveis modelos preditivos e gerais sob certas circunstâncias).

Na realidade, o analista de redes sociais deve utilizar alguma combinação das diversas medidas de rede a partir do problema empírico estudado. Assim, várias medidas descritivas são complementares quando se trata de identificar ou comparar o grau de inserção (*embeddedness*) de diferentes atores em uma rede.

Segundo Degenne e Forsé (1999), a análise de redes pessoais é o primeiro tipo e o mais básico da análise de redes sociais, pois captura os dados relacionais em torno de um indivíduo. Por outro lado, para uma compreensão profunda das estruturas e de estrangimentos emergentes nas interações, as redes totais (rede total é aquela na qual se apresentam diversos atores conectados uns aos outros sem haver necessariamente um único centro, formando agrupamentos mais ou menos homogêneos (cliques) e revelando a estrutura social completa de uma coletividade de atores ou nós) se fazem necessárias, dado o efeito de “multiplexidade” – quando diferentes conteúdos de relações (parentesco, amizade, contato formal, relações financeiras, etc.) se sobrepõem ao campo de interação social.

Em seu aspecto geral, a análise de redes sociais não se preocupa com indivíduos isolados, mas com suas conexões em uma coletividade. Por isso, a abordagem das redes totais é mais indicada para se atingir o objetivo principal da

análise de redes sociais: explicar a ocorrência de diferentes estruturas e como essas estruturas interferem ou não nos comportamentos dos atores.

Como afirmam Knoke e Kuklinsky (1982), a estrutura das relações entre os atores e sua localização em uma rede gera conseqüências sobre os comportamentos, as percepções e as atitudes individuais tanto quanto sobre o sistema total.

Portanto, os estudos de rede sociais normalmente utilizam uma lógica combinatória e não simplesmente agregada das relações, ou seja, muitas vezes em uma rede o mais importante que a adição ou a subtração de um laço é a conformação global e real de toda a rede.

Finalmente, segundo Degenne e Forsé (1999), quatro pontos fundamentais sintetizam a análise estrutural das redes sociais, são eles:

1) a estrutura afeta a ação formalmente através de um fraco determinismo: concentração ou disposição de determinadas estruturas favorece ou facilita a ação para esse caminho;

2) a estrutura afeta as percepções de auto-interesse: o ator percebe mais facilmente as alternativas pessoais de escolha que fazem parte de seus relacionamentos e por isso fazem mais sentido, objetiva e (inter)subjetivamente;

3) princípio da racionalidade: indivíduos racionais tomam suas decisões como função de interesses pessoais (em uma escala de preferências), o que induz efetivamente à ação; e

4) a estrutura é um efeito emergente das interações sociais: cada interação em um sentido reforça o arranjo estrutural desse sentido.

2.3 Redes de pesquisa

Mudanças nas condições econômicas, sociais e políticas demandam alterações profundas nos paradigmas que vinham regendo as atividades humanas nas últimas décadas. Até então a aceleração nas mudanças científicas e tecnológicas tornou necessário o encurtamento dos ciclos pesquisa–desenvolvimento–produto. A necessidade de uma base científica ampla para a produção e a abertura dos

mercados à competição global bem como a globalização da ciência resultaram em novas exigências para a pesquisa.

Novas idéias científicas, a criação de novas tecnologias e a formação de bons profissionais suplantaram a possibilidade de existência de instituições isoladas. Ademais, crescentes restrições orçamentárias, por um lado, a existência de diferentes centros apresentando capacidades semelhantes e a submissão internacional da ciência e tecnologia em acréscimo a melhores meios de comunicação, por outro lado, resultaram na formação de redes de atividades de pesquisa. Os principais objetivos dessas redes são a pesquisa, a educação e o desenvolvimento tecnológico cooperativo, a cooperação externa e as atividades interdisciplinares (WEISZ; ROCO, 1996).

Ultimamente um número crescente de projetos de pesquisa e ensino tem sido concebido e executado na forma de cooperação, em que diferentes instituições assumem tarefas distintas com vistas a um dado resultado. Essa organização de coesão tênue formada por diferentes indivíduos ou grupos ligados entre si por vínculos de naturezas diversas pode então ser definida como uma rede de pesquisa (WEISZ; ROCO, 1996).

Os fatores que levam à formação dessas redes de colaboração científica são os mais diversos, os quais são abordados a seguir.

2.3.1 Fatores que contribuem para a colaboração em Redes de Pesquisa

Podendo ser encontrada nas mais diferentes formas, a colaboração científica se dá, freqüentemente, no âmbito dos chamados colégios invisíveis (PRICE, 1963). Esses grupos são caracterizados por sua alta produtividade, por compartilharem prioridades de pesquisa, por treinarem estudantes, por produzirem e monitorarem o conhecimento em seu campo (CRANE, 1972). Encontram-se freqüentemente em congressos, conferências, reuniões sobre suas especialidades, visitam-se por meio de intercâmbios institucionais ou realizam trabalhos em colaboração. Esse tipo de organização transcende aos 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 (PRICE; BEAVER, 1966).

Dada a variedade de colaborações que podem ser estabelecidas, não é de se estranhar que as razões que levam os cientistas a colaborar entre si sejam também das mais diferentes naturezas.

Nos itens a seguir, serão detalhados os fatores que de alguma forma contribuem para a colaboração entre pesquisadores bem como aqueles fatores que não exercem nenhuma influência na colaboração.

a) Colaboração de formação (orientador- orientando)

Em 1965 surge o primeiro estudo com preocupação na relação de formação (HAGSTROM, 1965), onde constatou que alguns orientadores não consideravam seus orientandos como seus colaboradores.

A relação entre orientador e orientando é com certeza um dos tipos mais evidentes de colaboração, visto que a necessidade de contribuição especializada de outrem para alcançar os objetivos da pesquisa e também a necessidade de se trabalhar próximo de outrem para adquirir novas habilidades e conhecimento tácito são os fatores que contribuem para a colaboração, como é tipicamente o caso das relações entre mestre e aprendiz (BEAVER; ROSEN, 1979).

b) Colaboração teórica e experimental

Em geral, os trabalhos teóricos produzem artigos com poucos co-autores em comparação aos trabalhos experimentais, ou seja, experimentalistas tendem a colaborar mais do que teóricos (KATZ; MARTIN, 1997).

c) Proximidade na colaboração

A proximidade regional parece encorajar a colaboração, já que os pesquisadores tendem a gerar mais comunicação informal. Isso mostra que as colaborações decrescem exponencialmente com o aumento da distância que separa os pesquisadores institucionais (KATZ, 1993). Contudo, essa pesquisa é anterior à Internet. O surgimento da Web não modifica a conclusão e sim a definição de distância, que não é necessariamente geográfica e sim cultural, de idioma, de interesses, de afinidades e oportunidade.

d) Produtividade e colaboração

Pesquisas que buscam responder à questão “autores produtivos tendem a colaborar mais do que autores menos produtivos?” parecem indicar que a alta

produtividade (em termos de publicação) é de fato correlata com os altos níveis de colaboração. A saída científica como medida de publicações é intimamente dependente da frequência de colaboração entre autores. A natureza do efeito sobre produtividade depende do tipo de ligação; enquanto colaborações com pesquisadores de alta produtividade tendem a aumentar a produtividade pessoal, colaborações com pesquisadores de baixa produtividade tendem a diminuir a colaboração pessoal (KATZ; MARTIN, 1997).

e) Quantidade de colaboradores inspira maior confiança

Um estudo de Nudelman e Landers (1972) sugere que o total de crédito dado por uma comunidade científica para todos os autores de um artigo é em média maior do que o crédito alocado para um único autor de um artigo. O número de co-autores também parece ser fortemente correlacionado com o impacto de um artigo.

Goffman e Warren (1980) mostram que pesquisas por grandes grupos tendem a ter mais influência; e Narin e Whitlow (1990) encontraram evidências de que artigos de co-autoria internacional são citados duas vezes mais do que um artigo de um único país.

f) Interdisciplinaridade

Mais um fator preponderante que tem motivado a colaboração é o importante aumento dos campos interdisciplinares. Está se tornando claro que alguns dos mais significativos avanços científicos surgiram da integração ou fusão dos campos anteriormente separados (KODAMA, 1992).

g) Nível de especialização

Outro fator relatado tem sido o aumento da necessidade de especialização dentro de certos campos, especialmente, nos quais a instrumentação requerida é muito complexa, isto é, quando um determinado campo da ciência precisa juntar vários experts de cada área, pois um único indivíduo não poderia desenvolver todas as tarefas específicas (PRICE, 1963).

h) Compartilhamento de recursos

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

A necessidade de compartilhar equipamentos e estruturas pode ser a primeira base para grupos de pesquisa multi-institucionais. Assim, a informação sobre infraestrutura pode ser base dessas consultas. A Plataforma de Informação em C&T (Plataforma Lattes, inclusive) tem essa informação no Data Mart de Fomento.

i) Reconhecimento pelos pares

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 (LAWANI, 1986; PRAVDIC; OLUIC-VUKOVIC, 1986; NARIN; WHITLOW, 1991).

Tais fatores – freqüentemente agrupados em cognitivos, econômicos e sociais – têm importância relativa variada para explicar as freqüências de colaboração observadas nas diferentes áreas do conhecimento e nos diferentes países (LUUKONEN et al., 1992).

Não apenas as razões para se colaborar em ciência, mas também os níveis de agregação em que a colaboração pode ocorrer variam. Dessa forma, na próxima subseção serão abordados os níveis de colaboração que podem ocorrer.

2.3.2 Níveis de colaboração

A colaboração pode dar-se entre indivíduos, grupos, departamentos, instituições e setores, nas mais diferentes combinações dessas unidades, dentro de uma mesma nação ou envolvendo nações diferentes (KATZ; MARTIN, 1997).

No nível mais básico, são as pessoas que colaboram e não as instituições. Cooperação direta entre dois ou mais pesquisadores é a unidade fundamental de colaboração. Contudo, há outros níveis de colaboração, como, por exemplo, entre grupos de pesquisa dentro de um departamento, entre departamentos dentro da mesma instituição, entre instituições, entre setores e entre regiões geográficas e países.

Além de se diferenciar a maioria dos níveis de colaboração, precisa-se também reconhecer que a colaboração ocorre entre diferentes níveis e dentro deles. Os prefixos “inter” e “intra”, respectivamente, têm sido adotados para distinguir essas categorias. Por exemplo, colaboração internacional significa colaboração entre

nações, enquanto que colaboração intranacional significa colaboração dentro de uma única nação.

Algumas vezes, a colaboração não pode ser claramente classificada, já que pode pertencer a ambas as categorias. Por exemplo, ao se considerar uma colaboração que envolve duas instituições nacionais e uma instituição estrangeira, pode-se dizer que essa colaboração é claramente interinstitucional. Todavia, em uma outra perspectiva, essa colaboração constitui uma colaboração internacional, enquanto em outra ela pode ser considerada uma mistura de colaboração inter e intranacional. Assim, Katz e Martin (1997) classificam as colaborações como homogênea (i.e. quando não há mistura das formas de colaboração) e heterogênea (i.e. quando há mistura das formas de inter e intracolaboração).

Os diferentes níveis de colaboração, nas formas inter e intra, estão sumarizados na tabela 2.2.

	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 departamentos 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

Tabela 2.2 - Diferentes níveis de colaboração e distinção entre as formas inter e intra (KATZ; MARTIN, 1997)

Dentre os níveis de colaboração elencados, pode-se destacar a colaboração intergrupo. Ela acontece quando, por exemplo, dois líderes de grupo podem concordar que seus grupos foquem o mesmo objetivo de pesquisa, e então eles começam a trabalhar juntos, cada grupo apontando seu conhecimento coletivo para resolver o problema. Outra forma ocorre quando pesquisadores de um grupo também integram outro grupo, onde realiza intercâmbio de experiências e promove produção científica com autores de mais de um grupo. Outra forma de colaboração intergrupo se dá quando estudantes de doutorado em grupos de instituições consolidadas integram (ou lideram) grupos de pesquisa emergentes em suas

instituições de origem. Essas colaborações são gênese indutoras de capacidade em pesquisa, com impacto significativo no desenvolvimento regional.

Outro nível de colaboração que se pode salientar é o interdepartamental. Ele pode ser o resultado de uma decisão formal de dois departamentos na tentativa de oferecer uma situação de junção para o indivíduo. Em outras palavras, é uma explícita iniciativa de produzir ligações entre dois departamentos, o que não será sempre o caso. Embora o grau de colaboração de pesquisa interdepartamental possa ser muito pequeno. O que geralmente ocorre é o pesquisador ter de trabalhar em sua pesquisa em um determinado departamento e ter de dar aulas em outro. A colaboração interdepartamental também pode ocorrer quando cursos de formação em uma área do conhecimento têm disciplinas de outros cursos, como a formação em Cálculo ou Física da Engenharia ou a formação em disciplinas da Administração em cursos de Sistemas de Informação.

Alguns cenários podem ser apresentados para se discutir a colaboração institucional. Supõe-se que, num primeiro caso, um autor liste dois endereços institucionais num artigo, no segundo caso, outro autor liste dois departamentos de uma mesma instituição, e, no terceiro caso, que um outro autor liste dois departamentos de uma mesma instituição juntamente com uma terceira instituição diferente. Todos esses casos podem ser considerados legítimos de colaboração interinstitucional? Katz e Martin (1997) respondem a essa questão afirmando que uma colaboração internacional ou interinstitucional não necessariamente exige uma colaboração interindividual. Alguém poderia argumentar que não há nenhum tipo de colaboração quando um único autor está envolvido, contudo, tais publicações podem ser perfeitamente colaborações institucionais válidas, desde que os autores façam um acordo entre departamentos ou instituições para compartilhar o pesquisador; a colaboração é então manifestada como uma publicação de um pesquisador “compartilhado” listando vários endereços institucionais. A colaboração interinstitucional pe comum quando docentes de uma instituição passam por formação (normalmente doutorado) ou estágio sabático em outra instituição e, durante esse período, produzem com pesquisadores da instituição que o acolheu.

Ao lado dos vários níveis de colaboração apresentados, a colaboração internacional é provavelmente a forma mais familiar. Um cenário denominado A é apresentado para exemplificar quando pesquisadores de vários países trabalham

juntos numa única instituição. Assim, todos os artigos produzidos listam somente aquela instituição, ou seja, somente um país sempre aparecerá endereçado. Outro cenário, denominado B, ocorre quando um pesquisador tem duas afiliações institucionais localizadas em países diferentes, por exemplo, um departamento de uma universidade e um hospital. A análise bibliométrica conclui que houve colaboração internacional no cenário B e não no cenário A, mesmo que o cenário B possua somente um indivíduo, enquanto que no cenário A há o envolvimento de várias pessoas de diferentes países trabalhando juntas. Isso significa que uma colaboração é considerada internacional não pela nacionalidade dos pesquisadores envolvidos e sim pela origem de suas instituições. Na prática, uma colaboração internacional exige que haja uma colaboração institucional.

2.3.3 Taxonomias para o estudo de Redes de Pesquisa

Nas últimas décadas, acompanhando-se a expansão da ciência e da tecnologia, tornou-se cada vez mais evidente a necessidade de avaliar tais avanços e de determinar os desenvolvimentos alcançados pelas diversas disciplinas do conhecimento. Nesse sentido, apontou-se para a medição das taxas de produtividade dos centros de pesquisa e dos investigadores individuais, para a detecção daquelas instituições e áreas com maiores potencialidades e para o estabelecimento das prioridades no momento da alocação de recursos públicos. Existem diversas formas de medição voltadas para avaliar a ciência e os fluxos da informação.

Nesta subseção serão apresentadas duas taxonomias voltadas à avaliação do fluxo de colaboração entre pesquisadores, a primeira é a de redes de citação (via produção científica), e a segunda é a de redes de co-autoria (via produção científica).

a) Redes de Citação

O meio mais comum pelo qual se atribuem crédito e reconhecimento na ciência é a citação. O artigo de periódico com a sua lista de citações é, e provavelmente assim permanecerá, o meio universalmente aceito pelo qual a instituição científica registra e divulga os resultados de suas investigações. Com o surgimento de índices de citação nas últimas três décadas, a importância das citações ganhou uma nova

dimensão nas vidas dos cientistas, departamentos de pesquisa, das universidades e nações (MACIAS-CHAPULA, 1998).

A indexação de citações está baseada na premissa de que uma citação bibliográfica é a expressão de uma relação entre dois documentos, aquele que cita e aquele que é citado. Haverá invariavelmente uma lacuna entre o motivo pelo qual o autor citou e o motivo pelo qual se pensa que ele citou. As razões que levam um autor a citar de uma determinada forma podem ser sérias ou frívolas.

A citação é influenciada por uma multiplicidade de fatores, mas nem todos estão relacionados com as convenções reconhecidas pela publicação acadêmica. Os fatores sociais e psicológicos, por exemplo, têm aí uma função significativa, juntamente com as lembranças e os esquecimentos subconscientes (ARONSON, 1975). Há também fatores extrínsecos, por exemplo, os leitores ou a percepção do autor em relação às necessidades e expectativas dos leitores. O perfil e o status da revista na qual o artigo será publicado, a abrangência, o formato e a extensão do artigo, o conhecimento do autor sobre a área e a sua habilidade/disposição em usar as fontes e os serviços apropriados de informação, são outros fatores envolvidos.

É necessário pensar a citação como um processo. Os resultados desse processo são as listas de citações que acompanham os trabalhos acadêmicos. O tipo e a composição dessas listas refletem a personalidade do autor e seu meio profissional. Não existe uma teoria única da citação capaz de explicar por que os autores citam de uma determinada maneira (CRONIN, 1984). A maioria das análises de citação tem apresentado mais características internas do que externas, concentrando-se em dados quantitativos e em distribuições de frequência, em vez de tratar os contextos nos quais os autores utilizam as citações.

Hauffe (1994) diz que há, em muitos casos, uma relação entre o artigo e a referência citada por seu(s) autore(s). Essa relação foi o ponto de início da idéia básica de Eugene Garfield, o fundador do Instituto para Informação Científica (The Institute for Scientific Information – ISI), cujo produto é o conhecido Índice de Citação Científica (The Science Citation Index – SCI). Este é um dos métodos mais fáceis de análise de citação, baseado na contagem de citação, cujo princípio é determinar quantas citações recebeu um dado documento ou um conjunto de documentos por um determinado período de tempo (SMITH, 1981).

Análise de Citação pode ser uma boa medida de estudo sociológico na ciência (e o impacto das contribuições científicas é um caso sociológico), contudo, a medida está muito distante da perfeita avaliação de qualidade nos trabalhos científicos. Assim sendo, um pesquisador que recebeu “somente” 100 (cem) citações não pode dizer que é menos qualificado, neste caso, pela metade, em relação a outro que recebeu 200 (duzentas) citações, mas, de fato, ele pode em tese dizer que é mais qualificado que outro pesquisador que trabalhou no mesmo campo de pesquisa que recebeu somente 5 (cinco) citações (HAUFFLE, 1994). Outra análise que se pode fazer sobre o uso do Índice de Citação para contribuição da avaliação científica indica que um fator pode falsificar os resultados em virtude de somente artigos serem avaliados, de nenhum livro ser avaliado, portanto se trata de uma avaliação incompleta.

Antes de se discutirem as redes de citação, faz-se necessária uma apresentação dos indicadores da atividade científica para o entendimento das análises de citação. Atualmente, os indicadores da atividade científica estão no centro dos debates, sob a perspectiva das relações entre o avanço da ciência e da tecnologia por um lado e o progresso econômico e social por outro. Revisões de políticas científicas pareceriam inconcebíveis hoje sem se recorrer aos indicadores existentes (MACIAS-CHAPULA, 1998).

Em tudo que se refere à ciência, os indicadores bibliométricos e cienciométricos tornaram-se essenciais. Contudo, entre os cientistas da informação da Europa e dos Estados Unidos, o termo “informetria” tornou-se comum somente nos últimos dez anos como um campo geral de estudo que inclui as áreas mais antigas da bibliometria e da ciencimetria. Tague-Sutcliffe (1992) definiu essas disciplinas da forma descrita a seguir.

Bibliometria é o estudo dos aspectos quantitativos da produção, disseminação e do uso da informação registrada. Usada pela primeira vez por Pritchard em 1969, a bibliometria mede esses processos usando seus resultados para elaborar previsões e apoiar tomadas de decisão.

Ciencimetria é o estudo dos aspectos quantitativos da ciência como uma disciplina ou atividade econômica. A ciencimetria é um segmento da sociologia da ciência, sendo aplicada no desenvolvimento de políticas científicas. Envolve estudos

quantitativos das atividades científicas, incluindo a publicação, e, portanto, sobrepõe-se à bibliometria.

Informetria é o estudo dos aspectos quantitativos da informação em qualquer formato, e não apenas dos registros catalográficos ou bibliográficos ou bibliografias, referentes a qualquer grupo social, e não apenas aos cientistas. A informetria pode incorporar, utilizar e ampliar os muitos estudos de avaliação da informação que estão fora dos limites tanto da bibliometria como da cienciometria.

A Figura 2.1 representa graficamente a abrangência dos objetos de estudo destas três disciplinas.

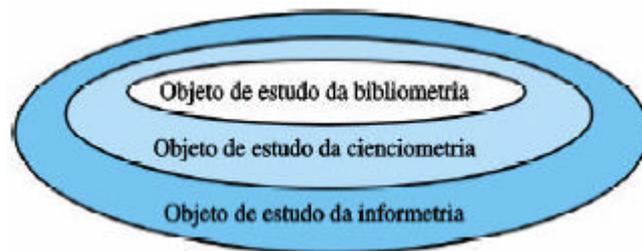


Figura 2.1 - Abrangência dos objetos de estudo da bibliometria, cienciometria e informetria (PACHECO; KERN, 2001)

No contexto das redes de citação, pode-se defini-las como um conjunto de nós interconectados por uma ligação, em que os nós são os pesquisadores, e a ligação entre eles são as citações, ou seja, dois pesquisadores podem ser considerados conectados ou ligados se um dos pesquisadores citou o outro em alguma produção científica.

Na Figura 2.2 ilustra uma rede de citação. A análise para configurar uma rede de citação se dá pela avaliação de artigos (textos) em que se procura mostrar a ligação entre os autores do artigo e os autores dos trabalhos citados.

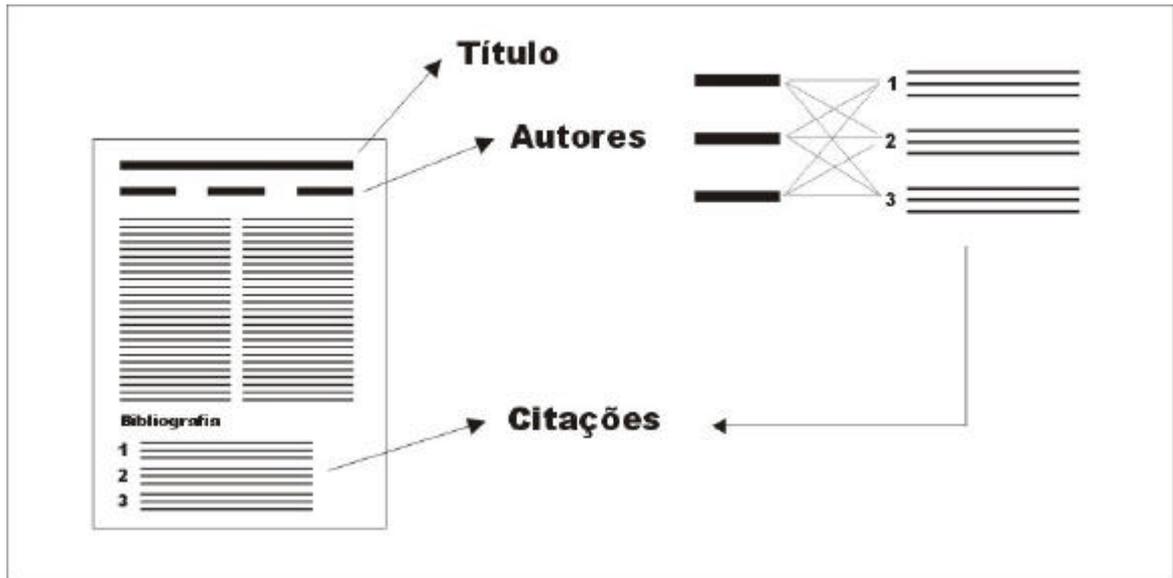


Figura 2.2 - Rede de Citação

Esse tipo de estudo é mais adequado para análise do fluxo de conhecimento, ou seja, permite verificar o espectro de influência entre os autores. Na literatura tem sido pouco encontrado, visto que, para a viabilização desse tipo de análise, são necessárias as informações bibliométricas (que têm base em documentos completos), mas a maioria das informações disponíveis nas bases de dados são informações informétricas (que têm base em informações acerca da atividade científica e tecnológica). Não estando disponíveis os documentos completos, não se pode viabilizar a análise das citações dos trabalhos. Além disso, mesmo quando os documentos estão disponíveis é necessária a extração das informações sobre os autores do artigo e dos autores dos trabalhos citados. Para tal, é necessário considerar as diferentes formas com que um mesmo autor aparece citado e as diferentes normas de citação.

b) Redes de Co-Autoria

Sob o aspecto de redes de pesquisa, dois pesquisadores podem ser considerados conectados ou ligados se eles forem co-autores em um ou mais artigos juntos. Assim, nas redes de pesquisa formadas pelo relacionamento de co-autoria, os atores ou nós são os pesquisadores, e as ligações entre eles são as colaborações científicas através de co-autoria.

A idéia de se construir uma rede de co-autoria não é nova. O conceito do número Erdős, nome dado em homenagem a Paul Erdős, um matemático húngaro e um dos pais da Teoria dos Grafos, representa muito bem isso, pois Erdős publicou 1.401

artigos em colaboração com outros pesquisadores. No jargão das redes sociais, o número Erdős é a distância geodésica (menor distância) entre um dado pesquisador e Paul Erdős nas redes de co-autoria (NEWMAN, 2000).

Se um dado pesquisador tivesse publicado um artigo com Erdős, receberia o número Erdős de 1 (um); já, se outro pesquisador tivesse publicado um artigo com outro que tivesse publicado com Erdős, a esse pesquisador seria atribuído o número 2 (dois), e assim por diante. Em geral, se dentre os números de Erdős de cada pesquisador com quem este escreveu artigos, o menor número é N , então o número de Erdős deste pesquisador é $N+1$. Em recentes estudos, descobriu-se que a média do número Erdős de um pesquisador é aproximadamente 4.7 (NEWMAN, 2000).

Por décadas a publicação de múltiplos autores, freqüentemente referenciada como publicação de co-autoria, tem sido usada como base de unidade de contagem para atividades colaborativas (KATZ; MARTIN, 1997).

Dentre várias propriedades estatísticas para redes de co-autoria, pode-se listar as seguintes: número de artigos escritos por cada autor, número de autores por artigo, número de colaboradores dos pesquisadores, a distância entre a rede de um pesquisador e a outra rede, e uma variedade de medidas de conectividade dentro da rede, tais como proximidade e intermediação.

Para Newman (2001), as redes de co-autoria mostram um alto grau de transitividade, o que significa que há uma maior probabilidade de duas pessoas estarem relacionadas se elas tiverem uma ou mais co-autorias em comum, ou seja, a probabilidade de pesquisadores colaborarem aumenta com o número de outros colaboradores em comum, sendo que a probabilidade de um pesquisador, em particular, adquirir novos colaboradores aumenta com o número de pesquisadores com ao quais ele colaborou no passado.

Quaisquer que sejam as razões que levam à colaboração científica, 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, vários autores definem a cooperação científica como o conjunto de trabalhos cooperativos desenvolvidos entre dois ou mais pesquisadores e identificados por meio de artigos co-assinados (MEDOWS; O'CONNOR, 1971; BEAVER; ROSEN, 1979; LECLERC et al., 1992). Essa 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).

Entretanto, mesmo reconhecendo as limitações acima apontadas, a contagem de artigos co-autorados tem sido a medida mais comumente usada para detectar a ocorrência, a abrangência e os atores envolvidos em parcerias científicas. Há também um corpo substancial da literatura em ciência da informação sobre extração de padrões de co-autoria usando dados extraídos de registros de publicações. Isso ocorre por várias razões. Primeiramente, porque boa parte dos outros produtos da colaboração científica – conhecimento tácito, aprendizagem de determinada técnica e outros – são intangíveis, isto é, quase impossíveis de serem quantificados. Segundo, a contagem de artigos em co-autoria é invariante e verificável, ou seja, dado o acesso à mesma base de dados, outros investigadores deveriam encontrar resultados idênticos. Além disso, é um método relativamente barato e prático para quantificar a colaboração e permite a análise de amostras bastante grandes, gerando resultados estatisticamente mais significativos do que os estudos de caso.

2.3.4 Cadeias de Suprimento (*Supply Chain*) e Redes de Pesquisa

O objetivo desta subseção é fazer a correspondência entre as cadeias de suprimento e as redes de pesquisa. Assim, serão apresentados inicialmente os conceitos para cadeias de suprimento e, em seguida, fazer a associação com as redes de pesquisa.

A expressão “cadeia de suprimento” é uma metáfora usada para descrever as empresas que estão envolvidas no fornecimento de um produto ou serviço (GIANNAKIS, 2001).⁴

O conceito de cadeia de suprimento é relativamente novo no pensamento da administração de empresas. Seu desenvolvimento deve-se à constatação, nos anos de 1970 e de 1980, de que os fabricantes japoneses de automóveis administravam o

⁴ Para evitar que conduzam à linha de raciocínio por caminhos perigosos e distantes da busca da verdade a que o cientista deve se dedicar, o uso de metáforas deve repousar num método científico que permita usá-las, mas que não exclua a experimentação e o teste de hipóteses. VAILLANT, V. George E. *The Wisdom of the Ego*. Cambridge, Mass: Harvard University Press, p. 118.

fornecimento de insumos, além dos simples relacionamentos contratuais com os fornecedores diretos (NISHIGUCHI, 1994; COX et al., 2001).^{5 6}

Beamon (1998) define cadeia de suprimento como:

processo integrado dentro do qual um número de várias entidades de negócios (i.e., fornecedores, fabricantes, distribuidores e varejistas) trabalham juntos num esforço para 1) adquirir matérias-primas, 2) converter estas matérias-primas em determinados produtos finais e 3) oferecer estes produtos nos pontos de venda. (p. 281)

Em Cox et al. (2001) afirma-se que uma cadeia de suprimento pode ser descrita como:

rede ampliada de relacionamentos diádicos de permuta que devem existir para a criação de qualquer produto ou serviço que é fornecido a um cliente final. Para que um produto ou serviço seja fornecido para um cliente, ele deve começar como matéria-prima de alguma espécie e então passar por uma série de estágios de permuta entre compradores e fornecedores. Em cada estágio deve ocorrer alguma forma de transformação, ou intermediação, destinada a agregar valor ao produto ou serviço como percebido pelo comprador naquele estágio e, em última instância, pelo cliente final. (p. 28)

Dentro desse contexto surgem na literatura os conceitos de *supply chain management* (SCM) ou Gestão da Cadeia de Suprimentos. Esse conceito pode expandir o conceito de cadeia de produção, pois pressupõe a integração de todas as atividades da cadeia mediante à melhoria nos relacionamentos entre os diversos elos ou agentes - organizações de diferentes tipos interagindo - em busca da construção de vantagens competitivas sustentáveis para a cadeia. A cadeia de suprimentos deve ser vista como uma rede de empresas independentes que agem em sintonia de forma a criar valor para o usuário final através da distribuição de produtos. Essa sintonia é exatamente o que a gestão de cadeias de suprimentos objetiva (BATALHA, 1999).

Betchel e Jayaram (1997) apresentam o SCM como um novo conceito que ainda carece de uma definição mais clara. Os autores citam algumas escolas de pensamento que procuram definir o conceito de SCM. Duas escolas aqui destacadas para tratar a questão do SCM referem-se à escola da informação e à escola da integração/processos.

⁵ Usaram-se "fornecedores diretos" como tradução de "first-tier suppliers".

⁶ No original, a referência é NISHIGUCHI, T. *Strategic Industrial Sourcing: The Japanese Advantage*. Oxford, England: Oxford University Press, 1994.

A escola da informação enfatiza o fluxo de informações entre os membros da cadeia de suprimentos e considera a informação a “espinha dorsal” do SCM efetivo, não somente o fluxo de informação unidirecional do comprador ao fornecedor é considerado, mas também o fluxo bidirecional. Além disso, as informações devem fluir não apenas de um membro da cadeia para o seguinte, mas também entre todos os membros da cadeia que precisam do feedback de como sua performance está sendo percebida por seus clientes e usuários finais. Segundo os autores dessa escola de pensamento (JOHANNSON, 1994; TOWILL, Naim; WIKNER, 1992; MANRODT; HARRINGTON, 1995 apud BETCHEL; JAYARAM, 1997), companhias que estão prosperando e obtendo vantagens competitivas fazem uso da tecnologia da informação em diversos níveis (agregando mais serviços, como, por exemplo, os Bancos).

Outra linha de pensamento, a escola da integração/processos, foca a integração entre as áreas funcionais da cadeia de suprimento, conduzindo-a a um sistema definido como um conjunto de processos que busca um melhor desempenho global que adicione valor. As decisões de integração são livres para explorar configurações alternativas para a cadeia de suprimentos, objetivando a eliminação de redundâncias quando da correta realização de atividades. O exemplo mais representativo dessa escola é a indústria automobilística. Clientes que tradicionalmente estão no final da cadeia de suprimentos estão sendo usados em coordenação com fornecedores de autopeças para criarem melhores partes de veículos. Segundo os autores, os representantes dessa escola de pensamento são os trabalhos de Cooper e Ellram (1990), de Ellram e Cooper (1993), e de Hewitt (1992) apud Betchel e Jayaram (1997).

Como o objetivo desta subseção é fazer a correspondência entre as cadeias de suprimento e as redes de pesquisa. Precisa-se retomar o conceito dado por Cox et al. (2001), em que eles afirmam que uma cadeia de suprimento pode ser descrita como uma rede ampliada de relacionamentos de um grupo de troca que devem existir para a criação de qualquer produto ou serviço fornecido a um cliente final. Tomando-se esse conceito para a ótica das redes de pesquisa, pode-se dizer que a rede é formada pelos pesquisadores, que trocam, entre si, informação e conhecimento e que têm como produto final algum tipo de produção: bibliográfica, tecnológica, artística ou cultural. Dentro da gestão de cadeia de suprimentos, a

escola da informação é de correspondência mais direta, pois enfatiza justamente a troca de informações entre os membros (pesquisadores) da cadeia. Assim, uma forma de ver esse tipo de cadeia de suprimento é tratando-a como uma cadeia de troca de conhecimento, em que os pesquisadores estão suprindo seus colaboradores, e vice-versa. A troca de informação é a analogia mais direta, porque é parte do processo de geração do conhecimento. No entanto, a metáfora do fluxo de processos também é possível onde as configurações são diferentes conformações de redes (de grupos, instituições, etc.) e os processos são as comunicações para geração de produção e novos conhecimentos.

Dentro desse contexto é possível analisar de várias formas os relacionamentos entre pesquisadores, como, por exemplo, a relação entre um primeiro pesquisador orientador que orientou (no passado) um segundo pesquisador, o qual está orientando um terceiro. Esse fluxo de informação pode representar a transferência de conhecimento entre os pesquisadores em cada etapa de formação dos pesquisadores envolvidos.

Outro exemplo são os relacionamentos por citação, em que um pesquisador cita um outro pesquisador em seu artigo. Uma forma de ver esse tipo de relacionamento é tomar as referências bibliográficas como “matéria-prima” para o novo artigo.

A correspondência entre as cadeias de suprimento e as redes de pesquisa apresentadas aqui tem como propósito analisar as redes de cooperação científica como as cadeias de suprimento analisam os fluxos de mercadorias, relações e negócios entre as cadeias de empresa e organizações, pois dessa forma pode-se ter insights sobre análises simples de relações de cooperação em CT&I.

Diante dessa correspondência é possível estudar algumas perspectivas que serão propostas para estudos futuros, como a associação de diferentes bases de dados e estudos a partir das redes de colaborações, unidas com vetores de contexto para analisarem as cadeias de troca de conhecimento.

2.4 O Mundo Pequeno (*Six Degrees Separation*)

Uma cena comum de acontecer é a seguinte: numa festa com dezenas de pessoas, naturalmente, é possível que se conheçam algumas delas e não outras. E de repente, pode-se dar conta de que dois conhecidos de uma mesma pessoa

também se conhecem. Ou, em vez disso, ao conversar com alguém a quem acabou de ser apresentado, descobre-se que esse alguém afinal é cunhado de um dos melhores amigos que vive numa cidade distante. Dificilmente se consegue escapar ao lugar comum, as pessoas repetem constantemente em circunstâncias semelhantes: “O mundo é pequeno!”.

Curiosamente, resultados recentes de investigação matemática vêm mostrar que esse lugar comum é muito mais complexo do que parece. Num certo sentido muito preciso, pois as teias de relações humanas estão quase fatalmente - e surpreendentemente - condenadas a evidenciar um fenômeno desse tipo, conhecido na literatura científica precisamente como fenômeno dos “pequenos mundos” ou popularmente como “seis graus de separação”. Essa idéia consiste no fato de qualquer ser humano estar apenas a seis pessoas de distância de qualquer outra em virtude de que uma pessoa conhece alguém, que conhece alguém, que... até chegar no máximo em seis passos.

2.4.1 Breve histórico do Mundo Pequeno

A história começa no final dos anos 60, com a experiência do psicólogo norte-americano Stanley Milgram (célebre pelas suas experiências muito polêmicas sobre os efeitos da autoridade no comportamento humano). Milgram (1967) enviou um conjunto de pacotes a cidadãos do Nebraska e Kansas com instruções para que os destinatários reenviassem esses pacotes a conhecidos seus com o objetivo final de fazê-los chegar a uma dada pessoa em Massachusetts, com a qual não tinham qualquer relação. Os seus resultados foram surpreendentes: os pacotes atravessaram os Estados Unidos e chegaram ao destino através dessas cadeias, em média, ao final de cinco passos. Realizando correções estatísticas (houve pacotes que nunca chegaram ao destino), Milgram (1967) concluiu que o comprimento médio de uma cadeia de conhecimentos que leva uma pessoa a qualquer outra é seis, ou seja, qualquer pessoa está separada de outra por seis apertos de mão - seis graus de separação.

É claro que esse estudo empírico em si não é prova científica suficiente. No entanto, essa idéia perturbadora enraizou-se na cultura popular. É essa a origem da expressão “seis graus de separação”, que dá nome à peça de teatro de 1990 do dramaturgo americano John Guare e ao filme de 1993 de Fred Schepisi. Uma afirmação da personagem do filme retrata essa descoberta:

[...] todas as pessoas neste planeta estão separadas entre si por apenas seis outras pessoas. Seis graus de separação. Entre nós e qualquer outra pessoa do planeta. O Presidente dos Estados Unidos. Um gondoleiro de Veneza. Um esquimó [...] estou ligado a qualquer pessoa do planeta por uma cadeia de apenas seis pessoas. É uma idéia profunda... cada pessoa é uma porta que se abre para novos mundos. (RE:MOTE INDUCTION, 2004)

Por outro lado, a idéia deu origem a variantes curiosas como, por exemplo, ao jogo *Six Degrees of Kevin Bacon*, implementado na Internet pelo Departamento de Ciência de Computação da Universidade da Virgínia através de um programa chamado O Oráculo de Bacon. Dessa forma, dá-se a cada ator que contracenou com Bacon em algum filme o “número de Bacon” 1 (um), a um ator que nunca contracenou com Bacon mas que contracenou com alguém que contracenou com Bacon o número 2 (dois), etc. O número de Bacon mede, assim, a distância cinematográfica de qualquer ator a Bacon. Existem atores completamente separados de Bacon? A resposta é não! O número de Bacon máximo é 6 (seis), contudo, o número de Bacon médio, num universo de quase 300.000 (trezentos mil) atores, é surpreendentemente baixo, cerca de 2,9.

Os experimentos são conhecidos, embora folclóricos ou curiosos e tem pouca ou nenhuma substância científica. Até 1998, quando foi publicado na *Nature* um artigo dos matemáticos Duncan Watts e Steven Strogatz, da Universidade de Cornell, intitulado de *Collective dynamics of 'small-world' networks*. Mais tarde Duncan Watts forneceu uma versão mais profunda e pormenorizada do seu trabalho numa sucessão de artigos, culminando no seu livro *Small Worlds*, publicado pela Princeton University Press em 1999.

O trabalho de Watts e Strogatz (1998) é curioso, porque não pode ser catalogado em nenhum ramo fixo da Matemática. Pelo contrário, ele vai buscar idéias e métodos em ramos diferentes, alguns resultados são teoremas formais, outros são obtidos através de simulação por computador. O conjunto, no entanto, abre as portas a um novo campo de investigação e assume relevância científica e tecnológica.

Como é possível abordar matematicamente o fenômeno dos pequenos mundos em que a partir de uma teia de relações (seja de conhecimento social, ou por contracenar em filmes ou por escrever em conjunto) pretende-se descobrir o caminho com menos ligações que vai de um ponto a outro? A resposta está no

conceito de grafos introduzido pelo matemático suíço Leonhard Euler no século XVIII.

Deste modo, pode-se construir o grafo da colaboração matemática, em que a relação entre vértices (os matemáticos) é ter publicado um artigo em co-autoria, ou o grafo de Bacon, em que a relação entre vértices (os atores) é ter contracenado num filme, etc. A análise de grafos é precisamente o elemento matemático de base considerado por Watts e Strogatz (1998).

Já é possível perceber que o problema dos seis graus de separação não é trivial. Pensando em um grafo das relações de conhecimento entre pessoas, uma abordagem ingênua poderia abordar a questão da seguinte forma: cada pessoa no mundo conhece, por exemplo, cem outras. Assim, ao fim de seis iterações, o número de conhecimentos cresceu exponencialmente para 10^{12} , o que é aproximadamente cem vezes superior à população mundial.

A falácia, evidentemente, está no fato de as relações de amizade não serem independentes. É impossível que todos os amigos de um determinado amigo o sejam desconhecidos. Matematicamente, o grafo das relações sociais não tem uma estrutura em árvore, pelo contrário, as arestas entrecruzam-se e sobrepõem-se. É um grafo complexo, cuja estrutura exata é impossível de se determinar (e, além disso, varia com o tempo). Assim, quaisquer resultados matemáticos sobre esse problema têm de ser resultados sobre classes de grafos.

2.4.2 “Mundo Paroquiano” e “Mundo Estilhaçado”

Na abordagem de Watts e Strogatz (1998) são descritos o mundo paroquiano e o mundo estilhaçado. Considerando um conjunto com n elementos em que cada elemento se relaciona com k outros elementos, que tipos de grafos podem surgir? Num extremo estão os grafos regulares, de tipo cristalino, nos quais cada elemento relaciona-se apenas com os k vizinhos mais próximos, conforme pode ser visto na Figura 2.3.

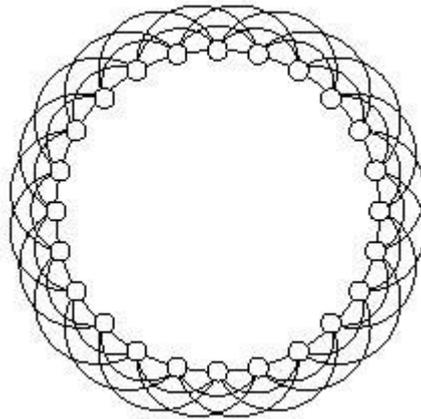


Figura 2.3 - Grafo Regular: "Mundo Paroquiano" (WATTS; STROGATZ, 1998)

Do ponto de vista social isso representa um extremo absolutamente paroquiano, cada pessoa conhece apenas as outras que vivem nas casas mais próximas - e mais ninguém. Pode-se chamar a este "o mundo paroquiano". No outro extremo, surge um grafo totalmente aleatório, as k ligações de cada vértice distribuem-se de forma totalmente aleatória por entre os outros $(n - 1)$ vértices possíveis, conforme pode ser visto na Figura 2.4.

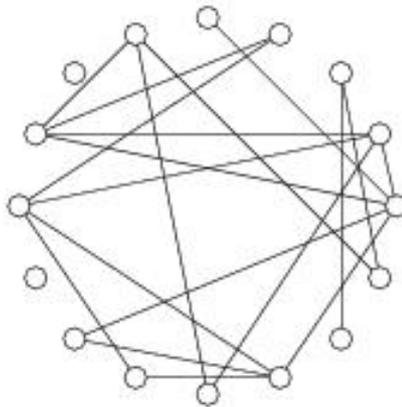


Figura 2.4 - Grafo Aleatório: "Mundo Estilhaçado" (WATTS; STROGATZ, 1998)

Em termos de redes sociais, a rede corresponde à probabilidade de dois amigos de uma mesma pessoa se conhecerem - nomeadamente, essa probabilidade é independente de se ter um amigo em comum. Se n for suficientemente grande, numa festa nenhum convidado conhece mais ninguém do que o anfitrião. A este mundo dá-se o nome de estilhaçado.

Watts e Strogatz (1998) definiram duas grandezas básicas com as quais vão caracterizar o fenómeno dos pequenos mundos.

a) Comprimento característico – corresponde à média tomada sobre todos os pares de vértices, do comprimento do percurso de um vértice a outro. O comprimento característico é, portanto, uma medida estatística do grau de separação entre vértices.

b) Coeficiente de agregação – corresponde à média sobre todos os vértices da fração de vértices, que, estando relacionados com um vértice comum, estão relacionados entre si. Este coeficiente mede a probabilidade de amigos de uma mesma pessoa serem amigos entre si ou, mais geralmente, a tendência para a formação de “cliques” ou subsociedades.

Nesse sentido, o mundo paroquiano e o mundo estilhaçado são extremos. Para o mundo paroquiano, o coeficiente de agregação é muito grande, no entanto, o comprimento característico também. Sendo assim, para se chegar ao extremo oposto da cidade é preciso bater de porta em porta. Para o mundo estilhaçado, o comprimento característico é pequeno, mas o coeficiente de agregação é mínimo, no nível do acaso. Pode-se chegar a qualquer pessoa rapidamente, contudo a pessoa será provavelmente forçada a percorrer a cidade várias vezes no processo.

Um modelo realista de sociedade é evidentemente o intermediário entre o paroquiano e o estilhaçado. A estratégia de Watts e Strogatz (1998) foi a de que, partindo de um grafo paroquiano, permitam que, com probabilidade p , cada aresta de cada vértice se desfaça e reconstitua ligando o vértice original a outro escolhido ao acaso, ou seja, cada ligação é refeita ao acaso com probabilidade p . Para $p = 0$ (zero), obtém-se o mundo paroquiano; para $p = 1$ (um), o mundo estilhaçado. Para p entre 0 e 1 obtêm-se modelos que se esperam realistas para o nosso mundo. Abaixo é apresentada a Figura 2.5 do modelo do mundo pequeno proposto por Watts e Strogatz (1998).

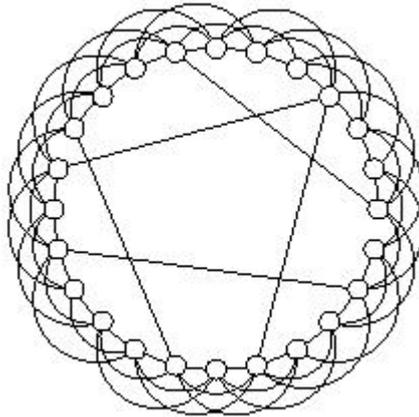


Figura 2.5 - Modelo do “Mundo Pequeno” (WATTS; STROGATZ, 1998)

As conclusões são surpreendentes. O comprimento característico e o coeficiente de agregação dependem de p de formas radicalmente opostas. Enquanto o comprimento característico sofre uma queda quase na vertical em função de p , o coeficiente de agregação mantém-se praticamente constante e igual a 1 (um) até se atingirem valores relativamente elevados de p , ou seja, basta um pequeno elemento aleatório (exemplo, $p = 10^{-5}$) nas relações para que o grafo correspondente represente um mundo em que simultaneamente há um grande grau de coesão social (amigos dos amigos conhecem-se entre si), embora simultaneamente o mundo seja pequeno. Essa situação é expansiva, persistindo por mais de quatro ordens de grandeza em p e englobando todo o intervalo de valores realista para p .

Esses tipos de resultados têm implicações extraordinárias. O fato de um mundo ser pequeno não é determinável a partir de características locais (coeficiente de agregação), isto é, uma pessoa não pode saber se vive ou não num mundo pequeno analisando apenas o que se passa à sua volta. Um mundo pode ser pequeno mesmo quando existem cliques indistinguíveis localmente dos cliques de um “mundo grande”.

Por outro lado, o fato de bastar em média apenas uma ligação aleatória em, por exemplo, 100 (cem) ligações para transformar um mundo paroquiano em pequeno significa que é necessário um número muito pequeno de atalhos para realizar essa transição. Basta um pequeno conjunto de viajantes regulares para transformar um conjunto de paróquias num mundo pequeno. Por exemplo, Watts e Strogatz (1998) constroem um modelo simples de propagação de doenças contagiosas que dá origem a um pequeno mundo, sendo suficiente para isso um número pequeno de atalhos para que a propagação seja epidêmica.

Um fenômeno semelhante a este pode ter ocorrido entre 1347 e 1348, quando ocorreu a propagação pandêmica da peste negra à escala global - da Ásia à Europa Ocidental em dois anos – num mundo em que a comunicação física era muito difícil. De fato, a introdução da peste na Europa foi feita por uma única pequena frota Genovesa proveniente do porto asiático de Kaffa em 1348. À medida que essa frota ia sendo expulsa dos vários portos em que aportava, espalhava-se a peste. Bastou um pequeno número de viajantes para transformar um mundo de comunidades medievais isoladas num pequeno mundo, com conseqüências devastadoras, como o desaparecimento de cerca de um terço (ou até metade) da população europeia em dois anos.⁷

Outro exemplo interessante é o da conectividade da World Wide Web. Estudos recentes de Barabási (2003), em que são aplicados métodos da Física Estatística, mostram que o mundo virtual de centenas de milhões de páginas Web com links que levam os usuários de umas páginas para as outras é um pequeno mundo. Por várias décadas, acreditou-se que as redes eram fundamentalmente aleatórias, isto é, presumia-se que os nós, como são as páginas da Web, as pessoas na sociedade ou as substâncias que existem na célula vinculavam-se aleatoriamente umas às outras. Mas, à medida que se começou a observar redes reais, notou-se que havia elementos recorrentes em todas elas, o que fragiliza crescentemente a hipótese aleatória. Na Web, o grupo de pesquisa de Barabási (2003) documentou a existência de uns poucos sites, como o Yahoo (<http://www.yahoo.com>), em que está dirigida para si uma extraordinária quantidade de links. Na sociedade, os sociólogos observaram a existência de conectores, isto é, uns poucos indivíduos com um número extraordinário de conhecidos. Na célula, Barabási e outros notaram a existência de umas poucas moléculas que participam em praticamente todas as reações químicas. Esses “hubs” (ou centrais), como vieram a ser chamados, simplesmente não podiam ser explicados a partir da hipótese aleatória. Os autores diziam que algumas leis comuns a todas as redes deviam existir, responsáveis pelo surgimento desses "hubs".

De acordo com os estudos de Barabási (2003), o grau de separação entre quaisquer duas páginas Web é apenas de 18,59 links, isto é, partindo de qualquer

⁷ Esse fato não pode deixar de nos preocupar, ao passo que o mundo é menor do que nunca e o risco de terrorismo por meio de armas biológicas não é uma hipótese acadêmica.

página Web, pode-se, em média, atingir qualquer outra no máximo com 19 cliques. Esse é um bom exemplo de uma rede em evolução, através da qual estão permanentemente sendo adicionados à Web novas páginas (vértices do grafo correspondente) e links (arestas do grafo). É interessante ainda notar que a popularidade do motor de busca Google (<http://www.google.com>) é alta porque ele sabe tirar proveito da Web por esta se tratar de um pequeno mundo, visto que o buscador classifica os documentos de acordo com a posição topológica que ocupam no grafo de ligações da Web.

Outro exemplo ainda é o da evolução da linguagem. A rede das palavras utilizadas na linguagem comum pode ser pensada como um gigantesco grafo (que evolui com o tempo, visto que a linguagem é dinâmica), em que palavras relacionadas estão ligadas por vértices. O físico José Fernando Mendes, da Universidade de Aveiro, provavelmente o grande especialista português em pequenos mundos e investigador ativo nessa área, mostrou recentemente num artigo nos *Proceedings of the Royal Society* que a rede de palavras na linguagem é quantitativamente um dos melhores exemplos de pequenos mundos.

Todos esses exemplos revelam que as conseqüências científicas e tecnológicas do trabalho iniciado por Watts e Strogatz (1998) são difíceis de serem imaginadas, embora haja ainda muita matemática por se fazer. Mas, como se vê, o espectro de potenciais aplicações é quase infindável. Essencialmente, todas as aplicações que dependem de redes interligadas podem se beneficiar desses resultados. Alguns dos exemplos possíveis são redes de transmissão elétrica, telecomunicações, Internet, modelos de propagação de doenças infecciosas ou de vírus de computador.

Retrospectivamente, a quantidade de problemas científicos sérios que podem se beneficiar da investigação sobre o fenômeno dos pequenos mundos. Ainda na sua infância, talvez esse fenômeno venha metaforicamente a se revelar um atalho no grafo dos problemas científicos e contribua para transformar o mundo da ciência num pequeno mundo, com poucos graus de separação.

2.5 Redes de Colaboração e Redes de Pesquisa: Um breve histórico

Nesta subseção será apresentado um breve histórico das áreas de redes de colaboração e de redes de pesquisa. As tabelas a seguir foram separadas por

períodos, em cada uma delas são descritos o ano e um breve comentário da publicação. É válido ressaltar que este histórico não pretende analisar todos os estudos existentes e sim os principais sob o ponto de vista das áreas de redes de colaboração científica.

2.5.1 Período até a década de 60

Ano	Referências	Comentário
1958	(SMITH, 1958)	Michael Smith foi um dos primeiros autores a observar um crescimento na incidência de artigos em co-autoria e a sugerir que tais artigos poderiam ser usados como uma medida aproximada da colaboração entre grupos de pesquisadores. Ele observou, também, que independentemente da natureza de um projeto de cooperação científica, seja ele interpessoal, interinstitucional ou internacional, seus resultados finais serão freqüentemente expressos por meio da publicação dos resultados produzidos pelos pesquisadores envolvidos.
1963	(PRICE, 1963)	As observações de Smith foram testadas empiricamente por Price, 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 dá, freqüentemente, no âmbito dos chamados “colégios invisíveis”.
1965	(HAGSTROM, 1965)	Em 1965, Warren Hagstrom em seus estudos fez uma curiosa observação sobre o relacionamento entre professores e estudantes. Ele constatou que alguns professores não consideravam seus orientandos como colaboradores. Prova disso, foi que durante suas entrevistas, Hagstrom perguntava aos professores que tinham escrito artigos se o trabalho realizado foi em colaboração com outros, e um número considerável de professores respondia que “não”, embora a maioria ou todos os seus artigos tivessem sido escritos juntamente com estudantes.
1966	(PRICE; BEAVER, 1966)	Price e Beaver desenvolveram estudos mostrando que os pesquisadores encontram-se em congressos, conferências, reuniões sobre as suas especialidades, visitam-se por meio de intercâmbios institucionais ou realizam trabalhos em colaboração. Esse tipo de organização transcende aos limites do departamento, da instituição, de um país e abrange cientistas de todos os lugares do mundo onde tiver atividade científica relevante na área, ou na especialidade em questão. Eles apresentaram que a maior parte das colaborações em pesquisa iniciam-se com relações informais, várias delas são estabelecidas durante o período de treinamento do pesquisador.
1967	(MILGRAM, 1967)	Stanley Milgram apresenta o “problema do mundo pequeno” (Small World), também estudado pela “estatística física”. O modelo do “mundo pequeno” afirma que os atores em uma rede (independente do seu tamanho e da densidade) podem encontrar outro ator através de 6 (seis) passos em média – também conhecido como “seis graus de separação”.

Tabela 2.3 - Publicações do período até a década de 1960

O período até a década de 60 é marcado pelo início dos estudos na área de colaboração científica. Através de estudos teóricos e testes empíricos, iniciou-se a investigação das formas com que se davam os relacionamentos de colaboração, identificando-se que os mais freqüentes aconteciam no âmbito dos “colégios invisíveis”. Esses colégios constituíram-se em comunidades informais de

pesquisadores, que se comunicam, trocam informações e experiências, e publicam formalmente seus resultados no campo do conhecimento científico.

Nesse período, observou-se também que essas publicações, na maioria das vezes, aconteciam na forma de co-autoria e que o início para esse tipo de colaboração começou nas relações entre orientador e orientando. O período foi marcado também pela teoria do “mundo pequeno”, que afirmava que o mundo era composto de uma rede de pessoas e que a distância que separava uma pessoa de qualquer outra no mundo era de no máximo 6 (seis) outras pessoas – também conhecido como “seis graus de separação”. Dentro da colaboração científica essa teoria teve significativa relevância, pois a partir dela se começou a pensar sobre as distâncias entre um pesquisador e outro, formadas pelas redes de pesquisadores que tinham artigos escritos em co-autoria.

2.5.2 Período da década de 70

Ano	Referências	Comentário
1970	(STORER, 1970)	Norman Storer afirma que boa parte do que se sabe sobre colaboração científica foi investigada através do uso de técnicas bibliométricas. Entre várias outras revelações, tais estudos têm fornecido 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; demonstrando, por exemplo, que as ciências básicas e as ciências naturais apresentam um índice maior de cooperação do que as ciências aplicadas e as ciências sociais.
1971	(MEDOWS; O'CONNOR, 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, vários autores definem a cooperação científica como sendo o conjunto de trabalhos cooperativos desenvolvidos entre dois ou mais pesquisadores e identificados por meio de artigos co-assinados.
1972	(CRANE, 1972) (LODAHL; GORDON, 1972) (NUDELMAN; LANDERS, 1972)	Diana Crane mostra que os colégios invisíveis são caracterizados por sua alta produtividade, por compartilharem prioridades de pesquisa, por treinarem estudantes, por produzirem e monitorarem o conhecimento em seu campo. Janice Lodahl e Gerald Gordon, em suas pesquisas, mostram que o fato de as ciências básicas e naturais apresentarem um índice maior de cooperação do que as ciências aplicadas e sociais se deve ao motivo de que geralmente é atribuído o caráter universal às ciências básicas e o grau de maturidade e consenso paradigmático às ciências naturais. O grau de maturidade e o consenso paradigmático se contrapõem ao caráter localizado e contingente das ciências aplicadas, que geram resultados mais facilmente apropriáveis, e à falta de consenso paradigmático das ciências sociais, que dificultam o processo de negociação entre os pesquisadores e a tomada de decisão.

		Um estudo de Nudelman e Landers sugere que o total de crédito dado por uma comunidade científica para todos os autores de um artigo é em média maior do que o crédito alocado para um único autor de um artigo, ou seja, o número de co-autores parece ser fortemente correlato com o impacto de um artigo.
1978	(BEAVER; ROSEN, 1978)	Segundo relatam Beaver e Rosen, o primeiro artigo científico escrito em colaboração, entre diferentes pesquisadores, surgiu em 1678.
1979	(BEAVER; ROSEN, 1979) (FRAME; CARPENTER, 1979)	Beaver e Rosen apresentaram que, dada a variedade de colaborações que podem se estabelecer, não é de se estranhar que as razões que levam os cientistas a colaborarem 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 adquirir novas habilidades e conhecimento tácito, como é tipicamente o caso das relações entre mestre e aprendiz. No mesmo ano, Frame e Carpenter reafirmam aquilo que Storer (1970) já tinha mostrado, que as ciências básicas e naturais cooperam mais do que as ciências aplicadas e sociais.

Tabela 2.4 - Publicações do período da década de 1970

Na década de 70 as pesquisas desenvolvidas estavam direcionadas para a descoberta de qual(is) área(s) do conhecimento mais cooperava(m) e que fundamentava(m) a dinâmica dessas relações. Algumas dessas pesquisas só foram possíveis através do uso de técnicas bibliométricas. Tais estudos evidenciaram que o grau de cooperação científica varia significativamente entre as várias áreas do conhecimento em função de suas características cognitivas e organizacionais. Esses estudos revelaram que as ciências básicas e naturais cooperam mais do que as ciências aplicadas e sociais.

Ainda nesse período, os estudos assumiram que, na grande maioria, os resultados finais de uma pesquisa seriam publicados de forma a identificar os pesquisadores envolvidos, assim como as suas instituições e seus países de origem.

Uma outra pesquisa sugeriu que uma comunidade científica costuma referenciar mais artigos em múltiplos autores do que um artigo com um único autor. Nessa fase demonstrou-se que os “colégios invisíveis” são caracterizados por sua alta produtividade dentro das pesquisas científicas. Com base em todas essas premissas, fortaleceu-se a idéia de cooperação científica por co-autoria.

2.5.3 Período da década de 80

Ano	Referências	Comentário
1980	(GOFFMAN; WARREN, 1980)	William Goffman e Kenneth Warren mostram que pesquisas por grandes grupos tendem a ter mais influência no meio científico.
1982	(STEFANIAK,	Stefaniak afirmou que a relação direta entre colaboração e co-

	1982)	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.
1983	(SUBRAMANYAM, 1983)	<p>Subramanyam mostrou que as interações entre os pesquisadores também variam em intensidade, podendo ser desde substanciais ou até praticamente vazias. Algumas vezes um pesquisador pode ser considerado como 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 através do 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.</p> <p>No mesmo estudo, Subramanyam mostrou que a contagem de artigos co-autorados tem sido a medida mais comumente usada para detectar a ocorrência, a abrangência e os participantes das colaborações científicas. Isso ocorre por várias razões. Primeiramente, porque boa parte dos outros produtos da colaboração – conhecimento tácito, aprendizagem de determinada técnica, etc. – são intangíveis, isto é, quase impossíveis de serem quantificados. Segundo, a contagem de artigos co-autorados é invariante e verificável, ou seja, dado o acesso à mesma base de dados, diferentes investigadores devem encontrar resultados idênticos. Além disso, é um método relativamente barato e prático para quantificar a colaboração e permite a análise de amostras, gerando resultados estatisticamente mais significativos do que os estudos de caso.</p>
1986	(LAWANI, 1986) (PRAVDIC; OLUIC-VUKOVIC, 1986)	Pesquisas como as de Lawani, Pravdic e Oluic-Vukovic 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.

Tabela 2.5 - Publicações do período da década de 1980

A década de 80 foi marcada pelos questionamentos relativos à definição de colaboração, pois, de acordo com os estudos realizados em vários países, houve tradições e critérios significativamente diferentes para qualificar alguém como colaborador. Nesses estudos, o objetivo era buscar os parâmetros para definição de como alguém seria classificado ou não como colaborador. No entanto, os estudos nessa fase apontaram, também, que as publicações realizadas por várias pessoas tendem a ter mais influência do que as publicações realizadas por uma única pessoa. Assim sendo, pesquisas apontaram que a contagem de artigos co-autorados tem sido a medida mais comumente usada para detectar a ocorrência, a abrangência e os participantes das colaborações científicas. Os fatores mais comumente mencionados como estimulantes para a colaboração científica eram o

aumento da visibilidade e o reconhecimento pelos pares. Deste modo, enquanto alguns buscavam definir efetivamente quem eram os colaboradores, outros tentavam medir o impacto das publicações realizadas por vários colaboradores, mesmo não tendo certeza de que os colaboradores listados no artigo eram ou não verdadeiros colaboradores.

2.5.4 Período a partir de 1990

Ano	Referências	Comentário
1990	(NARIN; WHITLOW, 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.
1991	(NARIN; STEVENS; WHITLOW, 1991) (SCOTT, 1991)	Narin, Stevens e Whitlow apontavam que em temas de pesquisa nos quais 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 significativo em artigos co-autorados por pesquisadores de diferentes países maior do que em outras áreas. Ainda durante este período (fase Pré-Internet), um trabalho muito importante para a área de Redes Sociais foi o de John Scott. Ele realizou um estudo apontando que a análise das redes sociais é uma orientação ao mundo social, que herda um conjunto particular de métodos e não uma teoria formal da teoria social. Para isso, ele faz um resgate dos estudos realizados desde 1930 com a Teoria da Gestalt, em que alemães migrados aos EUA trabalhavam com psicologia cognitiva e psicologia social. Na seqüência ele aborda a sociometria e dinâmica de grupos, visto que através de estudos de casos os laboratórios observavam como os grupos se estruturavam, quais eram os fluxos de informações e quais eram as idéias ⁸ . Seguido disso, Scott estuda a Teoria dos Grafos, iniciando-se pela teoria dos conjuntos para analisar as propriedades do espaço de interações sociais até os estudos que demonstravam que indivíduos conseguiam trabalho com melhores salários por contato social ⁹ .
1992	(LECLERC ET AL, 1992) (KODAMA, 1992) (LUUKONEN ET AL, 1992) (LIPNACK; STAMP, 1992)	Leclerc et al. apresentam um exemplo do uso da contagem de co-autoria de artigos como medida de colaboração científica. 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. Luukonen et al. mostram que os fatores que levam os pesquisadores a colaborar são agrupados em cognitivos, econômicos e sociais, dentro dos quais cada um deles tem importância relativa variada para explicar as diferenças nas taxas de colaboração das áreas do conhecimento e dos diversos países. Lipnack e Stamp mostram que o que é novo no trabalho em redes de conexões é sua promessa como uma forma global de organização com raízes na participação individual. Uma forma que

⁸ Jacob Moreno foi o primeiro autor a representar esses tipos de relações em Sociogramas (MORENO, Jacob. *Who Shall Survive?* New York: Beacon Press, 1934).

⁹ GRANOVETTER, M. *Getting a Job*. Cambridge, MA: Harvard University Press, 1974.

		reconhece a independência enquanto apóia a interdependência. O trabalho em redes de conexões pode conduzir a uma perspectiva global baseada na experiência pessoal.
1993	(KATZ, 1993)	Katz mostra que as colaborações decrescem exponencialmente com a distância geográfica que separa os pesquisadores institucionais. Assim sendo, a proximidade regional parece encorajar a colaboração, já que os pesquisadores tendem a gerar mais comunicação informal.
1994	(PAKDAMAN, 1994) (KATZ, 1994) (DEGENE; FORSÉ, 1994)	Logo após a Segunda Guerra Mundial, várias agências de cooperação para o desenvolvimento, todas elas com um forte componente em cooperação em C&T, foram criadas com o intuito de “ajudar a integração dos países pobres do sul à arena internacional”. Pakdaman apresenta uma revisão recente desses instrumentos e agências. Ainda que a maneira de atuar das agências e programas contemporâneos tenha se modificado, a racionalidade de programas cooperativos, que obrigatoriamente envolvem pesquisadores europeus e latino-americanos, ainda é basicamente a mesma. No mesmo ano, Katz mostra 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. Outro estudo realizado neste ano foi de Degene e Forsé, em que mostram que a análise de redes não constitui um fim em si mesma. Ela é o meio para se realizar uma análise estrutural cujo objetivo é mostrar que a forma da rede dos fenômenos analisados é explicativa. O intuito é demonstrar que a análise de uma díade (interação entre duas pessoas ou dois grupos) só tem sentido em relação ao conjunto das outras díades da rede, porque a sua posição estrutural tem necessariamente um efeito sobre sua forma, seu conteúdo e sua função. Portanto, a função de uma relação depende da posição estrutural dos elos, e o mesmo ocorre com o status e o papel de um ator. Uma rede não se reduz a uma simples soma de relações, e a sua forma exerce uma influência sobre cada relação.
1995	(COLONOMOS, 1995)	Colonomos mostra que houve um grande investimento acadêmico nos estudos de redes a partir do campo das relações internacionais, e isso se manifestou sobre a história recente das ciências políticas. A origem da reflexão se dá logo no 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 através da reflexão de ordem e desordem nos sistemas políticos, ligados à elaboração da noção de globalização.
1996	(WEISZ; ROCO, 1996)	Weisz e Roco afirmam que um número crescente de projetos de pesquisa e ensino tem sido concebido e executado na forma de cooperação, em que diferentes instituições assumem tarefas distintas com vistas a um dado resultado. Assim, uma rede pode ser então definida como uma organização de coesão tênue, consistindo em 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 promoverem a interação de participantes com qualificações complementares. O grau de participação de cada unidade participante é flexível.
1997	(KATZ; MARTIN, 1997)	Katz e Martin mostram que, em geral, os trabalhos teóricos produzem artigos com poucos co-autores em comparação aos trabalhos experimentais, ou seja, experimentalistas tendem a colaborar mais do que teóricos. Ainda neste mesmo trabalho, Katz e Martin apresentaram os diferentes níveis de colaboração, entre

		indivíduos, grupos, departamentos, instituições e setores, nas mais diferentes combinações dessas unidades, dentro de uma mesma nação ou envolvendo nações diferentes.
1999	(CASTELLS, 1999)	Castells define rede como um conjunto de nós interconectados. Nó é o ponto no qual uma curva se entrecorta. Redes são estruturas abertas capazes de se expandirem de forma ilimitada integrando novos nós, desde que consigam comunicar-se dentro da rede, ou seja, desde que compartilhem os mesmos códigos de comunicação.
2000	(NEWMAN, 2000)	De acordo com Newman, uma rede, sob o aspecto social, é 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 rede social, as pessoas ou os grupos são chamados de “atores”, e as conexões, de “nós”. Ambos, atores e nós, podem ser definidos em diferentes caminhos dependendo da questão de interesse. Um ator pode ser uma única pessoa, um grupo ou uma empresa. Um nó 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.
2001	(NEWMAN, 2001)	Sob o aspecto de redes de pesquisa, para Newman, dois pesquisadores podem ser considerados conectados ou ligados se eles tiverem co-autoria em um ou mais artigos juntos. Assim, os atores são os pesquisadores, e os nós entre eles são as colaborações científicas. As propriedades estatísticas para essas redes são: número de artigos escritos por cada 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, e uma variedade de medidas de conexão dentro da rede, tais como proximidade e intermediação. Essas redes de relacionamento mostram um alto grau de transitividade, o que significa que há uma maior probabilidade de duas pessoas estarem relacionadas se elas tiverem um ou mais relacionamentos em comum, ou seja, a possibilidade de pesquisadores colaborarem aumenta com o número de outros colaboradores em comum. Outra conclusão de Newman é de que a probabilidade de um pesquisador, em particular, adquirir novos colaboradores aumenta com o número de pesquisadores com quem ele colaborou no passado.
2003	(BARABÁSI, 2003) (ERDÖS; RÉNYI, 1951) (MILGRAM, 1967) (GRANOVETTER, 1973) (ALBERT; JEONG; BARABÁSI, 1999)	Albert-László Barabási em seu livro <i>Linked</i> faz um estudo muito interessante sobre a influência das Redes Sociais. Ele afirma que é possível ir além do estudo da estrutura e da topologia das redes sociais para sua dinâmica. Para isso, ele realiza um levantamento sobre as teorias de redes sociais, iniciando pela Teoria das Redes Randômicas de Erdős e Rényi, em que os autores afirmam que as redes sociais teriam origem randômica a partir do comportamento dos nós da rede ¹⁰ . Em seguida, ele apresenta a Teoria dos Seis Graus de Separação, na qual Stanley Milgram (<i>The Small World Problem</i>) com base em uma pesquisa direta chegou à conclusão de que a distância média entre duas pessoas nos EUA é de 5,5. Depois ele apresenta os estudos de Mark Granovetter, que diz que a sociedade é formada por uma rede completamente conectada de clusters (com nós fortemente conectados), que se comunicam entre clusters por ligações fracas, ou seja, que as sociedades não são randômicas ¹¹ . Em seguida, ele exibe os estudos que fez em 1998 utilizando como exemplo a Web, com 800 milhões de nós e 18,59

¹⁰ ERDÖS, Paul; RÉNYI, Álfred. Connectivity of Random Nets. *Bulletin of Mathematical Biophysics*, n. 13, p. 107-227, 1951.

¹¹ GRANOVETTER, Mark S. The Strength of Weak Ties. *American Journal of Sociology*, n. 78, p. 1360-1380, 1973.

		links de distância; e conclui que cientistas em campos diferentes da ciência, em termos de co-autoria, são separados em distâncias de 4 a 6 autores. Utilizando 203 milhões de Web pages, ele deduz que 90% de todos os documentos têm 10 ou menos links direcionados a eles e que 3% são referenciados por quase 1 milhão de outras páginas (hubs), originando a Teoria dos Hubs e Connectors. Por último, Barabási mostra os estudos que fez em 1999, apresentando as Redes Reais, que têm crescimento de aderência preferencial, ou seja, uma rede começa com um pequeno núcleo, e a cada período de tempo um novo nó é acrescentado (1 nó por vez), e os nós conectam-se preferencialmente aos mais conectados ¹² . Assim, diante de todos esses estudos realizados por Barabási, pode-se verificar que as técnicas existentes são aplicáveis à própria Web.
--	--	---

Tabela 2.6 - Publicações a partir de 1990

A partir de 1990, estudos mostram a importância das colaborações científicas internacionais, indicando que artigos co-autorados por países diferentes são citados duas vezes mais que artigos de um único país. Por outro lado, nesse mesmo período, foi apresentado um estudo que constatou o fato de que as colaborações decresciam exponencialmente com o aumento da distância geográfica que separa os pesquisadores institucionais, sendo um dos motivos a pouca colaboração internacional (é importante ressaltar que essa visão vem antes do período Pré-Internet). Outra pesquisa mostrou que em geral os trabalhos teóricos produzem artigos com poucos co-autores em comparação aos trabalhos experimentais. Nesse período, a colaboração científica é vista sob diferentes níveis de colaboração, entre pesquisadores, grupos, departamentos, instituições e setores, nas mais diferentes combinações dessas unidades, dentro de um mesmo país ou envolvendo países diferentes. Outra característica presente nesse período é a junção de várias áreas (Link Analysis, Teoria dos Grafos, ciência da informação e redes sociais) para o entendimento e/ou a visualização das redes de relacionamentos formadas pela colaboração científica.

A Figura 2.6 representa essas características cronológicas da área de redes de colaboração científica.

¹² ALBERT, Réka; JEONG, Hawoong; BARABÁSI, Albert-Lászlo. Diameter of the World Wide Web. *Nature*, n. 401, p. 130-131, 1999.

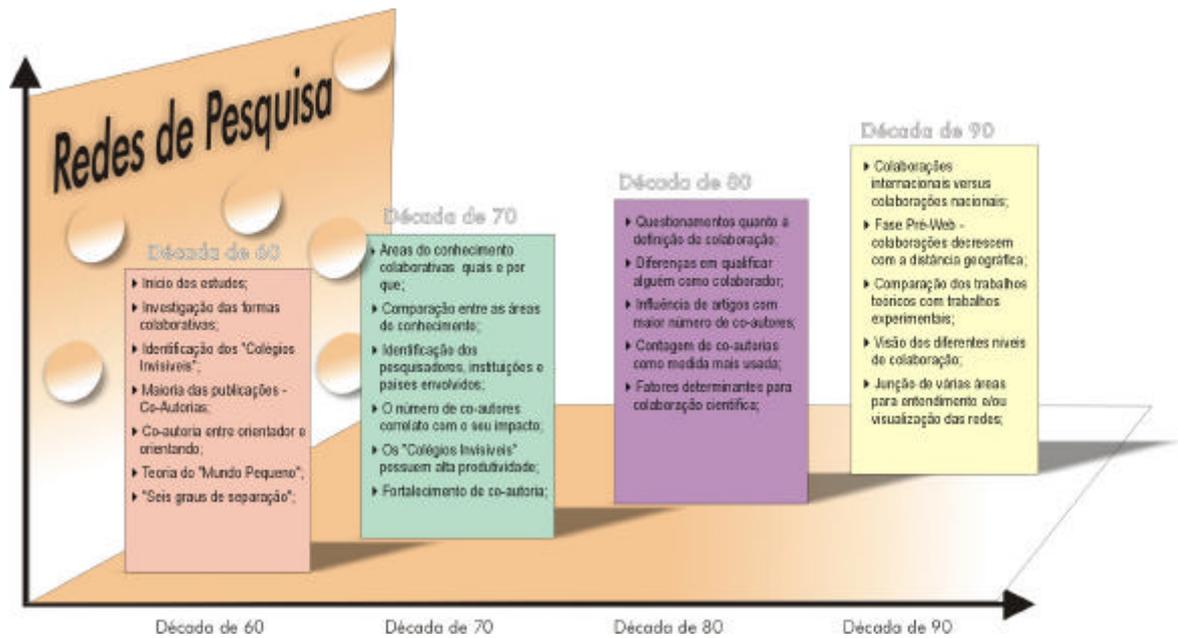


Figura 2.6 - Cronologia de Redes de Colaboração Científica

2.6 Considerações Finais

Este capítulo abordou aspectos referentes à área de redes de colaboração, suas diferentes abordagens e áreas de estudo, dando ênfase às redes de colaboração científica. O estudo mostra os diferentes fatores que contribuem para a colaboração em redes de pesquisa, os níveis da colaboração e as taxonomias propostas para este trabalho. Também é apresentada uma correspondência entre as cadeias de suprimento (*supply chain*) e as redes de pesquisa, possibilitando analisar os relacionamentos entre os atores de CT&I como as cadeias de suprimento analisam os fluxos de mercadorias, os negócios entre as cadeias de empresas e as organizações. Finalmente, exibe-se um breve histórico, separado por décadas, do estudo das redes de colaboração e da evolução cronológica das pesquisas realizadas nessa área.

No próximo capítulo será mostrado que todos os estudos de redes de pesquisa têm, em sua base, relação com uma plataforma de gestão de CT&I, ou seja, incluem a escolha de uma fonte de informação e de uma possibilidade de análise sobre como redes de pesquisa se estruturam. Dessa forma, o objetivo do próximo capítulo é apresentar uma plataforma de gestão de CT&I e dar destaque aos aspectos normalmente presentes nas análises de redes de pesquisa.

3 PLATAFORMA PARA GESTÃO DE CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Hoje, no entanto, as fontes de poder não são mais o controle de território, os recursos naturais ou a produção industrial, mas sim o controle sobre a informação, tecnologia e conhecimento.

Dedijer Stevan

3.1 Introdução

A gestão de ciência, tecnologia e inovação é essencial para consolidar e ampliar a capacitação científica e tecnológica no País. Consiste em administrar o desenvolvimento de estratégias e instrumentos organizacionais que envolvem aspectos estruturais, culturais, políticos, tecnológicos, gerenciais e de serviços, de forma a promover a pesquisa viável e relevante, além de estimular a parceria de setores público e privado para a transferência de tecnologia que favoreça o processo de inovação (SEMINÁRIO VIRTUAL, 2003).

Alguns objetivos principais podem ser alcançados pela gestão de CT&I, tais como

- ? a escolha de linhas de pesquisa prioritárias quanto à relevância para o desenvolvimento socioeconômico e cultural;
- ? a execução mais eficiente das pesquisas; e
- ? a conversão mais rápida de resultados obtidos em contribuições concretas para a sociedade.

Esses objetivos devem ser considerados em três níveis de gestão:

- ? o das políticas públicas;
- ? o institucional (universidades, institutos de pesquisa, empresas, etc.); e
- ? o de programas e projetos específicos de pesquisa.

No nível de políticas públicas destaca-se a gestão de investimentos em CT&I para:

- ? a criação e manutenção de infra-estrutura;
- ? a capacitação de recursos humanos; e
- ? o desenvolvimento de mecanismos de acesso a recursos e informações específicas necessárias à pesquisa e à inovação.

No Brasil, essa política está inscrita no Plano Básico de Desenvolvimento Científico e Tecnológico (PBDCT), para o qual foram criados o Fundo Nacional de Desenvolvimento Científico e Tecnológico (FNDCT) e o Programa de Apoio ao Desenvolvimento Científico e Tecnológico (PADCT) do Ministério de Ciência e Tecnologia. Há dois órgãos principais de financiamento nesse sistema: o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Financiadora de Estudos e Projetos (FINEP).

A gestão de CT&I em cada instituição do sistema deve se acoplar ao nível de gestão das políticas públicas. Nesse sentido, cada instituição deve definir uma política própria e clara quanto a projetos científicos e tecnológicos viáveis e relevantes. Para isso, a instituição precisa:

- ? identificar suas capacidades específicas e combinações de recursos e competências;
- ? aproveitar bem suas características próprias, além das contextuais; e
- ? estabelecer formas de parcerias com outras instituições do sistema de CT&I.

No que se refere aos programas e projetos de pesquisa, a gestão abrange um processo constituído das seguintes atividades: planejamento, captação de recursos, gerenciamento de recursos (infra-estrutura, pessoal, orçamentário), gerenciamento de execução das atividades, divulgação científica, difusão tecnológica e reprogramação.

Em todos os níveis de gestão de CT&I há que se desenvolverem os processos de planejamento, monitoramento e avaliação, de forma a assegurar relevância, viabilidade e execução adequadas e eficientes de pesquisas (SEMINÁRIO VIRTUAL, 2003).

Na Era da Informação, é de suma importância a utilização pela sociedade e principalmente pelos organismos governamentais de sistemas de informações que forneçam subsídios para a tomada de decisão diária e para o planejamento de ações futuras. Um instrumento capaz de fornecer as respostas para um governo eletrônico são as Plataformas de Gestão de CT&I.

3.2 Plataforma de Gestão de CT&I

Uma plataforma de gestão de CT&I é composta de sistemas de captura de dados, bases de dados e portais Web de intercâmbio de informações. Seus recursos servem aos diversos atores ligados à CT&I e, no campo institucional, atendem às missões de Organizações Nacionais de CT&I – ONCYT (e.g., CNPq e Capes no Brasil, Conicyt no Chile, OCT em Portugal), instituições de ensino e pesquisa e Organizações Internacionais de CT&I – OICYT (e.g., OPS, Bireme, Unesco), mas, além desses atores naturais, ou seja, atores ligados diretamente a CT&I, existe uma grande quantidade de informações que podem ser aproveitadas por outros órgãos governamentais (Ministérios, Secretarias, Prefeituras e Governos Estaduais e Federal) (MONTENEGRO, 2002).

Atores de produção de conhecimento (i.e., pesquisadores, docentes e estudantes) e organizações que oferecem serviços a esses usuários (e.g., bibliotecas) utilizam as informações segundo os conceitos de bibliometria (i.e., “estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada”). Gestores e tomadores de decisão ligados a ONCYT e OICYT necessitam de recursos da cienciometria e informetria, que estudam, respectivamente, os aspectos quantitativos da ciência como atividade econômica e da informação em qualquer formato (MACIAS-CHAPULA, 1998).

Respostas para as perguntas “quais os principais pesquisadores e/ou grupos de pesquisa que podem ajudar na resolução do problema do fornecimento de energia elétrica?” e “onde se está trabalhando com soluções alternativas para o problema da

fome?” podem ser encontradas dentro de uma plataforma de CT&I. É também possível ajudar governos estaduais a encontrarem soluções para problemas na área de saúde, por exemplo, e essa solução pode ser encontrada em seu próprio Estado, ou em outro local (MONTENEGRO, 2002).

O maior desafio para uma plataforma de sistemas de gestão de CT&I caracteriza-se pela necessidade de combinar essas visões complementares em Sistemas Gerenciadores de Banco de Dados (SGDB), comuns ou integrados, de forma a permitir maior transparência nos processos de gestão da informação em CT&I. Além disso, a plataforma deve viabilizar a divulgação de suas informações em sites e portais de CT&I, respeitando os interesses específicos de cada um dos atores ligados às atividades científicas e tecnológicas (MONTENEGRO, 2002).

A consequência natural da disponibilização das informações em sites e portais é o aumento da visibilidade, e isso faz com que cresça a preocupação com a qualidade das informações, pois estas vão estar disponíveis a uma quantidade crescente de pessoas, instituições e empresas interessadas. A qualidade da informação deve estar no ponto de saída dessas informações, ou seja, em sistemas de captura de dados fáceis de serem utilizados.

A seguir, será apresentada uma arquitetura de uma plataforma de gestão de CT&I que dará ênfase aos aspectos normalmente presentes nas análises de redes. A apresentação de uma plataforma em todas as suas camadas permitirá mostrar que estudos já realizados em redes sociais podem ser levados para redes de pesquisa quando disponíveis os componentes necessários (ex.: um estudo em redes sociais que se baseou em trocas de mensagens eletrônicas ou em utilização de fóruns pode ser feito no âmbito de redes de pesquisa se a plataforma de gestão incluir esses recursos na camada de portais e em serviços Web).

3.3 Plataforma Lattes

A Plataforma Lattes, projeto do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil (CNPq), é um conjunto de sistemas de informações, bases de dados e portais Web voltados para a gestão de CT&I. Foi concebida para integrar os sistemas de informações das agências federais, racionalizando o processo de gestão de CT&I.

3.3.1 Breve histórico

Lançada em 16 de agosto de 1999 com a primeira versão do Sistema CV-Lattes, nas versões Windows (off-line) e Web (on-line), com o objetivo de estabelecer um ambiente integrado para o conjunto de sistemas de informações em CT&I, utilizados pela comunidade científica do País, proporcionou um aumento significativo no número de currículos enviados ao CNPq, que chegou a mais de 100 novos currículos Lattes ao dia.

No ano de 2000 foi a vez do lançamento do Sistema do Diretório de Grupos de Pesquisa versão 4.0, e durante este mesmo ano foram disponibilizados no site do CNPq os primeiros mecanismos de busca de currículo. No ano de 2001 foram aprimorados os sites de busca através da implantação do Data Warehouse da Plataforma Lattes.

Até agosto de 1999, o sistema de informação do CNPq continha um total aproximado de 35 mil currículos de pesquisadores, atualmente, mais especificamente em janeiro de 2004, conta com cerca de 346 mil currículos. Nessa primeira fase da Plataforma Lattes, a prioridade encontrava-se no desenvolvimento de sistemas de captura de dados (CV-Lattes e Grupo de Pesquisa do Brasil), com interfaces de fácil navegação e oferecendo recursos adicionais de gerenciamento e visualização das informações, em que o usuário do sistema, dono de sua informação, tenha possibilidade de mantê-las com a qualidade necessária ao sistema de informação da Plataforma Lattes.

No ano de 2001 foi firmado um convênio com a Organização Pan-Americana de Saúde (OPS), com sede em Washington, e com a Bireme, organização brasileira vinculada à OPAS, para o desenvolvimento da plataforma latino-americana de currículos, Plataforma LAC, compatível com a Plataforma Lattes. Inicialmente este projeto contava com a participação de 5 (cinco) países: Chile, Colômbia, México, Venezuela e Cuba. Atualmente já aderiram Argentina, Equador, Panamá e Peru. Em 2002 foi a vez de Portugal firmar convênio com o CNPq para utilizar a tecnologia da Plataforma Lattes, batizada neste país de Plataforma De Góis. Com a participação desses países, foi formalizada a Rede Internacional de Fontes de Informação e Conhecimento para Gestão de Ciência, Tecnologia e Inovação - Rede ScienTI, que corresponde à unificação das Plataformas Lattes, LAC e De Góis, além de outras iniciativas dos países participantes.

A Plataforma Lattes mudou gradativamente a face das informações em C&T do País e transformou-se em recurso para a comunidade científica brasileira (SABBATINI, 2001).

Em termos evolutivos, a Plataforma partiu da agregação de informações disponíveis nas diversas bases relacionadas à CT&I do País, promovendo posteriormente a padronização de cada unidade, a construção e a divulgação de serviços de informação e, mais recentemente, o intercâmbio internacional para a comunidade científica do Brasil, dado que se transformou em tecnologia e metodologia de referência para a Rede ScienTI (REDE SCIENTI, 2003).

A Plataforma Lattes contempla atualmente conjunto de bases de dados, sistemas de informação, diretórios de serviços e portais Web, compartilhados por gestores, técnicos de governo, comunidade científica e sociedade em geral. Sua construção é fruto do trabalho multidisciplinar e multistitucional, coordenado e patrocinado pelo CNPq (PACHECO, 2003).

Contudo, a Plataforma Lattes não se trata de um conjunto independente de sistemas. Sua construção foi realizada seguindo a metodologia integrada para desenvolvimento de plataformas de governo proposta por Pacheco (2003).

3.3.2 Arquitetura da Plataforma Lattes

Na Figura 3.1 é apresentada a arquitetura da Plataforma Lattes, representada em forma de pirâmide, sendo a base composta de unidades de informação da Plataforma. Depois da base são exibidas a padronização, sistematização e publicação de informações e serviços, e, por último, é mostrado o topo, em que estão os instrumentos de gestão, produção e publicação de conhecimento.

A estrutura arquitetônica da Plataforma Lattes é composta de níveis conceituais 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).

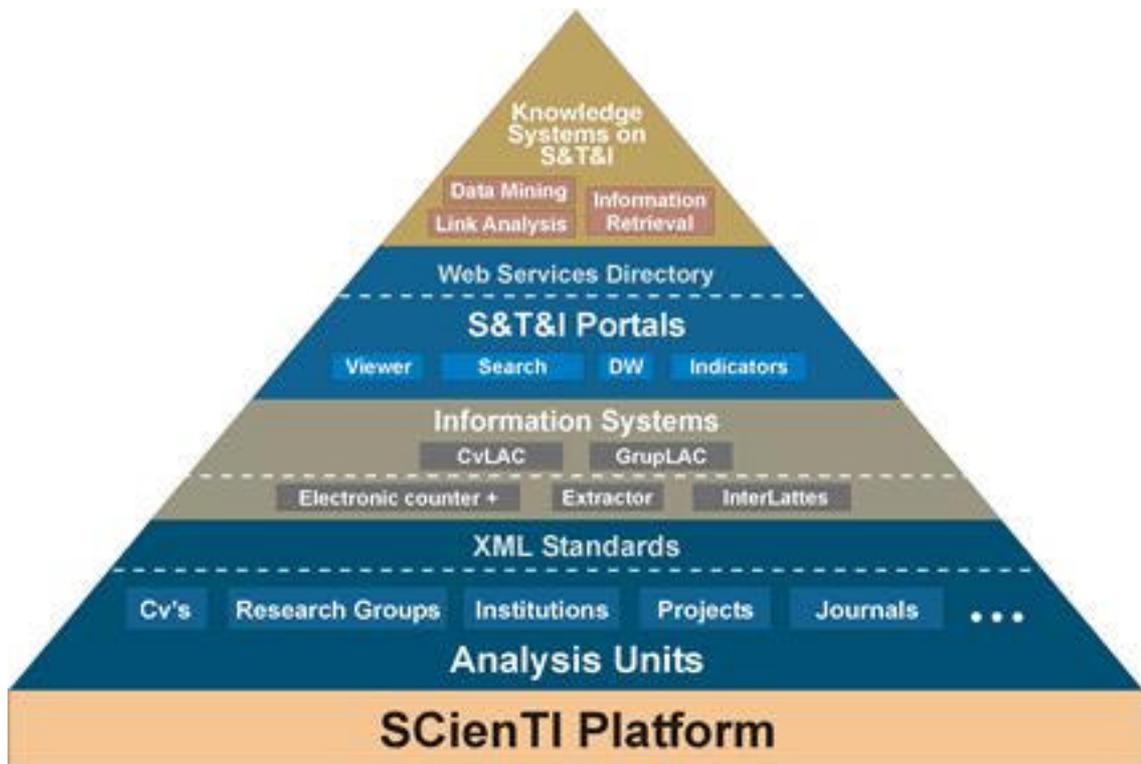


Figura 3.1 - Arquitetura Conceitual da Plataforma Lattes (Fontes: STELA, 2002)

As quatro camadas conceituais da arquitetura da Plataforma Lattes formam um conjunto de elementos tecnológicos e metodológicos: unidades de análise e normalização da informação de CT&I, sistemas e ferramentas de infra-estrutura de informação, diretórios de informação sobre CT&I, portais relacionados com CT&I e sistemas de extração de conhecimento de CT&I. A seguir é descrito individualmente cada elemento pertencente à arquitetura.

3.3.2.1 Unidades de Informação

Na primeira camada dessa arquitetura estão situadas as Unidades de Informação, as quais descrevem subdomínios da área-fim à qual se destina a Plataforma. As unidades de informação estão associadas a conteúdos, processos e serviços específicos. Segundo Pacheco (2003),

[...] uma unidade de informação não pode ser genérica a ponto de não especificar conteúdo e processos independentes nem ser específica na descrição ou funcionamento (casos em que provavelmente seja um elemento da unidade de informação).

Ainda, segundo o autor, uma unidade tem identificação unívoca para cada registro da plataforma de informação e possui relação direta com outras unidades de informação do domínio da Plataforma.

As unidades de informação da Plataforma Lattes são denominadas unidades de análise, taxonomia que melhor representa as especificações e os estudos do domínio da CT&I. As tarefas de identificação e especificação das unidades de análise da Plataforma Lattes são resultado da análise de cada subdomínio da área de CT&I. Como resultado, a Plataforma Lattes registra as seguintes unidades de informação: currículo, grupo de pesquisa, projeto de pesquisa e instituição.

- ✍ **Currículo.** Conjunto de informações que descrevem a atividade profissional dos usuários da Plataforma, incluindo-se todas as pessoas relacionadas à CT&I (pesquisadores, estudantes, docentes, gestores de CT&I, técnicos de governo, administradores e representantes da sociedade interessados no tema) (PACHECO, 2003).
- ✍ **Grupos de Pesquisa.** Esta unidade representa as informações referentes aos grupos de pesquisa, definidos como conjuntos de indivíduos organizados hierarquicamente, de acordo com hierarquia de experiência e liderança no terreno científico ou tecnológico em que atuam. Congrega profissionais envolvidos permanentemente com atividades de pesquisa, cujo trabalho é organizado em torno de linhas comuns de pesquisa, compartilhando instalações e equipamentos (GUIMARÃES et al., 1999).
- ✍ **Projetos de Pesquisa.** Esta unidade contempla a descrição das atividades de pesquisa, do desenvolvimento ou da extensão, realizada por um pesquisador ou por uma equipe de pesquisa, tendo como base um tema ou objeto específico, com objetivos, metodologia e duração definidos (i.e., projetos de pesquisa), e os pedidos de auxílios enviados ao CNPq (PACHECO, 2003).
- ✍ **Instituições.** Esta unidade representa as informações de organizações, institutos, empresas, universidades e demais organismos ligados à CT&I, referenciados nas demais unidades como local de lotação profissional ou de pesquisa, ou representando atores institucionais que interagem com o CNPq (PACHECO, 2003).

As unidades de análise são resultado direto do trabalho realizado nas etapas de levantamento de requisitos, planejamento e estudo de projetos correlatos. O levantamento de requisitos tem sido realizado de forma contínua junto a

representantes da comunidade científica (inicialmente por consulta direta na etapa de concepção de sistemas e posteriormente por análise freqüente das críticas e sugestões enviadas, bem como pelas informações fornecidas pela central de atendimento ao usuário da Plataforma), junto a técnicos do CNPq e de outras agências e, mais recentemente, através de revisões junto a especialistas e grupos de desenvolvimento internacionais (em função do trabalho cooperativo realizado na Rede ScienTI) (PACHECO, 2003).

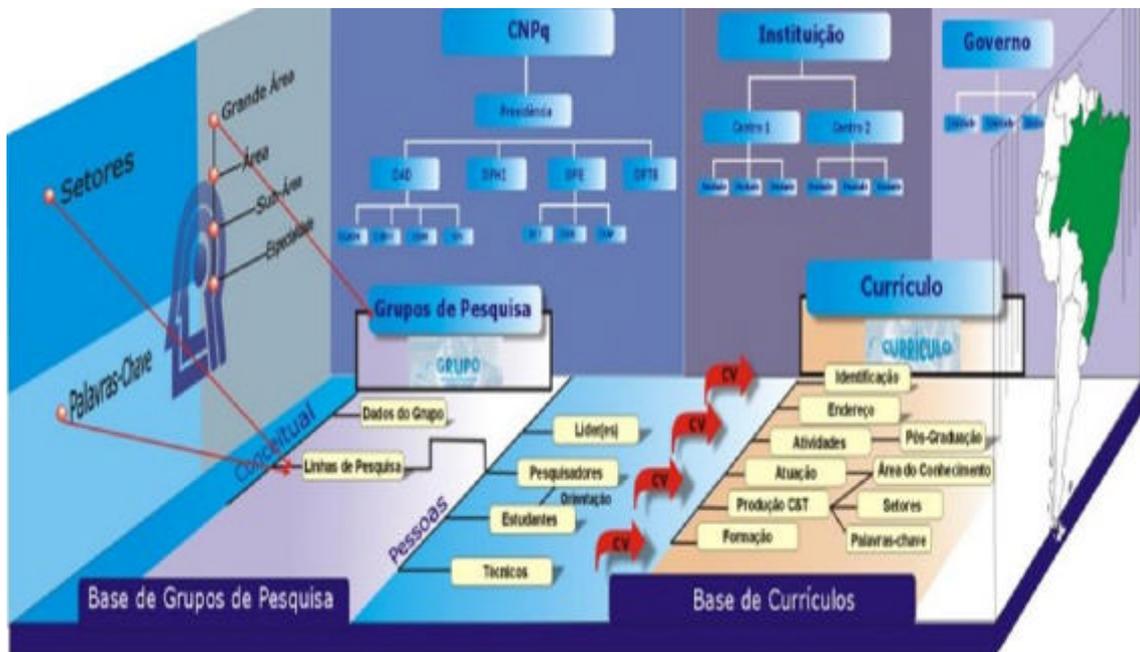


Figura 3.2 - Unidades de Análise da Plataforma Lattes (Fonte: STELA, 2002).

Na Figura 3.2 representa-se de forma esquemática a relação entre as unidades de análise da Plataforma Lattes (currículos e grupos de pesquisa), as diferentes formas de classificá-las e de consultas a elas, bem como as principais comunidades usuárias (instituições).

As unidades de análise formam, portanto, a base conceitual da Plataforma Lattes. Os processos de construção de padrões para cada unidade e de atualização contínua desses padrões têm sido efetivados por comunidade virtual específica.

3.3.2.1.1 Abertura e Padronização XML

Com a disseminação da Plataforma Lattes como principal recurso de informações em gestão de CT&I do País, o CNPq e os grupos responsáveis passaram a ser solicitados para a abertura dos sistemas e, sobretudo, para a padronização das

informações geradas pela Plataforma. As demandas tinham como raiz o problema da interação dos sistemas Lattes com os sistemas corporativos das instituições ligadas à CT&I no Brasil (PACHECO, 2003).

No ano de 2000, as Instituições Federais de Ensino Superior reuniram suas equipes de informática no Workshop de Sistemas de Informações das IFES (UFOP - Ouro Preto) e convidaram as agências federais para a construção de um modelo único de informação, visando racionalizar o processo de captura de dados no Sistema Federal de Educação em Ciência e Tecnologia (CNPq, 2003).

Em fevereiro de 2001, como resposta a essa demanda, formava-se a Comunidade de Linguagem de Marcação da Plataforma Lattes - LMPL, com o objetivo de se estabelecerem padrões para cada unidade de análise Lattes, de forma colaborativa e interativa. Entre os integrantes da Comunidade incluem-se pesquisadores e técnicos de informática das instituições-membro, e especialistas conselheiros. Todos são responsáveis pela padronização das unidades de informação da Plataforma e, por meio de workshops e interação virtual na Web, têm mantido contínuo trabalho de construção e divulgação dos padrões Lattes (PACHECO, 2003).

O resultado da formação da Comunidade Virtual LMPL foi a definição do modelo DTD (*Data Type Definition*) XML do Currículo Lattes, que faz parte da versão 1.4. Com esse modelo, as universidades brasileiras podem extrair informações do currículo Lattes e/ou gerar informações para o currículo a partir dos seus sistemas corporativos. O projeto viabilizou a abertura da Plataforma Lattes, do ponto de vista de conteúdo dos dados, e manteve inalterado o acesso técnico às informações, preservando a segurança dos pesquisadores (CNPq, 2003).

A definição de padrões Lattes e, principalmente, sua disseminação e inclusão nos sistemas de informação têm promovido pesquisas específicas da comunidade científica nacional na proposição de ontologias para a Plataforma Lattes (e.g., BONIFÁCIO, 2002).

A construção de ontologias comuns tem sido proposta como abordagem promissora para a interoperabilidade de sistemas. Ontologia é usualmente definida como “a especificação explícita de uma conceitualização” (GRUBER, 1994). Uma ontologia comum é uma formalização compartilhada de um certo domínio de

aplicação. Assim, por exemplo, uma formalização de conceitos sobre informações em ciência e tecnologia poderia permitir que diversos aplicativos desse domínio compartilhassem um vocabulário comum sobre o assunto. Essa é a característica da Linguagem de Marcação da Plataforma Lattes (PACHECO; KERN, 2001).

A LMPL foi escrita em XML, que é uma linguagem de marcação. Bax (2000) delineou as principais características de linguagens de marcação. A linguagem HTML (*hypertext markup language*) foi projetada para marcar documentos com conteúdo fixo. Em contraste, a linguagem XML foi projetada para a marcação de documentos com conteúdo variável. Desta forma, pode-se encarar um banco de dados como um texto de conteúdo variável ao qual se aplica uma marcação que segue regras gramaticais bem definidas.

Bray et al. (1998) afirmam que, no XML 1.0, um documento bem formado deve ajustar-se a uma estrutura sintática bem definida. Assim, é documentada a forma de uso de tags e restrições aplicáveis a um documento XML. Para que um documento XML seja também válido para intercâmbio, deve aderir a uma estrutura sintática definida no DTD.

a) Currículo

A Figura 3.3 ilustra de forma esquemática a estrutura da informação sobre currículos de pesquisadores conforme a LMPL. Segundo o esquema, as informações presentes em qualquer curriculum vitae são categorizadas em dados gerais, produção bibliográfica, produção técnica, outra produção e dados complementares. Essa definição é o resultado do consenso dos membros da Comunidade LMPL.

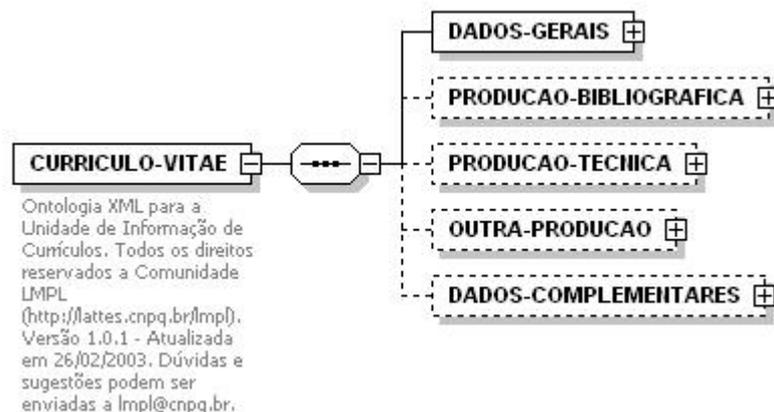


Figura 3.3 - Diagrama esquemático da representação de informações sobre currículos segundo a LMPL (Fonte: CNPq, 2003)

A Figura 3.4 apresenta um esquema mais detalhado do que o da Figura 3.3, mostrando como a informação curricular referente a “dados gerais” é organizada. Os dados gerais de cada pesquisador são organizados em: endereço, formação acadêmica, atuações profissionais, áreas de atuação, idiomas e prêmios e títulos.

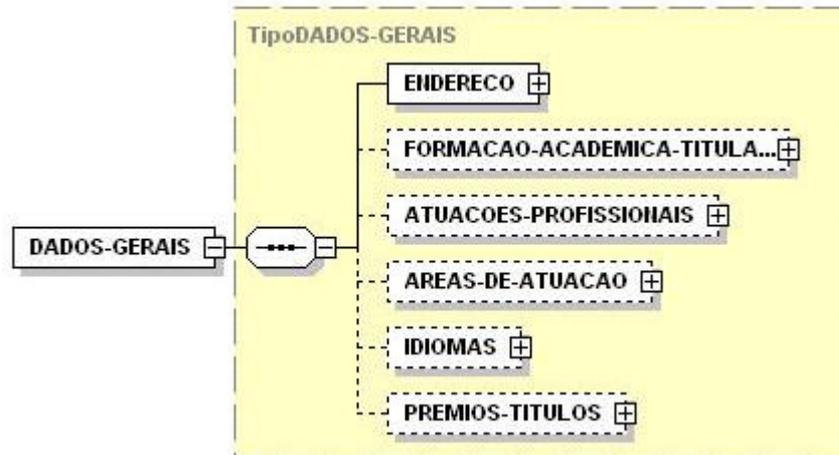


Figura 3.4 - Detalhamento do diagrama da Figura 3.3, mostrando a organização dos dados gerais de qualquer currículo (Fonte: CNPq, 2003)

A LMPL utiliza-se de sucessivos níveis de detalhamento como os apresentados nas Figuras 3.3 e 3.4 para representar a estrutura da informação curricular dos pesquisadores.

b) Grupo de Pesquisa

A Figura 3.5 apresenta de forma esquemática a estrutura da informação sobre a unidade de análise grupo de pesquisa. Segundo o esquema, as informações presentes em um grupo de pesquisa são categorizadas em identificação do grupo, pesquisadores, estudantes, técnicos, linhas de pesquisa e empresas.

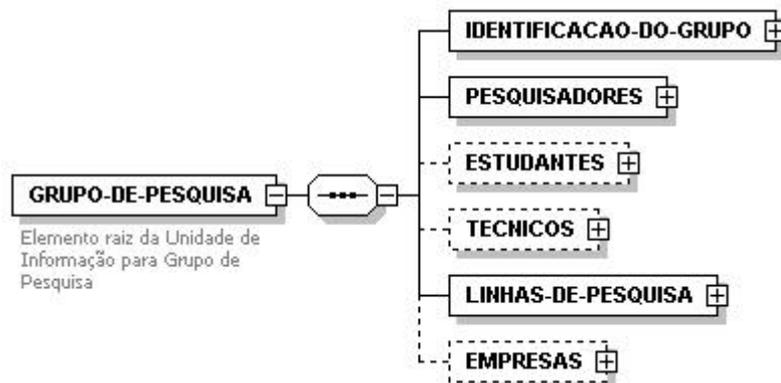


Figura 3.5 - Diagrama esquemático da representação de informações sobre grupo de pesquisa segundo a LMPL (Fonte: CNPq, 2003)

c) Projeto de Pesquisa

A Figura 3.6 mostra de forma esquemática a estrutura da informação sobre a unidade de análise projeto de pesquisa. Segundo o esquema, as informações presentes em um projeto de pesquisa são categorizadas em equipe do projeto, financiadores do projeto, produções C&T do projeto e orientações.

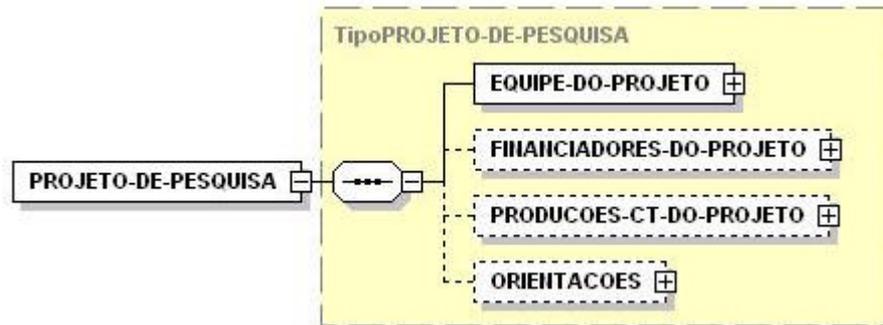


Figura 3.6 - Diagrama esquemático da representação de informações sobre projeto de pesquisa segundo a LMPL (Fonte: CNPq, 2003)

Essas visões esquemáticas são uma abstração dos documentos efetivamente usados no intercâmbio e compartilhamento de informações: DTDs e arquivos de conteúdo XML. A especificação contida na LMPL favorece a escrita de componentes de software reusáveis para a geração e interpretação dos arquivos XML que contêm informações sobre ciência e tecnologia. A partir da LMPL, não apenas as agências podem acessar a informação de seu interesse alimentada por pesquisadores e gestores, mas também a comunidade de instituições de ensino e pesquisa, parceiros na construção da LMPL (PACHECO; KERN, 2001).

3.3.2.2 Sistemas e Fontes de Informação

A segunda camada é composta do conjunto de instrumentos de gestão de informação (particularmente sistemas de informação). Esses instrumentos incluem desde sistemas disponíveis internacionalmente até ferramentas nacionais criadas pelos seus atores, compatíveis com as normas XML da primeira camada. Seu principal objetivo consiste em viabilizar a formação de bases nacionais de informação de CT&I e, ao mesmo tempo, permitir a gestão individual e coletiva dessa informação entre os atores encarregados das atividades de CT&I.

Para capturar as informações em cada unidade de análise da camada anterior, têm-se à disposição instrumentos, sistemas e ferramentas utilizados para tornar efetivo o fluxo de informação, desde a reunião e atualização até a publicação de dados nacionais. Entre eles, vale citar os sistemas de envio, recebimento e

certificação de arquivos, os sistemas de cargas nas bases de dados, os sistemas de visualização de informação e os sistemas extratores de informação.

Esses sistemas incluem as regras de controle e qualidade específicas a cada unidade de análise (identificação unívoca, segurança, campos obrigatórios e relações esperadas entre campos), sendo projetados de forma a atender tanto às necessidades de informações do governo como a permitir a inclusão contínua e modular de novos recursos para os usuários.

3.3.2.3 Portais e Serviços Web

A terceira camada é formada pelo conjunto de portais e serviços Web cuja base conceitual de operação é a informação gerada pela Plataforma Lattes. Compreende desde serviços orientados à divulgação de indicadores nacionais de CT&I até instrumentos de gestão de comunidades virtuais. As atividades de concepção, desenvolvimento e implantação dos elementos dessa camada constituem a responsabilidade dos mentores dos projetos.

3.3.2.4 Sistemas de Conhecimento

A quarta e última camada da estrutura arquitetônica da Plataforma Lattes é constituída de instrumentos de gestão e extração de conhecimentos de CT&I. É possível indicar dois campos de abordagens para esses componentes: (a) algoritmos e sistemas desenvolvidos no âmbito de projetos; e (b) estudos propostos pela comunidade científica que revelam novos conhecimentos sobre a atividade científica e tecnológica no País a partir de fontes de informação ou que propõem novos modelos para as áreas de gestão de CT&I, testando-os através das bases de dados.

Dentro desses campos de abordagem, para a realização de estudos sobre os atores do processo científico, é necessário que se utilizem técnicas que empreguem algoritmos e métodos capazes de extrair conhecimento de dados, os quais evidenciam o comportamento e a relação entre as diversas unidades de análise em CT&I.

Os instrumentos projetados nesta camada deverão utilizar métodos de descoberta de conhecimento, como, por exemplo, Diagramas de Venn, Link Analysis e Grafos, que serão apresentados a seguir, como também podem usar técnicas de recuperação da informação, estatística, etc. Além disso, os desenvolvimentos

realizados nesta camada guardam relação com todas as outras camadas da arquitetura.

a) Diagramas de Venn

Em 1880, o lógico inglês (e posteriormente o historiador) John Venn publicou um artigo com o título *Sobre representação diagramática e mecânica de proposições e raciocínios*, no qual trabalha a recém-criada área de Álgebra de Boole, associando-a com a nova visão da Teoria de Conjuntos desenvolvida por George Cantor.

Venn (1880) propôs a idéia de representar as relações entre conjuntos através de configurações de figuras no plano, o objetivo dele estava claramente formulado no artigo “[...] antes de mais nada os diagramas servem para auxiliar o olho e a mente graças a natureza intuitiva do seu testemunho [...]” e foi plenamente alcançado, já que 120 anos mais tarde os livros elementares de Matemática usam esse caminho para introduzir alunos em Teoria de Conjuntos.

Um Diagrama de Venn é um caminho para graficamente representar conjuntos e operações de conjuntos. Cada diagrama começa com a representação retangular do conjunto universo, em seguida cada conjunto é representado por um círculo (PENNA; ARROYO, 2003). A seguir serão apresentados os Diagramas de Venn com um, dois, três e quatro conjuntos.

? Diagramas de Venn com um conjunto

Os diagramas de Venn podem ser representados também com um conjunto. A fórmula apresentada por Venn para determinar o número de áreas é 2^n . De acordo com essa fórmula, são apenas duas áreas para um conjunto: uma área interna e uma área externa (Figura 3.7 a seguir).

Para um conjunto: $2^n = 2^1 = 2$ áreas (interna e externa).

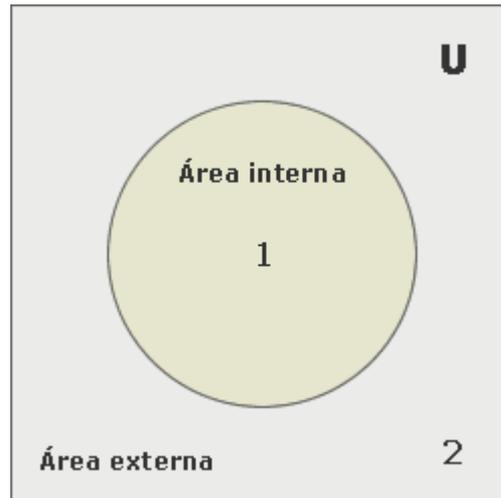


Figura 3.7 - Diagrama de Venn usando um conjunto

? Diagramas de Venn com dois conjuntos

Usando a fórmula 2^n , têm-se quatro áreas, representadas na Figura 3.8 a seguir.

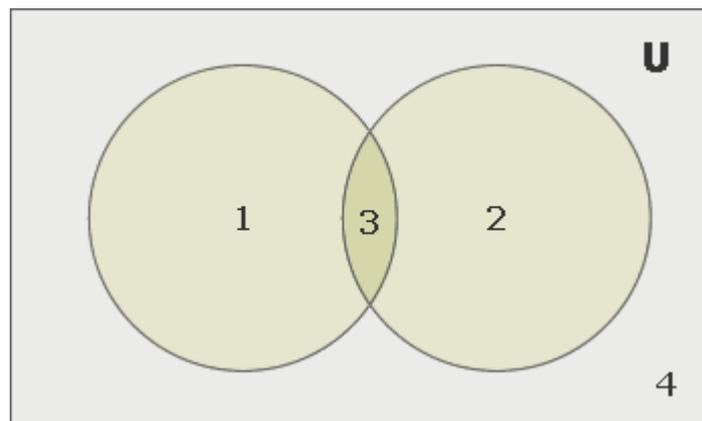


Figura 3.8 - Diagrama de Venn usando dois conjuntos

? Diagramas de Venn com três conjuntos

Para três conjuntos, são oito áreas, que são representadas na Figura 3.9 na seqüência. Pode-se notar que a complexidade para se desenharem os círculos e o número de áreas é exponencial.

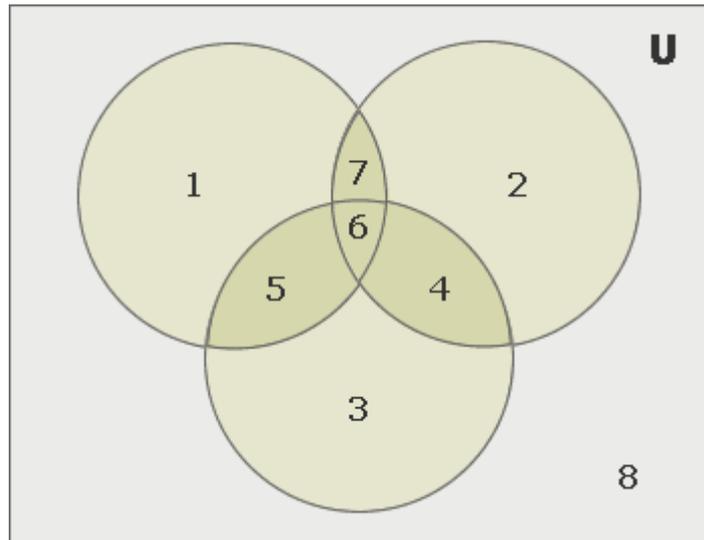


Figura 3.9 - Diagrama de Venn usando três conjuntos

? Diagramas de Venn com quatro conjuntos

Para diagramas de Venn usando um, dois ou três conjuntos, as elipses e os círculos têm a mesma representação, porque eles não têm grande complexidade, mas ao se tentar desenhar um diagrama de Venn usando quatro conjuntos, as elipses podem representar melhor o diagrama do que os círculos (Figuras 3.10 e 3.11 na seqüência).

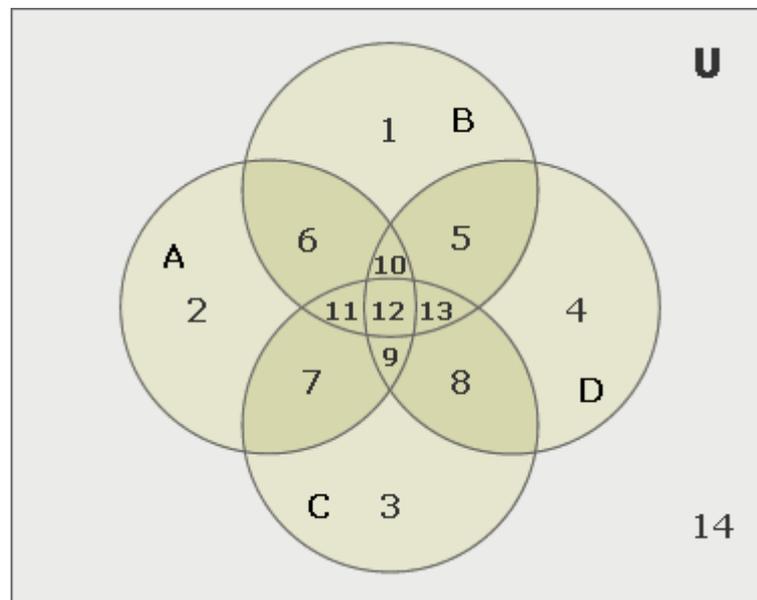


Figura 3.10 - Diagrama de Venn usando quatro conjuntos em formato de círculo

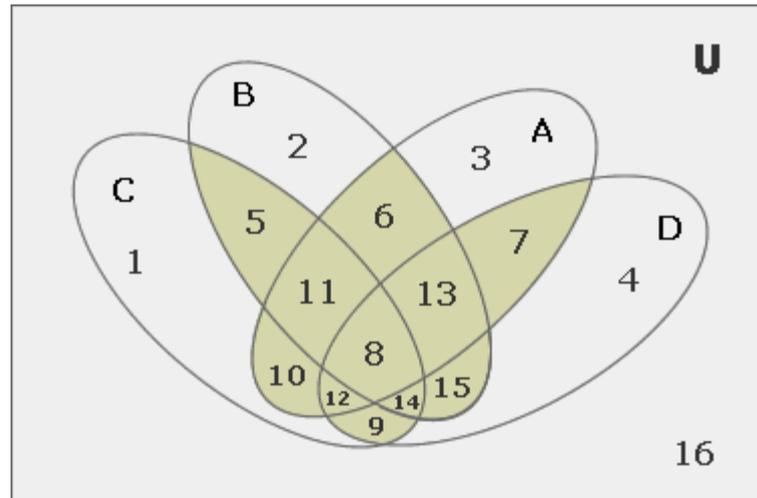


Figura 3.11 - Diagrama de Venn usando quatro conjuntos em formato de elipse

Na teoria, as relações entre os conjuntos nos diagramas de Venn e o máximo número de conjuntos que pode ser incluído em um diagrama podem ser infinitas (PENNA; ARROYO, 2003).

A diferença entre a Figura 3.10 e a 3.11 é a de que esta última em formato de elipse serve como um recurso para representações em três dimensões, não alcançadas com a Figura 3.10.

Um exemplo de aplicação dos diagramas de Venn pode ser verificado nas análises de co-participação de pesquisadores em mais de um grupo de pesquisa, apresentando os grupos de pesquisa como conjuntos, os pesquisadores como elementos do conjunto e as intersecções como co-participações.

A seguir, será apresentada outra ferramenta da tecnologia da informação capaz de auxiliar na manipulação de dados para a extração de conhecimentos.

b) Link Analysis

Em várias situações é possível identificar relacionamentos entre pessoas, lugares, objetos ou mesmo conceitos, como o existente entre pessoas que conversam por meio de ligações telefônicas, ou entre pesquisadores quando estes são co-autores em publicações, ou ainda entre instituições de ensino quando um egresso de uma instituição é um docente de outra. Relacionamentos são também observados nas coleções de documentos que são interconectados por referências e/ou hyperlinks (JENSEN, 1998). Para estudar tais relacionamentos, surgiu a técnica de Link Analysis.

Trata-se de uma técnica baseada em um ramo da Matemática, denominado Teoria dos Grafos, que tem por finalidade revelar a estrutura e o conteúdo de um conjunto de informações por meio de unidades (entidades ou objetos) interconectadas entre si, conforme mostra a Figura 3.12. Nessa figura, são mostrados os relacionamentos entre pessoas que participam de atividades de CT&I. Podem ser observadas uma relação de orientação indicando quem orientou quem e uma relação de quais pessoas publicaram ou trabalharam juntas.

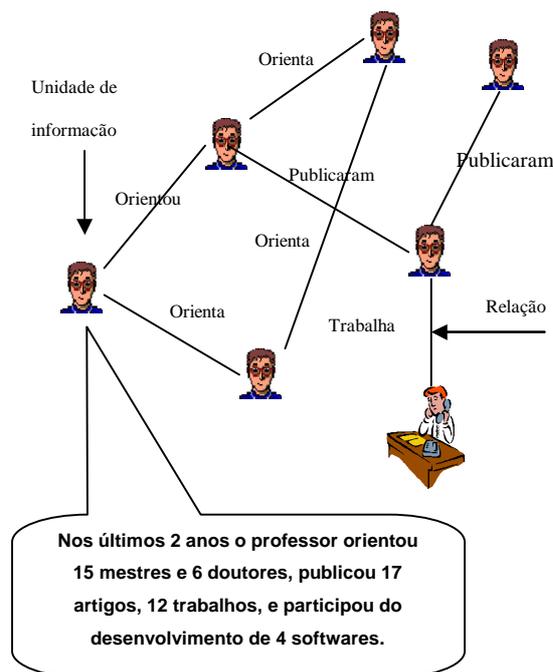


Figura 3.12 - Exemplo de aplicação de Link Analysis em CT&I

Sendo as unidades de informação e suas relações existentes exibidas na forma de um grafo, como mostra a Figura 3.12, pretende-se, ao analisá-lo, compreender o significado das interconexões e descobrir relações que não eram percebidas ou cuja explicação é mais bem apresentada por meio de nós interconectados. Para viabilizar esse procedimento, torna-se necessária a construção de redes em um processo chamado de formação de links. As técnicas para formação, análise, busca e apresentação das redes são geralmente referidas como Link Analysis (GOLDBERG; SENATOR, 1998). Segundo Berry e Linoff (1997), a ferramenta Link Analysis é útil em análises investigativas, mas não é aplicável a todo tipo de dado ou problema. Essa ferramenta tem produzido bons resultados nas áreas de: a) análise de padrões de chamadas telefônicas, sendo cada chamada de telefone um relacionamento entre

dois pontos; b) análise de padrão do receituário médico. Médicos podem receitar alguns medicamentos e não outros, o que cria um padrão que pode associá-los a determinado laboratório farmacêutico; c) investigações criminais. O FBI possui um sistema para auxiliar na combinação de informações de fontes distintas para solucionar crimes.

Além dessas áreas, a Link Analysis também tem sido aplicada em classificação de sites, análise de redes de telecomunicações e de rotas de transporte, epidemiologia e outras aplicações especializadas. Essa técnica explora associações entre grandes quantidades de objetos de diferentes tipos e é útil nas operações de mineração de dados para problemas que naturalmente envolvam links. O modo de visualização permite que o conhecimento implícito em determinada base seja mais facilmente observado. Segundo Lions e Tseytin (1998), algumas questões são tradicionalmente consideradas quando se trata de Link Analysis, tais como as apresentadas na Tabela 3.1.

1) Quais nós são chaves ou centrais na rede formada?
2) Quais links podem ser reforçados para aumentar a eficiência das operações da rede?
3) É possível descobrir links ou nós não detectados a partir dos dados conhecidos?
4) Existem similaridades na estrutura de subpartes da rede, as quais podem indicar um relacionamento não conhecido?
5) Quais são as sub-redes relevantes dentro de uma rede com muitos nós?
6) Quais modelos de dados e níveis de agregação revelam com mais facilidade certos tipos de relações e sub-redes?

Tabela 3.1 - Algumas questões sobre Link Analysis

Entretanto, a Link Analysis possui suas limitações, pode ser aplicada a poucos tipos de dados, é implementada por poucas ferramentas, as estruturas de dados em geral demandam um processamento elevado, uma vez que implementações utilizando somente seleções encadeadas em bancos relacionais podem ser ineficientes.

A seguir, será apresentada a ferramenta da tecnologia da informação capaz de auxiliar na visualização de dados.

c) Grafos

A primeira referência à Teoria dos Grafos foi feita por Leonard Euler (1707-1783) em 1735 sobre o problema das Pontes de Königsberg (Kaliningrad). “Será possível a um cidadão sair de sua casa, passar por cada uma das 7 (sete) pontes apenas uma só vez, e retornar a sua casa?”. O problema foi codificado como um grafo representando as áreas terrestres como vértices e as pontes como as ligações entre os vértices. Dessa forma, a Teoria dos Grafos nasceu com estudos para solucionar problemas desse tipo (HAYES, 2000).

Um grafo é representado matematicamente por $G = (V, E)$, sendo V o conjunto de vértices e E um conjunto de arestas ou relações entre os vértices (*Edges*, em inglês), ou seja, cada aresta é um par de vértices (ex.: (v, w)). Um grafo é representado graficamente através de bolinhas para os vértices e de retas ou curvas para as arestas. É uma noção simples, abstrata e intuitiva usada para representar a idéia de alguma espécie de relação entre os “objetos”.

Exemplo:

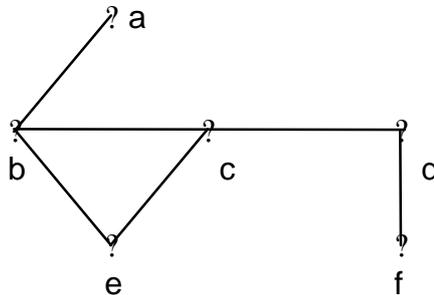


Figura 3.13 - Exemplo de um grafo

O grafo da Figura 3.13 possui $V = \{a, b, c, d, e, f\}$ e $E = \{(a, b), (b, c), (b, e), (c, e), (c, d), (d, f)\}$, sendo (a, b) uma aresta entre o vértice a e b . Normalmente, as arestas do tipo (a, a) são chamadas de laço e, dependendo do problema que o grafo representa, perdem o sentido e são ignoradas.

? Subgrafo

Um subgrafo $G' = (V', E')$ de um grafo $G = (V, E)$ é um grafo tal que $V' \subseteq V$ e $E' \subseteq E$.

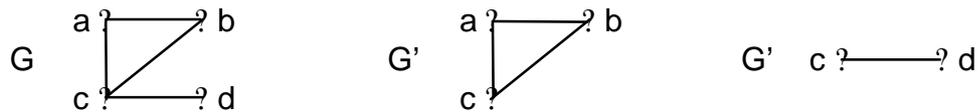


Figura 3.14 - Exemplo de subgrafos

? Grafo Orientado (ou Dígrafo)

Um grafo é dito orientado ou dígrafo quando é considerado o sentido das ligações entre os vértices. Neste caso, denomina-se de arco a aresta direcionada.

Um dígrafo é representado matematicamente também por $G = (V, E)$, sendo V o conjunto de vértices, e E , uma relação binária em V (i.e., um conjunto de pares ordenados das ligações).

Em um grafo orientado, a ordem entre os vértices de uma aresta (v, w) é importante. Essa aresta é diferente da aresta (w, v) e é representada com uma flecha de v para w .

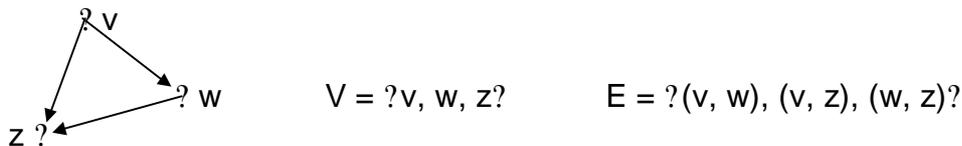


Figura 3.15 - Exemplo de Grafo Orientado

? Grafo Valorado

Uma rede é um grafo orientado (ou um dígrafo) no qual um número real é associado aos vértices e/ou às ligações. Esse número é freqüentemente referido como o peso da ligação. Essa classificação é dada de acordo com a necessidade, ou não, da indicação do fluxo entre os vértices. Na prática esse número pode representar custos, distâncias, capacidades e/ou suprimentos e demandas, tempo (trânsito, permanência), confiabilidade de transmissão, probabilidade de ocorrerem falhas, capacidade de carga e outros.

Uma rede é representada matematicamente por $G = (V, E, w)$, sendo V o conjunto de vértices, E o conjunto de ligações e w o peso associado aos vértices e/ou às ligações.

? Caminho, Percurso, Ciclo, Circuito e Comprimento

Um caminho de um vértice $vi0$ para o vértice vik é uma seqüência de arestas $\langle vi0, vi1 \rangle, \langle vi1, vi2 \rangle, \dots, \langle vi,k-1, vik \rangle$. Um caminho é dito elementar se passa exatamente uma vez por cada vértice, e é simples se passa exatamente uma vez por cada aresta. Quando o grafo é não orientado, o conceito de caminho é substituído por cadeia, que pode ser representada pela seqüência de arestas que a formam ou dos vértices nela contidos. Alguns autores usam o termo *percurso* para denominar genericamente um caminho.

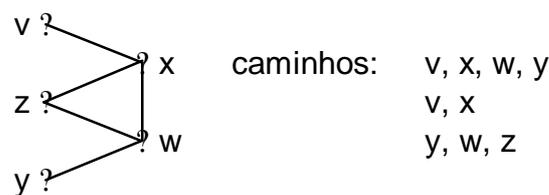


Figura 3.16 - Exemplo de um grafo com Caminhos

Se os vértices inicial e final são coincidentes ($vi0 = vik$), dizemos que o caminho é fechado e forma um ciclo que é chamado de circuito se o grafo for orientado. Um circuito é um caminho onde $v_1 = v_n$, como b, c, d, e, f, d, b .

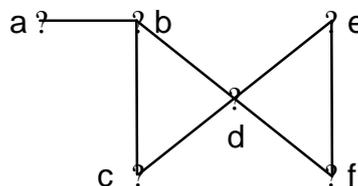


Figura 3.17 - Exemplo de um grafo com um ciclo

O comprimento de um percurso num grafo valorado é a soma dos custos para percorrer cada aresta, e num grafo não valorado o comprimento é igual ao número de arestas que o compõem.

? Ciclo euleriano e Circuito hamiltoniano

Um ciclo que passa por todas as arestas de um grafo é dito euleriano, e um circuito elementar que passa por todos os vértices é chamado de hamiltoniano. O famoso problema do caixeiro-viajante consiste em analisar todos circuitos hamiltonianos existentes para $(n+1)$ pontos.

? **Conexidade**

Um grafo $G = (V, E)$ é conexo se para todo par de vértices existir pelo menos uma cadeia (grafo não orientado) ou um caminho (grafo orientado) entre eles, por outro lado, se existir pelo menos um par de vértices que não é unido por nenhuma cadeia/caminho, diz-se que o grafo é não conexo ou desconexo.

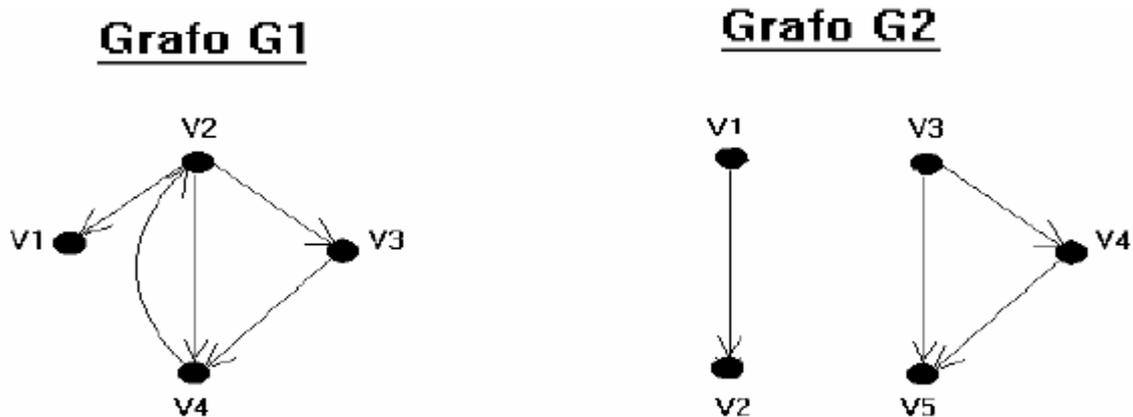


Figura 3.18 - Exemplo de Grafo Conexos (G1) e Grafo Desconexo (G2)

O conceito de conexidade em grafos orientados não exige um caminho ligando qualquer par de vértices, se isso acontecer, diz-se que o grafo é fortemente conexo, ou seja, dados dois vértices v e w quaisquer, cada um pode ser atingido a partir do outro, partindo-se de v pode-se chegar a w , e vice-versa. A Figura 3.19 mostra um exemplo de grafo orientado fortemente conexo.

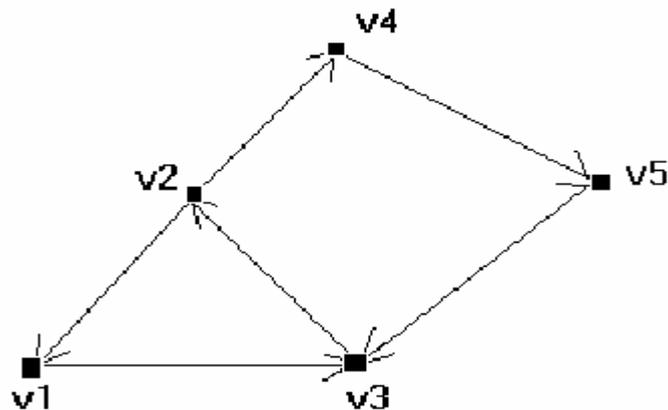


Figura 3.19 - Exemplo de Grafo Orientado fortemente conexo

? **Relação de Adjacência e de Incidência**

Dois vértices $v \in V$ e $w \in V$ de um grafo $G = (V, E)$ são ditos adjacentes se existe a aresta (v, w) , ou seja, $(v, w) \in E$. Duas arestas são ditas adjacentes se possuem uma extremidade (vértice) comum.

Uma aresta é incidente sobre um vértice se este vértice for uma de suas extremidades. Assim, as arestas (u,v) e (v,w) são incidentes sobre o vértice v .

? Grau de um Vértice

O grau de um vértice é o número de arestas adjacentes a ele. Em um grafo orientado, o grau de entrada de um vértice v é o número de arestas (w, v) , e o grau de saída é o número de arestas (v, w) . No exemplo a seguir, v possui grau de entrada 2 e grau de saída 1.

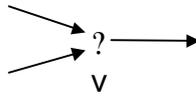


Figura 3.20 - Exemplo de grau de entrada 2 (dois) e grau de saída 1 (um)

? Grafo Completo

Grafo completo ou clique é um grafo, ou subgrafo, em que seus vértices são interligados ou adjacentes dois a dois, de forma que o caminho mais curto entre quaisquer dois vértices v e w seja a aresta (v, w) .

A Figura 3.21 mostra um grafo com dois cliques.

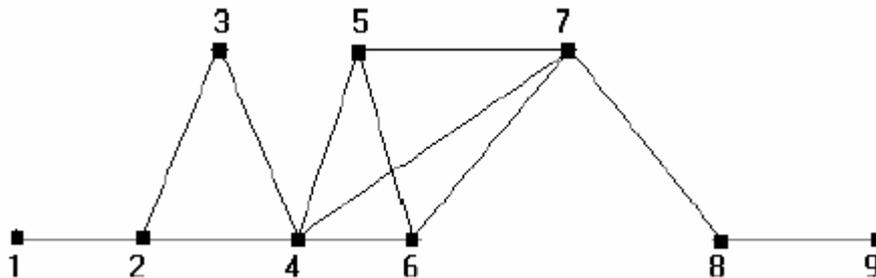


Figura 3.21 - Grafo com dois cliques: $\{2,3,4\}$ e $\{4,5,6,7\}$

? Grafos Planares

Um grafo é planar se os seus vértices dispostos num plano não apresentam cruzamento de arestas.

Pode-se fazer uma associação entre um grafo planar e um mapa em que cada região do mapa corresponde a um vértice, e as fronteiras entre as regiões

correspondem às arestas. Na seqüência pode ser visto um exemplo de um grafo planar.

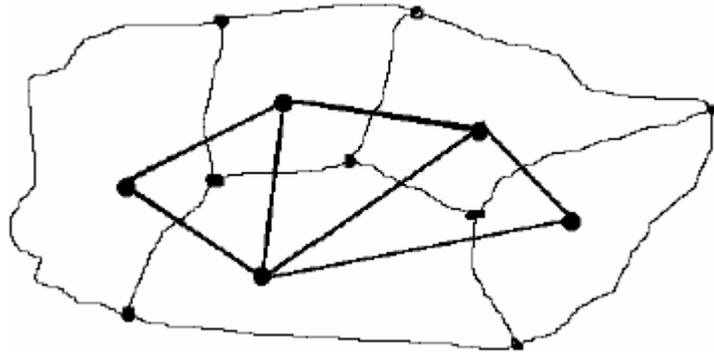


Figura 3.22 - Grafo Planar com 5 (cinco) regiões

3.4 Considerações Finais

Este capítulo abordou aspectos relativos a uma plataforma de gestão de CT&I e deu ênfase aos aspectos normalmente presentes nas análises de redes de pesquisa.

Entre os aspectos abordados estão as fontes de informação utilizadas e as formas de análise, destacando que todos aqueles estudos de redes de pesquisa, apresentados no capítulo anterior, têm, em sua base, relação com uma plataforma de gestão de CT&I, ou seja, incluem a escolha de uma fonte de informação e de uma possibilidade de análise sobre como redes se estruturam.

A ferramentas de Tecnologia da Informação foram apresentadas para identificar o melhor método de estudo para as análises de redes de pesquisa que serão apresentadas no Capítulo 4.

4 INSTRUMENTOS PARA ANÁLISE DE REDES DE PESQUISA EM CT&I NA PLATAFORMA LATTES DO CNPQ

Há três séculos, o conhecimento científico não faz mais do que provar suas virtudes de verificação e de descoberta em relação a todos os outros modos de conhecimento.

Edgar Morin

4.1 Introdução

O objetivo do capítulo é apresentar a viabilidade de aplicação da metodologia de análise de redes de pesquisa para identificar áreas de estudo de relações sociais ou redes de cooperação na área de CT&I quando se tem a disponibilidade de uma plataforma de informações.

Para obter conhecimento sobre CT&I são necessárias bases de dados sobre o parque científico e tecnológico do Brasil. A base adotada para extrair esse conhecimento é a Plataforma Lattes, que inclui o banco de currículos dos pesquisadores e o Diretório de Grupos de Pesquisa do Brasil.

Essas bases serão adotadas como estudo de caso para aplicação de instrumentos de análise de redes de pesquisa, mostrando tanto a viabilidade de replicação de estudos encontrados na literatura como a possibilidade de novas análises não encontradas na literatura de redes de pesquisa, mas comentadas como possíveis na área de redes de relacionamentos.

4.2 Unidades de Análise e Redes de Pesquisa

Nesta subseção será apresentada a relação entre as unidades de análise identificadas em uma plataforma de gestão de CT&I (Subseção 3.3.2.1) e os estudos na área de Redes de Pesquisa. Com isso, da relação de unidades de análise identificadas em plataformas de gestão de CT&I e das visões das redes de pesquisa, foi possível montar a Tabela 4.1.

Unidades de Análise	Redes de Pesquisa
Currículo	<p>Redes de Co-Autoria – Nestas redes a unidade de informação <i>currículo</i> é incluída por meio de seu módulo “colaboradores” (pesquisadores, estudantes, docentes, etc.). Este tipo de rede acontece quando um dado colaborador produz vários artigos em conjunto com outro(s) colaborador(es).</p> <p>Redes de Citação – Assim como as redes descritas acima, a unidade de informação <i>currículo</i> também é incluída por meio de seu módulo “colaboradores”. O estudo dessas redes é mais adequado para a análise do fluxo do conhecimento, ou seja, permite verificar o espectro de influência entre os autores. O meio mais comum de atribuir crédito e reconhecimento na ciência é a citação.</p> <p>Redes de Formação – A análise das informações relativas à formação acadêmica da unidade de análise <i>currículo</i> possibilita a construção dessas redes. Estas tratam as relações entre indivíduos que se formaram no mesmo período (ex.: colegas de curso de formação, colegas de área de formação, colegas de instituição de formação).</p> <p>Redes de Orientação – Este tipo de rede diz respeito às relações de orientação (entre orientadores, entre orientados e entre orientador–orientado). De acordo com Hagstrom (1965), ocorre relação de orientação entre professores e estudantes, mesmo quando não há uma relação formal.</p>
Grupos de Pesquisa	<p>Redes de Grupos de Pesquisa – A análise deste tipo de rede trabalha com a unidade de informação <i>grupo de pesquisa</i>, podendo tanto privilegiar as relações institucionais dos grupos (instituição do grupo, instituição e/ou programas de pós-graduação em que atuam seus pesquisadores ou instituição e/ou cursos em que estudam seus estudantes) como os relacionamentos dos indivíduos intra e intergrupos de pesquisa.</p>
Projetos de Pesquisa	<p>Redes de Equipes de Projetos de Pesquisa – Este tipo de rede não foi encontrado na literatura, contudo o que acontece são pesquisadores e/ou grupos de pesquisa que se juntam para executar determinado projeto de pesquisa, formando assim uma rede de equipes de projetos de pesquisa.</p>
Instituições	<p>Redes Institucionais – Nestas redes a unidade de informação <i>instituição</i> é incluída por meio das informações de organizações, institutos, empresas, universidades e demais organismos ligados à CT&I, referenciados nas demais unidades como local de lotação profissional ou de pesquisa ou representando atores institucionais que interagem com agências de fomento. Permitem medir o grau de relações entre indivíduos ou departamentos na mesma instituição ou entre instituições e/ou cursos diferentes a partir da atividade de pesquisa.</p>

Tabela 4.1 - Redes de Pesquisa sob o aspecto das unidades de análise

Na Tabela 4.2 são apresentadas algumas possibilidades de relacionamentos entre as unidades de informação de uma plataforma de gestão em CT&I. Dentro desses relacionamentos, pode-se identificar dois tipos: os homogêneos e os heterogêneos. Estes são entre unidades de informação diferentes, por exemplo,

currículo e instituição; aqueles são entre unidades iguais, por exemplo, grupo de pesquisa e grupo de pesquisa. Tanto nos relacionamentos homogêneos quanto nos heterogêneos, o que se busca identificar são “visões” de como as unidades de informação podem se relacionar.

	Currículo	Grupo de Pesquisa	Projeto de Pesquisa	Instituição
Currículo	De pessoas que produziram um artigo em co-autoria com outras pessoas; ou que citaram outras pessoas; ou que se formaram com outras pessoas; ou que orientaram outras pessoas.	De pessoas que são integrantes de grupos de pesquisa com determinada natureza (área, instituição, ano de formação, etc).	De pessoas que participam da equipe do projeto de pesquisa.	De pessoas que têm endereço, formação, atuação ou produção em uma instituição.
Grupo de Pesquisa	Com integrantes com formação, atuação ou produção em área, instituição e curso.	Com integrantes, área, instituição, curso e linhas de pesquisa comuns a outro grupo de pesquisa.	Com participação na execução de um determinado projeto de pesquisa.	Com lotação em uma dada instituição.
Projeto de Pesquisa	Com integrantes que participam no desenvolvimento.	Com execução por um determinado grupo de pesquisa.	Com equipe de projeto que se une a outra equipe de projeto pertencente a outro grupo para desenvolver determinado projeto de pesquisa.	Com vinculação em uma instituição.
Instituição	Com indivíduos lotados pela formação, participação e/ou produção.	Com grupos de pesquisa lotados.	Com projetos de pesquisa vinculados.	Com indivíduos, cursos, departamentos ligados a outra instituição por meio de produção, atuação, colaboração e/ou parcerias.

Tabela 4.2 - Relacionamentos entre as unidades de informação de uma plataforma de gestão de CT&I

4.3 Metodologia de Análise de Redes de Pesquisa em uma Plataforma de Gestão de CT&I

Como se pôde ver, a definição de redes de pesquisa encontra, na literatura, uma gama de possibilidades que se são ampliadas quando se dispõe de uma plataforma integrada de sistemas e unidades de informação em CT&I. O resultado dessa gama

de possibilidades de análise é o fato de não haver uma definição única para redes de pesquisa e, sobretudo, o fato de que dependerá do analista, de seu propósito de análise e da disponibilidade de informações, o processo de estudos de redes de pesquisa. Deste modo, o primeiro passo no processo de análise de redes de pesquisa consiste na definição do objeto da análise, ou seja, na configuração dos elementos que farão parte dos estudos da rede de pesquisa. Isso implica determinar a unidade (ou unidades) disponível e desejável na análise da rede, explicitar o conceito de rede de pesquisa pelo qual essa análise melhor se caracteriza (conceito que deve ser revisto quando dos estudos dos resultados da análise, por exemplo, no caso de indução ou prospecção de redes, deve-se enfatizar qual o tipo de rede que interessa) e, finalmente, a selecionar as instâncias de informação para essas unidades (ou seja, seleção de dados ou informações para cada unidade que será projetada no estudo da rede de pesquisa com base na plataforma disponível e no interesse da análise). Em resumo, a análise de redes de pesquisa com base em uma plataforma de gestão em CT&I deve contemplar as fases descritas a seguir.

FASE I – DEFINIÇÃO DO OBJETIVO DA ANÁLISE DE REDES DE PESQUISA

Dada a variedade de formas de se analisarem redes de pesquisa, é fundamental que o analista responsável explicithe seus objetivos antes de proceder à análise. Esses objetivos serão referenciais para a verificação do alcance dos resultados esperados com a análise realizada. Entre os objetivos pode-se incluir a indução de redes de pesquisa com determinado fim (ex.: tratar de problemas de saúde coletiva, com urgência de resultados necessários) ou a prospecção de redes já ativas, porém não mapeadas diretamente pelo analista (ex.: verificar o grau de interação em rede de especialistas ou de grupos de pesquisa em um determinado domínio de interesse estratégico para a gestão de CT&I).

FASE II - DETERMINAÇÃO DA(S) UNIDADE(S) DE ANÁLISE ALVO DO ESTUDO

Neste caso, com base na unidade ou em unidades disponíveis, o analista deve, *a priori*, selecionar aquela(s) que deverá(ão) fundamentar suas análises em redes de pesquisa. O objetivo é destacar nos estudos o que se pretende realizar, deixando claros o escopo e a definição de rede que se aplicam ao problema.

FASE III - DEFINIÇÃO DO CONCEITO DE “REDE DE PESQUISA”

Esta definição será a base para a análise dos resultados da aplicação dos instrumentos disponíveis. Para explicitar o papel da definição na análise da rede, é importante que o analista identifique os elementos-chave da definição, ou seja, as palavras que caracterizam unidades de informação ou relacionamentos disponíveis na plataforma de informações, e que serão alvo de aplicação para os instrumentos de análise de redes.

FASE IV - ESCOLHA DO SUBCONJUNTO DE INSTÂNCIAS DA(S) UNIDADE(S) DEFINIDA(S)

Trata-se da seleção dos dados e das informações que devem fazer parte da análise de redes de pesquisa. Para tal, é útil a aplicação de classificadores de informação disponíveis nas unidades de análise. Por exemplo, pode-se selecionar os grupos de pesquisa da região Sudeste que tenham pesquisadores com formação em Direito e que tenham projetos no setor de petróleo com a finalidade de induzir a formação de uma rede de pesquisa que estude a legislação em petróleo.

FASE V - APLICAÇÃO DE TÉCNICA OU FERRAMENTA DE ANÁLISE

Com base nas escolhas anteriores, o analista deve selecionar a técnica, ferramenta ou metodologia mais adequada à análise de redes de pesquisa que pretende realizar. Entre as metodologias mais utilizadas, estão Diagramas de Venn, Link Analysis e Teoria dos Grafos.

4.4 Possibilidades de Análise de Redes de Pesquisa na Plataforma Lattes

Este item visa apresentar as várias possibilidades de análise de relacionamentos na Plataforma Lattes a partir de cada unidade de análise Lattes.

Do ponto de vista de unidade de análise, os relacionamentos podem ser homogêneos e heterogêneos. O primeiro tipo corresponde ao envolvimento de

apenas uma unidade de informação na análise; já o segundo tipo corresponde ao envolvimento de mais de uma unidade de informação na análise.

Na Tabela 4.3 é possível visualizar os relacionamentos homogêneos e heterogêneos, as unidades de análise, a informação curricular envolvida (campos) e os exemplos de análise de redes de pesquisa na Plataforma Lattes. É importante salientar que os exemplos apresentados não são únicos, ou seja, é possível extrair muitos outros exemplos de redes de pesquisa da Plataforma. Esses exemplos servem para mostrar as diversas possibilidades de análise que podem ser realizadas quando se tem a disponibilidade de uma plataforma de informações.

Relacionamentos Homogêneos		
Unidade(s)	Campos	Exemplo
Currículo	Identificação (Endereço)	Pode-se verificar as redes formadas por conterrâneos (mesma UF ou mesma cidade), como também se são vizinhos (pela UF, cidade e CEP).
Currículo	Identificação Atuação Profissional	Pode-se verificar o grau de colaboração entre pesquisadores nascidos em um mesmo Estado e atuando em uma mesma área profissional.
Currículo	Formação Atuação Profissional Produção	Pode-se verificar o grau de co-autoria entre pesquisadores em que um foi orientando do outro comparando-se o grau de colaboração por área do conhecimento em que esses pesquisadores atuam.
Currículo	Identificação Produção	Pode-se verificar a colaboração por co-autorias com base nos dicionários dos autores e identificação do currículo.
Currículo	Identificação Formação Área predominante	Pode-se verificar as relações de colegas de turmas (mesmo curso na mesma instituição no mesmo período). Aqui é possível obter relações por colegas de área de formação, colegas de instituição de formação e colegas de curso de formação (i.e., colegas de pós-graduação).
Currículo	Identificação Atuação Profissional Área predominante	Pode-se verificar as relações de colegas de trabalho (mesmo endereço profissional no mesmo período). Aqui é possível obter relações por colegas de área de atuação.
Currículo	Identificação Dados Complementares	Pode-se verificar as relações de co-participantes em banca de conclusão de curso (mesmo candidato, mesma instituição e mesmo curso).
Grupo de Pesquisa	Identificação Pesquisadores Estudantes	Pode-se verificar o grau de colaboração entre os grupos de pesquisa, cujo relacionamento se baseia em pesquisadores/estudantes comuns aos grupos de pesquisa, ou seja, quando um mesmo pesquisador/estudante participa de mais de um grupo de pesquisa ou quando é pesquisador em um grupo e estudante em outro. Outra análise possível pode ser realizada nas relações institucionais dos grupos (instituição do grupo, instituição em que atuam seus pesquisadores ou em que estudam seus alunos).

Grupo de Pesquisa	Identificação Área de Conhecimento	Pode-se verificar o grau de co-autoria entre grupos de pesquisa comparando-se o grau de colaboração por área do conhecimento a que pertencem esses grupos de pesquisa.
Relacionamentos Heterogêneos		
Unidade(s)	Campos	Exemplo
Currículo Grupo de Pesquisa	Identificação (Currículo) Área predominante (Grupo de Pesquisa) Pesquisadores (Grupo de Pesquisa)	Pode-se verificar o índice de participação de pesquisadores nascidos em outros Estados em cada grupo, segundo área predominante e UF do grupo. Isso pode permitir verificar, por exemplo, que em determinada área os grupos têm maioria de pesquisadores nascidos no próprio Estado e não atuam entre si (às vezes formando <i>clusters</i> conceituais nas áreas de trabalho, como no caso da Filosofia do Brasil, dos pesquisadores cariocas).
Currículo Grupo de Pesquisa	Identificação (Currículo) Produção (Currículo) Pesquisadores (Grupo de Pesquisa)	Pode-se verificar o relacionamento entre pesquisadores pertencentes a um mesmo grupo. Através das produções, os trabalhos realizados em co-autorias poderiam ser a ligação entre esses pesquisadores ou a relação de orientação, em que um pesquisador de um grupo pode orientar um estudante de um outro grupo. Nessas análises é possível verificar qual é o grau de colaboração entre os pesquisadores do mesmo grupo.
Currículo Projeto de Pesquisa	Identificação (Currículo) Atuação Profissional (Currículo) Identificação (Projeto de Pesquisa)	Pode-se verificar as colaborações interinstitucionais de um projeto de pesquisa (participantes de mais de uma instituição). Outra forma de colaboração seria as co-participações dos integrantes da equipe de um projeto de pesquisa (um integrante participa de mais de um projeto de pesquisa).

Tabela 4.3 - Possibilidades de análise de redes de pesquisa na Plataforma Lattes

4.4.1 Instrumentos de análise de redes de pesquisa na Plataforma Lattes

Nesta subseção serão apresentados três instrumentos de análise de redes de pesquisa na Plataforma Lattes. Os conjuntos de dados utilizados para essas análises foram extraídos das bases de dados do CV-Lattes e do Diretório de Grupos de Pesquisa no Brasil.

O primeiro instrumento apresentado a seguir aplica a metodologia proposta na Subseção 4.3, já o segundo instrumento se baseia em dados estatísticos retirados do banco de dados através de consultas específicas mostrando a formação de redes de cooperação entre os grupos de pesquisa por grande área do conhecimento. Por último, é apresentado um software desenvolvido para visualizar redes de pesquisa segundo diversos critérios.

4.4.1.1 Aplicação do Método Proposto: Redes de Co-Participação em Grupos de Pesquisa por Instituição

Este exemplo trata as redes institucionais formadas pelas colaborações entre membros participantes dos grupos de pesquisa vinculados a uma determinada instituição.

FASE I: O objetivo da análise de redes de pesquisa, neste exemplo, é verificar o grau de colaboração entre os grupos de pesquisa.

FASE II: A unidade de análise alvo é o grupo de pesquisa.

FASE III: O conceito de rede de pesquisa na análise realizada é o seguinte:

Redes de pesquisa do tipo “rede de grupos” é um conjunto de pesquisadores com trabalhos em grupos de pesquisa que utilizam o intercâmbio de informações, conhecimento e de recursos humanos de seus grupos para executar novas pesquisas.

Dessa definição é possível identificar alguns conceitos-chave que podem ser classificados como elementos de informação e elementos de relacionamento. Dentro dos elementos de informação, têm-se os conjuntos de pesquisadores (currículos) e grupos de pesquisa. Já os elementos de relacionamento são o intercâmbio de informações, o intercâmbio de conhecimento e o intercâmbio de recursos humanos.

Conceitos-chave		
Elementos de Informação	Pesquisadores (Currículos)	Representado pela unidade de informação <i>Currículo</i> .
	Grupos de Pesquisa	Agrupamento de um conjunto de pesquisadores que se unem para desenvolver projetos comuns.
Elementos de Relacionamento	Intercâmbio de informações	Exemplo desse tipo de relacionamento são as trocas de mensagens eletrônicas (e-mails).
	Intercâmbio de conhecimento	O meio mais comum de exemplificar a troca de conhecimento é a citação.
	Intercâmbio de recursos humanos	Co-participações de pessoas nos grupos de pesquisa.

Tabela 4.4 - Conceitos-chave do exemplo de Redes Institucionais de Grupos de Pesquisa

FASE IV: Para efeito de exemplificação, será apresentada neste passo uma aplicação realizada para extrair conhecimento das Redes de Pesquisa formadas na base de dados do Diretório de Grupos de Pesquisa no Brasil.

A seleção dos dados e informações dos grupos de pesquisa (subconjunto de instâncias) ocorreu, inicialmente, através da escolha de uma amostra de grupos de pesquisa. O critério para a escolha dessa amostra foi selecionar as 50 instituições de ensino superior do Brasil com maior número de grupos de pesquisa, por representarem a grande maioria dos dados presentes na base.

O objetivo desse exemplo foi identificar algumas redes de pesquisa entre instituições, formadas a partir da participação de integrantes dos grupos de pesquisa. Os vértices desse grafo representam as instituições dos grupos de pesquisa, e as arestas representam a existência de membros em comum entre as instituições. A representação dessa rede de pesquisa escolhida foi a dos grafos.

Após a geração do grafo, aplicou-se um algoritmo da Teoria dos Grafos (Apêndice A), que tem como objetivo identificar os conjuntos independentes maximais. De uma forma sucinta, em um grafo $G = (X, E)$, um conjunto de vértices independentes (também conhecido como conjunto internamente estável) é um conjunto de vértices tal que não há dois vértices adjacentes dentro desse conjunto. Se o grafo G tiver o conjunto $X = \{x_0, x_1, x_2, x_3, x_4, x_5\}$ e $E = \{\{x_0, x_1\}, \{x_0, x_3\}, \{x_1, x_2\}, \{x_1, x_4\}, \{x_2, x_3\}, \{x_2, x_5\}, \{x_2, x_4\}, \{x_4, x_5\}\}$, então, os conjuntos $\{x_0, x_2\}$, $\{x_1, x_3\}$ e $\{x_1, x_3, x_5\}$ são exemplos de conjuntos independentes.

Um conjunto independente é chamado de maximal quando não houver outro conjunto independente que o contenha. Por exemplo, o conjunto $\{x_1, x_3\}$ é independente, mas não maximal. Já os conjuntos $\{x_0, x_2\}$ e $\{x_1, x_3, x_5\}$ são independentes maximais, porque não há outro conjunto independente que os contenha.

Com a aplicação do algoritmo (Apêndice A), encontraram-se vários conjuntos independentes maximais. Selecionaram-se os menores conjuntos entre os encontrados, que são os núcleos da rede analisada. Assim, tem-se um “centro de comunicação” da rede, que é um conjunto de instituições com grupos de pesquisa que permite a “comunicação” com todas as outras instituições (grupos de pesquisa) da rede. Entre os menores conjuntos, restaram apenas dois núcleos com somente um vértice, ou seja, com uma instituição cada. Isso significa que, na rede formada por instituições que possuem grupos de pesquisa em qualquer área, duas instituições estão ligadas diretamente a todas as outras. No exemplo, essas duas

instituições são a USP (Universidade de São Paulo) e a UNESP (Universidade Estadual Paulista Júlio de Mesquita Filho).

FASE V: Neste exemplo, as técnicas de análise escolhidas foram a Teoria dos Grafos e a de Link Analysis, pelo fato de haver um certo interesse em se estudarem alguns tipos de subconjuntos de vértices, definidos por determinadas características de suas relações de adjacência. Uma aplicação comum desse conceito é o cálculo de todos os conjuntos independentes maximais de um grafo.

Para mais detalhes desse exemplo, pode-se consultar o Apêndice A.

4.4.1.2 Dados Estatísticos: Redes de Co-Participação em Grupos de Pesquisa por Grande Área do Conhecimento

Este exemplo mostra as redes formadas pelas colaborações entre membros participantes dos grupos de pesquisa pertencentes a uma grande área do conhecimento.

Os conjuntos de dados estatísticos que serão apresentados têm como objetivo básico identificar como são os relacionamentos de co-participação dos membros do grupo por grande área do conhecimento. Para isso, realizou-se um cálculo para gerar um índice de co-participação, o qual leva em consideração o total de co-participações de cada grande área, dividido pelo total de grupos certificados da mesma grande área. Assim, para se ter o maior índice é preciso que o número de co-participações seja proporcional ao número de grupos de pesquisa, ou seja, para grandes áreas com o mesmo número de grupos de pesquisa, terá índice mais alto aquela com o maior número de co-participação, ou para grandes áreas que tenham o mesmo número de co-participação, terá melhor índice aquela com o menor número de grupos de pesquisa.

De acordo com a Tabela 4.5, pode-se verificar que a grande área com o maior índice de co-participação são as Ciências Agrárias, no entanto, nota-se que, em termos de número de grupos de pesquisa, outras grandes áreas possuem números maiores, contudo o total de co-participação não é proporcionalmente maior. Já a grande área que obteve o menor índice de co-participação foi a Lingüística, Letras e Artes, isso se deve ao fato do total de co-participações ser quase o mesmo do total de grupos de pesquisa.

Grande Área	Total de Co-Participação	Total de Grupos	Índice
Ciências Agrárias	3240	1653	1,96
Ciências Biológicas	3865	2126	1,81
Ciências Exatas e da Terra	3123	2051	1,52
Ciências Humanas	3430	2399	1,42
Ciências Sociais Aplicadas	1810	1429	1,26
Ciências da Saúde	3550	2513	1,41
Engenharias	3458	2243	1,54
Lingüística, Letras e Artes	882	744	1,18

Tabela 4.5 - Índice de co-participações por Grande Área de Conhecimento

Na Tabela 4.6 são apresentados os totais de co-participação, separadamente, por grande área do conhecimento. É possível analisar os relacionamentos de co-participação entre todas as grandes áreas, inclusive os relacionamentos entre os membros da mesma grande área.

Uma falsa idéia que se pode ter é de que o maior número de co-participação ocorre entre os membros da mesma grande área do conhecimento, no entanto, isso não é verdade. Entre as oito grandes áreas do conhecimento, quatro delas (Ciências Agrárias, Ciências Biológicas, Ciências Exatas e da Terra e Ciências Sociais Aplicadas) colaboram menos consigo do que com as outras. Já as quatro grandes áreas restantes (Ciências da Saúde, Ciências Humanas, Engenharias e Lingüística, Letras e Artes) colaboram mais consigo do que com as outras. Os dados que comprovam essas afirmações estão disponíveis na Tabela 4.7.

Grande Área	Total de Co-Participação	Grande Área
Ciências Agrárias	1494	Ciências Agrárias
Ciências Agrárias	716	Ciências Biológicas
Ciências Agrárias	184	Ciências da Saúde
Ciências Agrárias	348	Ciências Exatas e da Terra
Ciências Agrárias	74	Ciências Humanas
Ciências Agrárias	133	Ciências Sociais Aplicadas
Ciências Agrárias	291	Engenharias
Ciências Biológicas	617	Ciências Agrárias
Ciências Biológicas	1728	Ciências Biológicas
Ciências Biológicas	657	Ciências da Saúde
Ciências Biológicas	459	Ciências Exatas e da Terra
Ciências Biológicas	144	Ciências Humanas
Ciências Biológicas	34	Ciências Sociais Aplicadas
Ciências Biológicas	221	Engenharias
Ciências Biológicas	5	Lingüística, Letras e Artes
Ciências da Saúde	150	Ciências Agrárias
Ciências da Saúde	650	Ciências Biológicas

Ciências da Saúde	1904	Ciências da Saúde
Ciências da Saúde	207	Ciências Exatas e da Terra
Ciências da Saúde	336	Ciências Humanas
Ciências da Saúde	95	Ciências Sociais Aplicadas
Ciências da Saúde	170	Engenharias
Ciências da Saúde	38	Linguística, Letras e Artes
Ciências Exatas e da Terra	256	Ciências Agrárias
Ciências Exatas e da Terra	393	Ciências Biológicas
Ciências Exatas e da Terra	193	Ciências da Saúde
Ciências Exatas e da Terra	1486	Ciências Exatas e da Terra
Ciências Exatas e da Terra	199	Ciências Humanas
Ciências Exatas e da Terra	70	Ciências Sociais Aplicadas
Ciências Exatas e da Terra	513	Engenharias
Ciências Exatas e da Terra	13	Linguística, Letras e Artes
Ciências Humanas	71	Ciências Agrárias
Ciências Humanas	117	Ciências Biológicas
Ciências Humanas	323	Ciências da Saúde
Ciências Humanas	179	Ciências Exatas e da Terra
Ciências Humanas	1885	Ciências Humanas
Ciências Humanas	462	Ciências Sociais Aplicadas
Ciências Humanas	143	Engenharias
Ciências Humanas	250	Linguística, Letras e Artes
Ciências Sociais Aplicadas	80	Ciências Agrárias
Ciências Sociais Aplicadas	24	Ciências Biológicas
Ciências Sociais Aplicadas	92	Ciências da Saúde
Ciências Sociais Aplicadas	71	Ciências Exatas e da Terra
Ciências Sociais Aplicadas	396	Ciências Humanas
Ciências Sociais Aplicadas	889	Ciências Sociais Aplicadas
Ciências Sociais Aplicadas	188	Engenharias
Ciências Sociais Aplicadas	70	Linguística, Letras e Artes
Engenharias	228	Ciências Agrárias
Engenharias	188	Ciências Biológicas
Engenharias	190	Ciências da Saúde
Engenharias	555	Ciências Exatas e da Terra
Engenharias	158	Ciências Humanas
Engenharias	224	Ciências Sociais Aplicadas
Engenharias	1863	Engenharias
Engenharias	52	Linguística, Letras e Artes
Linguística, Letras e Artes	4	Ciências Biológicas
Linguística, Letras e Artes	24	Ciências da Saúde
Linguística, Letras e Artes	9	Ciências Exatas e da Terra
Linguística, Letras e Artes	197	Ciências Humanas
Linguística, Letras e Artes	91	Ciências Sociais Aplicadas
Linguística, Letras e Artes	36	Engenharias
Linguística, Letras e Artes	521	Linguística, Letras e Artes

Tabela 4.6 - Total de co-participações entre as Grandes Áreas

Grande Área	Mesma Grande Área	Outra Grande Área
Ciências Agrárias	46,11%	53,89%
Ciências Biológicas	44,70%	55,30%
Ciências da Saúde	53,63%	46,37%

Ciências Exatas e da Terra	47,58%	52,43%
Ciências Humanas	54,95%	45,05%
Ciências Sociais Aplicadas	49,11%	50,89%
Engenharias	53,87%	46,13%
Linguística, Letras e Artes	59,07%	40,93%

Tabela 4.7 - Percentual de co-participações entre as Grandes Áreas

4.4.1.3 Sistema Lattes Redes

a) Objetivo

O objetivo do sistema é investigar, induzir, avaliar e analisar as redes de pesquisa sob a ótica do sistema CT&I, através dos dados extraídos da Plataforma Lattes, segundo diversas possibilidades de visualização. Com isso, os atores do sistema CT&I são levados da condição de fornecedores de dados a beneficiários da Tecnologia da Informação.

b) Funcionamento do Sistema

Dado um tema de interesse ou aplicando os filtros existentes, o sistema faz uma busca nos currículos da base de dados do CV-Lattes. São os parâmetros determinados inicialmente que delimitarão o conjunto de pessoas às quais serão aplicadas as análises sob o aspecto de redes de pesquisa.

Depois de realizada a busca, o sistema apresenta um conjunto de pessoas com as respectivas quantidades de produções de cada pesquisador. Feito isso, basta escolher qual tipo de rede pretende-se analisar. As possibilidades de análise de redes de pesquisa que o sistema apresenta são: co-autoria, orientação, projetos de pesquisa, colegas de formação, colegas de trabalho, co-participante em banca, colegas de grupos de pesquisa, colegas de pós-graduação, conterrâneos (UF, Cidade) e vizinhos (UF, Cidade e CEP). Após escolher o tipo de rede que se quer analisar, o sistema apresenta a rede de relacionamentos formada entre todos os pesquisadores pertencentes ao conjunto inicial.

c) Benefícios

Acredita-se que este instrumento trará diversos benefícios à atividade de gestão em CT&I. Entre esses, pode-se destacar a criação de meios para que analistas de informação em CT&I possam estudar as formas de organização da pesquisa brasileira, segundo redes de atuação.

O sistema, também, poderá servir de base para atividades de inspeção de oportunidades, indução e formação de redes, análise de resultados de investimentos, sob a ótica de redes de trabalho, e investigação de possibilidades de incentivo ao intercâmbio acadêmico e científico.

d) As interfaces do Sistema

A Figura 4.1 apresenta a página inicial do site Lattes Redes. Esta interface tem como finalidade a entrada do conjunto de parâmetros para filtrar os currículos existentes na base CV-Lattes. A título de exemplo foi realizada uma consulta buscando todos os pesquisadores que têm como área de atuação a Engenharia de Produção.

Depois de configurados os parâmetros do filtro e executada a busca, o sistema apresenta a lista de pessoas encontradas na base CV-Lattes, ou seja, que possuem currículos Lattes e que atuam na área de conhecimento Engenharia de Produção. Para visualizar o currículo do pesquisador, basta clicar no nome dele. O número na frente do nome do pesquisador corresponde à quantidade de produções, conforme pode ser visto na Figura 4.2.

Redes Lattes - Microsoft Internet Explorer provided by Grupo Stela

Arquivo Editar Exibir Favoritos Ferramentas Ajuda

Links

Lattes Redes GrupoStela

Consulta Currículos Lattes Lattes Redes

Para recuperar assuntos iniciados por um termo radical, coloque * após o radical

Nome do Pesquisador opções

Pontos em que já publicou ou produziu opções

Aplicar filtros

Formação Acadêmica opções

Instituição opções

Área de Atuação opções

Atuação Profissional Natureza da atividade opções

Instituição opções

Exibir atividades atuais

Idiomas opções

© 2004 Grupo Stela - UFSC. Todos os direitos reservados.

Concluído Meu computador

Figura 4.1 - Interface Principal do Sistema Lattes Redes



Figura 4.2 - Resultado da busca por área de atuação

O tipo de rede escolhido para servir de exemplo foi a rede de co-autoria. Esse tipo de rede mostra os relacionamentos entre os pesquisadores através da co-autoria em suas produções. Na Figura 4.3 é possível visualizar a rede formada, que possui 36 vértices (pesquisadores), várias arestas (relacionamentos de co-autoria) e a legenda com os nomes dos pesquisadores.

De acordo com abordagem do fenômeno dos mundos pequenos de Watts e Strogatz (1998), que considera um conjunto com n elementos em que cada elemento se relaciona com k outros elementos, a rede de co-autoria da Figura 4.3 assemelha-se com o “mundo estilhaçado” da Figura 4.4. É apresentado um grafo totalmente aleatório, em que as k ligações de cada vértice distribuem-se de forma totalmente randômica por entre os outros $(n - 1)$ vértices possíveis.

As grandezas básicas que caracterizam o mundo estilhaçado mostram que o comprimento característico é pequeno e que o coeficiente de agregação é mínimo, aproximando-se ao acaso. Isso que dizer que se pode chegar a qualquer pesquisador rapidamente, embora a pessoa tenha provavelmente que percorrer os pesquisadores várias vezes no processo.

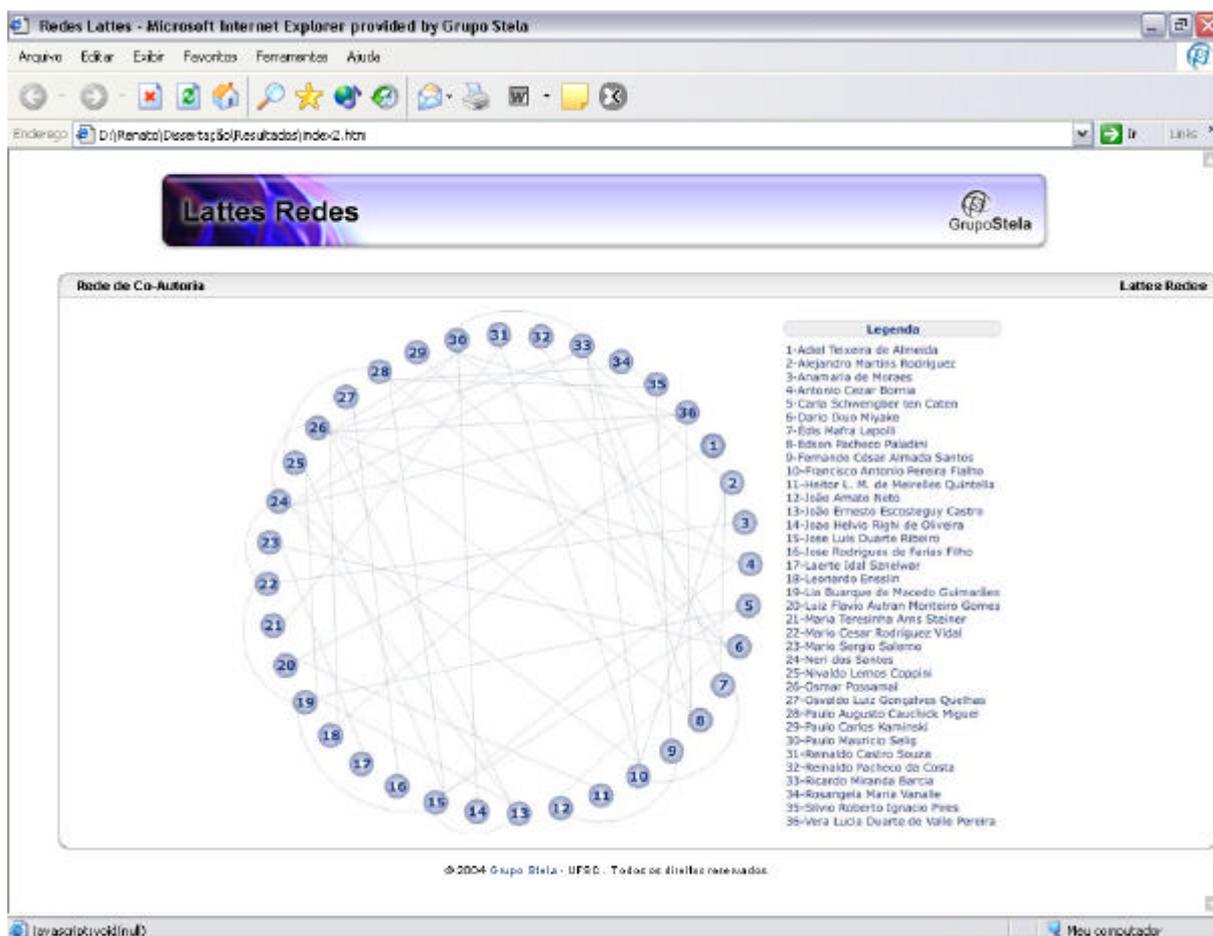


Figura 4.3 - Visualização de uma Rede de Co-Autoria com todos os relacionamentos

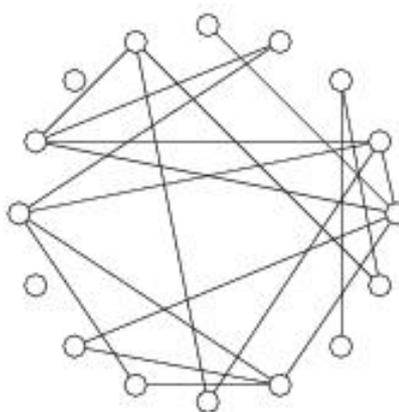


Figura 4.4 - Grafo Aleatório: “Mundo Estilhaçado” (WATTS; STROGATZ, 1998)

Na Figura 4.5 é apresentada a mesma rede de co-autoria anterior, no entanto mostrando apenas as arestas que saem dos vértices de grau maior ou igual a sete, ou seja, apenas os pesquisadores que tiverem sete ou mais co-autorias com outros pesquisadores serão apresentados. No exemplo, apenas três pesquisadores foram selecionados: Neri dos Santos, Osmar Possamai e Ricardo Miranda Barcia. Esses

pesquisadores são os conectores (“hubs”) identificados por Barabási (2003) no fenômeno do Mundo Pequeno. Os conectores são os responsáveis por um número pequeno de atalhos entre um pesquisador e outro numa teia de relações.

Outras redes de co-autoria como a rede da Figura 4.5 foram geradas, as quais variaram de vértices de grau maior ou igual a cinco até vértices de grau maior ou igual a oito. A visualização dessas redes se encontra no Apêndice B.

Na Figura 4.6 a rede de co-autoria apresentada exhibe vértices maiores ou iguais a oito. Nessa rede somente um pesquisador, Osmar Possamai, tem oito produções em co-autorias com outros membros da lista. Partindo desse pesquisador, pode-se em média percorrer três arestas para se chegar a qualquer outro pesquisador da rede. Assim, Osmar Possamai é o melhor “hub” dessa rede de co-autoria, pois é possível navegar por toda a rede partindo dele pelo menor caminho. O conhecimento extraído dessa análise é muito importante, visto que numa rede em que se têm tantos relacionamentos, pode-se saber qual o menor caminho que vai de um ponto a outro da rede.

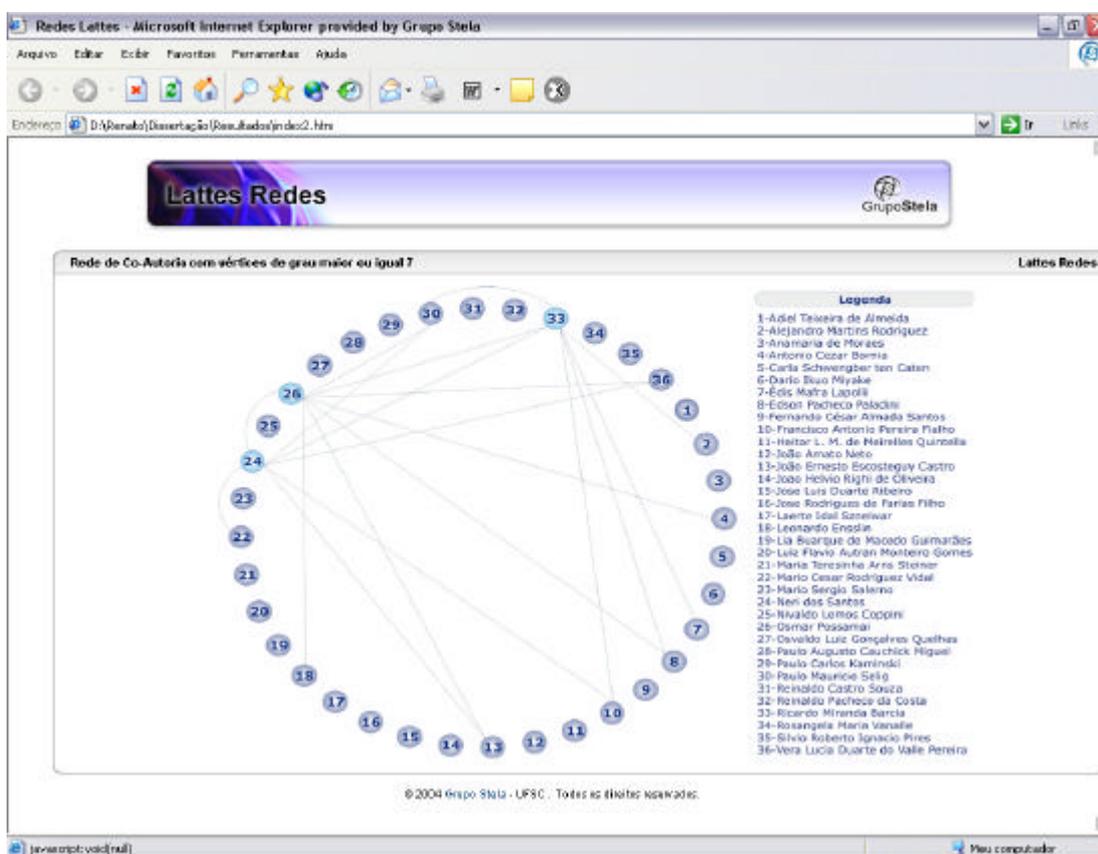


Figura 4.5 - Visualização de uma Rede de Co-Autoria com vértices de grau 7

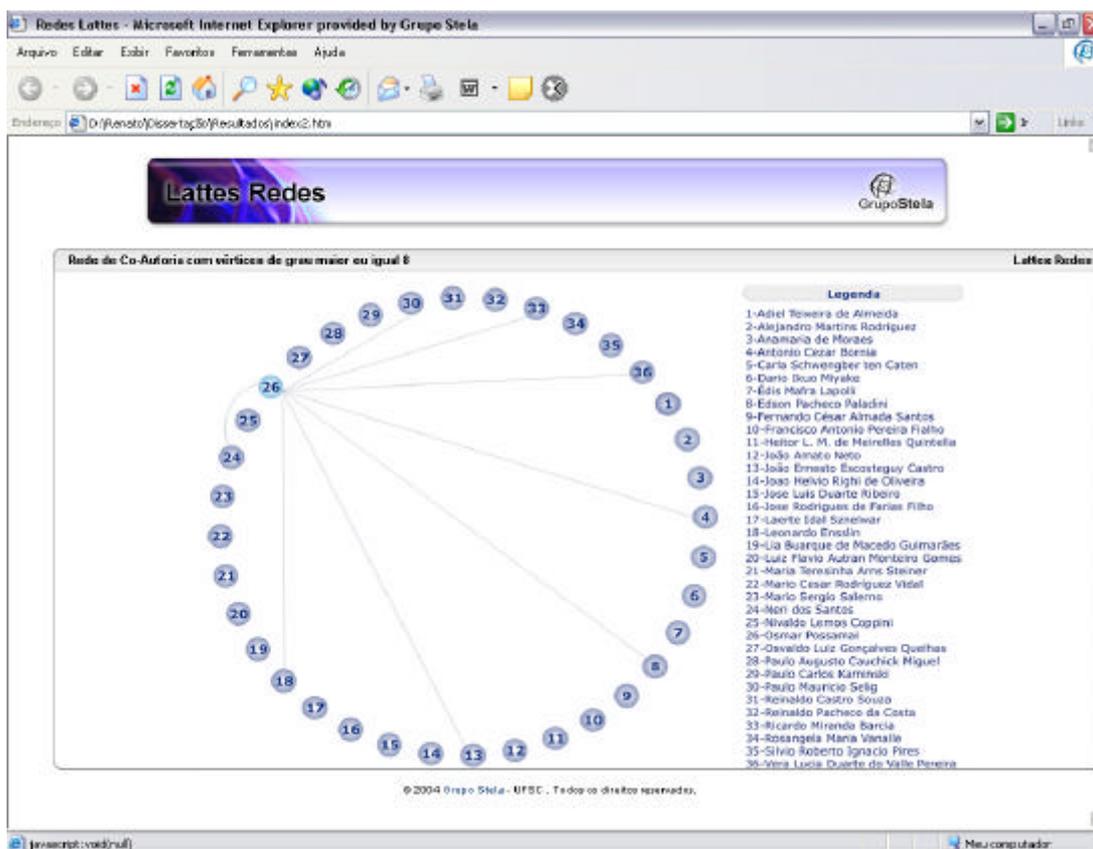


Figura 4.6 - Visualização de uma Rede de Co-Autoria com vértices de grau 8

4.5 Conclusão

Neste capítulo pôde-se ver a importância da análise de redes, principalmente quando aplicadas em uma base de informações de CT&I. O grande desenvolvimento recente desse campo se deve ao fato de que a Internet tem-se mostrado um excelente campo de teste e desenvolvimento para os seus conceitos.

As redes de relacionamento nos permitem elucidar as relações em grupos, redes de distribuição, redes de canais de comunicação, fluxos de informação e diversas outras áreas. É necessário, contudo, um plano estratégico de aplicação para analisar da forma correta o que se deve analisar, qual a disponibilidade de fontes de informação e o conhecimento do fenômeno estudado, para adotar a técnica mais adequada.

5 CONCLUSÕES E RECOMENDAÇÕES

Não basta a leitura sem unção, não basta a especulação sem a devoção,
não basta a pesquisa sem maravilhar-se; não basta a circunspecção sem o júbilo,
o trabalho sem a piedade, a ciência sem a caridade,
a inteligência sem a humildade, o estudo sem a graça.
São Boaventura

Por trás da maior parte dos sistemas complexos há uma rede intrincada. A vida está codificada numa complexa rede de moléculas ocultas na célula. A Internet é uma rede complexa de computadores ligados por fios. A economia é uma rede complexa de empresas, consumidores e agências reguladoras. A sociedade é uma rede complexa de pessoas ligadas por amizade, laços familiares e profissionais.

Só nos últimos anos constatou-se quão importante é o papel dessas redes em moldar o comportamento da maior parte dos sistemas complexos. Aprendeu-se que a compreensão das redes é o pré-requisito crucial para o entendimento da complexidade.

Portanto, muitos cientistas de várias disciplinas diferentes deflagraram uma agressiva abordagem que visa entender as teias de relacionamentos com as quais a natureza nos cerca. A rede social não é diferente daquela rede química de quatro bilhões de anos de idade que está em nossas células ou da Web, que só conta com uma década.

5.1 Conclusões

Constatou-se que os métodos disponíveis na literatura para análise de redes de pesquisa podem ser aplicados em uma plataforma de gestão de CT&I a partir das unidades de informação disponíveis. Portanto, a principal contribuição deste trabalho diz respeito à sistematização de uma abordagem de estudo de redes de colaboração, com ferramenta, sobre uma base extensa de informação de CT&I.

Assim, por exemplo, se o ONCYT de um determinado país tiver uma base de dados em CT&I, será possível aplicar a metodologia utilizada nesta dissertação. Isso se deve ao fato de esta base de dados responder à gestão de CT&I (mais

especificamente, como as informações na base respeitam o conceito de unidades de análise) e disponibilizar os dados da base em termos estruturais e semânticos (ex.: se estes dados estiverem em XML classificáveis em unidades de informação conhecidas em CT&I, pode-se aplicar um instrumento de análise, ainda que não haja uma plataforma de gestão naquele ONCYT).

A disponibilidade de uma plataforma de gestão de CT&I amplia as possibilidades de análises de redes, além de os instrumentos construídos com base em uma plataforma de informações ajudarem no processo de entendimento sobre como a atividade científica e tecnológica se organiza na forma de coletividade.

A aplicação dos instrumentos de análise de redes de pesquisa realizada neste trabalho auxiliou nas conclusões sobre a importância das metodologias e técnicas da área de redes de colaboração, bem como na extração de conhecimento das redes de relacionamentos geradas a partir de bases de informações em CT&I.

As redes de relacionamento nos permitem elucidar as relações em grupos, redes de distribuição, redes de canais de comunicação, fluxos informação e diversas outras áreas. Contudo, é necessário possuir um plano estratégico de aplicação para análise, ter disponibilidade de fontes de informação e conhecimento do fenômeno estudado para então adotar a técnica mais adequada.

Essencialmente, as aplicações que dependem de redes interligadas podem se beneficiar da investigação sobre o fenômeno dos pequenos mundos. Esse novo conhecimento sobre redes pode ser usado para catalisar (aumentar a velocidade) da solução de problemas, alguns dos exemplos são as redes de transmissão elétrica, de telecomunicações, a Internet, os modelos de propagação de doenças infecciosas ou de vírus de computador.

A escolha do ponto de corte (quantidade de hubs) é função do plano estratégico de análise. A interpretação dos resultados depende da experiência e dos objetivos do analista. Na Figura 1d (Apêndice B), por exemplo, os três hubs identificados pertencem ao mesmo programa de pós-graduação, mostrando que através desses pesquisadores é possível “caminhar” pela rede com o menor número de passos, partindo de um único programa de pós-graduação do Brasil.

Por fim, a revisão crítica da literatura na área de redes de colaboração com a construção de sumários e conclusões proporciona aos futuros pesquisadores da área um panorama que não existia na literatura.

5.2 Trabalhos futuros

É possível enxergar algumas perspectivas na correspondência entre *supply chain* (cadeia de suprimento) e redes de pesquisa, como a associação de diferentes bases de dados e estudos a partir das redes de colaborações, unidas com vetores de contexto para analisar as cadeias de troca de conhecimento.

Para pesquisas futuras, dado que a estrutura dos relacionamentos afeta as decisões, poder-se-ia estudar como as diferentes redes de co-autoria influenciariam a capacidade produtiva e, sobretudo, inovativa dos pesquisadores, e se as redes maiores implicariam em mais facilidade de interação com o mercado produtivo. Acredita-se que essas respostas mudariam com a área de conhecimento. Para isso, seria necessário um trabalho que investigasse se a natureza das redes de colaboração mudaria e de que forma de dariam essas mudanças entre as áreas científicas. Dentro desse aspecto, seria necessário calcular as medidas de redes sociais para todas as redes possíveis e fazer análises por área do conhecimento para responder a essa questão.

O fato de um indivíduo estar na rede de co-autoria de um outro determinado indivíduo, estudar o quanto isso exerce influência nas suas decisões futuras, como, por exemplo, se as redes de colaboração influenciam o tipo de atividade do pesquisador, se quem tem mais projetos tem maiores redes, e vice-versa, se quem tem maiores redes é mais citado ou se quem tem maiores redes é mais produtivo. Além disso, poderia-se verificar correspondências entre as redes de mesma dimensão para pesquisadores de produtividade diferente.

APÊNDICE A - APLICAÇÃO PARA EXTRAÇÃO DE CONHECIMENTO EM REDES DE PESQUISA DA BASE DE DADOS DO DIRETÓRIO DOS GRUPOS DE PESQUISA DO BRASIL

O objetivo desse trabalho foi identificar as redes de pesquisa entre instituições, formadas a partir da co-participação de seus membros, e representar essa rede através de um grafo. Os vértices desse grafo representam as instituições dos grupos de pesquisa, e as arestas representam existência de membros em comum entre as instituições. Após a geração desse grafo, aplicou-se o algoritmo, descrito abaixo, para encontrar todos os conjuntos independentes maximais. Depois disso, selecionou-se os menores conjuntos, dentre os encontrados, que são os núcleos da rede analisada. Assim, tem-se um “centro de comunicação” da rede, que é um conjunto de instituições que permite a “comunicação” com todas as outras instituições da rede.

1.1 Algoritmo para cálculo de conjuntos independentes maximais de um grafo (BOAVENTURA NETTO, 1996; BERGE, 1985)

Inicialização:

Passo 1: Faça $S_0 = Q_0^- = \emptyset$, $Q_0^+ = X$, $k = 0$.

Passo 2: Escolha um vértice $x_{ik} \in Q_k^+$ e forme S_{k+1} , Q_{k+1}^- e Q_{k+1}^+ , como descrito a seguir, mantendo Q_k^- e Q_k^+ intactos.

$$S_{k+1} = S_k \cup \{ x_{ik} \}$$

$$Q_{k+1}^- = Q_k^- - G(x_{ik})$$

$$Q_{k+1}^+ = Q_k^+ - G(x_{ik}) - \{ x_{ik} \}$$

Faça $k = k + 1$.

Passo 3: Se a condição abaixo for verdadeira, vá para o passo 5, senão vá para o passo 4.

$$\text{Existe algum } x \in Q_k^- \text{ tal que } G(x) \cap Q_k^+ = \emptyset$$

Passo 4. Se $Q_k^+ = Q_k^- = \emptyset$, mostre o conjunto independente maximal S_k e vá para o passo 5. Se $Q_k^+ = \emptyset$ mas $Q_k^- \neq \emptyset$ vá para o passo 5. Senão vá para o passo 2.

Passo 5. Faça $k = k - 1$. Remova x_{ik} de S_{k+1} para produzir S_k . Recupere Q_k^- e Q_k^+ , remova x_{ik} de Q_k^+ e o adicione em Q_k^- . Se $k = 0$ e $Q_k^+ = \emptyset$, pare. Senão vá para o passo 3.

1.2 A Base de Dados do Diretório dos Grupos de Pesquisa do Brasil

Na base de dados do Diretório dos Grupos de Pesquisa têm-se os dados sobre os grupos de pesquisa, tais como as instituições as quais eles pertencem, suas áreas de atuação, e seus membros, entre outros. Assim, utilizou-se desses dados para poder identificar as redes de pesquisa institucionais segundo co-participações dos seus membros. Para isso, selecionou-se as 50 instituições com maior número de grupos de pesquisa, por estas representarem a grande maioria dos dados presentes na base.

Cada grupo de pesquisa possui somente uma área de pesquisa, a qual pertence a uma grande área. Assim, identificaram-se três redes: a primeira foi formada analisando as instituições com grupos de pesquisa de todas as áreas, a segunda utilizando as instituições que possuíam grupos com pesquisa na grande área “Engenharias”; e, a última, com instituições que tinham grupos que realizam pesquisa na grande área “Ciências Agrárias”. A seguir são apresentados os resultados obtidos com aplicação do algoritmo em cada rede.

1.3 Todas as grandes áreas

Na primeira rede formada, de instituições com grupos de todas as grandes áreas, obteve-se um grafo com 50 vértices, sendo que todos eles possuíam pelo menos uma aresta com algum outro vértice. Com a aplicação do algoritmo, encontraram-se vários conjuntos independentes maximais. Dentre eles selecionaram-se os menores para se ter os núcleos da rede. Assim, conseguiu-se encontrar dois núcleos com somente um vértice, ou seja, uma instituição cada. Isso significa que cada uma dessas duas instituições está ligada a todas as outras 49. Essas duas instituições são: USP (Universidade de São Paulo) e UNESP (Universidade Estadual Paulista Júlio de Mesquita Filho).

1.4 Grande área “Engenharias”

A segunda rede foi formada com instituições que possuem grupos na grande área “Engenharias”. Dessa rede, retirou-se seis instituições das cinquenta selecionadas previamente, pois não estavam ligadas à pelo menos uma outra instituição. Portanto, os seis vértices que representavam essas instituições foram retirados do grafo. O grafo resultante ficou com 44 vértices, representando uma rede de co-participação em pesquisa, da grande área “Engenharias”, com 44 instituições.

Com a aplicação do algoritmo encontrou-se também vários conjuntos independentes maximais. Dentre eles selecionou-se os menores para se ter os núcleos da rede. Dessa forma, conseguiu-se encontrar 10 núcleos com 12 vértices cada, ou seja, 12 instituições. Isso significa que cada um desses conjuntos de 12 instituições permite acesso às demais 32 instituições. Um desses conjuntos é mostrado a seguir: UFMA (Universidade Federal do Maranhão), UFMS (Universidade Federal de Mato Grosso do Sul), UFRRJ (Universidade Federal Rural do Rio de Janeiro), UFAM (Universidade Federal do Amazonas), UFPEL (Universidade Federal de Pelotas), ULBRA (Universidade Luterana do Brasil), PUC-SP (Pontifícia Universidade Católica de São Paulo), PUC-RJ (Pontifícia Universidade Católica do Rio de Janeiro), UFSCAR (Universidade Federal de São Carlos), FIOCRUZ (Fundação Oswaldo Cruz), UFBA (Universidade Federal da Bahia) e (UFSC – Universidade Federal de Santa Catarina).

1.5 Grande área “Ciências Agrárias”

Na terceira rede, utilizou-se somente as instituições com grupos na grande área “Ciências Agrárias”. Nessa rede encontrou-se 9 instituições, das 50 selecionadas previamente, que não estavam ligadas a pelo menos uma outra instituição. Portanto, os 9 vértices que representavam essas instituições foram retirados do grafo. O grafo final ficou com 41 vértices, representando uma rede de co-participação em pesquisa na grande área “Ciências Agrárias” com 41 instituições.

Com a aplicação do algoritmo encontrou-se vários conjuntos independentes maximais. Dentre eles selecionou-se os menores para se ter os núcleos da rede. Assim, conseguiu-se encontrar um núcleo com 7 vértices cada, ou seja, 7 instituições. Isso significa que esse conjunto de instituições permite acesso às demais 34

instituições. Esse conjunto é mostrado a seguir: UNISA (Universidade de Santo Amaro), UNISINOS (Universidade do Vale do Rio dos Sinos), EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária), UFRN (Universidade Federal do Rio Grande do Norte), ULBRA (Universidade Luterana do Brasil), EFES (Universidade Federal do Espírito Santo) e UNIPAR (Universidade Paranaense).

APÊNDICE B – LATTES REDES

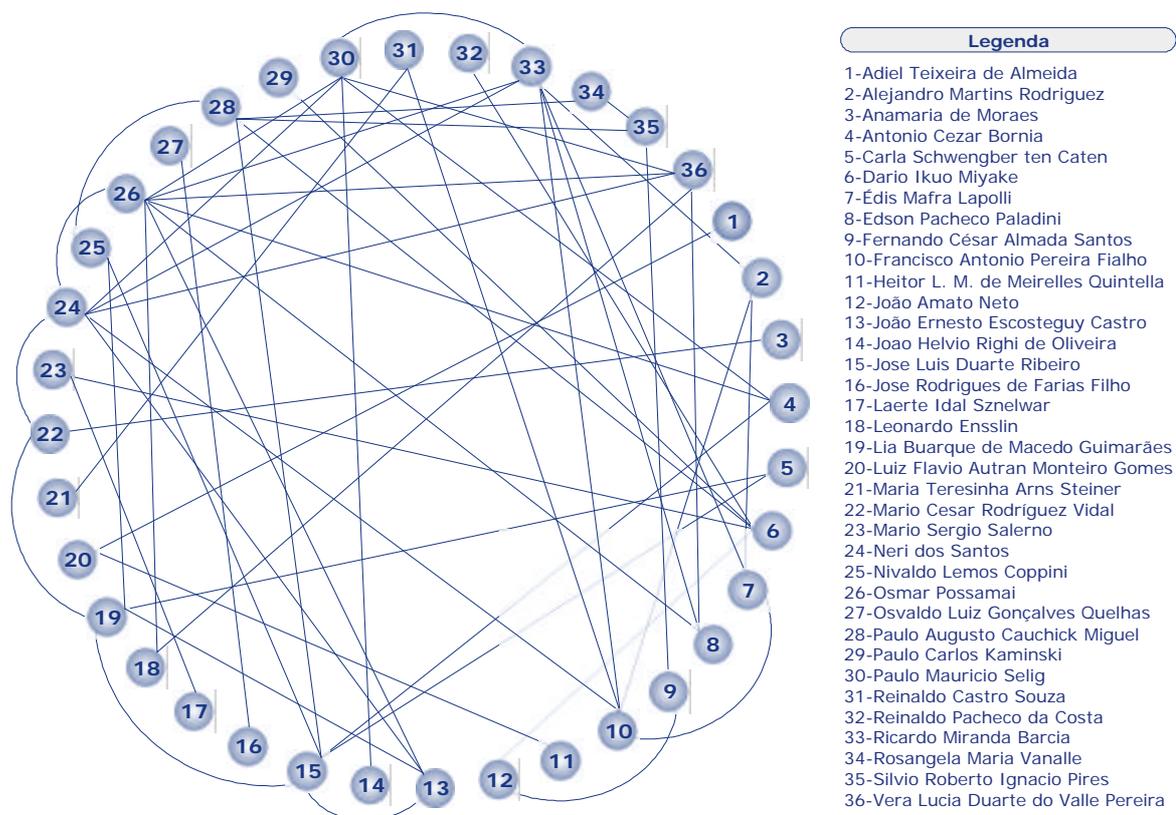


Figura 1a - Visualização de uma Rede de Co-Autoria

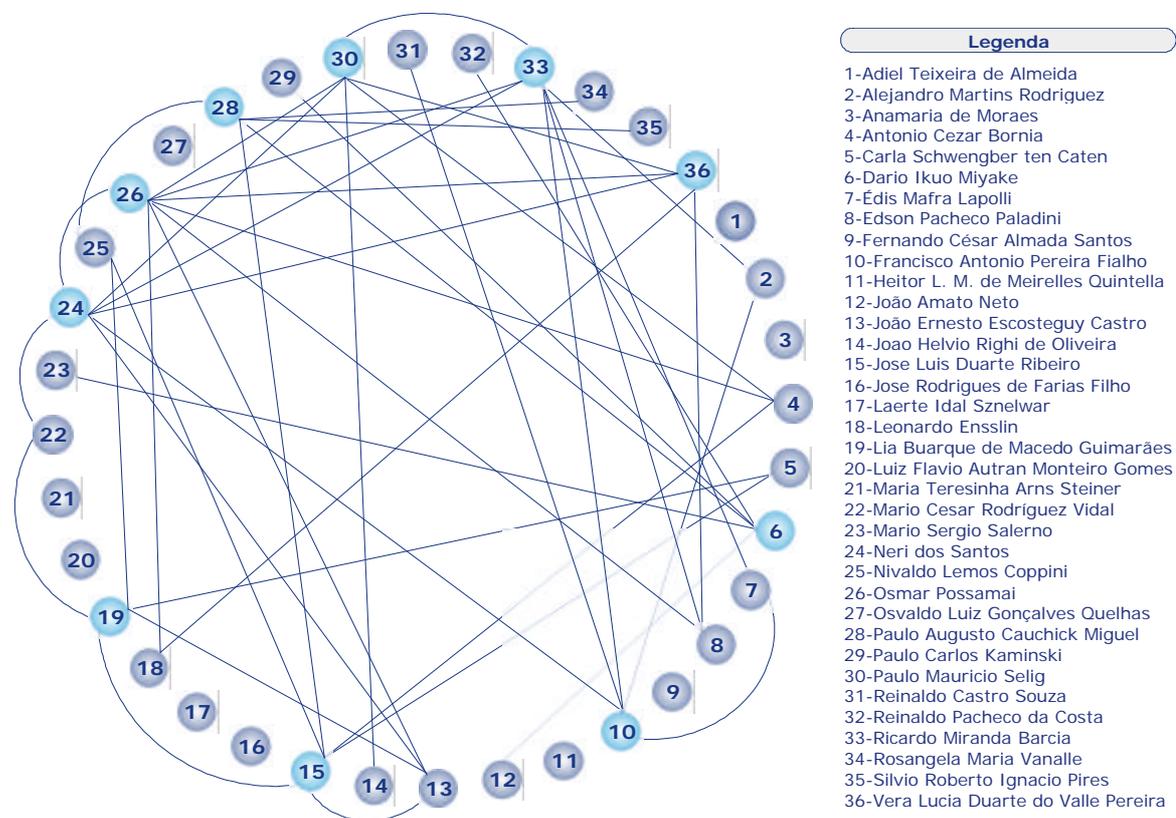


Figura 1b - Visualização de uma Rede de Co-Autoria com vértices de grau 5

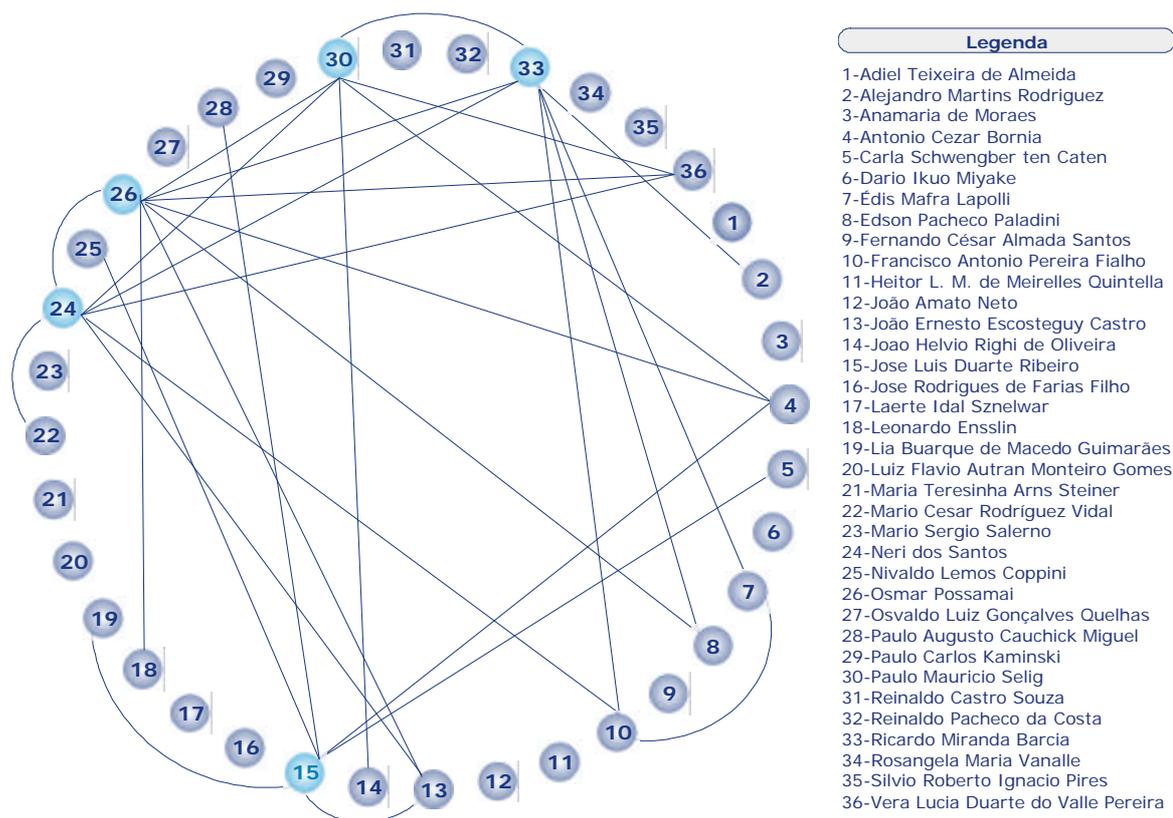


Figura 1c - Visualização de uma Rede de Co-Autoria com vértices de grau 6

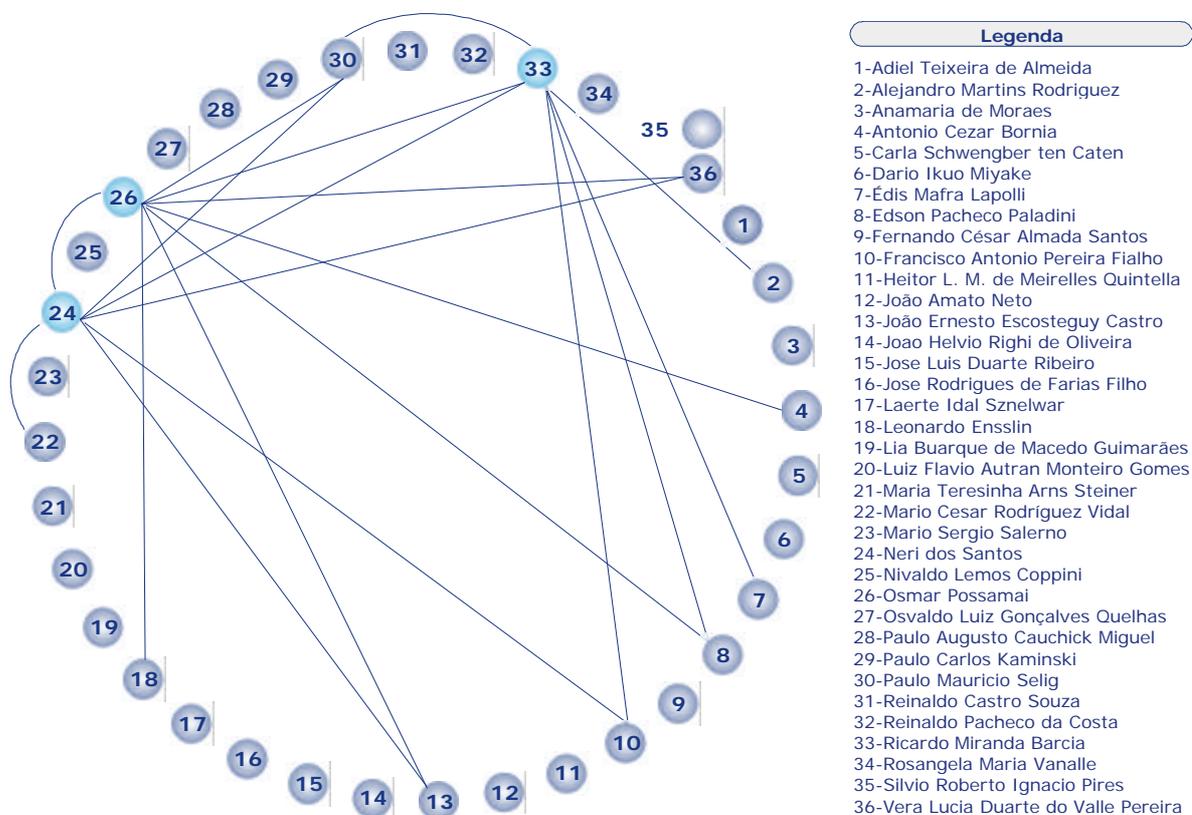


Figura 1d - Visualização de uma Rede de Co-Autoria com vértices de grau 7

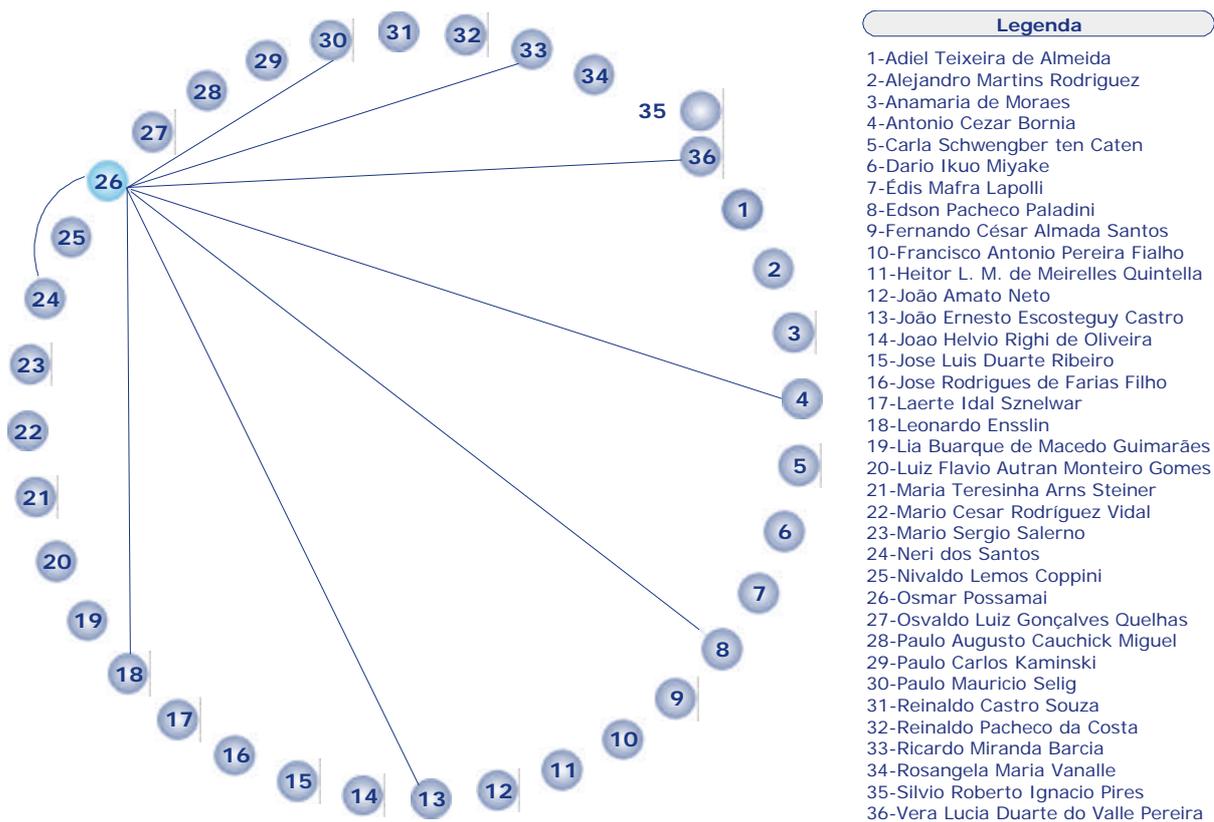


Figura 1e - Visualização de uma Rede de Co-Autoria com vértices de grau 8