



ROBERTO MARIO LOVÓN CANCHUMANI

Domínios Científicos na UFRJ: Mapeamento de Áreas de Conhecimento

**Tese de Doutorado
Agosto de 2015**



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO - UFRJ
ESCOLA DE COMUNICAÇÃO - ECO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - PPGCI

ROBERTO MARIO LOVÓN CANCHUMANI

DOMÍNIOS CIENTÍFICOS NA UFRJ: MAPEAMENTO DE ÁREAS DE
CONHECIMENTO

Rio de Janeiro
2015

ROBERTO MARIO LOVÓN CANCHUMANI

**DOMÍNIOS CIENTÍFICOS NA UFRJ: MAPEAMENTO DE ÁREAS DE
CONHECIMENTO**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da Informação.

Orientadora: Jacqueline Leta

Co-orientador: Antônio MacDowell de Figueiredo

Rio de Janeiro

2015

C215 Canchumani, Roberto Mário Lovón.
Domínios científicos na UFRJ: mapeamento de áreas de conhecimento / Roberto Mario Lovón Canchumani. 2015.
185f.

Orientadora: Jacqueline Leta.

Coorientador: Antônio MacDowell de Figueiredo.

Tese (doutorado) – Universidade Federal do Rio de Janeiro, Escola de Comunicação; Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós Graduação em Ciência da Informação, 2015.

1. Produtividade científica. 2. Universidade Federal do Rio de Janeiro – Pós-Graduação – Pesquisa. 3. Bibliometria. 4. Ensino superior – Pesquisa - Brasil. I. Leta, Jacqueline. II. Figueiredo, Antônio MacDowell de. III. Universidade Federal do Rio de Janeiro. Escola de Comunicação. IV. IBICT.

CDD: 001.4

ROBERTO MARIO LOVÓN CANCHUMANI

**DOMÍNIOS CIENTÍFICOS NA UFRJ: MAPEAMENTO DE ÁREAS DE
CONHECIMENTO**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da Informação.

Rio de Janeiro, 26 de agosto de 2015

Prof. Dra. Jacqueline Leta (Orientadora)
PPGCI/IBICT - ECO/UFRJ

Prof. Dr. Antonio MacDowell de Figueiredo (Co-orientador)
COPPE/UFRJ

Prof. Dra. Gilda Olinto de Oliveira
PPGCI/IBICT - ECO/UFRJ

Prof. Dra. Eloísa da Conceição Príncipe de Oliveira
PPGCI/IBICT - ECO/UFRJ

Prof. Dr. Fernando Alves Rochinha
COPPE/UFRJ

Prof. Dr. Jesús Pascual Mena Chalco
CMCC/UFABC

Prof. Dr. Pierre Ohayon
FACC/UFRJ

AGRADECIMENTOS

À professora Jacqueline Leta, por sua valiosa orientação e confiança depositada ao longo de todo este percurso, além dos incentivos para o desenvolvimento do doutorado sanduíche.

Ao professor Antonio MacDowell de Figueiredo, pela oportunidade de participar do Scire-COPPE/UFRJ, e pelo encorajamento e estímulo para levar adiante esta investigação.

Ao pessoal do Scire-COPPE/UFRJ, em particular a Antonella Latorraca e Marcos Sposito, pelo convívio, atenção e facilitação do acesso às informações indispensáveis para este trabalho.

Ao Dr. Ed Noyons, pela sua contribuição ao desenvolvimento da pesquisa como *supervisor* durante o período de realização do doutorado sanduíche no *Center for Science and Technology Studies (CWTS)*, Leiden University, Holanda.

À comunidade do CWTS, pelo acolhimento. Em especial a Clara Calero-Medina, pela oportunidade de participar do *Brazilian Research Ranking*.

Aos professores (as) e funcionários (as) do PPGCI-IBICT.

Aos professores (as) integrantes da banca, pelas valiosas contribuições.

Aos meus familiares, por terem compreendido a minha ausência.

A todos que, direta ou indiretamente, contribuíram para a realização desta tese.

À CAPES, pelo apoio financeiro através de bolsa de estudo e pela concessão da bolsa de doutorado sanduíche.

RESUMO

CANCHUMANI, Roberto Mario Lovón. **Domínios Científicos na UFRJ: mapeamento de áreas de conhecimento.** Tese (Doutorado em Ciência da Informação) – PPGCI, IBICT-ECO/UFRJ, Rio de Janeiro, 2015, 185 f.

No Brasil, a maior parte da atividade científica é desenvolvida nos programas de pós-graduação das universidades públicas. A diversidade temática e a complexidade estrutural dessas instituições fazem com que a identificação de seus perfis e padrões acadêmicos não seja uma tarefa fácil. O estudo da produção científica, a partir da análise de domínios científicos, permite conhecer a forma como estão configuradas/estruturadas as atividades científicas dessas instituições e os diversos atores nelas envolvidos. Considerando as seguintes questões de pesquisa: De que maneira estão configurados os domínios científicos da UFRJ? Eles são estáticos ou mudam ao longo do tempo? Os diversos domínios interagem entre si? Esta tese tem como principal objetivo mapear os domínios científicos de uma das maiores e mais importantes instituições de ensino e pesquisa do país, a Universidade Federal do Rio de Janeiro (UFRJ), identificando as principais áreas de conhecimento da instituição, bem como as interações entre os diferentes atores de seu corpo social. Trata-se de um estudo de natureza quantitativa, envolvendo informações de áreas vinculadas à produção científica e autores dessa produção dos programas de pós-graduação da UFRJ, que utiliza técnicas bibliométricas e de análise de redes a partir das medidas de co-ocorrência. Para tanto, este trabalho investigou 44.233 registros recuperados da base de dados institucional, EspaçoSIGMA.UFRJ, de artigos publicados em periódicos no período de 2001 a 2012. Os dados foram organizados e analisados com auxílio dos softwares Excel e Gephi. A partir de dois blocos de análises, considerando artigos em periódicos-áreas do conhecimento e artigos em periódicos-autores, os resultados mostram as principais áreas temáticas da UFRJ, destacando-se com maior quantidade de registros: Medicina, Química, Bioquímica, Zoologia, Microbiologia, Engenharia Química, Engenharia de Materiais e Metalúrgica. Além destas, encontramos também destacada presença de registros nas áreas de Letras, História, Educação, Psicologia. Outro aspecto observado foi o nível de interação entre as áreas: em torno de 30% dos artigos publicados são vinculados a mais de uma área de conhecimento. Verificou-se uma dinâmica de interação entre áreas de conhecimento, na qual se destacam: Ciências da Saúde, Engenharias e Ciências Humanas. O desdobramento das grandes áreas permitiu identificar aspectos mais específicos da interação entre as mesmas, possibilitando o mapeamento de comunidades de áreas de conhecimento. A análise por autores e as interações que se estabelecem a partir da medida de coautoria revelaram uma tendência de queda no número de artigos contendo entre 1 e 4 autores, enquanto que tende a crescer o número de artigos contendo 5 ou mais autores. A distribuição das autorias pelo tipo de vínculo institucional do autor e a relação entre produtividade, colaboração (coautoria) e ligação institucional também foram observados. As métricas de análise de redes sociais mostraram que, ao longo do período estudado, novas colaborações entre autores da UFRJ foram estabelecidas, formando novos agrupamentos/comunidades de autores. As maiores comunidades de autores identificadas foram detalhadas, revelando suas principais especificações.

Palavras-chave: Domínios Científicos; Coautoria; Colaboração; Bibliometria; Cientometria; Mapas da Ciência; Conhecimento; UFRJ.

ABSTRACT

CANCHUMANI, Roberto Mario Lovón. **Scientific Domains in UFRJ: mapping of fields of knowledge**. Thesis (PhD in Information Science) – PPGCI, IBICT-ECO/UFRJ, Rio de Janeiro, 2015, 185 f.

In Brazil, most of the scientific activity is developed by graduate programs from public universities. The thematic diversity and structural complexity of these institutions make the identification of their profiles and academic standards a hard task. The study of scientific production, using scientific domain analysis, allows knowing how the scientific activities are configured and the various actors involved. Considering the research questions: how are the scientific domains of the UFRJ configured? Do they change over time? Do different domains interact with each other? This thesis aims mapping the scientific domains of one of the largest and most important educational and research institutions of the country, the Federal University of Rio de Janeiro (UFRJ), identifying the its main fields of expertise as well as the interactions between the different actors of its social body. It is a quantitative study that encompasses information related to the fields of knowledge and authorship from the scientific production of UFRJ's graduate programs, which uses the network analysis technique based on the co-occurrence measures. The study investigated 44,233 records retrieved from the institutional database, EspaçoSIGMA.UFRJ, articles published in journals from 2001 to 2012. Data were organized and analyzed using Excel and Gephi software. The two sets of analysis, including articles in journals-fields of knowledge and articles in journal-authors, the results show the main thematic fields of UFRJ. Those with the highest number of records are: Medicine, Chemistry, Biochemistry, Zoology, Microbiology, Chemical Engineering, Materials Engineering and Metallurgy. We also find a strong presence of records in the fields of Letters, History, Education, Psychology. Another finding was the level of interaction among fields: around 30% of the published articles are linked to more than one field of knowledge. From this perspective, there was a dynamic interaction, over the period studied, among fields, in which the largest core fields are: Health Sciences, Engineering and Humanities. The analysis of large fields has identified more specific aspects of the interaction between them, allowing the mapping of clusters/communities of fields of knowledge. The analysis by authorship and the interactions established from co-authorship measure showed a downward trend in the number of articles with 1 to 4 authors, whereas tends to increase the number of articles with 5 or more authors. The distribution of authorship by the type of institutional link of the author and the relationship between productivity, collaboration (co-author) and institutional link were also observed. The social network analysis showed that, over the study period, new collaborations among authors of UFRJ were established, forming new clusters/communities of authors. The largest authors' clusters/communities were detailed and revealed their main specifications. It is, therefore, a singular study, with a comprehensive view of the scientific activity of the entire social body of one of the leading teaching and research institutions in Brazil.

Key-words: Scientific Domains; Co-authorship; Collaboration; Bibliometrics; Scientometrics; Maps of Science; Knowledge; UFRJ.

LISTA DE FIGURAS

Figura 1:	Crescimento do número de periódicos de acordo com Solla Price (1963).....	29
Figura 2:	Esquema da análise de redes sociais.....	55
Figura 3:	Redes de autorias e de <i>papers</i>	60
Figura 4:	Representações visuais utilizadas por White e McCain (1998).....	67
Figura 5:	Exemplo de mapeamento de grandes domínios científicos EUA-2000 a partir da co-citação de periódicos da base <i>ISI-WoS</i>	69
Figura 6:	Página principal do EspaçoSIGMA.UFRJ: SIGMA.Foco.....	74
Figura 7:	EspaçoSIGMA.UFRJ: atividades-fim, estrutura e funções, vínculos e associações.....	76
Figura 8:	EspaçoSIGMA.UFRJ: vínculos dos resultados dos programas de pós-graduação.....	77
Figura 9:	Exemplo de uma rede com diferentes algoritmos de visualização...	81
Figura 10:	Mapa da produção de artigos da UFRJ, agrupado por grande área do conhecimento (2001-2012).....	87
Figura 11:	UFRJ – Artigos publicados em periódicos e percentuais de artigos vinculados a mais de uma área do conhecimento por triênios.....	88
Figura 12:	UFRJ - Percentual de artigos publicados em periódicos por número de áreas do conhecimento a eles vinculados.....	89
Figura 13:	Mapa da interação entre grandes áreas do conhecimento a partir das informações de artigos publicações em periódicos UFRJ (2001-2012).....	90
Figura 14:	Mapa da interação entre grandes áreas do conhecimento a partir das informações de artigos publicações em periódicos UFRJ, por triênios.....	94

Figura 15:	Mapa da produção de artigos da UFRJ, agrupado pelas 75 áreas que compõem o nível áreas do conhecimento (2001-2012).....	95
Figura 16:	Mapa da interação entre as 75 áreas do conhecimento a partir das informações de artigos em periódicos UFRJ (2001-2012).....	107
Figura 17:	Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos, UFRJ (2001-2003).....	109
Figura 18:	Grafo que destaca a Comunidade 1, identificada no período 2001-2003.....	112
Figura 19:	Grafo que destaca a Comunidade 2 identificada no período 2001-2003.....	114
Figura 20:	Grafo que destaca a Comunidade 3 identificada no período 2001-2003.....	115
Figura 21:	Grafo que destaca a Comunidade 4 identificada no período 2001-2003.....	117
Figura 22:	Grafo que destaca a Comunidade 5 identificada no período 2001-2003.....	118
Figura 23:	Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas no período 2001-2003.....	120
Figura 24:	Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos, UFRJ (2010-2012).....	121
Figura 25:	Grafo que destaca a Comunidade 1 identificada no período 2010-2012.....	123
Figura 26:	Grafo que destaca a Comunidade 2 identificada no período 2010-2012.....	124
Figura 27:	Grafo que destaca a terceira comunidade de áreas identificada no período 2010-2012.....	126
Figura 28:	Grafo que destaca a Comunidade 4 identificada no período 2010-2012.....	127
Figura 29:	Grafo que destaca a comunidade 5 identificada no período 2010-2012.....	129

Figura 30:	Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas no período 2010-2012.....	131
Figura 31:	UFRJ – Artigos publicados em periódicos e percentuais de Artigos em coautoria por triênios.....	136
Figura 32:	Distribuição (%) de autores com três ou mais artigos publicados em colaboração, de acordo com o vínculo institucional com a UFRJ.....	138
Figura 33:	Distribuição (%) de autores por número de artigos da UFRJ publicados em colaboração e vínculo institucional.....	140
Figura 34:	Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos em periódicos no período 2001-2010.....	142
Figura 35:	Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos em periódicos por triênios.....	143
Figura 36:	Grafo das comunidades de autores da UFRJ de publicações em periódicos identificadas no período 2001-2003.....	145
Figura 37:	Grafo da maior comunidade de autores de artigos em periódicos da UFRJ no período 2001-2003.....	146
Figura 38:	Grafo dos autores mais centrais da maior comunidade de autores de publicações da UFRJ identificada em 2001-2003.....	148
Figura 39:	Grafo que mostra as 45 comunidades de autores da UFRJ identificadas no período 2004-2006.....	149
Figura 40:	Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2004-2006.....	150
Figura 41:	Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2004-2006.....	152
Figura 42:	Grafo que mostra as 50 comunidades de autores da UFRJ identificadas no período 2007-2009.....	153
Figura 43:	Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2007-2009.....	154
Figura 44:	Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2007-2009.....	156
Figura 45:	Grafo que mostra as 40 comunidades de autores da UFRJ identificadas no período 2010-2012.....	157

Figura 46:	Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2010-2012.....	158
Figura 47:	Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2010-2012.....	159

LISTA DE TABELAS

Tabela 1:	UFRJ – Distribuição percentual de artigos publicados em periódicos por grande área do conhecimento.....	85
Tabela 2:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas, de acordo com as grandes áreas do conhecimento por triênios.....	91
Tabela 3:	UFRJ – Centralidade de grau de grandes áreas do Conhecimento.....	93
Tabela 4:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências da Saúde.....	97
Tabela 5:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Biológicas.....	98
Tabela 6:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Exatas e da Terra....	99
Tabela 7:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas de Engenharias.....	101
Tabela 8:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas de Ciências Humanas.....	102
Tabela 9:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Linguística, Letras e Artes.....	103
Tabela 10:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Sociais Aplicadas.....	104
Tabela 11:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Agrárias.....	105
Tabela 12:	UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas nas Outras (Grandes Áreas).....	106
Tabela 13:	Relação de áreas e medidas de centralidade da Comunidade 1 identificada em 2001-2003.....	110
Tabela 14:	Relação de áreas e medidas de centralidade da Comunidade 2 identificada em 2001-2003.....	113

Tabela 15:	Relação de áreas e medidas de centralidade da Comunidade 3 identificada em 2001-2003.....	115
Tabela 16:	Relação de áreas e medidas de centralidade da Comunidade 4 identificada em 2001-2003.....	116
Tabela 17:	Relação de áreas e medidas de centralidade da Comunidade 5 identificada em 2001-2003.....	117
Tabela 18:	Relação de áreas, e medidas de centralidade, responsáveis por conectar as comunidades identificadas em 2001-2003.....	119
Tabela 19:	Relação de áreas e medidas de centralidade da Comunidade 1 identificada em 2010-2012.....	122
Tabela 20:	Relação de áreas e medidas de centralidade da Comunidade 2 identificada em 2010-2012.....	124
Tabela 21:	Relação de áreas e medidas de centralidade da Comunidade 3 identificada em 2010-2012.....	125
Tabela 22:	Relação de áreas e medidas de centralidade da Comunidade 4 identificada em 2010-2012.....	127
Tabela 23:	Relação de áreas e medidas de centralidade da Comunidade 5 identificada em 2010-2012.....	128
Tabela 24:	Relação de áreas, e medidas de centralidade, responsáveis por conectar as comunidades identificadas em 2010-2012.....	130
Tabela 25:	Distribuição (%) da totalidade de artigos em periódicos da UFRJ de acordo com o número de autores e o período de publicação.....	133
Tabela 26:	Distribuição (%) dos autores dos artigos em periódicos da UFRJ de acordo com o vínculo institucional e o triênio da publicação.....	134
Tabela 27:	Distribuição (%) de autores em artigos com e sem colaboração da UFRJ de acordo com o vínculo institucional e período da publicação.....	137
Tabela 28:	Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2001-2003.....	147
Tabela 29:	UFRJ - Autores com maiores medidas de centralidade da maior comunidade identificada em 2004-2006.....	151

Tabela 30:	UFRJ - Autores com maiores medidas de centralidade da maior comunidade identificada em 2007-2009.....	155
Tabela 31:	UFRJ - Autores com maiores medidas de centralidade da maior comunidade identificada em 2010-2012.....	159

SUMÁRIO

1	INTRODUÇÃO.....	19
2	INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA PARA A GESTÃO INSTITUCIONAL.....	25
2.1	DA CIÊNCIA MODERNA À POLÍTICA CIENTÍFICA E TECNOLÓGICA.....	26
2.2	OS MECANISMOS FORMAIS DE AVALIAÇÃO E O DESEMPENHO DA CIÊNCIA.....	31
2.3	A BIBLIOMETRIA E A AVALIAÇÃO DA CIÊNCIA POR <i>EXPERTS</i>	38
3	DOMÍNIOS DA CIÊNCIA E VISUALIZAÇÃO DA INFORMAÇÃO.....	44
3.1	CONSIDERAÇÕES PRELIMINARES.....	45
3.2	ANÁLISE DE DOMÍNIO.....	46
3.2.1	Comunidades de discurso: foco da análise de domínio.....	48
3.2.2	Bibliometria e análise de domínio: algumas aplicações.....	50
3.3	ANÁLISE DE REDES.....	52
3.3.1	Aspectos conceituais da análise de redes sociais.....	53
3.3.1.1	Medidas de análise de redes sociais.....	55
3.3.2	Análise de redes sociais no âmbito de atividade científica.....	58
3.3.2.1	Redes de colaboração a partir das informações de autoria.....	58
3.4	VISUALIZAÇÃO DA INFORMAÇÃO CIENTÍFICA.....	64
3.4.1	Visualização da informação e análise de domínio: geração de mapas da ciência.....	66
3.4.2	Visualização da informação e o conceito de redes sociais.....	69
4	METODOLOGIA.....	71
4.1	OBJETIVOS E ESTRATÉGIA METODOLÓGICA.....	72
4.2	COLETA DE DADOS.....	74
4.2.1	Definição da unidade de análise.....	75
4.2.2	Coletando os dados.....	78

4.3	ANÁLISE DOS DADOS.....	78
5	DOMÍNIOS INSTITUCIONAIS DA UFRJ.....	83
5.1	PRODUÇÃO CIENTÍFICA E AS GRANDES ÁREAS DO CONHECIMENTO.....	84
5.1.1	A interação entre as grandes áreas do conhecimento.....	87
5.1.2	A centralidade das grandes áreas do conhecimento.....	92
5.2	PRODUÇÃO CIENTÍFICA E AS ÁREAS DO CONHECIMENTO.....	95
5.2.1	Conhecendo a composição das áreas do conhecimento.....	95
5.2.1.1	As áreas nas Ciências da Saúde.....	96
5.2.1.2	As áreas nas Ciências Biológicas.....	97
5.2.1.3	As áreas nas Ciências Exatas e da Terra.....	99
5.2.1.4	As áreas nas Engenharias.....	100
5.2.1.5	As áreas nas Ciências Humanas.....	101
5.2.1.6	As áreas na Linguística, Letras e Artes.....	102
5.2.1.7	As áreas nas Ciências Sociais Aplicadas.....	103
5.2.1.8	As áreas nas Ciências Agrárias.....	105
5.2.1.9	As áreas nas Outras (Grandes Áreas).....	106
5.2.2	Redes de interação entre áreas do conhecimento: identificação de comunidades.....	107
5.2.2.1	Comunidades e principais áreas identificadas no triênio 2001-2003.....	109
5.2.2.1.a	A Comunidade 1.....	110
5.2.2.1.b	A Comunidade 2.....	112
5.2.2.1.c	A Comunidade 3.....	114
5.2.2.1.d	A Comunidade 4.....	116
5.2.2.1.e	A Comunidade 5.....	117
5.2.2.1.f	Interação entre as Comunidades.....	118
5.2.2.2	Comunidades e principais áreas identificadas no triênio 2010-2012.....	120
5.2.2.2.a	A Comunidade 1.....	121
5.2.2.2.b	A Comunidade 2.....	123
5.2.2.2.c	A Comunidade 3.....	125
5.2.2.2.d	A Comunidade 4.....	126
5.2.2.2.e	A Comunidade 5.....	128

5.2.2.2.f Interação entre as Comunidades.....	129
5.3 PRODUÇÃO CIENTÍFICA E AUTORIA.....	132
5.3.1 Contexto geral: autoria, produção e vínculo institucional.....	132
5.3.2 Coautoria e vínculo institucional.....	135
5.3.3 Coautoria, produtividade e vínculo institucional.....	138
5.3.4 Redes de interação entre autores: identificação de comunidades.....	141
5.3.4.1 Comunidades de autores no triênio 2001-2003.....	144
5.3.4.2 Comunidades de autores no triênio 2004-2006.....	148
5.3.4.3 Comunidades de autores no triênio 2007-2009.....	152
5.3.4.4 Comunidades de autores no triênio 2010-2012.....	156
6 DISCUSSÃO E CONCLUSÃO.....	161
REFERÊNCIAS.....	172

1 INTRODUÇÃO

Desde meados do século passado, a ciência e tecnologia passaram a ser entendidos como elementos estratégicos para o desenvolvimento econômico. Tanto a ciência e a tecnologia como as instituições a elas relacionadas são, hoje, componentes centrais da economia e da sociedade do conhecimento em escala global. Elas são motores fundamentais do progresso econômico e social, fatores chave para competitividade das empresas e determinantes do nível de emprego e da qualidade de vida (VELHO, 2011). Por outro lado, o acompanhamento e a análise dos resultados da atividade institucional na ciência e na tecnologia, tais como publicações, relatórios, patentes etc., constituem, atualmente, uma ferramenta essencial para o estudo do desempenho do setor (VAN RAAN, 2004).

Governos assim como especialistas em políticas de cunho científico-tecnológico estão, cada vez mais, interessados nas relações da ciência e tecnologia com alguns aspectos econômicos e sociais, tais como o nível de emprego, o crescimento econômico e, também, com assuntos ou temas de interesse amplo, como o meio ambiente (VIOTTI, 2008; VELHO, 2011). Demandam-se, neste contexto, metodologias e instrumentos de medida de desempenho do setor, que coloquem tais relações em destaque. Daí, a relevância e a necessidade de explorar e desenvolver técnicas, ferramentas e metodologias para esse fim, tendo em conta que os indicadores da produção científica são um dos elementos chave para uma boa gestão do setor assim como das instituições (MOED *et al.*, 1995).

No Brasil, tal como na maioria dos países, grande parte da atividade científica desenvolve-se nas universidades públicas, que dão origem à maior parte da produção científica, em determinadas disciplinas. Atualmente, o desempenho do setor assim como a elaboração de políticas científicas voltadas para as universidades parecem inconcebíveis sem indicadores que apontem os resultados por elas alcançados.

Os indicadores da produção científica têm sido amplamente usados para este fim (MOED *et al.*, 1995; NOYONS, 2004). De fato, estudos sobre a produção científica, baseados em indicadores bibliométricos e cientométricos têm adquirido destacada relevância nas últimas décadas, seja para a avaliação dos resultados de pesquisa seja para identificar e caracterizar o perfil científico de países e de suas instituições (LETA, 2011). Tais estudos tornam possível situar um país em relação ao mundo, uma instituição em relação ao país e, inclusive, grupos de pesquisa em relação a outros (OKUBO, 1997; VAN RAAN, 2004).

Estes indicadores são igualmente adequados para análise desde o nível macro (países, regiões, grandes áreas do conhecimento, por exemplo) até para estudos micro (pesquisador, periódico, universidade). Neste caso, sua aplicação constitui uma forma de identificar o estado da atividade científica de uma determinada comunidade científica ou de uma universidade. Não substituindo, mas proporcionando aos especialistas informações elaboradas, dados e análises gerados a partir de estudos desta natureza constituem-se como importantes recursos para aqueles com atribuição de planejar as ações direcionadas a implementar uma ágil e eficaz política científica para essas instituições (BRAUN, 1999).

Uma vantagem desses indicadores é que eles permitem observar, descrever, avaliar e monitorar o estado e a evolução de domínios científicos em diferentes níveis institucionais ou temáticos (MOYA-ANEGÓN *et al.*, 2004); aqui, entende-se domínios científicos a forma como estão configuradas/estruturadas as atividades científicas e os diversos atores nelas envolvidos (McCAIN *et al.*, 2006). Da mesma forma, estes indicadores são capazes de mostrar a trajetória de uma instituição ao longo de determinado período de tempo.

O interesse nos aspectos acima mencionados e minhas atividades no Laboratório Scire¹, da COPPE/UFRJ - Coordenação dos Programas de Pós-Graduação em Engenharia da Universidade Federal do Rio de Janeiro, motivaram o desenvolvimento da presente investigação. O Scire dedica-se à conceituação, desenvolvimento e operação de sistemas de informações relativas às atividades fins de instituições de ensino superior, especialmente, de seus programas de pós-graduação, com objetivos de acompanhamento, avaliação e planejamento nas áreas

¹ Disponível em: <http://www.scire.coppe.ufrj.br/>

de ensino e de pesquisa, assim como de apoio gerencial. Um dos sistemas operacionais desenvolvidos pelo Scire foi o EspaçoSIGMA.UFRJ, criado especialmente para gerenciar as informações acadêmicas da UFRJ.

Assim, a partir da minha experiência no Scire e interesse nos indicadores de desempenho acadêmico com fins em políticas institucionais, foi desenhada esta investigação que têm como campo de estudo a UFRJ – Universidade Federal do Rio de Janeiro e, como conceito chave, o domínio científico.

O estudo busca encontrar evidências que respondam às seguintes questões motivadoras: **De que maneira estão configurados os domínios científicos da UFRJ? Eles são estáticos ou mudam ao longo do tempo? Os diversos domínios interagem entre si?**

Partindo dessas questões, o objetivo principal desta tese é conhecer os domínios científicos da Universidade Federal do Rio de Janeiro (UFRJ), para identificar (1) conhecimentos e competências da instituição como reflexo das interações entre atores de seu corpo social e (2) das trajetórias da comunidade científica, sendo útil para representar a evolução e projeção de domínios do conhecimento da instituição.

O estudo tem, então, o foco na informação como instrumento para conhecer ou auxiliar na gestão de uma instituição, reconhecida por seu desempenho acadêmico e científico, a UFRJ. Tal como Spinak (2001), acreditamos que o trabalho científico deve ser observado e interpretado relativamente ao contexto social no qual está inserido. Dessa forma, o entendimento do desempenho científico, em qualquer nível de análise, deve ser sensível ao contexto conceitual, sócio-econômico e histórico da sociedade em que atuam os pesquisadores e suas instituições.

Sob este ponto de vista, entende-se que a atividade científica não pode ser medida numa escala absoluta, senão em relação às expectativas que a sociedade deposita em seus resultados. Assim, os indicadores adotados para sua caracterização devem ser capazes de captar uma quantidade de elementos que permita uma análise multidimensional dos processos que nela se manifestam.

Essa visão holística da atividade científica, de enfoque destacadamente social, tem sido utilizada por diversos autores nos últimos anos, principalmente a partir da análise de domínios proposta por Hjørland e Albrechtsen (1995). Estes

autores propõem a análise de domínios como um novo paradigma disciplinar, baseado na ideia de que a compreensão da ciência e de seus resultados deve realizar-se a partir das práticas sociais dos atores envolvidos. O enfoque descarta o estudo individualizado de um fenômeno, sendo contrário ao modelo cognitivo que exclui os entornos social e cultural em que participam os atores envolvidos (cientistas, estudantes, gestores, etc.). Dessa forma, a combinação de diferentes métodos (histórico, epistemológico, bibliométrico, dentre outros) resulta numa abrangente maneira de obter uma imagem suficientemente objetiva de determinado domínio (HJORLAND; ALBRECHTSEN, 1995).

Esta abordagem tem alcançado destacado impacto na Sociologia da Ciência e na Ciência da Informação, constituindo-se um suporte teórico para diversos estudos que empregam técnicas bibliométricas e cientométricas de representação da informação e resgatam as ideias de Small e Garfield sobre a possibilidade de construir mapas da ciência mundial baseados em indicadores de desempenho, tais como o número de publicações e de citações. Ela possibilita, então, o estudo de grandes domínios temáticos, assim como de domínios setoriais e institucionais (MOYA-ANEGÓN *et al.* 2004), sendo este último o foco do presente estudo, tal como expresso anteriormente nas perguntas de pesquisa.

No Brasil, como já mencionado, a maior parte da atividade científica é desenvolvida nas universidades públicas, de onde também se origina a grande maioria da produção científica, sobretudo nos programas de pós-graduação. As universidades caracterizam-se por apresentarem um perfil temático multidisciplinar e por estarem, desde o ponto de vista administrativo e funcional, fragmentadas em faculdades, departamentos, institutos, laboratórios etc., os quais realizam atividades acadêmicas e de pesquisa em uma ou mais áreas, disciplinas ou especialidades temáticas. A diversidade temática e a complexidade estrutural dessas instituições fazem com que a identificação de seus perfis e padrões acadêmicos não seja uma tarefa fácil.

Partindo-se do pressuposto de que a produção científica de uma dada instituição estaria representada pelo conjunto das produções científicas geradas em seu âmbito (BRAUN, 1999), é possível imaginar que o perfil e os padrões científicos da mesma estariam determinados pelo perfil e os padrões científicos de cada um

desses domínios institucionais específicos, dentro dos quais se desenvolvem as atividades de pesquisa. É neste contexto que o estudo da produção científica, na perspectiva da análise de domínios científicos ou domínios setoriais, torna-se relevante, pois pode permitir a cada instituição, neste caso a UFRJ, conhecer mais profundamente as potencialidades científicas de suas unidades constituintes e contar com informação objetiva e confiável de apoio à tomada de decisão, planejamento e avaliação de suas atividades.

Considerando tal pressuposto, desenhamos, então, este estudo, que tem a UFRJ como campo de estudo e domínio científico como conceito chave. Para operacionalizá-lo, foram utilizadas informações extraídas da base de dados do EspaçoSIGMA.UFRJ², que é um sistema de informações da UFRJ, cujas informações foram utilizadas nos processos de avaliação institucional entre os anos 1998 e 2012. Este sistema representa toda a sua estrutura organizacional e registra a forma de organização das atividades de ensino e pesquisa assim como os seus resultados (FIGUEIREDO, 2006). Esta base informacional ainda não tinha sido utilizada/explorada para estudos desta natureza.

O trabalho foi dividido em seis capítulos, organizados da seguinte forma. Após a introdução, o segundo capítulo aborda aspectos da ciência moderna à política científica e tecnológica, assim como os mecanismos formais de avaliação e desempenho da ciência, destacando o papel da Bibliometria nesse âmbito.

O terceiro capítulo apresenta aspectos relacionados à abordagem da análise de domínios e da análise de redes sociais no âmbito da atividade científica, assim como da visualização do conhecimento e os mapas da ciência.

O quarto capítulo apresenta a metodologia desta tese: objetivos geral e específicos, justificativa, natureza do estudo. Nesta parte discutem-se as etapas de coleta e análise de dados, além dos procedimentos e ferramentas utilizadas para a obtenção dos resultados.

O quinto capítulo contém os resultados da pesquisa, mostrando empiricamente a aplicação da metodologia proposta, assim como as análises elaboradas.

² Disponível em: <http://www.sigma.ufrj.br/site/espaco/index.htm>

No sexto capítulo está registrada a discussão e conclusão do trabalho, assim como sugestões para trabalhos futuros. Posteriormente, a relação de referências do estudo é apresentada.

2 A INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA PARA A GESTÃO INSTITUCIONAL

“A coisa mais estúpida que se pode fazer é não medir. A segunda coisa mais estúpida que se pode fazer é confiar totalmente nas medidas”.

(Michal Kalecki)

2.1 DA CIÊNCIA MODERNA À POLÍTICA CIENTÍFICA E TECNOLÓGICA

A ciência moderna, que assume hoje um papel central na descrição e compreensão dos fenômenos da natureza, teve início no século XVI. Uma figura central na história da ciência foi Galileu Galilei (1564-1642), que incorporava os dois atributos mais importantes da nova ciência: o conhecimento empírico, baseado na observação precisa, e o conhecimento teórico, baseado na capacidade de formular e manipular abstrações simbólicas de quantidade, número, relação, estrutura etc. (MARICONDA, 2006).

Galileu percebeu que o novo sistema do mundo exigia um novo modo de pensar e abordar o estudo da natureza. Dessa forma, ao estabelecer as bases de uma nova física, se lhe reconhece como principal personagem do surgimento da ciência moderna³.

Nesse período inicial da ciência moderna, alguns novos achados causaram uma ruptura absoluta com a tradição anterior. Por exemplo, o conceito de que o sol estava no centro do mundo (heliocentrismo), se opunha ao conceito predominante até então, de que a Terra estava no centro do mundo (geocentrismo). No meio desta disputa estavam as instituições, especialmente a Igreja, que viam neste novo conceito (e em outros) uma forma de abalar os dogmas religiosos que davam (e dão ainda) sustentação a esta instituição (MARICONDA, 2006).

Essas revoluções impactaram significativamente as áreas do conhecimento mais relevantes, surgindo uma nova forma de olhar, compreender e transformar a natureza. A partir de então, inicia-se a configuração do sistema científico moderno.

Um primeiro elemento desta configuração, que passou a ganhar destaque neste momento, foram as academias e sociedades científicas. As primeiras destas instituições datam do século XII, mas a maior parte delas foi criada a partir do século

³ A física moderna teve sua gênese numa dupla finalidade proposta por Galileu: (i) encontrar uma solução definitiva ao problema da física aristotélica: os movimentos violentos e a aceleração da queda dos corpos, e (ii) demonstrar que é fisicamente possível que a Terra se mexa e que isso está de acordo com os fatos observados (contrário ao que acreditavam os aristotélicos). O uso do telescópio lhe permitiu comprovar que o cosmos não está dividido em duas regiões diferentes, assim como também a existência dos satélites de Júpiter, provas fundamentais a favor do copernicanismo (Cf. MARICONDA; VASCONCELOS, 2006).

XVII. Entre as mais prestigiosas, a Royal Society de Londres fundada 1662 e a Académie de Science, em 1666, que a partir de 1816 passou a denominar-se Académie de Science, na França. Um século e meio depois, tinha-se uma rede de Academias ou Instituições Científicas que cobriam toda a Europa (ZIMAN, 1981).

As reuniões científicas passaram a ser cada vez mais frequentes nas academias e sociedades científicas espalhadas pela Europa, a tal ponto que surgiu a necessidade de guardar seus registros. Surgem, assim, a partir da segunda metade do século XVII, os primeiros periódicos científicos.

O primeiro periódico foi o Journal des Sçavans, fundado pela Académie des Sciences, em Paris em 1665 por Denis de Sallo, com o objetivo de oferecer resenhas de livros de cunho literário e científico, assim como também informar sobre a realidade cultural europeia por meio de um pequeno número de contribuições originais (MEADOWS, 1974).

De acordo com Stumpf (1996), a importância do Journal des Sçavans e sua influência sobre a cultura desse período foram tamanhas que até o fim do século XVII mais de 30 periódicos científicos surgiram em toda Europa.

Mas foi no mesmo ano, 1665, na Inglaterra, que se deu o nascimento do periódico mais influente para a nova tradição científica: o *Philosophical Transactions of the Royal Society of London*. Trata-se de um periódico científico que surge como emanção da *Royal Society*. Segundo Ziman (1981), os membros dessa sociedade perceberam que, ao contrário dos livros, as publicações periódicas possibilitavam uma rápida difusão das descobertas, especialmente no campo da medicina e das ciências naturais. Assim, apenas dois meses depois do surgimento do Journal des Sçavans, Henry Holdenburg, secretário geral da *Royal Society of London*, funda o *Philosophical Transactions*. Apesar de ser inspirado no periódico francês, o *Philosophical Transactions* nasce com uma proposta diferente: além de publicar numerosas contribuições originais relativas às mais importantes descobertas científicas, propõe chegar a ser um registro público desses artigos (MEADOWS, 1974; ZIMAN, 1981).

A intuição de Holdenburg faz do periódico o lugar privilegiado para a atribuição da autoria das descobertas, o que consolidará a importância desse veículo de comunicação nos séculos seguintes, transformando o periódico no

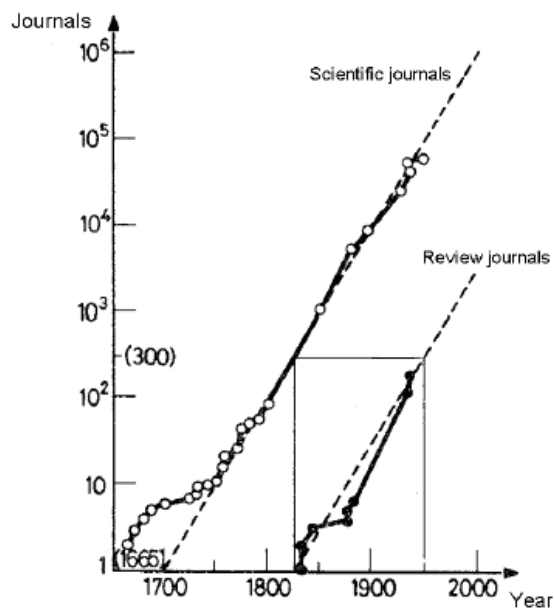
principal instrumento de difusão da informação científica (ZIMAN, 1981). Além disso, conforme destaca Stumpf (1996), os custos menores e a maior facilidade para transportar o periódico foram fatores decisivos para que este meio se popularizasse e, de certa forma, substituísse os livros na função de difusão do conhecimento científico.

Paralelamente, observamos um crescimento no número de cientistas, que se intensificou no século XIX, como consequência principalmente da incorporação da ciência como atividade fim das Universidades, aumentando o número de pessoas dedicadas à pesquisa e o nível de formação (BEN-DAVID; ZLOCZOWER, 1980).

Com a ciência ocupando, cada vez mais, um espaço socialmente reconhecido, como as universidades, o “fazer ciência” passou a ser percebido como uma atividade profissional e, portanto, houve um crescimento natural dos membros da comunidade científica. Criou-se, então, uma demanda crescente de transmissão e divulgação dos novos conhecimentos em uma escala cada vez maior, aumentando a relevância dos periódicos científicos (LÓPES-YEPES, 1989). E, no século XX, se consagram como o principal veículo de difusão do conhecimento científico.

Derek de Solla Price, historiador da ciência, em sua obra clássica *Little Science, Big Science*, publicada em 1963, destaca o incremento exponencial do número de periódicos (Figura 1), demonstrando como o desenvolvimento do conhecimento científico podia ser caracterizado e explicado através da contabilização de publicações.

Figura 1: Crescimento do número de periódicos de acordo com Solla Price



Fonte: Solla Price, 1963

Solla Price também apresenta uma série de considerações que sustentam sua tese da passagem da ciência moderna para a ciência atual. Essa transformação dividiu-se, segundo Barnes (1985), em diferentes fases, das quais podemos destacar: a profissionalização dos cientistas na Alemanha do século XIX, a consolidação da ciência em torno das universidades e o período entre as Grandes Guerras do século XX.

O ponto alto dessa transição veio com a Segunda Guerra Mundial e com os êxitos do *Projeto Manhattan*⁴, através do qual o governo dos Estados Unidos conscientiza-se da importância da pesquisa científica e de seus generosos retornos, sobretudo, do ponto de vista militar. De acordo com Galceran e Domingues (1997), o efeito demonstrativo do *Projeto Manhattan* serviu para justificar uma série de projetos de alta tecnologia, criando uma onda de esforços similares, não limitados à área militar, que também impactaram a tecnologia civil e determinadas prioridades nacionais.

É nesse cenário do pós-guerra em que são apresentadas as ideias de Vannevar Bush sobre a importância estratégica da ciência para o desenvolvimento dos EUA. Em seu destacado relatório intitulado *Science - The Endless Frontier*,

⁴ Caracterizados por determinados programas científico-militares dos Estados Unidos.

realizado a pedido do então presidente dos Estados Unidos, Franklin D. Roosevelt, Bush traz uma série de argumentos que sustentam os seguintes aspectos: a autonomia da ciência, a ciência básica como prioridade do setor público, necessidade de disponibilizar recursos públicos contínuos e a criação de uma agência responsável por gerenciar tais recursos (BUSH, 1945a; GODIN, 2005).

Já no início da década de 1950, foi criada a National Science Foundation (NSF), baseada nos conceitos propostos por Bush (1945), que destacava a necessidade de uma agência governamental que apoiasse as melhores pesquisas em universidades e institutos de investigação (Cf. EDWARDS, 1997).

Outro relatório, conhecido como *Steelman Report*, originalmente chamado de *Science and Public Policy*, de 1947, propunha o papel do governo federal dos Estados Unidos para apoiar a pesquisa nas universidades (LEYDESDORFF, 2005). Para Ziman (2003), a política científica e tecnológica é o fator mais importante na transição para o novo regime da ciência, na qual nos encontramos atualmente.

Outro fator determinante para a mudança apontada por Solla Price foi o lançamento, em 1957, em plena Guerra Fria do *Sputnik*, por parte da União Soviética. A partir de então, Estados Unidos e URSS empreendem uma corrida espacial sem precedentes, ocasionando um crescimento significativo de gasto público em pesquisa e desenvolvimento (P&D) (Cf. ZIMAN, 2003).

Os efeitos deste período, Guerra Fria, estenderam-se a outros países desenvolvidos com a implantação de novas políticas científicas e tecnológicas (GALCERAN; DOMINGUES, 1997). Essa nova virada levaria à denominada sociedade baseada no conhecimento, onde países e instituições que apresentam uma atividade científica e tecnológica forte e consolidada detêm vantagem competitiva em relação às outras onde esta atividade ainda é incipiente ou em desenvolvimento.

A partir dos anos de 1950 e 1960, os governos passaram a estruturar formalmente o setor de ciência e tecnologia e a destinar recursos públicos, cada vez maiores, para esse setor (ZIMAN, 1981).

No Brasil, segundo Schwartzman *et al.* (1995), a institucionalização da ciência teve início em 1951, com a criação do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, nomeado, inicialmente, de Conselho Nacional de

Pesquisas) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), cujas atuações se complementam até os dias atuais e são fundamentais para a organização da atividade científica no País.

Entretanto, Schwartzman *et al.* (1995) destacam que a maior parte do sistema científico e tecnológico brasileiro foi criado durante o regime militar. Nesse período, criaram-se instituições e fundos de investimento para a ciência e tecnologia, instituições coordenadoras das políticas científicas e tecnológicas, assim como planos de desenvolvimento para o setor.

2.2 OS MECANISMOS FORMAIS DE AVALIAÇÃO E O DESEMPENHO DA CIÊNCIA.

No século XX, o enorme desenvolvimento da ciência e as novas dimensões que passaram a caracterizá-la, dentre as quais o volume crescente de recursos públicos que os governos passaram a ser destinados para o setor, exigiram a implantação de mecanismos de avaliação dos resultados da atividade científica e dos atores que participam desse processo. Neste caso, a avaliação surge como uma estratégia fundamental para que os organismos gestores de políticas científicas e tecnológicas e agências de fomento pudessem distribuir os recursos de maneira (mais) eficiente, conhecendo os impactos da atividade científica na sociedade e a capacidade dos atores participantes (VAN RAAN, 2004).

Nesse contexto, a avaliação da ciência estende-se como um processo sistemático e objetivo que mensura a relevância, a eficiência e a efetividade das políticas e dos projetos de pesquisa em relação ao cumprimento dos objetivos originalmente estabelecidos (SANCHO, 2001).

Rip (2003) situa o surgimento da avaliação da ciência no período imediato pós Segunda Guerra Mundial. Nesse período, foram criadas as primeiras agências de financiamento das atividades científicas e, com elas, aparecem as primeiras práticas de avaliação, necessárias para determinar quais propostas de pesquisa deveriam ser financiadas. A avaliação enfatizava, nesse primeiro momento, o uso dos recursos públicos investidos na condução da atividade científica.

Segundo Godin (2002), a política científica dos anos de 1950 tratava de assegurar que os pesquisadores contassem com recursos públicos para produzir conhecimento, essencialmente de pesquisa básica. Assim, a avaliação dos resultados da atividade científica não foi considerada necessária durante alguns anos. Os próprios cientistas convenciam os governos da necessidade de financiar a pesquisa científica, ressaltando que este era um dos elementos que gerava progresso socioeconômico, exatamente como preconizado por Vannevar Bush (BUSH, 1945a).

Além disso, ainda de acordo com Godin (2002), o discurso de cientistas vinha reforçado pela visão que os economistas tinham sobre a atividade científica, concebida como um bem público gerador de benefícios, que, diferentemente dos bens privados, não poderia ser apropriado por seu produtor e cujos resultados eram de difícil mensuração. Esses argumentos estavam em sintonia com o modelo linear de organização e desenvolvimento do conhecimento científico e tecnológico, que emerge com grande ênfase nas palavras de Vannevar Bush (BUSH, 1945a), o qual se pode resumir em uma equação:

$$+ \textit{ci\^encia} = + \textit{tecnologia} = + \textit{riqueza} = + \textit{bem estar social}$$

Nesse modelo, a ciência e o conhecimento gerado pela pesquisa básica surgem como a base para o desenvolvimento de tecnologia e, conseqüentemente, para a geração de riqueza e de bem estar social. Assim, um aumento de insumo (*input*) na ciência resultaria, em última instância, em mais bem estar social. Desta forma, tornava-se urgente uma política científica e tecnológica que assegurasse recursos públicos, garantindo, assim, a justificação social da ciência e o investimento público nela.

No final da década de 1950, sem abandonar o modelo linear, os Estados Unidos propõem um relativo controle da educação científica e um processo de avaliação da produção de ciência e tecnologia, objetivando garantir sua qualidade. Criam-se, assim, instituições acadêmicas especializadas, institutos de estudos da ciência e revisam-se os programas curriculares. Por outro lado, o aumento da

demanda por recursos e espaço para atividades técnico científicas despertou a atenção de analistas e estudiosos.

Segundo Cabal (2000), tal demanda levou à Organização Europeia de Cooperação Econômica (OECE) a reunir, em 1957, um grupo de especialistas para examinar os gastos e definir métodos para a mensuração da atividade. Os estudos propostos buscavam uma padronização dos dados estatísticos, o que permitiria uma comparação do esforço das diferentes nações em pesquisa e desenvolvimento. Outro objetivo, talvez o mais relevante, era resolver o problema da coleta de dados para o planejamento de ações e políticas de ciência e tecnologia (OCDE, 1978, p.15).

Nesse período, é elaborado o Manual Frascati, que foi o primeiro manual internacional de normalização de indicadores voltados para as atividades de ciência e tecnologia, publicado em 1963, na Vila Falconieri de Frascati na Itália (OCDE, 1978). Os indicadores apresentados no Manual Frascati buscavam avaliar a eficácia da produção de ciência e tecnologia seja através do volume de investimentos e gastos realizados no setor, seja através do número de patentes, que indicariam o esforço na produção tecnológica.

O modelo linear entra em crise no final da década de 1970, quando se desenvolvem e consolidam uma série de movimentos de protestos contra certas linhas de desenvolvimento tecnológico e o papel tradicional dos especialistas na tomada de decisão. Os desastres relacionados com o desenvolvimento industrial contemporâneo, como os derramamentos de petróleo ou acidentes nucleares, contribuíram para a conscientização coletiva acerca dos riscos e impactos de uma ciência e tecnologia fora de controle. Formou-se, então, uma espécie de consenso básico:

[...] se bem a ciência e tecnologia proporcionam numerosos benefícios, também ocasionam impactos negativos, alguns imprevisíveis, mas todos eles refletem os valores, perspectivas e visões de quem esta em condições de tomar decisões concernentes ao conhecimento científico e tecnológico (GONZÁLES *et al.*, 1996, p.5).

Em etapa mais recente das políticas públicas de ciência e tecnologia começa a predominar o conceito de inovação, que se soma aos conceitos de ciência e tecnologia. Com base nas dimensões deste conceito, que diz respeito à novidade técnico-científica e o benefício derivado de sua introdução no mercado, trata-se agora de intervir no desenvolvimento científico e tecnológico para maximizar seu rendimento econômico.

De acordo com Gibbons e Georghiou (1987), a inovação tecnológica foi outro elemento que contribuiu para o aumento do interesse pela avaliação, já que começam a se priorizar disciplinas e pesquisadores com maiores possibilidades de contribuir à ulterior geração de tecnologia, considerada como elemento fundamental para o desenvolvimento da indústria, aumentando o interesse da avaliação de seu potencial econômico.

Esta fase é marcada pelos indicadores de inovação, os quais aparecem nos anos 1980 e se consolidam nos anos 1990 como um conjunto de indicadores de inovação no Manual de Oslo, cuja primeira edição foi publicada em 1992 e tinha o objetivo de servir de guia para a compilação de dados relativos à inovação tecnológica (OCDE, 1996). Estes indicadores são, basicamente, informações de empresários coletados através de questionários para mensurar o nível de aproveitamento no mercado das descobertas científicas. Isto supõe uma mudança substancial, já que se cobra uma repercussão de melhora no mercado derivada das políticas de ciência e tecnologia.

Constam ainda, organizados no âmbito da OCDE, o Manual de Patentes e o Manual de Canberra. O primeiro estabelece diretrizes sobre como os dados de patentes podem ser utilizados em análise que os relacionam com outras estatísticas sobre as atividades científicas, tecnológicas e econômicas. Já o Manual Canberra, publicado em 1995, tem o propósito de difundir a padronização de uma estrutura conceitual comum para a compilação, análise de perfís e tendências de dados relativos aos estoques e fluxos de recursos humanos em ciência e tecnologia.

Paralelamente aos Manuais, cabe mencionar que a OCDE investiu grande esforço também na editoração, desde 1984, de publicações e relatórios específicos de indicadores para o setor, tal como o *Science and Technology: Indicators Reports*,

posteriormente substituído pelo *Main Science and Technology Indicators* (OCDE, 2014).

Também a UNESCO tem estabelecido pautas e recomendações para a recopilação estatística de dados, por exemplo, por meio do trabalho *Recommendation Concerning the International Standardisation of Statistics on Science and Technology* (UNESCO, 1978), produzindo relevantes informes como os *World Science Report*.

Na América Latina, tradicionalmente, os indicadores de ciência e tecnologia têm sido utilizados como ferramenta para mensurar o logro de algumas políticas assim como para avaliar algumas áreas específicas de atividade científica e tecnológica ou para comparar o desempenho entre países.

As instituições que participam da Rede Íbero-Americana de Indicadores de Ciência e Tecnologia (RICYT)⁵, tal como a Organização dos Estados Americanos (OEA) entre outras instituições, após a realização de três seminários internacionais apresentaram à comunidade científica o Manual para a Normalização de Indicadores de Inovação Tecnológica para a América Latina e o Caribe, denominado Manual de Bogotá (RICYT, 2001). Apresentado em 2001, este manual, inspirado no Manual de Oslo para garantir a comparabilidade internacional, propõe algumas referências para a adequação dos indicadores de inovação às especificidades que caracterizam os sistemas de inovação e as empresas da América Latina e o Caribe.

O Brasil, seguindo a tendência dos outros países, inicia o desenvolvimento do seu sistema de indicadores analisando os recursos investidos em ciência e tecnologia. Em 1978, o CNPq, com auxílio da UNESCO, definiu a classificação das atividades de ciência e tecnologia financiadas com recursos públicos. Com esse levantamento obteve um indicador de gasto público em ciência e tecnologia, tal como destaca o texto a seguir, retirado do *site* oficial do Ministério da Ciência, Tecnologia e Inovação do Brasil:

⁵ Rede criada em 1994, auspiciada pela CYTED – programa pertencente à UNESCO – e a OEA. Tem como objetivo central apoiar tecnicamente aos países integrantes na melhoria em matéria de informação no âmbito da ciência, tecnologia e inovação (RICyT, 2001).

A partir dos anos 80, o CNPq iniciou a coleta e a publicação de informações sobre os recursos do Governo Federal aplicados em C&T, seguindo as primeiras recomendações do Manual Frascati da OCDE, para os gastos em P&D, e as sugestões da UNESCO, para as atividades científicas e tecnológicas correlatas. Uma década depois, os mesmos procedimentos passaram a ser utilizados na maioria dos estados brasileiros, permitindo a obtenção de um quadro abrangente dos recursos públicos aplicados em C&T. [...] Recentemente, o MCT passou a assumir a responsabilidade pela organização e a divulgação das informações de C&T do país, de forma centralizada. Para tanto, conta com a colaboração de inúmeras instituições públicas, no âmbito federal e estadual, e de organizações privadas que produzem informações de interesse para a construção de indicadores de C&T e para o desenvolvimento de estudos sobre o tema (BRASIL, 2002, p.2).

Cumprir mencionar, por outro lado, o papel do processo de avaliação, criado em 1976 pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), como mecanismo para analisar, de forma detalhada e sistemática, as atividades da pós-graduação *stricto sensu* no Brasil, objetivando garantir a qualidade dos cursos oferecidos e, ao mesmo tempo, induzindo a expansão do sistema (HORTA *et al.*, 2005).

Esse processo de avaliação, focado basicamente em resultados, tem produzido indicadores, principalmente de *outputs*, nos quais se baseiam as políticas governamentais de suporte e crescimento da pós-graduação *stricto sensu* brasileira, sendo atualmente o instrumento mais importante para os diversos tipos de fomento nesse âmbito (SALVÁ, 2015). Entretanto, os critérios utilizados por esse sistema de avaliação guardam um embate ideológico do controle da ciência, marcado pela disputa entre a ciência do pesquisador isolado e o modelo de produção de conhecimento de enfoque institucional, cujo campo de atuação envolve as grandes instituições de pesquisa e seus financiadores (GOLDANI *et al.*, 2010).

Nesse sentido, conforme destacam Goldani *et al.* (2010), as demandas por linhas de pesquisa e por resultados, antes de interesse pessoal, obedecem atualmente às influências de um cenário voltado para um mercado ainda indefinido, o que tem ocasionado mudanças no trabalho desenvolvido por pesquisadores integrantes dos programas de pós-graduação no Brasil.

Os indicadores de ciência e tecnologia que passaram a compor os manuais e também relatórios produzidos por órgãos internacionais, como OCDE e UNESCO, são originalmente oriundos da Bibliometria/Cientometria, um campo de

conhecimento das ciências sociais, que nasce com o intuito de explicar a dinâmica da ciência e da tecnologia a partir de indicadores de *outputs* e *inputs*. Os estudos neste campo passaram a ter maior destaque a partir de Derek de Solla Price, historiador da ciência. Ele junto com outros destacados nomes inauguram uma nova fase dos estudos sociais da ciência, a nova sociologia da ciência (GONZÁLES *et al.*, 1996).

Um dos autores de maior influência nesta corrente foi Thomas Kuhn, com a introdução de conceitos irredutivelmente sociais para explicar como a ciência se modifica, como é sua dinâmica e seu desenvolvimento. A obra de Kuhn (1962) dá lugar a uma conscientização sobre a dimensão social e o enfoque histórico da ciência, ao mesmo tempo em que inaugura um estilo interdisciplinar que tende a diminuir as fronteiras clássicas entre as especialidades acadêmicas, preparando o terreno para os estudos sociais da ciência.

Estes estudos, desde diversas tradições que incluem principalmente a filosofia da ciência pós kuhniana e a sociologia da ciência (superando a Escola de Robert Merton), propõem-se estudar os condicionantes sociais da geração do próprio conhecimento científico — é o caso de autores como Barnes, Bloor, Woolgar, Latour, Callon, Law, entre outros (GONZÁLES *et al.*, 1996). Há também os estudos históricos e de caso, onde a concepção tradicional se opõe à realidade, mostrando como os métodos de produção e validação do conhecimento não são sempre tão puros como o positivismo lógico postulava⁶.

Por outro lado, dentre os estudos sociais da ciência, surge uma disciplina focada na mensuração da atividade científica mediante técnicas estatísticas e quantitativas: a Bibliometria/Cientometria, da qual tratamos na seguinte seção deste capítulo.

⁶ Ver González *et al.* (1996) para um panorama completo sobre os estudos sociais da ciência, tecnologia e sociedade. Os autores distinguem entre duas tradições deste tipo de estudos: uma europeia, de corte essencialmente acadêmica, preocupada pela forma na qual o contexto social condiciona o desenvolvimento da ciência e a construção da tecnologia; e outra americana, de corte mais ativista e política, preocupada pelas consequências sociais e ambientais da mudança tecnológica.

2.3. A BIBLIOMETRIA E A AVALIAÇÃO DA CIÊNCIA POR *EXPERTS*

O termo Bibliometria foi proposto por primeira vez por Alan Pritchard em 1969 para referir-se à aplicação de métodos matemáticos e estatísticos a livros e outros meios de comunicação (PRITCHARD, 1969). Entretanto, nesse mesmo ano, Nalivov e Mulchenko publicaram o trabalho intitulado *Naukometrya: the study of the development of science as an information process*, que estava voltado à aplicação de métodos quantitativos aos estudos sobre o desenvolvimento da ciência considerada como um processo informativo (GRANOVSKY, 2001). Naukometrya foi, então, traduzido do russo para o inglês como Scientometrics.

Van Raan (1997) refere-se à Cientometria como os resultados quantitativos aplicados ao estudo da ciência e tecnologia, mencionando que seu objetivo é gerar conhecimento sobre o desenvolvimento da ciência e tecnologia e também em relação a questões sociais e políticas.

Hood e Wilson (2001), em estudo focado nas características históricas, de desenvolvimento e de relações entre Bibliometria, Cientometria e Informetria, concluem que esses termos são relacionados e utilizados para descrever parte ou toda uma disciplina. Os autores consideram, porém, interessante o fato de cada um desses termos terem uma origem histórica particular e em geral bem documentada. Os termos e definições usados por pesquisadores que atuam nesse campo do conhecimento geralmente se repetem ou sobrepõem em parte, mas não são necessariamente sinônimos.

A grande sobreposição, especialmente, de metodologias e a confusão terminológica são razões para o uso indistinto dos termos Bibliometria e Cientometria e também pelo uso mais frequente da expressão indicadores bibliometricos e não indicadores cientometricos, mesmo quando se trata do campo da ciência

Desde os anos de 1990, já havia intensa discussão em torno do melhor termo para dar identidade ao campo das métricas da ciência. E, em 1995, quando a sociedade internacional do campo passou a denominar-se *International Society of Scientometrics and Informetrics*, o termo Informetria ganhou destaque e passou,

então, a ser utilizado mais amplamente entre os especialistas para se referir tanto à Bibliometria como à Cientometria (EGGHE; ROUSSEAU, 1990).

Os estudos bibliométricos pioneiros têm início nas primeiras décadas do século XX, destacando-se: o realizado por Coles e Eales, em 1917, no qual analisaram as publicações sobre anatomia comparativa entre 1543 e 1860 (COLE; EALES, 1917); o estudo de Lotka, em 1926, sobre a produtividade dos autores que tinham publicado seus trabalhos no periódico *Chemical Abstracts* entre 1907 e 1916, a partir do qual enunciaria sua famosa lei sobre a distribuição regular da produtividade dos autores (LOTKA, 1926) e; a análise realizada por Gross e Gross, em 1927, a partir de referências bibliográficas incluídas num conjunto de artigos publicados no *The Journal of the American Chemical Society*, apontando que poucos periódicos eram citados muito frequentemente, sendo um destacado precedente dos índices de citações desenvolvidos de forma sistemática décadas depois (GROSS; GROSS, 1927).

Já na década seguinte, especificamente em 1934, Bradford analisou a distribuição de artigos em geofísica, publicados em 326 periódicos da área no período 1931-1933. Bradford descobriu que nove periódicos tinham publicado 429 trabalhos, outros 59 periódicos reuniam 499 trabalhos e os 258 restantes, outros 404 trabalhos. A partir de estudos semelhantes a estes, posteriormente, Bradford enunciaria o princípio da dispersão ou distribuição de frequências de artigos em periódicos (BRADFORD, 1934; 1948). Em 1935, Zipf realizou a primeira formulação sistemática de sua lei acerca da frequência de aparição das palavras nos textos, segundo a qual, um reduzido número de palavras são utilizadas frequentemente, enquanto que grande parte de palavras são pouco usadas (ZIPF, 1935).

Nos anos 1950 e 1960, aparecem as contribuições de Garfield no tocante ao índice de citações. Este autor desenvolveu, em 1961, um projeto para o Instituto Nacional de Saúde dos Estados Unidos, com o objetivo de produzir um índice de citações de genética (*Genetics Citation Index*), descobrindo que muitos dos artigos sobre genética eram publicados em periódicos não dedicados exclusivamente à genética (GARFIELD; SHER, 1963). Essa descoberta levou-o a associar o uso das citações como uma medida de importância dos periódicos, valorizando o interesse de dispor de uma ferramenta de busca de dados multidisciplinares. Tais ideias se

concretizaram no “*Citation Index*” e na medida do fator de impacto, índices que foram além das ferramentas bibliográficas tradicionais para tornarem-se ferramentas de avaliação da visibilidade e do impacto das publicações científicas (GARFIELD, 1955; 2006).

O ano 1963 destaca-se pela publicação de *Little Science, Big Science*, onde, seu autor, o historiador Derek J. de Solla Price analisa de maneira sistemática as características, a estrutura e o desenvolvimento da ciência moderna (SOLLA PRICE, 1963). O autor teve, adicionalmente, uma contribuição fundamental no desenvolvimento inicial dos estudos das redes de pesquisadores, a partir do conceito de “Colégios Invisíveis”, acenada na publicação mencionada e desenvolvida posteriormente, considerando a citação dos artigos, em *Networks of Scientific Paper* (SOLLA PRICE, 1965).

Estas ideias foram retomadas anos depois por Crane (1969) e muitos outros pesquisadores, os quais, a partir da década de 1970, têm se dedicado à análise das estruturas e inter-relações estabelecidas entre pesquisadores a partir das publicações, que são a base dos estudos de mapeamentos da ciência, que tiveram destacado desenvolvimento em décadas posteriores (ZUCCALA, 2006).

No final da década de 1970, proliferam os trabalhos e pesquisadores, consolidando-se a Bibliometria como uma disciplina científica com diversas sub-áreas ou especialidades. Nesse sentido, cabe ressaltar, o surgimento, em 1979, do periódico *Scientometrics*, primeira publicação especializada nos estudos bibliométricos (TAGUE-SUTCLIFFE, 1992).

Os anos 1980 caracterizam-se pela aplicação de novas técnicas estatísticas para a análise dos dados, além da Bibliometria consolidar-se academicamente nas universidades, com a criação de instituições e grupos de pesquisa especializados. Exemplo disto é o grupo do *Center for Science and Technology Studies* (CWTS), da Universidade de Leiden, em Leiden, na Holanda ou o *Information Science and Scientometrics Research Unit* da Academia Húngara de Ciências, em Budapeste, na Hungria.

Uma destacada contribuição proposta pelo CWTS foi uma metodologia para a avaliação de instituições como instrumento para orientar sua política científica, tornando-se um padrão. Mediante o projeto denominado *The Leiden Science*

Indicators Projects, o CWTS desenvolveu um novo sistema de financiamento da pesquisa que envolvia todos os níveis de um sistema universitário (universidade, faculdades e departamentos) a partir de indicadores bibliométricos (MOED; VAN RAAN, 1998).

No caso específico da Faculdade de Medicina da Universidade de Leiden, o processo se baseou em rodadas de discussão iniciais onde foram acordados os indicadores que iriam a ser empregados. Estes não se limitaram à produção publicada na base ISI (*WoS*), abrangendo diferentes resultados, como teses e livros, assim como *inputs* do sistema (financiamento externo e recursos humanos). Os indicadores obtidos foram considerados por parte do Decano da Faculdade de Medicina de grande utilidade para a concessão e acompanhamento de projetos de pesquisa (MOED; VAN RAAN, 1998).

De outras instituições também surgiram contribuições de grande relevância, como o *Science Policy Research Unit*, da Universidade de Sussex, na Inglaterra. Martin e Irvine (1983) propuseram um esquema *input-output* para avaliar nove centros de radioastronomia no mundo, analisando os elementos humanos e financeiros, a produção científica e a mensuração das citações para determinação do impacto. Assim, o estudo verificou a aplicabilidade dos métodos bibliométricos a nível institucional, destacando sua utilidade na tomada de decisão, sobre tudo na hora de distribuir recursos. Este esquema, durante a década de 1980 e início dos noventa, foi amplamente utilizado em diversos trabalhos (BECK; GASPAR, 1991; DELGADO; RUSSELL, 1992).

Contribuições como estas revelam como a Bibliometria foi expandindo seu campo de atuação, a uma menor escala, demandando um refinamento no tratamento da informação e ao desenho de indicadores mais objetivos e mais variados.

Também durante a década de 1980, aparecem novas técnicas de análise dos textos científicos, em especial, a metodologia desenvolvida pela *École Nationale Supérieure des Mines de Paris*, na França, que propôs, em 1983, a análise de palavras associadas (*Co-words Analysis*), baseada na ocorrência conjunta de palavras nos textos e sua representação por meio de diagramas e redes, e sua ponderação através de diferentes indicadores (CALLON *et al.*, 1995).

Outra aplicação prática da análise de *co-words*, como a realizada com a produção científica em arqueologia (RUIZ-BAÑOS, 1997), é sua capacidade de estabelecer os ciclos de vida da terminologia ou os mecanismos de tradução-translação. Esse método, entretanto, encontra maior uso aplicado à vigilância tecnológica em disciplinas técnico-científicas (BAILÓN-MORENO, 2003).

Ao longo da década de 1990, os estudos do campo são favorecidos devido ao surgimento da internet, que possibilitou a consulta às bases de dados de textos na íntegra ou parcialmente. No caso de textos científicos, as bases de dados *Medline* ou o *Science Citation Index*, por exemplo, facilitaram a acessibilidade, descarga, tratamento e manipulação dos dados. O massivo processamento da informação tornou possível desenhar indicadores nacionais mais robustos, com ampla margem de comparação, reduzindo o tempo e custo (MOED *et al.*, 1995).

Dentre os trabalhos que utilizam essas bases, podemos mencionar o realizado por Jiménez-Contreras (1996), dedicado à produção científica da Faculdade de Ciências, Medicina e Farmácia da Universidade de Granada no período de 1975 a 1987, e o conduzido por Moya-Anegón e Rodriguez (2006), que avaliam a produção da Universidade de Granada no *Science Citation Index*, empregando indicadores do fator de impacto.

A internet também favoreceu o nascimento da Webmetria, uma nova vertente do campo da Informetria, que tomou para si alguns conceitos e aplicações desenvolvidos pela Cientometria, como a análise de citações aplicadas aos links ou os rankings de visibilidade das instituições acadêmicas na web (INGWERSEN; BJONEBORN, 2004).

Nos anos 2000, surgem novas bases de informações que catalogam a produção científica e tecnologia, como *Scopus*, que indexa quase 18.000 títulos de mais de 5.000 editores internacionais envolvendo ciência, tecnologia, medicina, ciências sociais, artes e humanidades, e proporciona ampla cobertura global, pois mais da metade de seu conteúdo procede de Europa, América Latina e a região Ásia-Pacífico (ELSEVIER, 2010). A *Scopus* surge como uma alternativa a base da Thomson Reuters, *Web of Science*, originalmente fundada por Garfield. Além dessa base, surgem novos motores de busca na internet como *Scholar Google*, que se propõe a ser uma base informacional mais dinâmica (NOTESS, 2005).

As formas mais recentes de coletar dados e mensurar o desempenho da ciência têm caminhado em paralelo com algumas mudanças importantes relacionadas aos hábitos de comunicação entre os cientistas, como, por exemplo, a consolidação das revistas *on-line*, que colocam à disposição dos pesquisadores uma grande quantidade de periódicos com textos completos, e também o surgimento de novas tipologias documentais na web, como os blogs científicos ou os repositórios wiki, entre outras. Todas essas mudanças estão criando novos modos de comunicar a ciência, o que, certamente, levará a modificações na forma de investigar os diferentes aspectos do conhecimento, podendo contribuir para a geração de novos indicadores bibliométricos e de novas unidades de análises.

Um aspecto também a considerar é o fato que a tecnologia atual vem facilitando, cada vez mais, o desenvolvimento de ferramentas que permitem a visualização do conhecimento científico a partir de diferentes dimensões. É sobre este assunto que trata o seguinte Capítulo.

3 DOMÍNIOS DA CIÊNCIA E VISUALIZAÇÃO DA INFORMAÇÃO

“Aprender a ver é a mais longa aprendizagem de todas as artes”

(Jules de Goncourt)

3.1 CONSIDERAÇÕES PRELIMINARES

Embora a ideia de que a atividade científica poderia ser visualizada ou mapeada tenha sido inicialmente apontada por Bush (1945b) e Bradford (1948), o primeiro a articular essa necessidade foi Doyle (1961). Ele ressaltou a importância de sistemas computadorizados para produzir mapas similares aos gerados pelo cérebro, indicando como estes mapas podem ser projetados em espaços multidimensionais e comentados sobre os objetivos e a forma de como construí-los.

Possibilitar a visualização de um campo científico tem sido um objetivo perseguido durante muito tempo. Solla Price (1965) demonstrou pioneiramente que os padrões de citação usados pelos autores de artigos científicos definem as frentes de pesquisa, podendo ser aproveitados para delinear uma topologia que reflita a estrutura da produção científica de determinado campo. Small (1973) deu significativo avanço à construção de mapas, ao propor, de forma independente, a ocorrência de pares de referências citadas conjuntamente em uma terceira fonte (cocitação de documentos) como variável de vínculo da produção científica.

Os mapas da ciência, nos quais são mostradas todas as especialidades das ciências naturais (SMALL; GRIFFITH, 1974), usando como fonte de informação o *Science Citation Index (SCI)* e a cocitação como variável de relação ou vínculo, tornaram-se fundamentais para desenvolvimento da representação de campos científicos. O destaque principal da metodologia utilizada por Small e Griffith (1974) foi a possibilidade de identificar grupos de documentos que tinham interesses intelectuais comuns, indicando, assim, que a ciência é uma rede de especialidades interconectadas.

Já na década de 1990, as tecnologias da informação e comunicação facilitaram o acesso às bases de dados, tornando possível a descarga, o tratamento e o processamento de grandes volumes de dados, permitindo, dessa forma, a geração de novos, mais complexos e potentes indicadores bibliométricos.

Moed *et al.* (1995), revisando os métodos de avaliação da atividade científica, constatam que, no início da década de 1990, o desenvolvimento e a utilização dos procedimentos de avaliação bibliométrica, tanto a análise bibliométrica da produção científica como os mapas da ciência, ocorriam separadamente. Hoje, o uso combinado de ambas as técnicas para análises de campos científicos representa aperfeiçoamento e sofisticação dos métodos bibliométricos assim como um aumento da demanda desses indicadores para a mensuração da produção científica.

O desenvolvimento de metodologias para a comparação da produção científica no sistema global da ciência, nos níveis nacional e internacional, tem sido objeto de numerosos estudos (VAN RAAN, 2004; VAN LEEUWEN, 2004; GLÄNZEL; SCHUBERT, 2004; dentre outros).

Paralelamente, os estudos de análise de domínio vêm se consolidando. Eles incluem desde análises com enfoque holístico da atividade científica até os estudos de redes sociais, que se debruçam sobre a estrutura de relações que compõem o fluxo de informação entre os diferentes agentes envolvidos.

3.2 ANÁLISE DE DOMÍNIO

Estudos sobre análise de domínio são comuns na engenharia de *software*. Nesta área, os trabalhos sobre essa temática são estritamente vinculados à reutilização de *software*, procurando a construção de novos sistemas de informação a partir de componentes, especificações ou desenhos criados no passado (BERARD, 1992). Nesse sentido, o termo análise de domínio tem origem no início dos anos 1980, com o trabalho de James Neighbors, quem o definiu como “*a atividade que consiste em identificar os objetivos e operações de tipos de sistemas similares, dentro de um domínio de problema particular*” (NEIGHBORDS, 1981). Outra definição clássica é a de Pietro-Diaz, para quem a análise de domínio é “*o processo pelo qual a informação utilizada no desenvolvimento de sistemas de softwares é identificada, capturada e organizada com a finalidade de torna-la reutilizável na criação de novos sistemas*” (PIETRO-DIAZ, 1990, p.46).

Dentre as principais características presentes na análise de domínio, Pietro-Diaz (1990) destaca: (a) Domínio do problema, o qual compreende os itens de informação relacionados ao contexto do mundo real, inter-relacionados e que desperta o interesse de certa comunidade, abrindo possibilidades para a realização de processos baseados em conhecimento para a identificação e a aquisição de informações; (b) Modelo do domínio, o qual define entidades, operações, eventos e relações que abstraem similaridades e regularidades em um determinado domínio, formando uma arquitetura de componentes comuns às aplicações analisadas e criando modelos que tornam possível identificar, explicar e prever fatos difíceis de serem percebidos diretamente e; (c) Análise e modelagem do domínio, o qual compreende um conjunto de atividades cujo propósito é reduzir a complexidade da nossa percepção de um determinado domínio.

Além de ser estudada pela engenharia de *software*, a análise de domínio têm sido objeto de pesquisa em outras áreas do conhecimento (BEGHTOL, 1995). No âmbito da Ciência da Informação, um particular enfoque de análise de domínio foi proposto e introduzido por Hjørland e Albrechtsen (1995). Para estes autores, a análise de domínio compreende o entendimento das áreas do conhecimento por meio da análise das comunidades de discurso. Nesse sentido, caracteriza-se como uma perspectiva de análise que propõe que a melhor maneira de se entender um domínio de conhecimento é por meio da análise das comunidades de discurso de onde esse conhecimento se origina, já que a organização do conhecimento, sua estrutura, os padrões de cooperação, a linguagem, as formas de comunicação e os critérios de relevância de um determinado domínio são reflexo dessas comunidades e do papel que desempenham na sociedade (HJORLAND; ALBRECHTSEN, 1995).

Domínio, nesse contexto, pode ser entendido como a designação dada a uma comunidade de discurso, vinculada a um âmbito qualquer em que se desenvolve uma certa atividade (WENGER, 1998). Um domínio pode compreender um conjunto de atores que compartilham algo em comum, assim como o emaranhado de relações que entre estes se estabelece. No âmbito da atividade científica, *domínio* compreende uma disciplina ou uma especialidade científica. Os integrantes da comunidade que lhe é associada compartilham objetivos comuns, um mesmo corpo de conhecimentos especializados, uma linguagem específica, mecanismos de

interlocução semelhantes, usam os mesmos meios de comunicação, tais como periódicos científicos, eventos especializados etc. (McCain *et al.*, 2006).

Em domínios científicos setoriais e institucionais, como é o caso de universidades, as comunidades compartilham, entre outros aspectos, um mesmo contexto político, social e cultural. Isto faz com que, além da universalidade das formas fundamentais do pensamento e práticas disciplinares, cada instituição configure seu próprio estilo de desenvolver suas atividades, em especial a atividade científica, em função das peculiaridades das práticas que são condicionadas pelo contexto na qual estas se desenvolvem (Vessuri, 1995).

Dentre os aspectos mais relevantes da perspectiva de análise de domínio proposta por Hjørland e Albrechtsen (1995), podemos destacar:

- A visão holística do conhecimento como processo e produto social e cultural.
- A concepção e enfoque metodológico social-coletivista, em detrimento do cognitivo-individualista.
- A tentativa de compreender as características, tanto explícitas como implícitas, do comportamento da informação e comunicação.
- A análise focada na comunicação científica, nas publicações, nas disciplinas e especialidades científicas, nas estruturas do conhecimento e nos paradigmas.

3.2.1 Comunidades de discurso: foco da análise de domínio

As comunidades de discurso referem-se a um conjunto de atores sociais que compartilham uma visão de mundo e apresentam determinadas estruturas individuais de conhecimento, preferências, critérios de relevância e estilos cognitivos particulares, em manifesta inter-relação entre as estruturas de domínio e o conhecimento individual (Hjørland; Albrechtsen, 1995). Isto é, entre os níveis social e individual.

Assim sendo, ações e elementos que envolvem os atores (tais como a organização do conhecimento, a estrutura comunitária, os padrões de cooperação, a linguagem – conceitos e significados – e as formas de comunicação, os sistemas e

as necessidades de informação e os instrumentos) configuram e diferenciam cada comunidade ao constituir manifestações/reflexos dos objetivos de trabalho das mesmas e de seu próprio papel na sociedade (HJORLAND, 2002).

Esses aspectos inserem-se num processo comunicativo ordenado e limitado pela estrutura conceitual, o marco institucional e o governo do discurso. Este discurso resulta ser elemento conector da linguagem (tipo de linguagem específica desse domínio ou fala técnica transformada em fala cotidiana do grupo), da ação e dos membros da comunidade no processo construtivo, e reafirmador da identidade da comunidade, além do tempo e do espaço geográfico, proporcionado assim coerência e objeto social comum à comunidade em franca ligação entre o âmbito afetivo e cognitivo em nível individual e social (HJORLAND, 2002).

Mas que ferramentas, métodos ou técnicas são necessárias para levar adiante esse tipo de análise? Hjørland (2002) propõe onze métodos para a análise de domínio na Ciência da Informação:

- (1) Produção de guias de literatura;
- (2) Elaboração de classificações especiais e tesouros;
- (3) Indexação e recuperação da informação;
- (4) Estudos empíricos de usuários;
- (5) Estudos bibliométricos;
- (6) Estudos históricos;
- (7) Estudos de documentos e estilos;
- (8) Estudos epistemológicos e críticos;
- (9) Estudos terminológicos, linguagens para propósitos específicos (LSP), semântica de bases de dados e estudos de discurso;
- (10) Estruturas e instituições de comunicação científica;
- (11) Cognição científica, conhecimento especializado e inteligência artificial.

Para o autor, estes métodos podem ser aplicados em outras áreas para análise de todo tipo de domínios, proporcionando uma visão holística e objetiva dos mesmos.

Dos onze métodos propostos por Hjørland (2002), cabe destacar os estudos bibliométricos por constituir-se em uma forma de se iniciar uma análise de domínio, sendo que as outras dez técnicas podem ser utilizadas como apoio desta, complementando, assim, a visão holística (MOYA-ANEGÓN; HERRERO SOLANA, 2001; VARGAS QUESADA, 2005). No presente trabalho, focalizamos nossa atenção no método bibliométrico.

3.2.2 Bibliometria e análise de domínio: algumas aplicações

Considerando que um domínio é onde se desenvolve uma atividade, seja ela no nível temático ou institucional, a diferenciação e separação das partes do domínio em questão proporcionariam a possibilidade de conhecer seus princípios ou elementos.

Segundo McCain *et al.* (2006), para que um domínio científico se considere como tal, o grupo de indivíduos que o compõe deve reunir uma série de características: objetivos comuns, um corpo de conhecimento especializado, mecanismos de intercomunicação e participação, um gênero literário (um periódico científico, por exemplo), vocabulário especializado, dentre outras.

Os enfoques que permitem um pesquisador realizar uma análise de domínio são diversos. Por exemplo, a partir da análise das citações é possível estudar os padrões e frequência das citações realizadas e recebidas pelos autores, os periódicos, as disciplinas de pesquisa, etc. (SPINAK, 1996). Os dados que surgem destas técnicas podem ser direcionados para outras metodologias que vão além das contribuições da análise bibliométrica, proporcionando, por exemplo, a capacidade de delimitar a estrutura de relações existentes em uma determinada disciplina (MOYA-ANEGÓN; HERRERO SOLANA, 2001). O conhecimento das relações que se estabelecem entre os itens estudados possibilita uma carga cognitiva maior que a simples enumeração quantitativa de um conjunto de indicadores.

Vale lembrar que, segundo Hjørland (2002), a análise de domínio é:

[...] o estudo do campo (domínio de conhecimento) como um pensamento ou comunidade de discurso. Centra-se em temas tais como a organização do conhecimento, estrutura, padrões de cooperação, formas de linguagem e comunicação, sistemas de informação e critérios relevantes, como uma maneira de entender essa comunidade (HJORLAND, 2002, p.34).

A análise de domínio pode ser processada, então, utilizando-se as ferramentas bibliométricas ou as técnicas de representação da informação, que proporcionam uma imagem “caleidoscópica” da atividade ou domínio, objeto de estudo. À luz desta análise, floresce uma série de relações que conformam a estrutura do campo científico. O conjunto de conexões que mostra este tipo de análise pode não ser perceptível para o grupo da comunidade científica, mas a imagem que estabelece tais associações, representada com as técnicas acima mencionadas, pode trazer à tona relações, até então, desconhecidas e/ou negligenciadas.

Hjorland (2002), ao se referir aos estudos métricos, destaca que a Bibliometria pode ser usada de diversas formas como ferramenta e método para a análise de domínio como, por exemplo, na criação de mapas para a visualização de disciplinas científicas ou domínios institucionais mediante as técnicas de cocitação e coautoria.

Estudos desta natureza têm sido desenvolvidos por diversos autores, como White e McCain (1998), que estudaram o campo da Ciência da Informação através da análise comparativa da cocitação de autores em três períodos, 1972-1979, 1980-1987 e 1988-1995; Persson (1994) na Universidade de Umea, que seguindo a metodologia utilizada por Small e Griffith (1974), analisou artigos publicados no *Journal of the American Society for Information Science and Technology (JASIST)*, no período 1986-1990, para identificar principais áreas da Bibliometria e Recuperação da Informação; Borner *et al.* (2003) na Universidade de Indiana, que apresentam uma revisão das técnicas para análise, visualização e posicionamento espacial da informação e; Moya-Anegón *et al.* (2006) na Universidade de Granada, que desenvolvem uma análise da estrutura científica espanhola baseada na análise de cocitação de autores e de revistas; entre outros.

Entre os países latino-americanos, o desenvolvimento deste tipo de análise é dificultado pela carência de fontes de dados que registrem, com abrangência e

confiabilidade, a produção científica desses países. Daí que a maioria dos estudos realizados pelos países da região são feitos a partir das publicações registradas em bases de dados internacionais, particularmente do *Institute for Scientific Information (ISI)*. Nesse sentido, os países latino-americanos com maior presença internacional são, em ordem decrescente e quantidade de contribuições, Brasil, Argentina, México, Chile e Venezuela (RICYT, 2009).

Uma característica comum desses países é que a maioria das contribuições científicas é oriunda do setor universitário público e do setor governamental (RICYT, 2009). O peso que estes setores têm para o desenvolvimento científico e tecnológico da região aumentou significativamente o interesse pelos estudos da produção científica de domínios institucionais, especialmente, das universidades.

Exemplos disso são os trabalhos realizados por DeArenas e Cronin (1989) sobre as contribuições das instituições de educação superior para o desenvolvimento das ciências da saúde no México; Delgado e Russell (1992) sobre a visibilidade internacional da literatura publicada pelos pesquisadores da Universidade Nacional Autônoma do México (UNAM); o estudo comparativo realizado por Krauskopf (1992) sobre a produção científica das universidades chilenas; a análise feita por Russel *et al.* (1992) sobre a participação dos pesquisadores da UNAM na literatura biomédica internacional; a pesquisa realizada por Krauskopf *et al.* (1995) sobre o desempenho científico da Faculdade de Ciências Biológicas da Universidade Católica do Chile; a análise comparativa de Gutiérrez (1998) sobre os padrões de publicação e impacto da produção científica dos Institutos e Centros de Pesquisa ligados à UNAM e; a análise de domínio da Universidade Nacional de Mar del Plata realizada por Solana (2001), dentre outros.

3.3 ANÁLISE DE REDES

A origem teórica da análise de redes costuma ser identificada no ponto de convergência entre a Psicologia Social e a Antropologia (SCOTT, 2000). Pelo lado

da Psicologia Social, a Sociometria⁷ é considerada o principal ponto de partida da análise de redes, uma vez que apresenta a representação gráfica das relações entre atores por meio do conceito de estrela sociométrica, na qual se evidenciam as escolhas que fazem os membros de um grupo (estrutura). Tal representação permite, assim, uma visualização rápida da estrutura em seu conjunto.

Pelo lado da Antropologia, o grupo denominado de Manchester, liderado por Max Gluckman e cujos principais representantes foram Barnes, Mitchell e Bott, é considerado uma das bases fundamentais para o desenvolvimento ulterior dessa abordagem teórica (SCOTT, 2000).

Outros pilares antropológicos, nos quais se apoia a teoria de redes, tem sido os trabalhos de Elton Mayo e Lloyd Warner, desenvolvidos na década de 1920, onde observaram como o entorno determinava e influenciava as condutas dos atores (FREEMAN, 2004). Esses trabalhos proporcionaram ampla informação sobre os processos de agrupamento (formação de grupos), que posteriormente foram aperfeiçoados com o desenvolvimento dos processos algébricos para a compreensão dos processos de posicionamento na estrutura⁸.

Borgatti *et al.* (2009), ao descreverem a história e consolidação dos estudos de redes sociais, destacam o trabalho pioneiro do psiquiatra Jacob Moreno, que, em 1932, mapeou as redes de relações em uma escola em Hudson, EUA. Nesse trabalho, o autor utilizou técnicas sociométricas, que permitiram visualizar graficamente as relações sociais em um determinado conjunto de indivíduos.

Depois de Moreno, inúmeros autores se dedicaram ao tema rede social, seja para melhor caracterizá-lo teórica e conceitualmente, seja para investigá-lo empiricamente.

3.3.1 Aspectos conceituais da análise de redes sociais

Dos autores mencionados merece destaque Barnes (1954) que é, por muitos, considerado o primeiro a usar, de forma sistemática, o conceito de rede ao estudar

⁷ Estudo dos traços psicológicos de indivíduos de determinado grupo social, mediante a análise matemática de preferências e rejeições para deduzir modelos de inter-relações (MORENO, 1962).

⁸ Avanços nesse sentido foram realizados pelo grupo formado em Harvard, onde se destacam: Berkowits, White e Granovetter (Cf. SCOTT, 2000).

uma pequena comunidade norueguesa de pescadores e fazendeiros. Barnes propôs uma primeira definição do termo:

A imagem que tenho é de um conjunto de pontos alguns dos quais estão unidos por linhas. Os pontos da imagem são pessoas, ou às vezes grupos, e as linhas indicam que as pessoas interagem mutuamente. Podemos pensar que o conjunto da vida social gera uma rede deste tipo⁹.

Todo o arcabouço de conhecimento gerado em torno da temática “redes sociais” favoreceu sua disseminação, tornando-a um fenômeno usualmente compreensível, com potencial de retratar situações comuns, incluindo a interação social (de indivíduos e instituições), relações, conectividade, colaboração, ação coletiva, confiança e cooperação (PROVAN *et al.*, 2007).

Rede social é definida como um conjunto finito de elementos (atores) e de relações estabelecidas entre eles, onde a presença destas últimas é uma característica crítica e determinante da rede, já que o importante nas redes sociais não é o indivíduo, mas sua estrutura, que é definida como o conjunto de indivíduos e suas conexões (WASSERMAN; FAUST, 1994).

A representação de uma rede deve conter um conjunto de objetos ou sujeitos (em termos matemáticos, nós) e descreve as relações que existem entre eles. A rede mais simples contém pelo menos dois objetos ou sujeitos e uma relação entre eles. Considerando A e B como pessoas ou instituições e que entre eles há um vínculo que os une (por exemplo, ter trabalhos conjuntos), pode-se dizer que existe uma rede, neste caso uma rede social, uma vez que se trata de uma relação entre indivíduos.

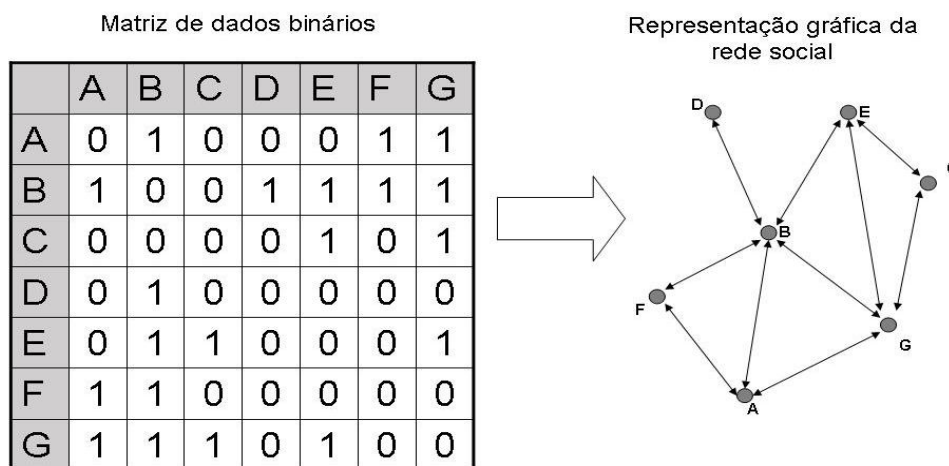
Para visualizar estas relações, as análises de redes utilizam a técnica de análise de correspondência, que se mostra adequada para analisar tabelas de duas ou múltiplas entradas, levando em conta medidas de correspondência entre linhas e colunas.

Para isto, elaboram-se matrizes, as quais podem ser quadradas ou retangulares em função do número de linhas e colunas e podem ser também

⁹ Texto original: “The image I have is of a set of points some of which are joined by lines. The points of the image are people, or sometimes groups, and the lines indicate which people interact with each other. We can of course think of the whole of social life as generating a network of this kind” (BARNES, 1954, p. 237).

idênticas, quando os nós são os mesmos para as linhas e colunas. Em função do tipo de fluxos, as matrizes podem ser normais (fluxos unidirecionais e bidirecionais) ou simétricas (unicamente fluxos bidirecionais). A Figura 2 apresenta um esquema simplificado que reproduz a análise de rede social.

Figura 2: Esquema da análise de redes sociais



Fonte: Adaptado de IRIBARREN-MAESTRO, 2006

Destaque-se que a análise de redes sociais proporciona um novo método para o exame de processos. Sua principal diferença em relação às análises tradicionais está no fato de não se basear numa análise individualista das características dos atores, senão, ao contrário, sustenta-se na informação das relações dos atores que compõem a estrutura da rede (MOLINA, 2001).

Este tipo de análise tem se configurado nas últimas décadas como uma área do conhecimento interdisciplinar com entidade própria, tendo-se nutrido das contribuições de diferentes pesquisadores e escolas que utilizaram as propriedades das interações sociais e os diferentes tipos de redes (NEWMAN, 2003).

3.3.1.1 Medidas de Análise de Redes Sociais

Segundo Knoke e Kuklinsky (1982), as medidas de redes sociais identificam as propriedades emergentes dos sistemas sociais que não podem ser avaliadas pela

simples agregação de atributos dos atores individuais. Além disso, tais propriedades emergentes podem impactar significativamente tanto o desempenho do sistema social quanto o comportamento dos atores de uma determinada rede.

O especialista em análise de redes sociais pode utilizar uma combinação das diferentes medidas de rede a partir do problema empírico estudado. Algumas das medidas de análise de redes sociais são apresentadas a seguir. As definições seguem as ideias originais, considerando o princípio que medidas com maior valor determinam vértices mais centrais (WASSERMAN; FAUST, 1994; FREEMAN, 2004). Seja $G = (V, E)$ um grafo onde $|V| = n$ e $|E| = m$. Seja $v \in V$ um vértice (nó) do grafo (NEWMAN, 2003).

Centralidade de grau (Degree Centrality)

A centralidade de grau é a forma mais simples de medir a importância de um vértice (nó, atores) no grafo e tem o mesmo valor numérico do grau do nó. Conta o número de arestas que cada ator tem, tratando diferentemente o grau de entrada e o grau de saída no caso de redes dirigidas (WASSERMAN; FAUST, 1994).

Esta medida esta baseada na seguinte ideia: vértices (atores) que possuem grau maior podem estar em uma posição mais privilegiada da rede. A centralidade de grau de um vértice v é definida como:

$$C_D(v) = \text{grau}(v)/(n-1)$$

Centralidade de proximidade (Closeness Centrality)

A centralidade de proximidade de um ator mede o quanto o vértice, que representa o ator, está próximo de todos os demais da rede (FREEMAN, 2004).

É definida a partir da média das distâncias geodésica do vértice v e todos os outros vértices alcançáveis por ele (NEWMAN, 2003):

$$C_C(v) = (n-1) / \sum_{\{t \neq v \in V\}} d_G(v, t)$$

Vértices com maior centralidade de proximidade assim definida ocupam uma posição próxima do centro do grafo.

Centralidade de intermediação (Betweenness Centrality)

Para medir o tráfego que passa em um dado vértice, é usada a medida chamada centralidade de intermediação (*betweenness centrality*), que mede o quanto um vértice está no caminho entre outros vértices (FREEMAN, 2004). Considera todos os caminhos mínimos entre pares de vértices de uma rede, e os vértices que pertencem a um número maior de caminhos mínimos são os que possuem maior intermediação (NEWMAN, 2003).

A centralidade de intermediação de um vértice v é definida como:

$$C_B(v) = \sum_{\{s \neq v \neq t \in V\}} \sigma_{st}(v) / \sigma_{st}$$

Onde $\sigma_{st}(v)$ denota o número de caminhos mínimos entre s e t , que passam por v e σ_{st} denota o número de caminhos mínimos entre s e t .

Centralidade de autovetor (Eigenvector Centrality)

Considerando um grafo G onde os vértices estão numerados de 1 até $n = |V|$ e seja $A = (a_{i,j})$ a matriz de adjacência do grafo G . Sendo x_i , $1 \leq i \leq n$, o *score* do i -ésimo vértice do grafo definido da maneira seguinte: $x_i = (\sum_{j=1, \dots, n} a_{i,j} x_j) / \lambda$, onde λ é o maior autovalor da matriz A . A centralidade do autovetor determina um *score* para cada vértice que é proporcional aos *scores* dos seus vértices vizinhos (NEWMAN, 2003). Dessa forma, a centralidade de autovetor é definida como:

$$C_A(v) = (\sum_{j=1, \dots, n} a_{i,j} x_j) / \lambda$$

Ou o componente do vetor de Perron da matriz A (autovetor associado ao maior autovalor, λ , da matriz de adjacência do grafo). Portanto, a centralidade do autovetor de um vértice é uma combinação linear das centralidades dos vértices que estão conectados a ele (FREEMAN, 2004).

3.3.2 Análise de redes sociais no âmbito de atividade científica

Considerando que a comunidade científica atual se estrutura cada vez mais em um sistema integrado, cooperativo ou de rede, análises de redes sociais podem ser úteis para identificar a natureza e intensidade das relações entre os membros da rede assim como a estrutura global da mesma através das características das interações entre os elementos (RODRIGUEZ, 1995). Quando realizadas no contexto dos estudos da Bibliometria, essas análises buscam a representação dos laços colaborativos a partir das informações contidas na produção científica, mais frequentemente a partir da comunicação em formato de artigos em periódicos (os denominados *papers*).

Esse tipo de comunicação não apenas apresenta uma nova contribuição em um dado assunto, mas revela as tendências e o dinamismo de uma atividade muitas vezes distante do grande público. Assim, a partir de uma série de informações bibliográficas, que são exclusivas do artigo em periódico (ou *paper*), é possível medir, estimar tendências, lideranças e carências e também visualizar diferentes aspectos da ciência, como as redes de colaboração.

3.3.2.1 Redes de colaboração a partir das informações de autoria

A base das análises de redes de colaboração entre indivíduos na ciência encontra-se nas informações de autoria, seja de pesquisadores que compartilham a autoria de uma publicação (coautoria), seja de pesquisadores autores que aparecem citados conjuntamente em uma mesma publicação (cocitação). Importante ter clareza que tal como qualquer outro estudo de rede social, as análises de coautoria e cocitação partem de informações sobre as relações existentes entre pares de autores.

Neste caso, gera-se uma matriz de valores quadrada ou simétrica, ou seja, com o mesmo número de linhas e colunas. Cada uma dessas linhas e colunas se inicia com os nomes dos autores recuperados nas respectivas consultas, produzindo uma coincidência recíproca e não direcional entre todos eles. A matriz gerada terá uma dimensão de N por N, onde N é o número de autores que compartilham autoria

na produção ou aparecem citados simultaneamente em uma publicação. Em ambos os casos, para evitar possíveis repetições do nome de cada autor no mesmo trabalho é necessária a eliminação dos valores da diagonal principal. Para a realização de cálculos ulteriores, análise e visualizações utilizam-se os valores absolutos de coautoria ou de cocitação que garantem a robustez da medida¹⁰.

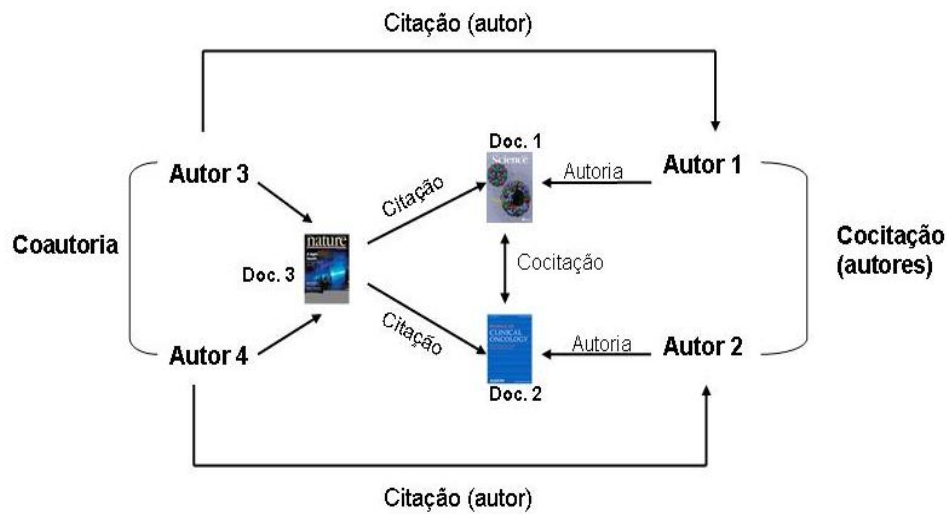
A Figura 3 permite visualizar os dois tipos de unidades de análise, a coautoria e a cocitação, utilizados nos estudos de redes de colaboração na ciência. Para a análise de coautoria, os autores 3 e 4 compartilham a autoria de uma publicação, o documento 3. Estes dois autores podem compartilhar muitas outras publicações e/ou cada um deles pode compartilhar a autoria de outras publicações com outros autores.

Assim, uma análise de rede de coautoria considerará o número de publicações que os pares de autores (autor 3 com autor 4; autor 3 com autor 5, etc.) aparecem juntos como autores. Importante mencionar que este tipo de análise pode também considerar não o autor, mas sua instituição de afiliação ou país de origem.

Para a análise de cocitação (de autores), a Figura 3 mostra que os autores 1 e 2 são citados simultaneamente no documento 3. Da mesma forma que na coautoria, este par de autores pode ser citado em conjunto em outros trabalhos, assim como cada autor individualmente pode ser citado com outros autores, formando novos pares de cocitações. Nessa análise, então, é considerado o número de publicações em que cada par de autores aparece citado conjuntamente.

¹⁰ Neste ponto, convém lembrar que são vários os autores que destacam que a utilização de medidas de normalização (correlação de Pearson, função cosseno etc.) produz distorções no momento de distribuir a informação nos espaços de informação (WHITE, 2003). Essas medidas de normalização são usadas para construir a matriz de proximidade no caso de dados que não são medidas de proximidade. Sendo uma matriz de co-ocorrência *per se* uma matriz de proximidade não requer nenhum tipo de tratamento que, se aplicado, ocasionará distorção dos dados, das análises e representações (LEYDESDORFF; VAUGHAN, 2006).

Figura 3: Redes de autorias e de papers



Fonte: Adaptado de RODRIGUEZ *et al*, 2010¹¹

Para Rodriguez *et al.* (2010), embora sejam semelhantes metodologicamente, os dois tipos de unidades de análise diferem quanto a sua natureza. Apesar das críticas e da pouca normatização sobre quem deve ser coautor de uma publicação, para estes autores as análises de rede a partir das coautorias indicam relações e laços sociais mais fortes entre os pesquisadores, onde existe um mínimo de conhecimento mútuo. O mesmo não ocorre para os estudos de rede a partir das autorias das citações (cocitações), uma vez que, segundo os autores, a citação de diferentes trabalhos não significa que seus autores se conheçam. De qualquer forma, há também de se reconhecer que deve existir alguma semelhança, temática ou de conteúdo, entre diferentes autores quando estes aparecem com grande frequência citados simultaneamente em outras publicações.

Sendo as análises de coautoria e cocitação os principais insumos utilizados para a identificação de redes de colaboração na ciência contemporânea, mencionamos, a seguir, algumas informações envolvendo essas temáticas.

¹¹Vale mencionar que na Figura 3, os autores utilizam o termo *co-citation* para designar o compartilhamento de informações (por exemplo, dois trabalhos têm uma referência em comum) entre dois documentos (1 e 2) citados simultaneamente no documento 3. Neste caso, o termo mais usualmente utilizado é *bibliographic coupling*, em português acoplamento bibliográfico. As diferenças entre *co-citation* e *bibliographic coupling* são amplamente descritas na literatura internacional, um exemplo é o trabalho de Garfield (2001).

Análises de coautoria

O trabalho de Smith (1958) foi pioneiro na análise de múltipla autoria em publicações na área de Psicologia. Com base no crescimento do número de publicações com mais de um autor no período de 1946 a 1957, Smith propõe que os artigos em coautoria sejam utilizados como dimensão aproximada da colaboração entre pesquisadores. Além disso, o autor relaciona esse aumento à nova forma de se estruturar da ciência:

A tendência de autoria múltipla em psicologia e na totalidade das ciências também pode se acelerar à medida que mais e mais recursos estão disponíveis para o desenvolvimento de pesquisas sobre os problemas mais complexos e mais difíceis que agora exigem atenção (SMITH, 1958, p.599)¹².

Solla Price (1965) verificou empiricamente as observações de Smith sobre o aumento da múltipla autoria em ciência. O autor notou que esta poderia ser identificada de diferentes formas e com regularidade no âmbito de comunidades informais de pesquisadores que se comunicavam, trocavam informações e experiências e publicavam formalmente seus resultados de pesquisa, evidenciando, dessa forma, uma tendência de aumento no número de trabalhos assinados em coautoria.

Já o estudo de Storer (1970) revelou que a colaboração é uma medida que varia significativamente nas diferentes áreas do conhecimento em função de suas características cognitivas e organizacionais, identificando um maior índice de cooperação nas ciências básicas e da natureza em relação às ciências aplicadas e sociais. Mas foram Frame e Carpenter (1979) os primeiros a usar a coautoria como indicador de colaboração, destacando os diversos tipos de atividade colaborativa, tais como, a participação de fontes de dados, a troca de ideias, estâncias em centros de pesquisa no exterior ou troca de artigos.

Na década de 1990, os estudos de coautoria ganharam nova dimensão: a identificação, apresentação e interpretação a partir do conceito de redes sociais.

¹² Texto original: "The trend toward multiple authorship in psychology and in the totality of the sciences may also accelerate as more and more funds become available to conduct research on the larger and more difficult problems that now demand attention" (SMITH, 1958, p. 599).

Esse novo direcionamento vem permitindo estabelecer a relação entre a estrutura social dos pesquisadores e o progresso do conhecimento científico assim como identificar de que forma se organizam as comunidades científicas (EATON *et al.*, 1999). Nesse sentido, estudos têm procurado identificar autores importantes em determinadas disciplinas (OTTE; ROUSSEAU, 2002), descrever as divergências entre diferentes comunidades científicas (MOLINA; DOMENECH, 2002), analisar as comunidades em relação ao gênero, situação geográfica e grau acadêmico (RIVELLINI *et al.*, 2006), caracterizar as redes de departamentos acadêmicos para determinar a estrutura interna e seus grupos (VANRAAN; PETERS, 1991).

Além destas abordagens, os estudos de redes na ciência, com base nas análises de coautoria, também se dedicam a investigar a rede mundial de colaboração entre países num determinado momento (PERSSON; MELIN, 1996) e aqueles que procuram mostrar a evolução dessa mesma rede em diferentes períodos cronológicos (GLÄNZEL, 2001; GLÄNZEL; SHUBERT, 2004) ou mesmo as redes institucionais de colaboração entre universidades (PERSSON *et al.*, 1997).

Análises de cocitação

Já nos estudos de cocitação foi Solla Price (1965) quem demonstrou pioneiramente que os padrões de citação usados pelos autores de artigos científicos podem definir as frentes de pesquisa, sendo aproveitados para delinear uma topologia que refletisse a estrutura da produção científica de determinado domínio. Small (1973) deu significativo avanço à construção de mapas da ciência, ao propor de forma independente a cocitação de documentos como variável de estudo na análise de citações da produção científica.

De acordo com Small (2003), os estudos de cocitação tiveram inicialmente como foco os documentos, identificando a precisão e concentração dos vínculos de cocitação entre documentos mais citados e, conseqüentemente, com mais robustez nos padrões obtidos ao compara-os com amostras históricas dos mesmos. Assim, os mapas da ciência, onde são mostradas as especialidades das ciências naturais e utilizam a cocitação como unidade de relação ou vínculo, tornaram-se fundamentais no desenvolvimento da representação de domínios científicos.

Os estudos de cocitação, a partir dos pares de autores, tiveram início na década de 1980, quando White e Griffith (1981), através da análise de cocitação de artigos extraídos da versão *on-line* da *Social SciSearch*, para o período 1972-1979 na área da *Information Science*, determinam diferentes grupos de autores, localizam esses grupos em relação a outros, mostram o grau de centralidade e periferia dos autores dentro do grupo, a proximidade dos autores dentro do grupo e em relação a suas fronteiras, e a posição dos autores em relação aos eixos do mapa. White e Griffith propõem, dessa forma, a análise de cocitação de autores como nova técnica para contribuir ao conhecimento da estrutura intelectual das disciplinas científicas, alcançando a mesma maior destaque no início da década de 1990.

Segundo McCain (1990a), este tipo de análise permite identificar a organização das comunidades de um conjunto de autores tal e como esta é percebida pelos citantes. Essa técnica tem sido usada para analisar determinadas áreas científicas como, por exemplo, a Macroeconomia (McCain, 1996), Genética (McCain, 1990b), Comportamento organizacional (PAISLEY, 1990), Matrimônio e família (BAYER *et al.*, 1990), Ciência da informação (McCain; WHITE, 1998), entre outras.

Fato é que estes estudos (tanto os de coautoria como os de cocitação) se diversificaram e passaram a constituir um novo sub-campo de conhecimento na Informetria, conhecido como *science mapping* ou *bibliographic mapping* (NOYONS *et al.*, 1999).

Em suma, pode-se dizer que a combinação do estudo sociológico das redes sociais, as pesquisas sobre redes científicas e a união de várias áreas de análises para a compreensão e representação das redes de colaboração (a partir das informações de autoria: coautoria e cocitação) constituem, atualmente, as bases de uma nova etapa de pesquisas onde a análise de domínio e de redes sociais, a teoria de grafos e a ciência da informação, juntamente com o desenvolvimento alcançado pelas tecnologias da informação e comunicação, tem proporcionado novas respostas para velhas questões.

3.4 VISUALIZAÇÃO DA INFORMAÇÃO CIENTÍFICA

A natureza complexa e multidimensional da informação científica exige a necessidade de complementar os métodos de análise com técnicas de visualização que permitam representar a realidade (KLOVDAHL, 1981).

A representação gráfica da informação para sua ulterior visualização tem se tornado uma atividade comum na maioria das disciplinas científicas nas últimas décadas (CROSBY, 1997). Entretanto, o uso combinado das representações gráficas com as tecnologias computacionais para conseguir uma adequada visualização da informação é uma tarefa relativamente recente, tornando-se um destacado objeto de estudo dos últimos anos.

O uso de imagens visuais como formas de representar a realidade é uma necessidade e característica comum a muitas áreas do conhecimento. No âmbito da Informetria, evidencia-se a procura por encontrar melhores formas de representação da informação, de interfaces entre os usuários e a literatura (WHITE; McCAIN, 1997).

Não obstante, a visualização não é o resultado implícito do ato de ver. Constitui, geralmente, atividade do comunicador visual que transforma dados abstratos e fenômenos complexos da realidade em mensagens visíveis, tornando possível ver com os próprios olhos dados e fenômenos que estão ocultos e que não são diretamente captados (COSTA, 1998). Ou seja, a visualização consiste em tornar visível para nossa mente aquilo que não é visível para nossos olhos.

O domínio de linguagens gráficas para visualizar efeitos invisíveis (informação imperceptível) configura uma nova ciência da comunicação visual, à qual Costa (1998) tem-se referido como a terceira linguagem (depois da imagem e do signo) e que tem sido estudada no âmbito da Ciência da Informação, com aplicações no desenvolvimento de interfaces baseadas em representações bidimensionais.

A visualização, portanto, é um processo de comunicação que se produz entre uma representação reduzida da realidade e quem a observa, a partir da qual é possível perceber através da visão, fatos e fenômenos da realidade multidimensional e mutante do mundo que de outra forma não seriam notados.

Por outro lado, a visualização da informação consiste na interseção da imagem, da palavra, do número e da arte, por meio da escrita e tipografia, da gestão de grandes quantidades de dados e da análise estatística, dos gráficos, e da sua distribuição e cor. Tudo isto com a finalidade de conseguir uma representação gráfica e reduzida da realidade multidimensional e em constante transformação que comunique ao observador, fatos e fenômenos de determinada realidade, os quais sem sua mediação seriam imperceptíveis e passariam totalmente despercebidos. (McCORMICK *et al.*,1987).

Destaque-se que a visualização da informação não é uma prática nova na Ciência da Informação. Proposta há quase sete décadas por Bush (1945b) e implementada a partir da década de 1960 por Garfield *et al.* (1964), esta prática tem sido utilizada por cientistas deste campo para evidenciar e divulgar a essência e estrutura da ciência.

Contudo, é a partir da década de 1990, com o surgimento de modernos métodos de recompilação da informação e novas técnicas para análise, visualização e posicionamento espacial da informação, que o trabalho de Borner *et al.* (2003) discute exemplarmente, começam a florescer estudos baseados em técnicas de visualização da estrutura do conhecimento científico.

Alguns exemplos são os trabalhos de Braam *et al.* (1991), que propõem o uso combinado da cocitação e análises de *co-words* para identificar a estrutura e aspectos dinâmicos de determinada área; Lin *et al.* (1991), que representam de forma gráfica e esquemática as relações semânticas entre documentos; White e McCain (1997), que indicam como tendência da área a combinação do uso de visualizações gráficas computarizadas com a recuperação da informação, entre outros.

Não obstante os anos de existência, ainda hoje a visualização da informação se encontra em fase de evolução e de desenvolvimento. Prova disso é sua aplicação em novas frentes de pesquisa, como por exemplo, na análise de domínio.

3.4.1 Visualização da informação e análise de domínio: geração de mapas da ciência

Segundo Chen *et al.* (2001), a proliferação das técnicas de visualização da informação tem levado a análise de domínio a configurar-se como uma das novas frentes de pesquisa que vem ajudando a revelar a essência do conhecimento científico. Por exemplo, ela tem sido utilizada para identificar o perfil temático de revistas científicas (PERSSON, 1994; MOYA-ANEGÓN, 2001; CHEN, 2002), para explorar e visualizar as bases intelectuais e a estrutura de relações de domínios geográficos (MOYA-ANEGÓN *et al.*, 2004) e institucionais (HERRERO-SOLANA, 2001; GUERRERO-BOTE, 2002; MIGUEL *et al.*, 2008), para explorar e ter acesso aos conteúdos das bibliotecas digitais (CHEN, 1999), assim como para estudar a evolução dos padrões de citação em patentes (CHEN; HICKS, 2004), entre outros.

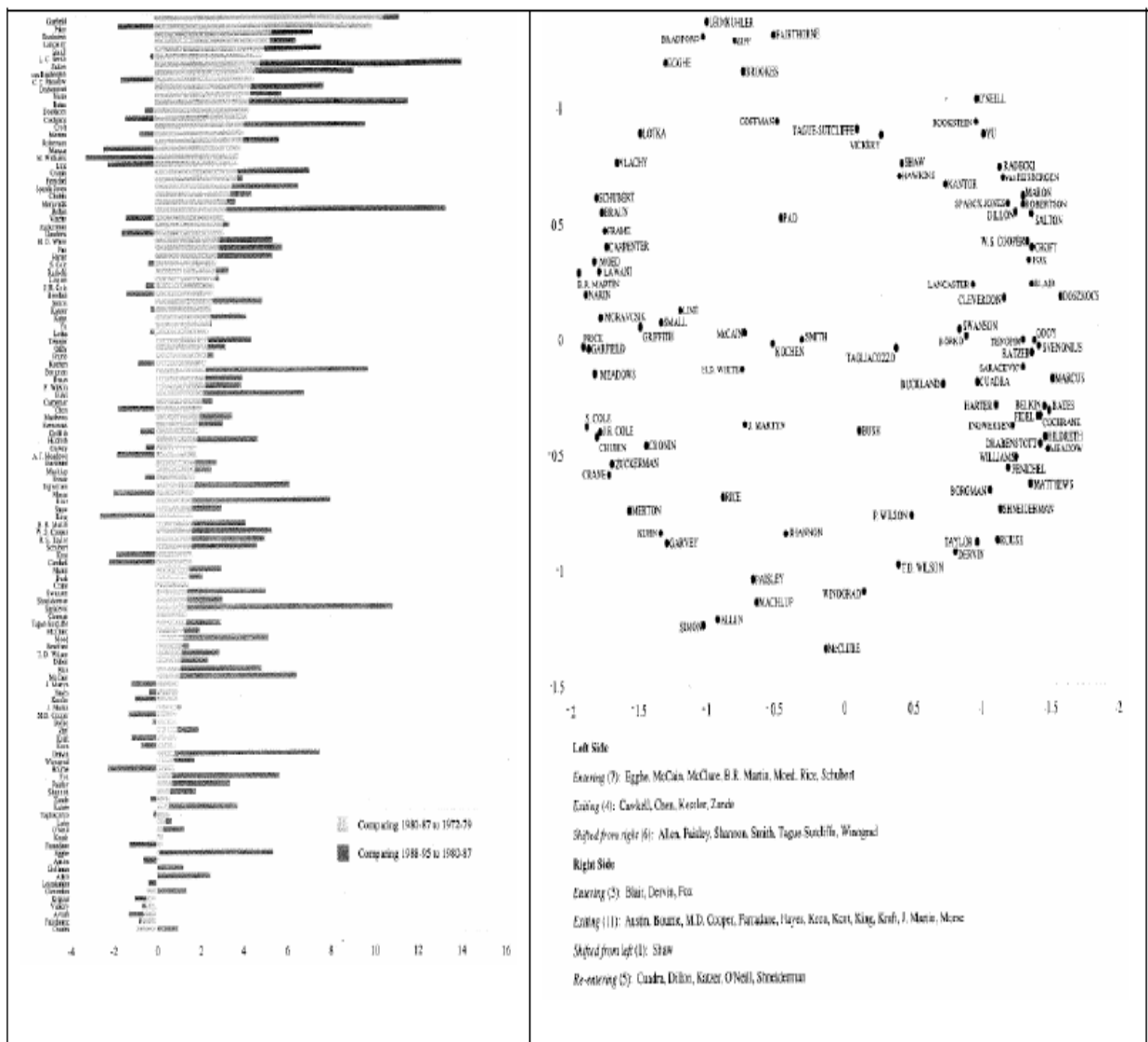
Há uma profunda conexão entre visualização de domínios e o que Hjørland e Albrechtsen (1995) denominam de análise de domínio. A visualização de domínios pode proporcionar técnicas de apoio para análise de domínio, especialmente nas áreas de conhecimento multidisciplinar e naquelas áreas que mudam e avançam rapidamente (BORNER *et al.*, 2003).

A partir da década 1990, verifica-se maior aproximação dos estudos de visualização e a bibliometria com foco nos estudos de conceitos como campos, disciplinas, áreas de especialização, estrutura do conhecimento e domínios do conhecimento (SCHLUETER; BORLUND, 2004), na tentativa de melhorar as análises da complexidade das inter-relações do conhecimento científico.

Em 1994, Garfield introduz o chamado mapeamento longitudinal, utilizando séries de mapas sequenciais para identificar os avanços temporais do conhecimento científico.

Em 1998, White e McCain desenvolvem metodologicamente uma análise do domínio da Ciência da Informação, baseada na análise de cocitação de autores. A Figura 4 mostra as projeções gráficas de barras para a comparação temporal da cocitação de autores, e de pontos no espaço (*scatterplots*) para a identificação de *clusters*, ou agrupamentos, no domínio, utilizadas como apoio visual da análise realizada por esses autores.

Figura 4: Representações visuais utilizadas por White e McCain



Fonte: WHITE; McCAIN, 1998

No mesmo ano, Garfield (1998) constata os avanços do campo da visualização da informação e estimula o uso das novas técnicas de visualização a partir da geração de mapas globais da ciência. Ressalta as possibilidades de interagir com eles (com os mapas) por meio do *zoom* para identificar frentes de pesquisa emergentes em várias especialidades, descobrir interesses de estudo de pesquisadores ou observar tendências nas temáticas de pesquisa.

Small (1999) destaca a tendência da conexão da Bibliometria com a visualização ao analisar os mapas da ciência baseados na cocitação. O autor define

os mapas como uma representação especial que reflete como disciplinas, áreas, especialidades, publicações ou autores relacionam-se uns com os outros de acordo com sua proximidade física ou lugares relativos, de forma análoga à forma como os mapas geográficos mostram as relações de características políticas e físicas na Terra.

Noyons *et al.* (1999) ressaltam que os mapas da ciência caminham principalmente em direção à análise de campos científicos para determinar suas estruturas cognitivas, sua evolução e os atores principais que os conformam.

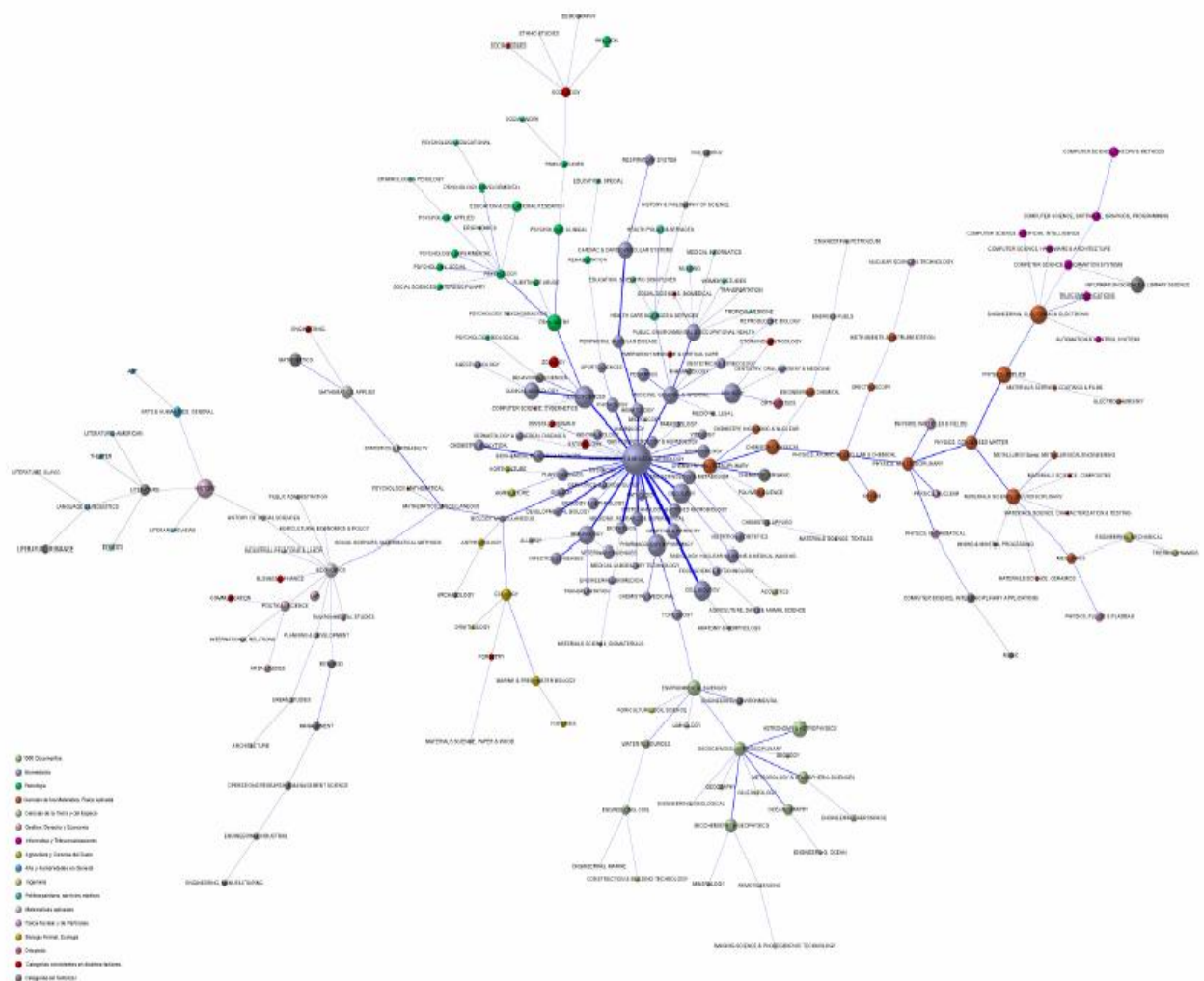
A partir dos anos 2000, identificar a natureza das especialidades em campos científicos tornou-se um desafio para especialistas da Ciência da Informação. O crescente interesse é motivado, de acordo com Chen *et al.* (2010), pelos seguintes aspectos:

- Grande volume de dados em diversas fontes.
- Presença de fontes de dados bibliográficas acessíveis e muitos repositórios de domínios específicos.
- Disponibilidade de programas (*softwares*) que permitem a visualização e análises de dados como UCINET, Pajek, CiteSpace, Gephi, entre outros.

Os aspectos citados por Chen *et al.* (2010), têm contribuído para criar uma comunidade interessada em revelar padrões e tendências a partir da análise de estruturas semânticas ligadas a domínios de conhecimentos da ciência, tendo como base teórica a análise de domínio.

A Figura 5 mostra um exemplo de mapa no qual se utiliza como unidade de análise as categorias temáticas do *ISI (WoS)*, sendo a relação entre duas categorias determinada pelo número de trabalhos em cujas referências coincidem artigos publicados em revistas pertencentes a uma ou outra categoria (co-citação de categorias). As categorias temáticas são representadas como nós, onde a área é determinada pelo número de artigos publicados em cada categoria. As relações de co-citação são representadas como vínculos (enlaces), cuja largura indica o grau ou peso da co-citação. Além disso, cada categoria é definida com uma cor que identifica a área temática geral à qual pertence (MOYA-ANEGON *et al.*, 2006).

Figura 5: Exemplo de Mapeamento de Grandes Domínios Científicos EUA-2000 a partir da Cocitação de Periódicos da Base ISI-WoS



Fonte: extraído de MOYA-ANEGÓN *et al.*, 2006

3.4.2 Visualização da informação e o conceito de Redes Sociais

A visualização ajuda a revelar, portanto, fluxos da comunicação científica, seja da perspectiva da colaboração e das redes de conhecimento geradas a partir do trabalho conjunto de pesquisadores ou grupos, como também de relações representadas na literatura através da coautoria e cocitação. Ambas conformam um emaranhado de documentos, autores, periódicos e disciplinas, que representam a

estrutura intelectual de domínios de conhecimento em diversos agregados ou níveis de análise.

Dessa forma, a representação gráfica das redes de interação oferece a possibilidade de descobrir e demonstrar teorias sobre os próprios gráficos e sobre os modelos que eles representam. Por isso, a representação gráfica tem sido muito utilizada em análises de redes sociais, como meio de representação formal das relações entre os agentes, assim como para identificar e quantificar as propriedades estruturais das mesmas (FREEMAN, 2004). O tipo de representação visual dos dados que oferece o grafo ou sociograma permite aos pesquisadores descobrir padrões e estruturas, que de outra maneira permaneceriam ocultas.

A análise de redes sociais e a visualização da informação científica por meio da representação gráfica tornaram-se uma das principais técnicas para evidenciar as relações entre os diferentes atores que compõem a estrutura intelectual de um campo ou domínio científico, potencializando as análises tradicionais e a explicação/compreensão das diversas disciplinas científicas (HOOK, 2007). Constitui-se também como destacada estratégia para observar a interação e evolução da ciência através das disciplinas e especialidades que a conformam.

Por conseguinte, mapas bibliométricos, onde são inferidas as redes sociais, são, ao mesmo tempo, um método para a análise da estrutura do conhecimento, uma vez que permitem visualizar as relações reais entre diversos autores, trabalhos, periódicos e disciplinas, e um subcampo em plena e ampla evolução na Bibliometria (SMALL, 2003). Da mesma forma, revelam o reconhecimento explícito que uns fazem de outros e as interações e evolução de determinado âmbito objeto de estudo.

Adicionalmente, cabe registrar que a visualização desses mapas do conhecimento é, na realidade, a representação de verdadeiros sistemas da informação científica, que podem ser utilizados como ferramenta de apoio à tomada de decisão por parte dos responsáveis pelo planejamento e avaliação da atividade científica de países e/ou instituições (MARTIN, 1996; NOYONS, 2004).

4 METODOLOGIA

“O êxito não se consegue só com qualidades especiais. É, sobretudo, um trabalho de constância, de método e de organização”.

(J.P. Sergent)

4.1 OBJETIVOS DO ESTUDO E ESTRATÉGIA METODOLÓGICA

Partindo das perguntas de pesquisa, já mencionadas na Introdução, o objetivo principal desta tese é mapear os domínios científicos da Universidade Federal do Rio de Janeiro (UFRJ), a fim de melhor conhecer tanto as competências da instituição no campo científico, as quais resultam das interações entre os diferentes atores de seu corpo social, como as trajetórias de sua comunidade científica.

Para isso, a pesquisa também buscará, como objetivos específicos, analisar o corpo social, particularmente o docente, da UFRJ em relação: (i) às atividades em colaboração e à produtividade, (ii) observar algumas características da produção intelectual, em relação ao número de autores por item de produção e o tipo e nível de colaboração, (iii) verificar a existência de núcleos de conhecimentos e o nível de interação entre eles ao longo do período investigado, e (iv) identificar a tendência de domínios de conhecimentos específicos ao longo do período estudado.

Um estudo com tais objetivos justifica-se inicialmente pela contribuição à área de ciência da informação, uma vez que estudos que versem sobre análise de domínios científicos, usando técnicas de visualização, são raros no Brasil, especialmente na área. Esse cenário se contrasta com o cenário Europeu. Estudos desta natureza vêm tendo significativos avanços na Europa nas últimas décadas, especialmente aqueles envolvendo redes colaborativas na ciência com base nas informações de autoria (co-autoria e co-citação). Tais estudos resultaram em uma nova área temática no campo da Informetria, conhecida como *Science Mapping* (NOYONS *et al.*, 1999).

O segundo aspecto que justifica o estudo é o foco institucional. Uma instituição como a UFRJ, com algumas particularidades - abrangência, dimensão e complexidade acadêmico-profissional - que a caracterizam, necessita conhecer as competências relativas ao seu corpo social, às atividades por estes desenvolvidas e seus respectivos resultados. É uma informação valiosa e necessária, dentre outros

aspectos, para a gestão interna, para representação de seus interesses no âmbito das relações interinstitucionais e para sua afirmação entre os pares das comunidades acadêmico-profissionais em que atua. Em sentido estrito, essa informação é importante para a atuação institucional nos âmbitos científico, artístico e tecnológico; em sentido amplo, para atuação nos contextos políticos e sociais mais gerais.

Assim, se por um lado, o presente trabalho busca preencher uma lacuna de informações e estudos desta natureza no Brasil, por outro, traz à tona a produção científica de uma instituição acadêmica, a UFRJ, uma das maiores universidades públicas do País.

Para viabilizar os objetivos deste estudo, que é essencialmente de natureza quantitativa, foram traçadas as estratégias metodológicas que se encontram no Esquema 1 e são apresentadas, em detalhe, nos tópicos a seguir.

Esquema 1: Etapas da coleta de dados e da análise dos domínios da UFRJ

→ → → →				
EXTRAÇÃO DE DADOS	DEFINIÇÃO DA UNIDADE DE ANÁLISE	SELEÇÃO DE MEDIDAS	REDUÇÃO E DISTRIBUIÇÃO DA INFORMAÇÃO	REPRESENTAÇÃO GRÁFICA
BUSCAS SIGMA.UFRJ SIGMA.UFRJ.Foco	INSTÂNCIAS Centros Programas de Pós-Graduação	VALORES/FREQUÊNCIAS Atributos (Conhecimentos) Coautoria Agrupamentos	REDUÇÃO DO ESPAÇO Análise de Cluster Análise Fatorial Multi-Dimensional Scaling (MDS) Pathfinder Networks (PFNETS) Redes Sociais	INTERAÇÃO Visualização de domínios <i>Software Gephi</i>

Fonte: Adaptado de BORNER *et al.*, 2003

4.2 COLETA DE DADOS

O trabalho utilizou a informação, envolvendo o período de 2001 a 2012, recompilada e armazenada na base de dados do EspaçoSIGMA.UFRJ (Figura 6).

A base EspaçoSIGMA.UFRJ foi criada em 1998, com a finalidade de ser um ambiente virtual e interativo de representação das atividades de ensino, de pesquisa e de extensão realizadas na Universidade Federal do Rio de Janeiro (UFRJ), construído sobre uma base de dados corporativa, onde são correntemente registradas estas atividades e evidenciadas suas múltiplas inter-relações.

Cabe destacar que até o ano de 2012, as informações da base EspaçoSIGMA.UFRJ foram utilizadas nos processos de avaliação dos programas de pós-graduação da instituição, já que a coleta de dados era realizada nesta base, para posteriormente ser exportada para a Coleta CAPES.

Figura 6: Página principal do EspaçoSIGMA.UFRJ: SIGMA.Foco

Desvendando a Universidade Federal do Rio de Janeiro

◆ Página Principal ◆ O que é? ◆ Como funciona? ◆ Como divulgar informações ◆ Contato

SIGMA.UFRJ FOCO

Produção Intelectual
Bibliográfica
Técnica
Artística
Todas as Áreas

Trabalhos de Conclusão
Graduação
Mestrado
Doutorado
Todos os Níveis

Atividades de Pesquisa
Linhas de Pesquisa
Projetos de Pesquisa
Projetos de Desenvolvimento

Atividades de Extensão
Cursos
Projetos de Extensão

Corpo Social
Docentes
Discentes de Graduação
Discentes de Pós-Graduação
Técnico Administrativos
Todas as categorias

Pós Graduação
Programas
Cursos
Áreas de concentração
Alunos
Teses e Dissertações
Concluintes

Estrutura Acadêmica
Pró-Reitorias
Centros
Órgãos e Unidades Suplementares
Departamentos
Visão Hierárquica

Cursos
Graduação
Mestrado
Doutorado
Especialização
Residência Médica
Aperfeiçoamento
Atualização
Treinamento Profissional
Extensão
Todos os Níveis

SIGMA.UFRJ Súmula
Números, Painéis e Indicadores

Conhecimentos & Competências
◆ Qualificação
◆ Currículo SIGMA

◆ Jornada
◆ PINC/Medicina

ATENÇÃO
Manutenção Preventiva: SIGMA - FOCO indisponível às 4a. feiras, das 17:00 às 22:00h.

SCIRE
COPPE - UFRJ

Disponível em: <http://www.sigma-foco.scire.coppe.ufrj.br/site/foco/index.htm>

Em 2013, por decisões da Pró-Reitora de Pós-graduação da UFRJ, essa coleta deixou de ser realizada para fins avaliativos. Apesar disso, a base vem sendo alimentada pelos integrantes do corpo social da instituição.

A escolha desta base como fonte exclusiva para recuperação da informação a ser usada na análise de domínio da UFRJ deve-se ao amplo modelo de representação da organização e das atividades institucionais da Universidade Federal do Rio de Janeiro (UFRJ), que ainda não foi explorada. O modelo de representação, configurado de acordo com Figueiredo (2006), inclui: pessoas (o Corpo Social), que se organizam para realizar “funções finalísticas” (Projetos e Cursos), das quais decorrem resultados (Concluintes, Trabalhos de Conclusão e Produção Intelectual), que podem ou não aparecer vinculados à instâncias institucionais (Departamentos, Cursos, Programas e Unidades Acadêmicas).

Neste modelo, cada um desses níveis organizacionais é caracterizado, em termos gerais, em função dos elementos categoria, situação, vínculo institucional e classificação. Por exemplo, integrantes do Corpo Social podem ser da categoria docente; sua situação pode ser de ativo ou inativo; sua relação/vínculo institucional pode ser a de estar lotado num dado Departamento; classificação refere-se à formação, em termos de nível de titulação e área do conhecimento, e de atuação, também em termos de áreas do conhecimento (FIGUEIREDO, 2006).

O SIGMA.Foco possibilitava diversas consultas à base EspaçoSIGMA.UFRJ, principalmente a partir de critérios de busca relativos aos elementos categoria e situação. Os dados relativos à relação institucional são apresentados como resultados das buscas. As classificações são apenas eventualmente utilizadas como critérios de buscas.

4.2.1 Definição da unidade de análise

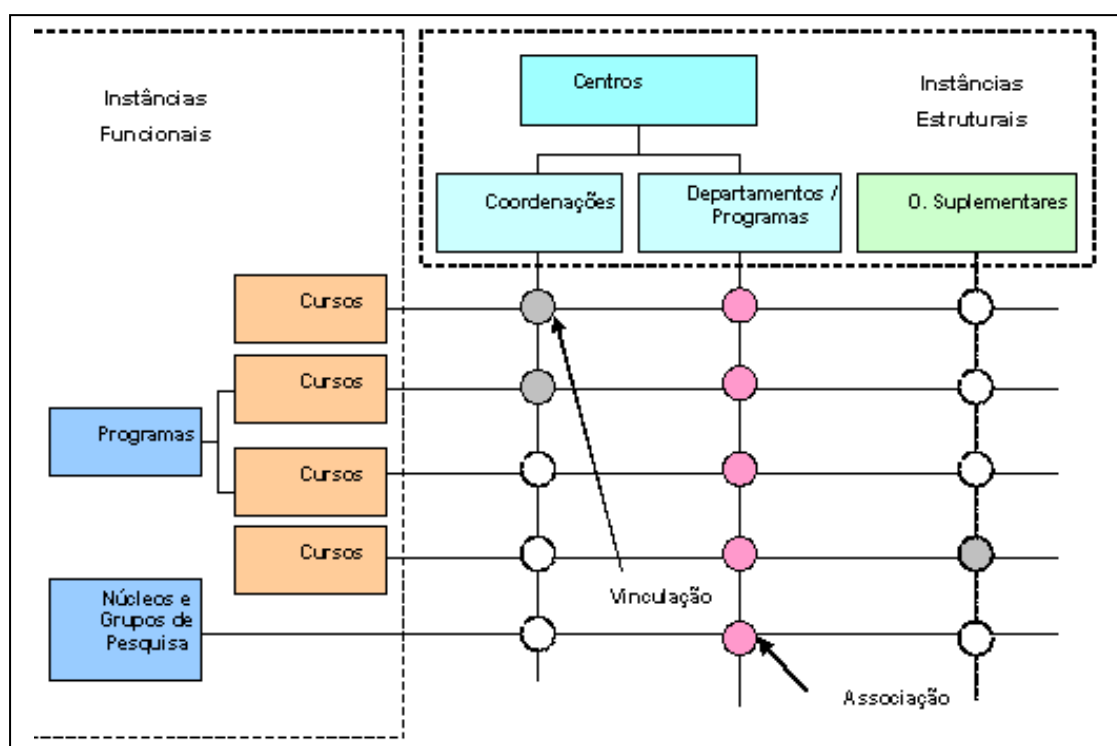
Antes de proceder à coleta de dados, propriamente dita, se fez necessária a definição da unidade de análise, já que a complexidade da base EspaçoSIGMA.UFRJ pode gerar unidades de diferentes naturezas.

Em sintonia com os objetivos do estudo, delimitamos duas unidades alvo de análise: produção científica, publicada em periódico, e os autores desta produção.

Para identificar os domínios temáticos, foram consideradas, então, as informações de área vinculadas à produção ou aos autores desta produção dos programas de pós-graduação das instâncias/centros da UFRJ¹³.

Importante saber que a estrutura da base de dados do EspaçoSIGMA.UFRJ está configurada de acordo com a organização específica da UFRJ, tanto em relação ao número de instâncias quanto à sua denominação. Essa organização compreende, portanto, a informação do vínculo de autores e de produção com um dos sete Centros da UFRJ, a saber: o Centro de Ciências da Saúde (CCS), o Centro de Tecnologia (CT), o Centro de Ciências Jurídicas e Econômicas (CCJE), o Centro de Ciências Matemáticas e da Natureza (CCMN), o Centro de Filosofia e Ciências Humanas (CFCH), o Centro de Letras e Artes (CLA) e o Fórum de Ciência e Cultura (FCC).

Figura 7: EspaçoSIGMA.UFRJ: Atividades-Fim, estrutura e funções, vínculos e associações



Fonte: Adaptado de FIGUEIREDO, 2006

¹³ Cf. SIGMA.Sumúla (2009). Disponível em: <http://www.sigma-sumula.scire.coppe.ufrj.br/>

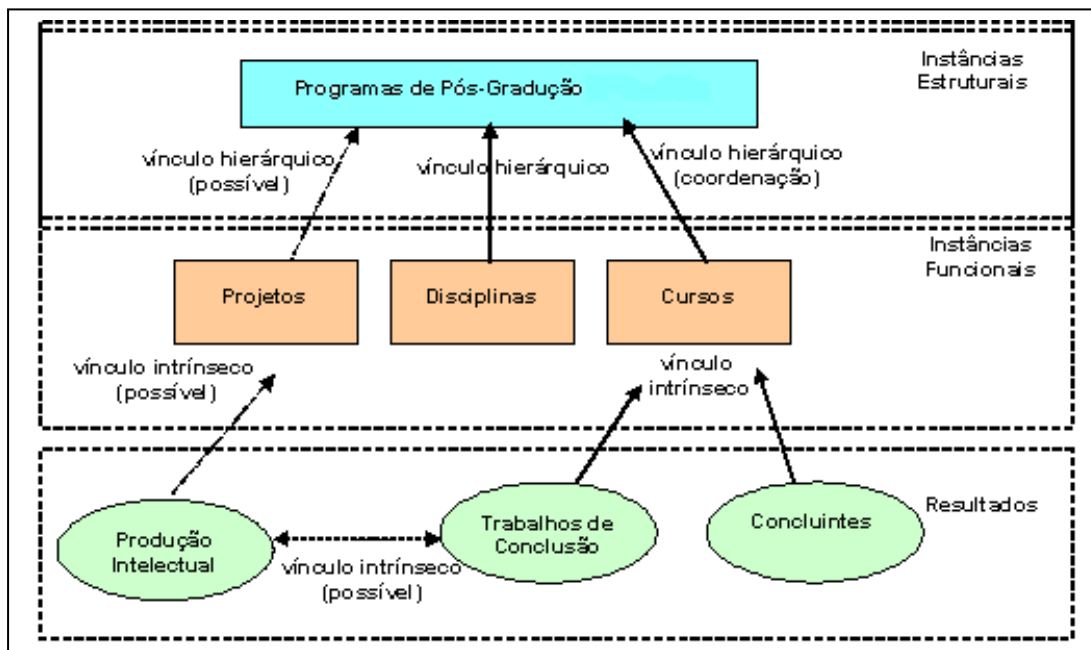
Além da associação com os Centros, a informação da produção e dos autores também aparece vinculada às atividades-fim que formam um conjunto multi-facetado de atuações e de inter-relações das instâncias estruturais e funcionais.

Essa múltipla inter-relação de informação que a base EspaçoSIGMA.UFRJ permite está apresentada na Figura 8.

Importante ressaltar que as relações podem ser intrínsecas, em virtude de sua própria natureza (ex.: mestres ou doutores são resultados dos cursos de pós-graduação) ou hierárquicas, quando estruturalmente definidas (ex.: uma coordenação de curso submete-se a uma Unidade Acadêmica).

Essa distinção é feita apenas para fins de elucidação, pois é suficiente identificar a existência de vínculo, conforme mostra a Figura 8.

Figura 8: EspaçoSIGMA.UFRJ: Vínculos dos resultados dos Programas de Pós-graduação



Fonte: Adaptado de FIGUEIREDO, 2006

4.2.2 Coletando os dados

As consultas e recuperação das informações sobre a produção em forma de publicações em periódicos e sobre os autores na base EspaçoSIGMA.UFRJ foram realizadas entre março e setembro de 2013. A busca na base EspaçoSIGMA.UFRJ considerou os seguintes critérios:

- Produção intelectual => artigos completos publicados em periódicos => autores
- Produção intelectual => artigos completos publicados em periódicos => áreas do conhecimento

A partir destes critérios, foi possível extrair 44.233 registros da produção da UFRJ de artigos publicados em periódicos no período de 2001 a 2012. Estes registros foram recuperados com as informações vinculadas a eles, por exemplo: tipos de autor, considerando seus respectivos vínculos institucionais; número de autores por item de produção; assim com também, número de áreas vinculadas à produção.

Os registros com suas informações foram exportados para matrizes multidimensionais, um processo que usou algumas ferramentas computacionais, como softwares para cruzamento de dados (Excel e Gephi).

A seguir, são apresentados os detalhes da análise destes dados.

4.3 ANÁLISE DE DADOS

Para identificar a configuração dos domínios científicos da UFRJ, foram considerados dois blocos de análise: (1) produção-área do conhecimento e (2) produção-autoria.

A análise por autores, ou seja, considerando a relação produção-autoria, foi realizada com o total de registros recuperados (44.233 artigos publicados em periódicos).

Com base nas informações originais da base EspaçoSIGMA.UFRJ, foram manualmente reclassificados os autores em seis grupos, de acordo com os respectivos vínculos formais com a instituição, da seguinte forma:

- 1) Docente (D): professores-pesquisadores;
- 2) Discente de Pós-graduação (P): nível de mestrado e doutorado;
- 3) Egresso de Pós-graduação (exP): nível de mestrado e doutorado;
- 4) Discente de Graduação (G);
- 5) Técnico (T);
- 6) Outros (O): inclui ex-docentes, ex-técnicos e ex-alunos de graduação.

Já a análise por áreas do conhecimento, isto é, considerando a relação produção-áreas do conhecimento, foi realizada com 69,33% do total de registros recuperados, ou seja, com 30.669 artigos publicados em periódicos, os quais possuíam registro de áreas do conhecimento vinculado a eles. Portanto, os 13.564 artigos que não apresentavam informações sobre a área do conhecimento da produção não foram incluídos nesta parte do estudo.

Para análise das áreas do conhecimento, utilizou-se a classificação do CNPq¹⁴ para agrupar as publicações com registro da área espontaneamente indicada pelos autores.

Considerando a análise de domínios, a literatura aponta algumas possibilidades de medidas utilizadas para evidenciar as relações existentes entre as unidades de análises e, assim, destacar a estrutura de conhecimentos e competências que a constituem. Neste estudo, optamos pela análise de redes a partir de medidas de co-ocorrência.

As matrizes multidimensionais, encarregadas de mostrar as relações entre as co-ocorrências que compõem um determinado domínio, e as técnicas de análise de redes, amplamente difundidas, permitem transformar um espaço de n-dimensões em

¹⁴ Convém lembrar que a Tabela de Áreas do Conhecimento do CNPq, divide-se em quatro níveis: (i) nível grande área, (ii) nível área, (iii) nível sub-área e (iv) nível especialidade. A análise por áreas do conhecimento considerou os dois primeiros níveis. Disponível em: <http://www.cnpq.br/documents/10157/186158/TabeladeAreasdoConhecimento.pdf>.

outro de duas ou três dimensões, o que facilita o entendimento e a visualização de tais relações.

As técnicas utilizadas para a redução do espaço n-dimensional que permitem representar a estrutura de um domínio, seja num papel ou na tela de um computador, classificam-se em três grandes grupos: os de natureza estatística multivariável, os de origem conexionista, e os baseados em análises de redes sociais (WASSERMAN; FAUST, 1994). Nosso estudo se concentrou nas técnicas de análise de rede.

De acordo com Camí *et al.* (2005), a análise de redes não é uma mera técnica, mais ou menos sofisticada para análise de fenômenos sociais, é também um novo enfoque teórico, tal como já apresentado no Capítulo 3. A diferença principal entre as explicações propostas pela análise estrutural de redes e às demais abordagens analíticas é a inclusão de conceitos e informações acerca das relações entre unidades. Assim, as relações que podem ser representadas dentro de determinada comunidade científica são muito diversas e, neste estudo, incluíram as redes de colaboração por tipo de vínculo institucional e por áreas do conhecimento.

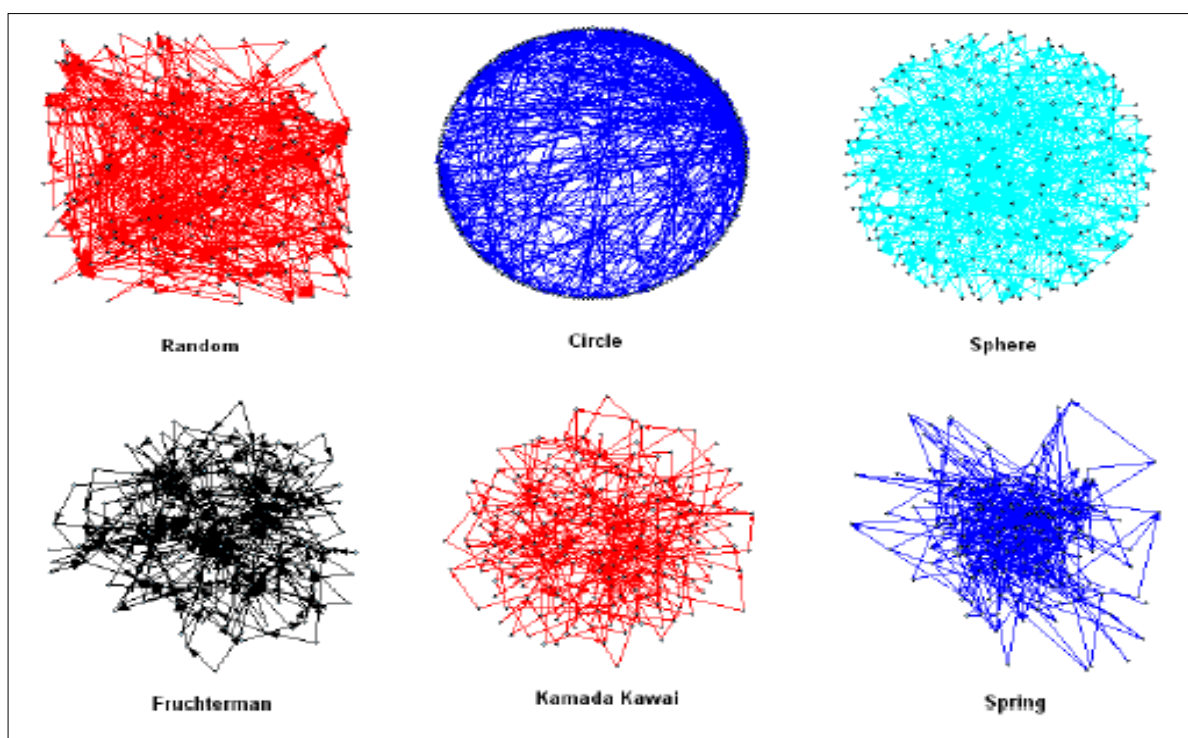
Os agrupamentos intrínsecos às redes são quantificados por meio do coeficiente de *clustering*, também conhecido como fenômeno da transitividade. Esse fenômeno ocorre quando um vértice A está conectado a um vértice B, e o vértice B está conectado ao vértice C, aumentando as chances do vértice A também estar conectado ao vértice C. O coeficiente de *clustering* do grafo é o valor médio dos coeficientes de *clustering* de todos os nós (WASSERMAN; FAUST, 1994).

Os atributos ou relações entre variáveis podem ser representados mediante matrizes de distância ou similaridade, por meio de técnicas de análises multivariável. Da mesma forma, as matrizes podem ser representadas através de procedimentos de distribuição espacial, de tal forma que as distâncias/similaridades entre variáveis similares apareçam juntas, enquanto que as variáveis diferentes separadas.

Na maioria dos casos, os resultados dessas técnicas de ordenamento são expressos em grafos, que, muitas vezes, apresentam um emaranhado de enlaces ao representar as relações (Figura 9).

Diversas técnicas e algoritmos (chamados de poda), entretanto, vêm sendo desenvolvidos com o objetivo de clarificar a rede mediante a eliminação de enlaces menos significativos ou importantes (BARBOSA *et al.*, 2011). O resultado do uso destes algoritmos é a geração de uma rede simplificada que, em função do método utilizado, representa com mais ou menos acerto, e com maior ou menor grau de inteligibilidade, a estrutura e essência da rede original.

Figura 9: Exemplo de uma rede com diferentes algoritmos de visualização



Fonte: Elaborado pelo autor

Em nosso caso, para a geração dos mapas para visualização dos domínios da UFRJ, com base nas análises de redes sociais, foram testados alguns *softwares* gratuitos e de uso não comercial, são eles: *Pajek*, *Gephi* e *VOSViewer*.

*Pajek*¹⁵, desenvolvido pela Universidade de Ljubljana, Slovenia, é um programa para visualização e análise de grandes redes sociais. O programa é gratuito para uso não comercial e é capaz de gerar redes sociais desde um ponto de vista estético, ou seja, utilizando o máximo de espaço disponível, minimizando o número de enlaces cruzados e forçando a separação dos nós.

¹⁵ Disponível em: <http://pajek.imfm.si/doku.php?id=pajek>

*Gephi*¹⁶, desenvolvido por um consórcio de instituições que inclui *SciencePo*, *Linkfluence*, *WebAtlas*, e *Quid*, é um programa gratuito para análise e visualização de redes. Funciona em Java, sendo, portanto, multiplataforma.

*VOSViewer*¹⁷ é um programa gratuito de visualização de redes, desenvolvido pelo *Center for Science and Technology Studies (CWTS)*, da Universidade de Leiden, Holanda.

Os testes realizados revelaram que o *software* que melhor se adaptou a nossa proposta de análises foi o *Gephi*. Portanto, foi o *software* utilizado nas análises.

¹⁶ Disponível em: <https://gephi.org/>

¹⁷ Disponível em: <http://www.vosviewer.com/>

5 DOMÍNIOS INSTITUCIONAIS DA UFRJ

“Se você quer resultados diferentes, não faça sempre o mesmo”

(Albert Einstein)

O presente capítulo mostra empiricamente a aplicação da metodologia apresentada a fim de atingir os objetivos propostos. Nesse sentido, para identificar a configuração dos domínios científicos da UFRJ, destacando os diversos padrões de interação a partir das informações de sua produção científica (especificamente de artigos publicados em periódicos), foram considerados dois blocos de análise: (1) por áreas do conhecimento e (2) por autores.

A continuação, procedemos à apresentação de cada um desses dois blocos de análise.

5.1 PRODUÇÃO CIENTÍFICA E AS GRANDES ÁREAS DO CONHECIMENTO

A produção científica da UFRJ foi analisada inicialmente a partir dos artigos, catalogados na base EspaçoSIGMA.UFRJ, que apresentavam o registro da área do conhecimento (um registro espontaneamente realizado pelos autores). Considerando, então, este recorte, foram encontrados 30.669 artigos completos publicados em periódicos no período de 2001 a 2012. Os artigos completos publicados em periódicos que não apresentaram a informação de área de conhecimento somaram 13.564 no mesmo período e não foram incluídos nesta análise, tal como descrito na metodologia.

Os 30.669 artigos com a informação da área de conhecimento foram, então, agrupados em nove grandes áreas de conhecimento, conforme descrito na metodologia.

A Tabela 1 mostra a fração de artigos por grande área do conhecimento ao longo do período estudado, expresso em quatro triênios.

Observa-se que as áreas mais experimentais, as quais tradicionalmente têm o artigo em periódico como principal veículo para comunicar seus resultados de pesquisa, são as que apresentam o maior número de artigos publicados. Assim, a

grande área Ciências da Saúde é a que registra maior número de artigos, seguida pelas Ciências Biológicas e Ciências Exatas e da Terra.

Muito embora estas áreas sejam as de maior representatividade, estas são também as que apresentam decréscimo na fração de artigos publicados entre o 1º e 4º triênio: Ciências da Saúde de 27,20% para 20,91%; Ciências Biológicas de 19,68% para 17%; Ciências Exatas e da Terra de 16,89% para 10,59%.

A grande área Engenharias apresenta produção constante de artigos publicados em periódicos, representando 13% do total em todos os períodos considerados. Já a grande área Ciências Humanas é a que registra o maior crescimento, no período todo, na fração de artigos publicados: de 9,95% para 21,61%.

Tabela 1: UFRJ – Distribuição percentual de artigos publicados em periódicos por grande área do conhecimento.

Grande Área*	Número de Artigos			
	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%
CIÊNCIAS DA SAÚDE	27,20	24,17	23,08	20,91
CIÊNCIAS BIOLÓGICAS	19,68	19,36	18,65	17,00
CIÊNCIAS EXATAS E DA TERRA	16,89	13,99	11,53	10,59
ENGENHARIAS	13,22	13,50	13,66	13,97
CIÊNCIAS HUMANAS	9,95	11,80	14,51	21,61
LINGUÍSTICA, LETRAS E ARTES	6,71	9,70	8,47	9,45
CIÊNCIAS SOCIAIS APLICADAS	5,47	6,01	8,25	5,41
CIÊNCIAS AGRÁRIAS	0,53	0,97	0,98	0,69
OUTROS	0,35	0,50	0,87	0,37
TOTAL %	100	100	100	100
TOTAL N°	8.187	7.868	7.790	6.824

Fonte: Dados da pesquisa

*Grandes áreas a partir da classificação do CNPq (<http://www.cnpq.br/>)

A área de Linguística, Letras e Artes mostra oscilação no número de artigos publicados ao longo dos quatro triênios, respondendo, em média, por 8% do total. Já Ciências Sociais Aplicadas registra aumento na fração de artigos nos três primeiros períodos, de 5% para 8%, e queda no último triênio, agora respondendo por 5% do

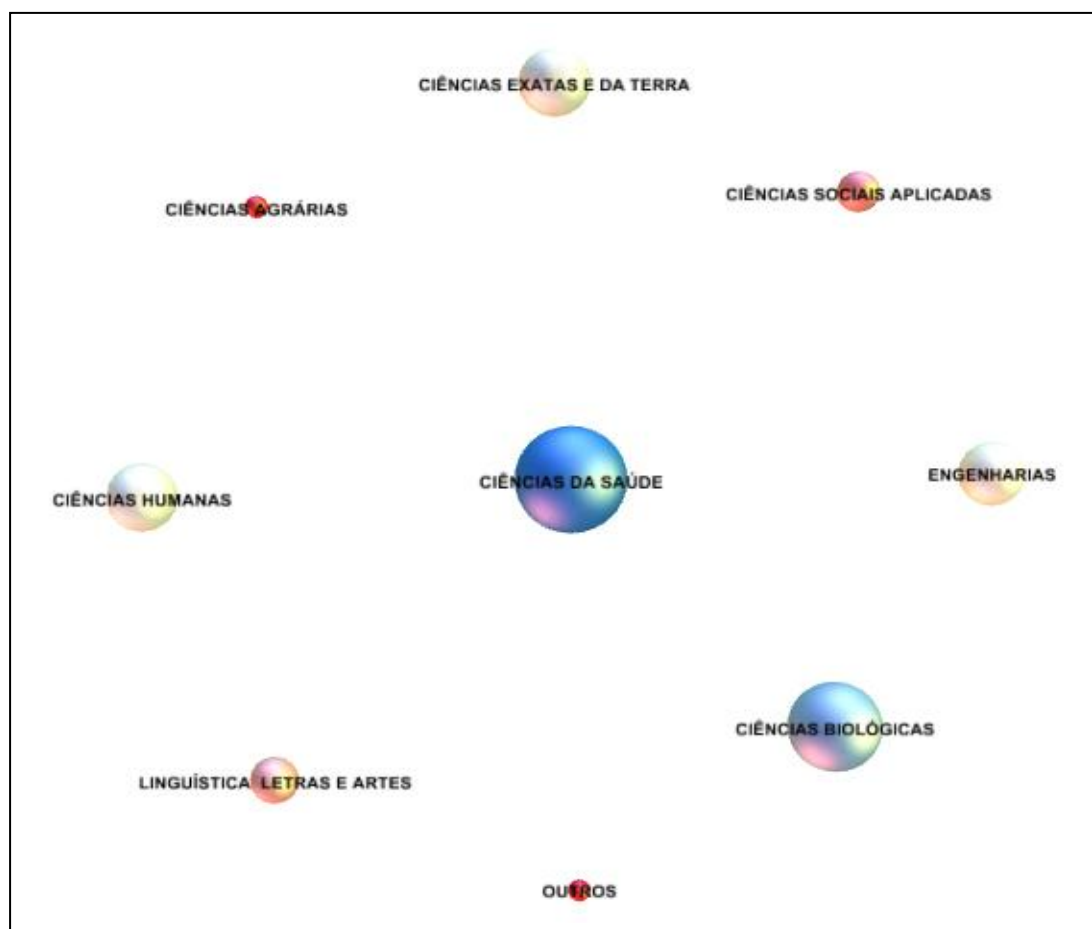
total de artigos. Outras grandes áreas, Ciências Agrárias e Outros, representam menos de 1%, do total de artigos publicados em todos os períodos considerados.

A distribuição da produção da UFRJ por grandes áreas segue um padrão encontrado em vários outros estudos, que também focaram em instituições acadêmicas. Um exemplo é o trabalho realizado por Huang e Chang (2008) que, observando a Universidade de Hong Kong, sugere que enquanto o periódico é o meio predominante de divulgação de pesquisas da maioria de disciplinas das ciências duras, nas ciências sociais e humanas, além dos periódicos, os pesquisadores utilizam uma gama mais ampla de veículos para divulgar suas pesquisas, como livros e capítulos de livros.

Essa conclusão também é coerente com os achados de Nederhof *et al.* (1989), que analisaram a produção científica de oito disciplinas das ciências sociais e humanas em universidades holandesas. Cabe destacar que alguns estudos mais recentes também apontam um aumento da proporção de artigos em periódicos em algumas disciplinas das ciências sociais e humanas (ARCHAMBAULT *et al.*, 2006; OSCA-LLUCH; HABA, 2005).

A fim de visualizar graficamente os 30.669 artigos, com indicação da área de conhecimento, nas nove grandes áreas do conhecimento, elaboramos a Figura 10. Cada circunferência do mapa representa uma grande área, sendo o tamanho de cada uma proporcional à quantidade de artigos dessa área.

Figura 10: Mapa da produção de artigos da UFRJ, agrupado por grande área do conhecimento (2001-2012)



Fonte: Elaborado pelo autor

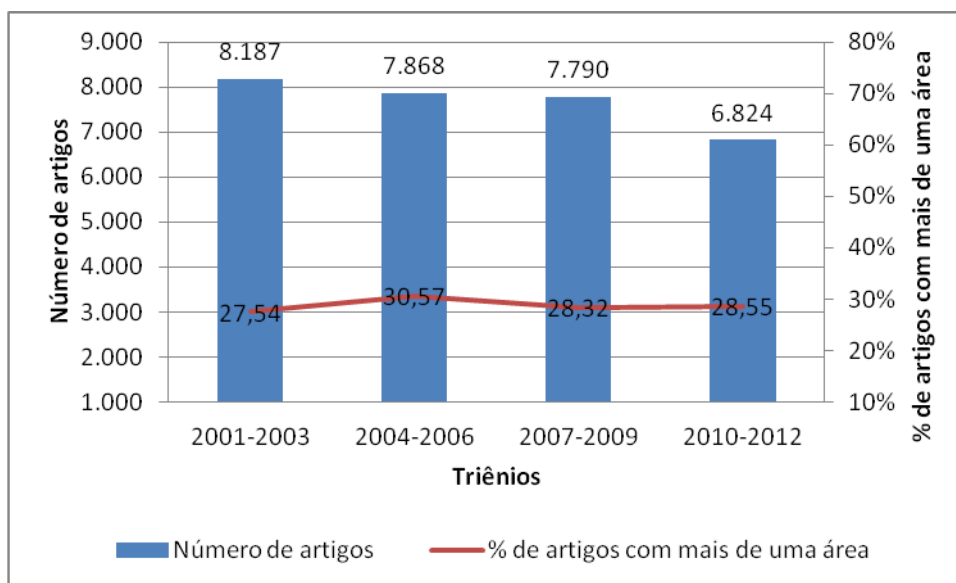
5.1.1. A interação entre as grandes áreas do conhecimento

Do total de 30.669 artigos publicados em periódicos no período de 2001 a 2012, em 8.806 artigos (o que representa 28,71% do total) foram encontradas mais de uma área do conhecimento associadas a eles, ou seja, os autores destas produções, ao cadastrá-las na base EspaçoSIGMA.UFRJ, entenderam que seus trabalhos não eram exclusivos de uma única área.

Na Figura 11, que mostra a distribuição dos artigos por triênios e o percentual de artigos associados a mais de uma área do conhecimento, verifica-se que no período 2001-2003, a quantidade total de artigos publicados em periódicos foi de 8.187, dos quais 2.247 (27,54%) eram artigos relacionados a mais de uma área do conhecimento. Em 2004-2006, foram 2.405 (30,57%) artigos associados a mais de

uma área do conhecimento. Nos triênios seguintes, esse grupo de artigos somou 2.206 (28,32%) e 1.948 (28,55%), respectivamente.

Figura 11: UFRJ – Artigos publicados em periódicos e percentuais de artigos vinculados a mais de uma área do conhecimento por triênios



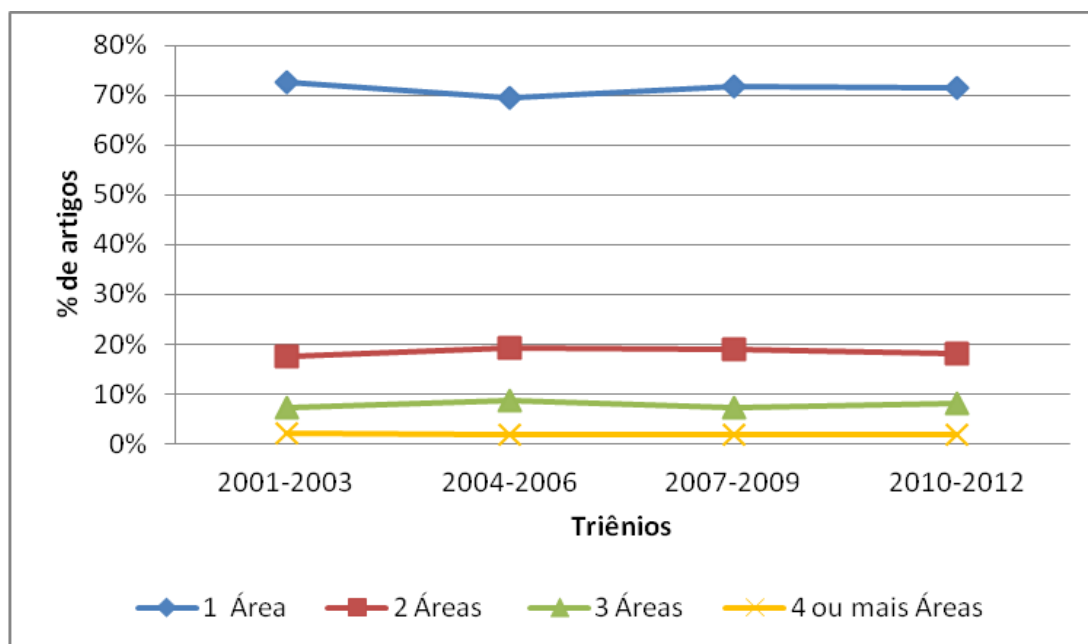
Fonte: Elaborado pelo autor

Embora a produção total de artigos em periódicos da UFRJ, observado a partir das informações de áreas do conhecimento, tenha decrescido de 8.187, no primeiro triênio, para 6.824 no último, o número de artigos que registram mais de uma área do conhecimento é quase constante no período estudado.

A Figura 12 mostra a distribuição do total de artigos (30.669), em termos percentuais e por triênios, por número de áreas do conhecimento a eles vinculados. Observa-se que o número de artigos com apenas uma única área do conhecimento (sem interação entre áreas) representa em torno de 70% do total, em todos os triênios.

No grupo de artigos vinculados a mais de uma área, aqueles vinculados a duas áreas do conhecimento são majoritários e representam em torno de 19% do total. Já o número de artigos que contam com três áreas do conhecimento associados a eles, representam, em média, 8% do total. Artigos com quatro ou mais áreas representam menos de 2% no período estudado.

Figura 12: UFRJ - Percentual de artigos publicados em periódicos por número de áreas do conhecimento a eles vinculados



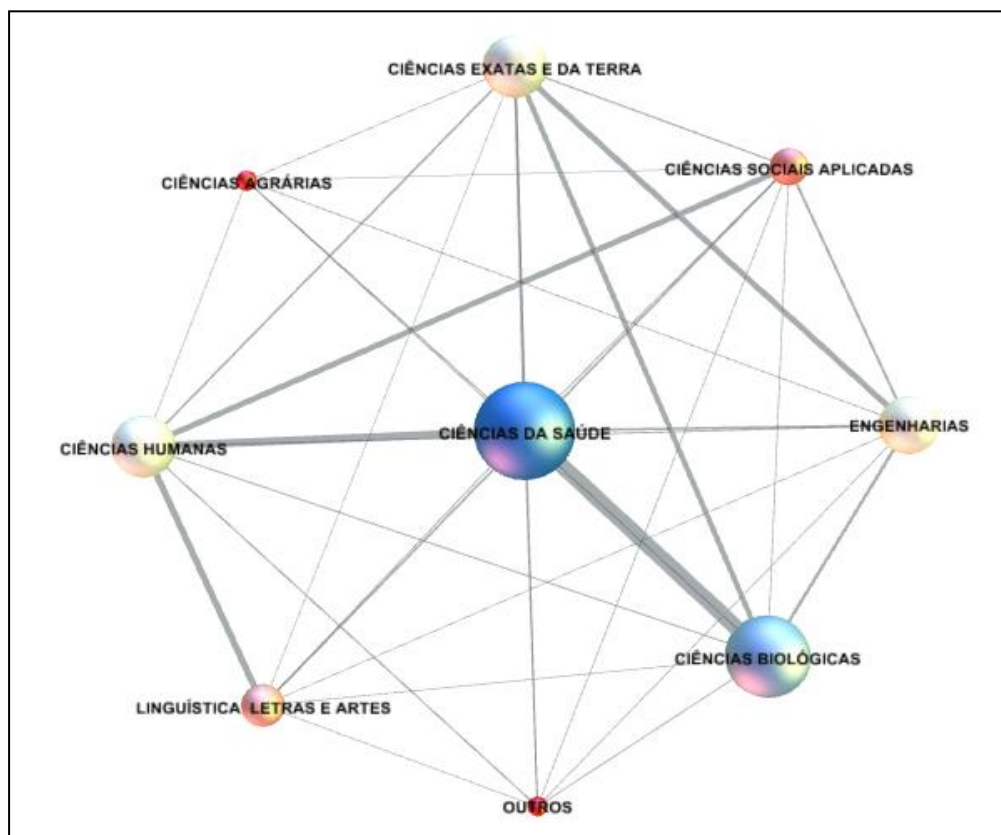
Fonte: Elaborado pelo autor

A interação espontânea entre áreas do conhecimento que observamos na produção científica da UFRJ, mais especificamente nos artigos completos publicados em periódicos, se dá, portanto, principalmente através de duas áreas do conhecimento por artigo.

A fim de visualizar esta interação, foi elaborada a Figura 13, que mostra a estrutura do domínio por meio de seus nós, vínculos (ou enlaces) e intensidade dos mesmos. Cada grande área do conhecimento está representada por um nó na rede com seu nome correspondente (como representado na Figura 10).

Os vínculos (ou enlaces), que unem os diferentes nós, indicam as interações que existem entre eles. A intensidade das interações está determinada pela largura dos enlaces. Por exemplo, podemos observar que o enlace que existe entre Ciências da Saúde e Ciências Biológicas tem largura maior que o enlace entre Ciências da Saúde e Linguística, Letras e Artes, o que indica, de maneira intuitiva, que a relação entre esses nós é mais intensa que as existentes entre os outros nós.

Figura 13: Mapa da interação entre grandes áreas do conhecimento, a partir das informações de artigos publicações em periódicos, UFRJ (2001-2012)



Fonte: Elaborado pelo autor

Se olharmos para outro ponto do mapa, veremos como ocorre fenômeno semelhante, mas aparentemente com menor intensidade, entre Ciências Exatas e da Terra e Engenharias, entre Ciências Sociais e Humanas e entre Ciências Humanas e Linguística. Dessa forma, ao mesmo tempo em que observamos vínculos (enlaces) de intensidades variadas, evidencia-se que o número de artigos que compartilham é maior que com os outros nós (tamanho das esferas, ou seja, das áreas do conhecimento).

Podemos, assim, considerar o compartilhamento de áreas uma espécie de medida de colaboração entre áreas do conhecimento, mesmo que, neste caso, essa intenção de estabelecer uma colaboração venha de um ato espontâneo daquele/a que registrou a produção na base. Informações como está, baseada na co-ocorrência de áreas, é uma relação reconhecida há décadas no campo da Bibliometria, servindo como medida de proximidade para a elaboração de mapas (GRIFFITH *et al.*, 1974; SMALL; GRIFFITH, 1974).

Estas representações, por outro lado, conforme destacado por Schlueter e Borlund (2004), tornam possível observar as mudanças que se produzem na estrutura de uma dada rede ao longo do tempo, permitindo visualizar dinâmicas das áreas do conhecimento e avaliar o grau de inter-relação entre especialidades.

Assim, por exemplo, a elaboração de séries históricas permite observar as mudanças nas tendências de pesquisa. Dessa forma, enquanto os mapas com dados anuais, ou período único, não possibilitam prever por si só mudanças ou avanços das atividades de pesquisa, a observação desses mapas em períodos consecutivos permitem mostrar as tendências, sendo um suporte visual inestimável para a interpretação de resultados (CHEN; CARR, 1999; NOYONS, 2001).

No sentido de verificar a dinâmica de artigos com uma área somente e com mais de uma área, elaboramos a Tabela 2, onde estão apresentadas as frações que estes dois grupos de artigos representaram em cada um dos quatro triênios. Observa-se que as Ciências da Saúde apresentam maior fração de artigos de ambos os tipos, variando entre 19-15% nos artigos com uma área e entre 9-6% nos artigos com mais de uma área, no período estudado.

Tabela 2: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas, de acordo com as grandes áreas do conhecimento por triênios

Grandes Áreas	Número de Artigos							
	Artigos com uma grande área				Artigos com mais de uma grande área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
CIÊNCIAS DA SAÚDE	18,97	14,76	16,21	14,7	8,23	9,42	6,87	6,21
CIÊNCIAS BIOLÓGICAS	13,11	12,07	11,34	12,38	6,57	7,28	7,32	4,62
CIÊNCIAS EXATAS E DA TERRA	11,86	10,21	8,79	7,02	5,03	3,79	2,73	3,58
ENGENHARIAS	10,77	10,52	10,76	9,91	2,44	2,97	2,90	4,06
CIÊNCIAS HUMANAS	7,61	8,41	10,42	15,75	2,35	3,39	4,08	5,86
LINGUÍSTICA, LETRAS E ARTES	5,61	8,05	6,51	6,64	1,10	1,65	1,96	2,81
CIÊNCIAS SOCIAIS APLICADAS	4,14	4,49	6,28	4,31	1,33	1,53	1,98	1,10
CIÊNCIAS AGRÁRIAS	0,40	0,66	0,87	0,56	0,12	0,31	0,10	0,13
OUTROS	0,09	0,27	0,50	0,19	0,27	0,23	0,37	0,18
TOTAL %	72,56	69,44	71,68	71,46	27,44	30,57	28,31	28,55
TOTAL N°	5.940	5.463	5.584	4.876	2.247	2.405	2.206	1.948

Fonte: Dados da pesquisa

A fração de artigos com uma área nas Ciências Biológicas é quase constante ao longo dos triênios (em torno de 12%). No tocante a artigos com mais de uma área, observa-se nível constante nos três primeiros triênios (em torno de 7%) e decréscimo, para 4%, no último.

Ciências Exatas e da Terra, por outro lado, apresenta diminuição na fração de artigos com uma área, de 11,86% em 2001-2003, para 7,02% em 2010-2012, e oscilação em relação aos artigos com mais de uma área (entre 5-2% no período estudado).

Engenharias apresentam nível de produção constante de artigos com uma área ao longo do tempo (em torno de 10%). Em relação aos artigos com mais de uma área, a fração é constante nos três primeiros triênios (2%), crescendo no último (passando para 4%).

As Ciências Humanas destacam-se pelo crescimento, tanto na fração de artigos com uma área (de 7,61% em 2001-2003, para 15,75% em 2010-2012), quanto daqueles com duas ou mais áreas (de 2,35% em 2001-2003, para 5,86% em 2010-2012).

Já para as demais áreas, onde este tipo de produção não é consolidado, como Linguística, Letras e Artes, observamos ligeiro crescimento na fração de artigos com mais de uma área, de 1,10% para 2,81% no período. Já em relação aos artigos com uma área, observa-se um nível constante de produção nos últimos triênios (6%). Também com baixa representação, observamos que as Ciências Sociais Aplicadas mostram nível constante de produção de ambos os tipos de artigos (em torno de 4% nos artigos com uma área e 1% nos artigos com mais de uma área) no período analisado.

As outras grandes áreas, Ciências Agrárias e Outros, registram número de artigos inferior a 1% de ambos os tipos.

5.1.2 A centralidade das grandes áreas do conhecimento

Para melhor interpretação da interação entre as grandes áreas, utilizaremos a medida de centralidade de grau para cada período analisado. Uma grande área do

conhecimento será mais central quanto maior é o seu grau nodal em relação às demais.

A Tabela 3 mostra a medida de centralidade de grau de cada uma das grandes áreas para cada triênio. Podemos observar que Ciências da Saúde apresenta grau de centralidade oito em todos os triênios, ou seja, ela interagiu com todas as outras grandes áreas no período estudado.

Tabela 3: UFRJ – Centralidade de grau de grandes áreas do conhecimento

Grande Área	2001-2003	2004-2006	2007-2009	2010-2012
	Grau	Grau	Grau	Grau
CIÊNCIAS DA SAÚDE	8	8	8	8
CIÊNCIAS BIOLÓGICAS	7	7	7	7
CIÊNCIAS EXATAS E DA TERRA	7	8	5	6
ENGENHARIAS	7	8	6	8
CIÊNCIAS HUMANAS	7	7	7	8
LINGÜÍSTICA, LETRAS E ARTES	4	6	6	5
CIÊNCIAS SOCIAIS APLICADAS	7	6	5	7
CIÊNCIAS AGRÁRIAS	4	4	2	6
OUTROS	7	6	6	5

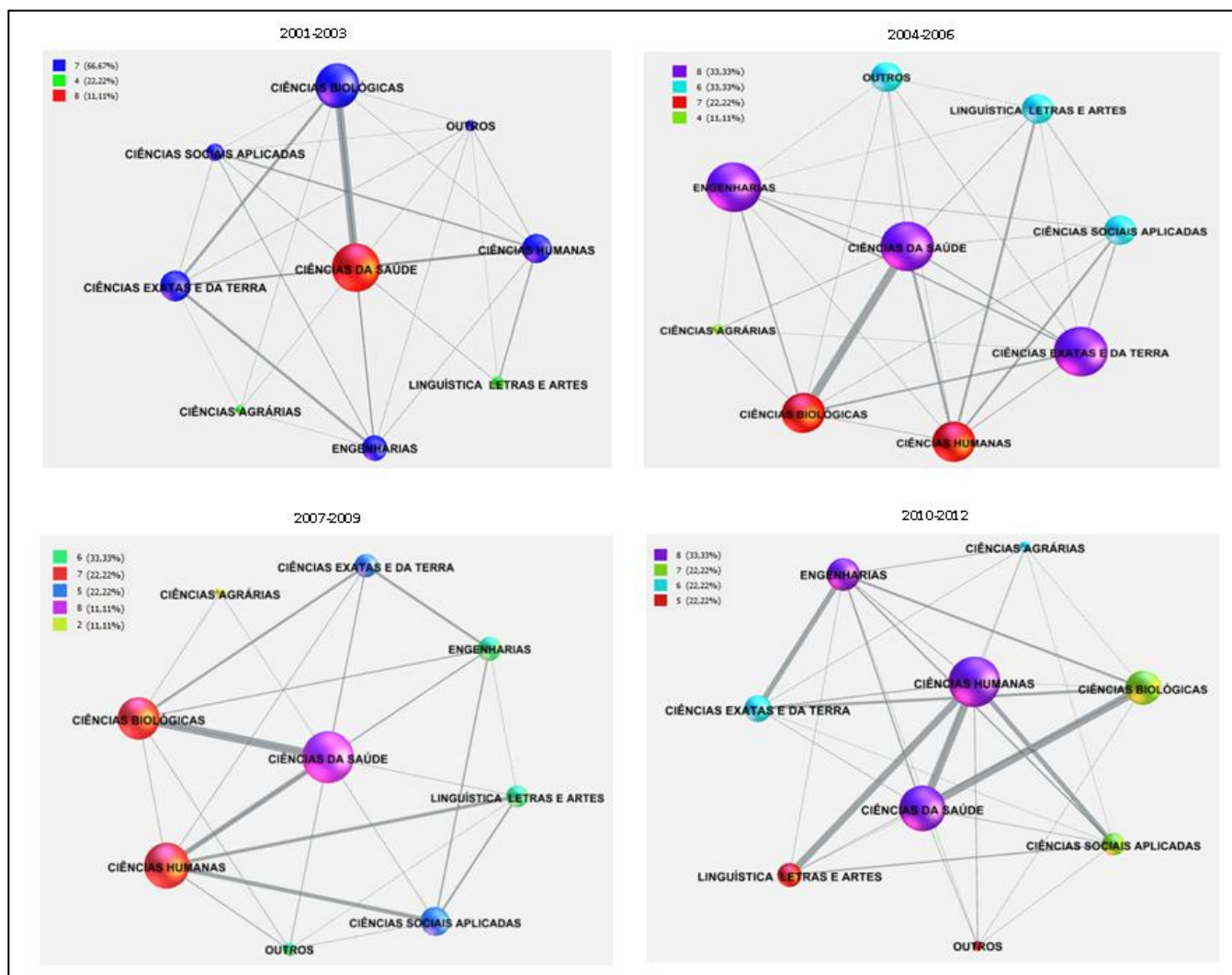
Fonte: Dados da pesquisa

Isso significa que Ciências da Saúde é a área mais universalmente compartilhada nas produções da UFRJ de outras áreas. Em outras palavras, autores de outras áreas associam, com maior frequência, seus trabalhos às suas áreas originais e também às Ciências da Saúde. Esse papel central das Ciências da Saúde no compartilhamento da informação sugere que esta é a área que mais intervém no desenvolvimento do macro domínio institucional.

Numa posição mais intermediária encontram-se Linguística, Letras e Artes e Ciências Sociais Aplicadas. Por outro lado, a grande área mais periférica e, conseqüentemente, a que menos recurso compartilha para o desenvolvimento do macro domínio é Ciências Agrárias.

Interessante notar que, de maneira geral, as grandes áreas têm apresentado uma dinâmica de interação ao longo do período estudado, o que pode ser visualizado na Figura 14, onde são agrupadas cada uma das grandes áreas por grau (representado por uma determinada cor) e período de tempo (triênios).

Figura 14: Mapa da interação entre grandes áreas do conhecimento, a partir das informações de artigos publicações em periódicos, UFRJ, por triênios



Fonte: Elaborado pelo autor

Por exemplo, no triênio 2004-2005, as grandes áreas de Engenharia e Ciências Exatas e da Terra tinham posição central junto com Ciências da Saúde, todas com grau 8. Já no último triênio (2010-2012), Ciências Humanas toma a posição de Ciências Exatas e da Terra, colocando-se, junto com Engenharia e Ciências da Saúde, como grande área central da instituição.

No sentido de ter uma visão mais detalhada do domínio por áreas do conhecimento, nas seções a seguir apresentamos uma aproximação a nível microestrutural das áreas. Isto nos permitirá observar mais detalhadamente o comportamento das diversas áreas que compõem cada uma das grandes áreas, descritas/analizadas até aqui, assim como verificar o grau de contribuição de cada uma delas para o domínio institucional da UFRJ ao longo do período estudado.

O tamanho de cada circunferência, e o tamanho da fonte do nome, mostrados na Figura 15, são proporcionais à quantidade de artigos de cada uma das áreas. Observa-se que Medicina ocupa papel central no mapa. Destacam-se, também, as áreas de Química, Bioquímica e Zoologia, assim como também Engenharia Química e Engenharia dos Materiais e Metalurgia. Por outro lado, Letras, História, Psicologia e Educação também se destacam no mapa.

A correspondência entre cada uma das 75 áreas e sua produção é especificada a seguir, considerando a grande área à qual pertencem. Após esta especificação, serão analisadas as redes de interação (colaboração) do conjunto das áreas, a fim de identificar a dinâmica das mesmas e suas respectivas características ao longo dos anos.

Teremos, dessa forma, uma visão mais ordenada e clara da participação de cada uma das 75 áreas no nível macro domínio institucional da UFRJ.

5.2.1.1 As áreas nas Ciências da Saúde

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Ciências da Saúde totalizou 7.354, dos quais 4.980 (67,72%) são artigos vinculados a uma única área e 2.374 (32,28%) são artigos vinculados a mais de uma área.

A Tabela 4 apresenta a distribuição do total desses artigos, em termos percentuais, pelas nove áreas que compõe Ciências da Saúde, a saber: Medicina, Odontologia, Enfermagem, Saúde Coletiva, Nutrição, Farmácia, Educação Física, Fisioterapia e Terapia Ocupacional e Fonoaudiologia.

Verifica-se que Medicina é, de longe, a área que registra maior fração de artigos, oscilando em torno de 30-40% nos artigos com uma área e 14-24% nos artigos com mais de uma área, no período estudado. Como as frações oscilam muito, não é possível perceber uma tendência, seja de crescimento ou redução em qualquer um dos dois tipos de publicação.

Tabela 4: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências da Saúde

Ciências da Saúde: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
MEDICINA	40,82	33,18	40,43	30,83	20,52	24,82	17,52	14,23
ODONTOLOGIA	11,00	8,25	7,68	13,10	1,21	0,79	1,61	1,96
ENFERMAGEM	4,58	5,26	7,73	13,59	1,44	3,47	1,89	4,06
SAÚDE COLETIVA	4,67	4,47	1,95	1,47	2,92	4,31	2,34	2,59
NUTRIÇÃO	2,96	4,10	4,34	3,08	0,85	1,52	1,06	1,61
FARMÁCIA	0,90	1,37	2,39	3,29	1,21	0,89	1,50	1,19
EDUCAÇÃO FÍSICA	0,45	1,00	2,22	0,63	0,13	0,42	1,89	2,38
FISIOTERAPIA E TERAPIA OCUPACIONAL	0,00	0,21	0,67	0,98	0,04	0,26	0,22	0,42
FONOAUDIOLOGIA	0,00	0,00	0,00	0,21	0,13	0,00	0,00	0,00
Sem área	4,36	3,21	2,84	3,08	1,80	2,47	1,72	1,26
TOTAL %	69,74	61,04	70,24	70,29	30,26	38,96	29,76	29,71
TOTAL N°	1.553	1.161	1.263	1.003	674	741	535	424

Fonte: Dados da pesquisa

Odontologia é a área que apresenta a segunda maior fração de artigos com uma área, mas também devido às oscilações não é possível identificar tendências, o que não ocorre na Enfermagem, que apresenta um crescimento contínuo deste tipo de produção.

Considerando os artigos com mais de uma área, Saúde Coletiva e Enfermagem são as áreas, depois de Medicina, que apresentam maior fração de artigos deste tipo, oscilando entre 2-4% e 1-4%, respectivamente. Ainda nos artigos com mais de uma área, observamos que os da Educação Física mostram tendência de crescimento deste tipo ao longo dos triênios considerados.

5.2.1.2 As áreas nas Ciências Biológicas

No período de 2001-2012, o número de artigos publicados em periódicos relacionados à grande área Ciências Biológicas totalizou 5.747, dos quais 3.751 (65,27%) são artigos com uma área e 1.996 (34,73%) são artigos com mais de uma área.

A Tabela 5 mostra a distribuição da totalidade desses artigos, em termos percentuais, pelas respectivas áreas que compõem as Ciências Biológicas, que

somam treze áreas. A saber: Zoologia, Bioquímica, Microbiologia, Morfologia, Fisiologia, Ecologia, Genética, Imunologia, Parasitologia, Botânica, Farmacologia, Biofísica e Biologia Geral.

Tabela 5: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Biológicas

Ciências Biológicas: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
ZOOLOGIA	11,24	13,59	5,71	6,90	6,89	9,78	7,09	2,59
BIOQUÍMICA	8,07	9,32	11,08	15,95	4,22	3,68	4,61	5,09
MICROBIOLOGIA	6,64	7,03	10,05	13,10	3,04	4,60	3,44	2,59
MORFOLOGIA	5,46	2,10	3,23	1,38	3,35	3,15	2,62	2,41
FISIOLOGIA	5,03	4,60	4,47	5,60	2,73	0,92	1,93	1,90
ECOLOGIA	4,72	2,43	5,37	5,00	1,30	1,90	3,58	1,38
GENÉTICA	4,03	3,48	2,48	3,19	2,42	3,81	4,40	2,07
IMUNOLOGIA	3,10	3,09	3,10	3,19	1,99	1,58	1,65	1,90
PARASITOLOGIA	3,35	2,76	4,27	3,19	0,81	0,92	1,58	1,38
BOTÂNICA	2,79	2,04	3,72	5,34	0,87	0,72	1,10	1,47
FARMACOLOGIA	3,04	3,09	2,55	4,40	3,60	2,69	2,68	2,59
BIOFÍSICA	2,05	2,23	1,72	2,76	0,87	1,38	1,03	0,78
BIOLOGIA GERAL	0,06	0,00	0,00	0,00	0,00	0,00	0,48	0,09
Sem área	7,01	6,63	3,03	2,84	1,30	2,50	3,03	0,95
TOTAL %	66,60	62,38	60,77	72,84	33,40	37,62	39,23	27,16
TOTAL N°	1.073	950	833	845	538	573	570	315

Fonte: Dados da pesquisa

No que se refere a artigos com mais de uma área, observa-se que, de modo geral, todas as áreas que compõem Ciências Biológicas apresentam oscilação no número de artigos desse tipo no período estudado. Portanto, não é possível identificar qualquer tendência nesta produção.

Por outro lado, no que concerne aos artigos vinculados a uma única área, destacam-se Bioquímica e Microbiologia pela evidente tendência de crescimento, praticamente dobrando o número de artigos deste tipo do primeiro triênio para o último.

5.2.1.3 As áreas nas Ciências Exatas e da Terra

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Ciências Exatas e da Terra totalizou 4.105, dos quais 2.938 (71,57%) são artigos com uma área e 1.167 (28,43%) são artigos com mais de uma área.

A Tabela 6 mostra a distribuição desses artigos, em termos percentuais, pelas respectivas áreas que compõe Ciências Exatas e da Terra, a qual é constituída por oito áreas. São elas: Química, Física, Ciência da Computação, Matemática, Geociências, Astronomia, Probabilidade e Estatística e Oceanografia.

A Química destaca-se como a área que representa a maior parcela de artigos de ambos os tipos. Mas, devido às oscilações, não é possível inferir qualquer tendência desta área.

Tabela 6: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Exatas e da Terra

Ciências Exatas e da Terra: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
QUÍMICA	25,67	29,97	26,73	22,13	10,85	8,90	8,46	10,79
FÍSICA	20,32	14,53	14,92	8,71	6,94	3,54	2,34	1,80
CIÊNCIA DA COMPUTAÇÃO	6,72	5,36	7,57	7,05	1,88	2,36	1,56	2,49
MATEMÁTICA	5,28	4,54	5,68	2,77	4,41	1,82	3,12	3,46
GEOCIÊNCIAS	3,33	7,27	10,13	11,20	3,18	7,99	5,57	8,44
ASTRONOMIA	1,66	2,63	4,68	8,30	0,87	0,82	1,56	4,98
PROBABILIDADE E ESTATÍSTICA	1,52	2,45	0,67	1,11	0,36	0,91	0,33	0,41
OCEANOGRAFIA	0,94	0,73	0,33	0,28	1,01	0,45	0,67	0,83
Sem área	4,77	5,45	5,57	4,70	0,29	0,27	0,11	0,55
TOTAL %	70,21	72,93	76,28	66,25	29,79	27,07	23,72	33,75
TOTAL N°	971	803	685	479	412	289	213	244

Fonte: Dados da pesquisa

No que se refere a artigos com mais de uma área, de maneira geral, o conjunto de áreas de Ciências Exatas e da Terra registra oscilação no número de artigos deste tipo ao longo do período estudado.

A área de Física, entretanto, revela tendência diferente. Verifica-se significativa redução percentual, de 6,94% em 2001-2003 para 1,80% em 2010-2012, da fração de artigos envolvendo mais de uma área, assim como também diminuição nos artigos com uma área.

Cabe destacar também a tendência de crescimento, apresentado pela área de Astronomia, nos artigos com mais de uma área, de 0,87% no primeiro triênio para 4,98% no último.

5.2.1.4 As áreas nas Engenharias

O conjunto das Engenharias registrou, no período de 2001-2012, um total de 4.161 artigos publicados em periódicos, dos quais 3.224 (77,48%) são artigos com uma área e 937 (22,52%) artigos são associados a mais de uma área do conhecimento.

A Tabela 7 mostra a distribuição do total desses artigos, em termos percentuais, pelas respectivas áreas que compõem a grande área Engenharias, que somam doze áreas: Engenharia de Materiais e Metalurgia, Engenharia Química, Engenharia de Produção, Engenharia Mecânica, Engenharia Elétrica, Engenharia Nuclear, Engenharia Civil, Engenharia Biomédica, Engenharia Naval e Oceânica, Engenharia Sanitária, Engenharia de Transportes e Engenharia Aeroespacial.

Verifica-se que a área com maior volume de artigos, Engenharia de Materiais e Metalurgia, apresenta tendência de diminuição do percentual de artigos com uma área (de 23,94% em 2001-2003, para 17,84% em 2010-2012) e, por outro lado, aumento do percentual de artigos com mais de uma área (de 2,03% em 2001, para 5,56% em 2010-2012).

Tabela 7: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Engenharias

Engenharias: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
ENGENHARIA DE MATERIAIS E METALÚRGICA	23,94	19,49	18,14	17,84	2,03	2,92	1,97	5,56
ENGENHARIA QUÍMICA	14,51	20,81	18,23	12,91	5,45	6,78	5,26	4,09
ENGENHARIA DE PRODUÇÃO	9,15	5,56	5,73	4,09	1,85	1,60	1,22	0,73
ENGENHARIA MECÂNICA	7,30	6,12	8,08	5,35	1,85	1,79	3,20	2,52
ENGENHARIA ELÉTRICA	6,65	5,56	6,20	7,24	2,22	4,52	3,67	8,92
ENGENHARIA NUCLEAR	4,71	4,52	7,89	11,23	0,55	1,51	0,56	0,42
ENGENHARIA CIVIL	4,53	4,24	2,54	0,84	0,83	1,13	1,69	1,47
ENGENHARIA BIOMÉDICA	3,42	4,52	6,95	8,08	2,13	0,66	1,41	1,47
ENGENHARIA NAVAL E OCEÂNICA	1,57	2,07	2,63	1,26	0,55	0,09	0,19	0,21
ENGENHARIA SANITÁRIA	1,20	0,28	1,22	0,94	0,55	0,38	1,22	0,31
ENGENHARIA DE TRANSPORTES	0,83	1,13	0,47	0,21	0,18	0,19	0,19	0,10
ENGENHARIA AEROESPACIAL	0,00	0,19	0,00	0,00	0,00	0,09	0,00	0,00
Sem área	3,70	3,48	0,66	0,94	0,28	0,38	0,66	3,25
TOTAL %	81,52	77,97	78,76	70,93	18,48	22,03	21,24	29,07
TOTAL N°	882	828	838	676	200	234	226	277

Fonte: Dados da pesquisa

As outras áreas, de modo geral, mostram oscilação, tanto nos artigos com uma área, como nos artigos vinculados a mais de uma área. Destacam-se, entretanto, Engenharia Nuclear, por apresentar clara tendência de crescimento nos artigos com uma área, e Engenharia Elétrica, por mostrar tendência de aumento de artigos com mais de uma área.

5.2.1.5 As áreas nas Ciências Humanas

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Ciências Humanas totalizou 4.349, dos quais 3.172 (72,94%) são artigos com uma área e 1.177 (27,06%) são artigos com mais de uma área.

A Tabela 8 mostra a distribuição do total desses artigos, em termos percentuais, pelas respectivas áreas que compõem Ciências Humanas. São elas:

Psicologia, História, Educação, Antropologia, Sociologia, Geografia, Filosofia, Ciência Política, Arqueologia e Teologia.

Observa-se que as áreas que registram maior fração de artigos com mais de uma área, no período estudado, são Psicologia, Educação e História, com percentuais oscilando entre 4-9%, 3-7% e 4-5%, respectivamente.

Tabela 8: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Humanas

Ciências Humanas: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
PSICOLOGIA	16,32	12,59	10,27	8,81	4,42	9,04	7,88	5,97
HISTÓRIA	9,94	12,06	14,87	19,80	4,05	5,06	3,98	4,27
EDUCAÇÃO	9,57	8,29	6,28	6,24	4,66	3,88	5,22	7,66
ANTROPOLOGIA	8,83	13,46	10,97	10,98	2,45	2,05	3,01	1,29
SOCIOLOGIA	6,87	5,60	6,11	3,93	1,35	1,94	1,50	1,76
GEOGRAFIA	5,77	3,98	1,50	4,34	0,98	1,94	1,33	1,29
FILOSOFIA	5,40	6,67	12,39	11,39	2,45	2,69	2,57	1,69
CIÊNCIA POLÍTICA	4,29	1,40	0,80	1,76	1,47	0,43	0,44	1,22
ARQUEOLOGIA	2,82	1,29	2,39	2,64	0,12	0,43	0,18	0,14
TEOLOGIA	0,00	0,11	0,00	0,00	0,12	0,11	0,09	0,00
Sem área	6,63	5,81	6,28	2,98	1,47	1,18	1,95	1,83
TOTAL %	76,44	71,26	71,86	72,88	23,56	28,74	28,14	27,12
TOTAL N°	623	662	812	1.075	192	267	318	400

Fonte: Dados da pesquisa

No que se refere a artigos com uma área, cabe ressaltar o aumento apresentado na História (de 9,94% em 2001-2003, para 19,80% em 2010-2012) e Filosofia (de 5,40% em 2001-2003, para 11,39% em 2010-2012).

5.2.1.6 As áreas na Linguística, Letras e Artes

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Linguística, Letras e Artes totalizou 2.617, dos quais 2.052 (78,41%) são artigos com uma área e 565 (21,59%) são artigos com mais de uma área.

A distribuição do total desses artigos, em termos percentuais, pelas respectivas áreas que compõem Linguística, Letras e Artes, é mostrada na Tabela 9.

Tabela 9: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Linguística, Letras e Artes

Linguística, Letras e Artes: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
LETRAS	49,18	61,73	46,52	43,26	8,56	9,96	11,82	15,50
ARTES	15,85	4,19	8,79	7,75	2,91	4,72	4,24	8,53
LINGUÍSTICA	11,11	13,63	12,42	12,71	2,00	2,10	5,76	3,72
Sem área	7,47	3,41	9,09	6,51	2,91	0,26	1,36	2,02
TOTAL %	83,61	82,96	76,82	70,23	16,39	17,04	23,18	29,77
TOTAL N°	459	633	507	453	90	130	153	192

Fonte: Dados da pesquisa

Nota-se que Letras é a área mais representativa em termos de volume de produção, tanto de artigos com mais de uma área, como de artigos com uma área apenas.

Em relação a artigos com mais de uma área, Letras e Artes revelam significativo crescimento, o primeiro de 8,56% em 2001-2003, para 15,50% em 2010-2012, e o último, de 2,91% em 2001-2003, para 8,53% em 2010-2012.

5.2.1.7 As áreas nas Ciências Sociais Aplicadas

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Ciências Sociais Aplicadas totalizou 1.933, dos quais 1.475 (76,31%) são artigos com uma área e 458 (23,69%) são artigos com mais de uma área.

A Tabela 10 mostra o total de artigos, em termos percentuais, pelas nove áreas que compõem Ciências Sociais Aplicadas, as quais são: Economia; Planejamento Urbano e Regional; Comunicação; Administração; Arquitetura e

Urbanismo; Serviço Social; Direito; Ciência da Informação; Turismo; Desenho Industrial e Museologia.

Tabela 10: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Sociais Aplicadas

Ciências Sociais Aplicadas: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
ECONOMIA	14,29	7,19	5,91	9,49	6,25	4,02	3,11	1,90
PLANEJAMENTO URBANO E REGIONAL	13,39	13,95	8,55	7,05	3,35	1,48	1,24	1,90
COMUNICAÇÃO	11,16	20,08	6,69	8,40	6,47	8,67	4,35	1,63
ADMINISTRAÇÃO	8,04	10,36	22,04	27,10	0,22	1,48	1,71	1,63
ARQUITETURA E URBANISMO	6,47	5,71	8,09	15,99	2,23	2,11	2,33	1,63
SERVIÇO SOCIAL	2,68	7,40	12,91	2,98	0,45	1,06	1,87	0,54
DIREITO	1,56	4,44	5,29	3,25	2,23	3,17	2,18	2,98
CIÊNCIA DA INFORMAÇÃO	0,00	0,42	0,93	1,63	1,34	0,63	3,42	3,79
TURISMO	0,22	0,00	0,00	0,00	0,00	0,21	0,16	1,90
DESENHO INDUSTRIAL	0,00	0,00	7,75	0,00	0,00	0,00	0,31	0,54
MUSEOLOGIA	0,00	0,00	0,00	0,00	0,00	0,00	0,16	0,27
Sem área	17,86	5,07	0,22	3,79	1,79	2,54	0,78	1,63
TOTAL %	75,67	74,63	78,38	79,67	24,33	25,37	21,62	20,33
TOTAL N°	339	353	489	294	109	120	154	75

Fonte: Dados da pesquisa

Podemos notar que, de modo geral, os percentuais de artigos com mais de uma área oscilam em praticamente todas essas áreas. A exceção é Economia, que apresenta significativa queda (de 6,25% em 2001-2003, para 1,90% em 2010-2012) na fração de artigos deste tipo.

Em relação a artigos com uma área, destacam-se pelo significativo aumento nos percentuais de artigos deste tipo, Administração (de 8% no primeiro triênio, para 27% no último) e Arquitetura e Urbanismo (de 6% em 2001-2003, para 15% em 2010-2012).

5.2.1.8 As áreas nas Ciências Agrárias

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Ciências Agrárias totalizou 242, dos quais 191 (78,93%) são artigos com uma área e 51 (21,07%) são artigos com mais de uma área.

A Tabela 11 mostra, em termos percentuais, a distribuição do total de artigos, considerando as respectivas áreas que compõem Ciências Agrárias, as quais são seis: Ciência e Tecnologia de Alimentos; Agronomia; Recursos Florestais e Engenharia Florestal; Medicina Veterinária; Recursos Pesqueiros e Engenharia de Pesca e Zootecnia. Observa-se que Ciência e Tecnologia de Alimentos é a área mais representativa, em termos de volume de produção, em ambos os tipos de artigos.

Tabela 11: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas de Ciências Agrárias

Ciências Agrárias	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
CIÊNCIA E TECNOLOGIA DE ALIMENTOS	55,81	64,47	77,63	63,83	13,95	18,42	2,63	17,02
AGRONOMIA	6,98	0,00	5,26	4,26	0,00	7,89	5,26	0,00
RECURSOS FLORESTAIS E ENGENHARIA FLORESTAL	4,65	2,63	0,00	0,00	4,65	3,95	0,00	2,13
MEDICINA VETERINÁRIA	0,00	1,32	1,32	0,00	2,33	1,32	1,32	0,00
RECURSOS PESQUEIROS E ENGENHARIA DE PESCA	0,00	0,00	0,00	4,26	0,00	0,00	0,00	0,00
ZOOTECNIA	0,00	0,00	0,00	0,00	0,00	0,00	1,32	0,00
Sem área	9,30	0,00	5,26	8,51	2,33	0,00	0,00	0,00
TOTAL %	76,74	68,42	89,47	80,85	23,26	31,58	10,53	19,15
TOTAL N°	33	52	68	38	10	24	8	9

Fonte: Dados da pesquisa

No tocante a artigos associados a uma única área, Ciência e Tecnologia de Alimentos registra mais de 50% dos artigos desse tipo em todos os períodos, chegando a 77% em 2007-2009.

Já em relação a artigos com mais de uma área, após apresentar crescimento de 2001-2003 para 2004-2006, de 13% do total de artigos, para 18%, respectivamente, Ciência e Tecnologia de Alimentos registra significativa queda em

2007-2009 (representando nesse período apenas 2%), voltando ao patamar de períodos anteriores no último triênio (agora com 17% do total de artigos).

5.2.1.9 As áreas nas Outras (Grandes Áreas)

O número de artigos publicados em periódicos, no período de 2001-2012, relacionados à grande área Outros totalizou 161, dos quais 80 (49,69%) são artigos com uma área e 81 (50,31%) são artigos com mais de uma área.

A Tabela 12 mostra o total de artigos, em termos percentuais, pelas duas áreas que compõem Outros: Multidisciplinar e Ensino.

Multidisciplinar responde por volta de 20% dos artigos com uma área apenas, com exceção do triênio 2004-2006, em que duplicou esse percentual. Entretanto, curiosamente, em relação a artigos com mais de uma área, essa área apresenta significativo decréscimo (de 48% do total de artigos em 2001-2003, para 12% em 2001-2012).

Tabela 12: UFRJ – Total e percentual de artigos com indicação de uma ou mais áreas segundo as áreas nas Outras (Grandes Áreas)

Outros: Áreas	Número de Artigos							
	Artigos com uma área				Artigos com mais de uma área			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
MULTIDISCIPLINAR	20,69	41,03	22,06	20,00	48,28	30,77	27,94	12,00
ENSINO	3,45	12,82	35,29	32,00	17,24	10,26	14,71	24,00
Sem área	0,00	0,00	0,00	0,00	10,34	5,13	0,00	12,00
TOTAL %	24,14	53,85	57,35	52,00	75,86	46,15	42,65	48,00
TOTAL N°	7	21	39	13	22	18	29	12

Fonte: Dados da pesquisa

Por outro lado, Ensino registra crescimento significativo de artigos com uma área (de 3% do total de artigos em 2001-2003, para 32% em 2010-2012) e, em relação a artigos com mais de uma área, apresenta oscilação nos três primeiros triênios e significativo crescimento no último.

Essas comunidades cumprem papel relevante nas propriedades de estruturas complexas, como a observada na Figura 16, que mostra a estrutura complexa da interação das 75 áreas do conhecimento, descritas anteriormente, a partir das informações da produção científica da UFRJ, especificamente de artigos publicados em periódicos no período de 2001 a 2012.

Identificar e analisar a natureza das comunidades ai ocultas é uma importante tarefa para revelar a organização informal e a natureza dos fluxos de informação que acontecem dentro desses sistemas complexos.

Importante destacar que tais comunidades são compostas por diferentes subconjuntos de nós, os quais se caracterizam por estarem densamente ligados mais entre si do que em relação ao resto da rede. A identificação dessas comunidades facilita a compreensão das redes.

Diversas abordagens foram desenvolvidas para identificar comunidades, nas quais se adotam funções específicas para descrever características dos grupos a serem determinados. Portanto, a escolha do algoritmo mais adequado para cada tipo de problema determinará o melhor caminho possível para obtenção de resultados (BARBOSA *et al.*, 2011).

A maioria dos algoritmos para identificação de comunidades utiliza técnicas de agrupamento hierárquico, os quais, além de mostrar os diferentes componentes da rede, possibilitam extrair sua hierarquia. Este é o caso, por exemplo, do algoritmo proposto por Clauset *et al.* (2004), que faz uma busca gulosa em um mapa hierárquico (*dendograma*) a fim de encontrar a divisão da rede que maximiza a medida de modularidade¹⁸, e o proposto por Blondel *et al.* (2008), que apresenta algoritmo similar, porém considerando uma otimização de modularidade local, ou seja, a cada nível do dendograma, nós (vértices) são agrupados com o vizinho local que mais contribui para a variação positiva da medida de modularidade.

O algoritmo utilizado pelo *software Gephi* para a identificação de comunidades, proposto por Blondel *et al.* (2008), é utilizado nas análises que serão

¹⁸ Medida que avalia o agrupamento do grafo, avaliando o número de conexões dentro de um agrupamento do grafo original.

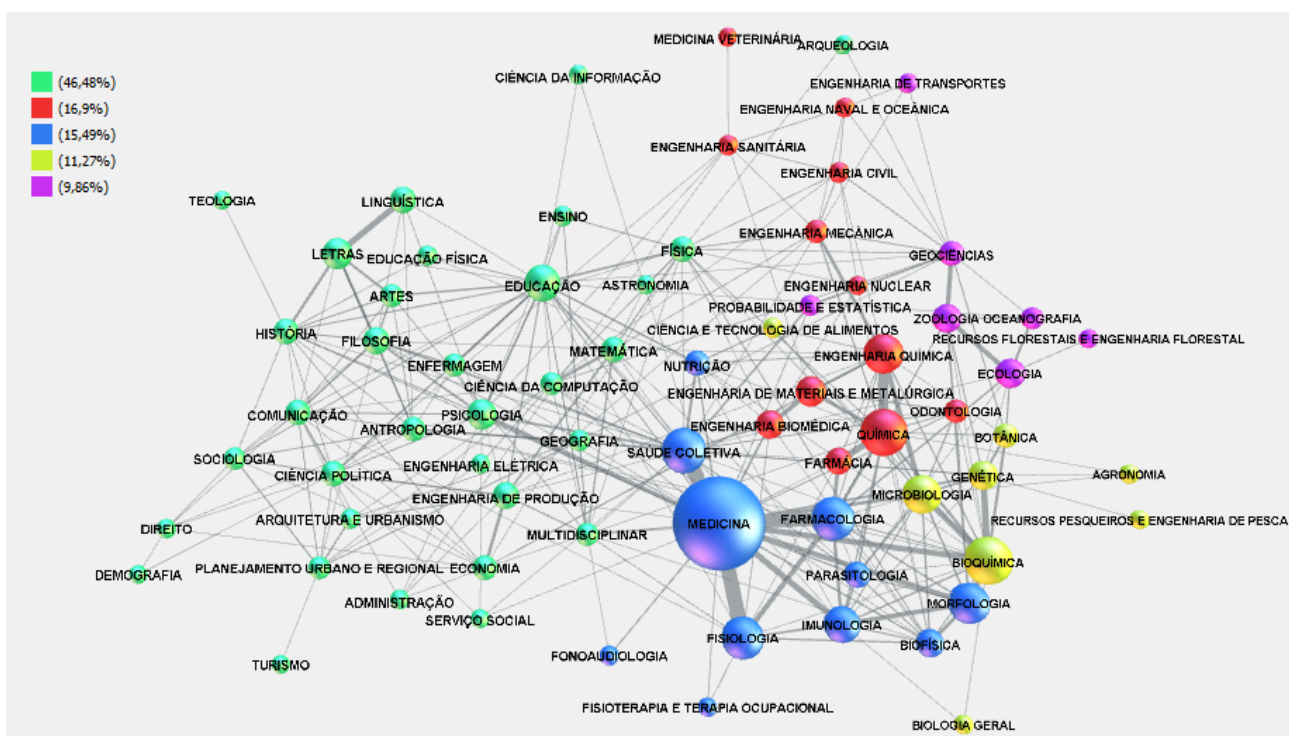
apresentadas a seguir. Como o objetivo é observar eventuais mudanças destas comunidades, os dados são apresentados em dois triênios: 2001-2003 e 2010-2012.

5.2.2.1 Comunidades e Principais Áreas Identificadas no Triênio 2001-2003

A rede de colaboração entre áreas do conhecimento, no período de 2001 a 2003, contém 71 nós (71 áreas), e 310 arestas (relacionamentos). A Figura 17 mostra a rede, destacando as cinco comunidades identificadas, as quais são caracterizadas por cores distintas.

Para melhor compreensão das características de cada comunidade, analisaremos cada uma delas, considerando as medidas de centralidade e, assim, identificar as áreas mais representativas.

Figura 17: Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos, UFRJ (2001-2003)



Fonte: Elaborado pelo autor

A análise parte das comunidades maiores (que contem maior número de áreas) para as comunidades menores, conforme indicado pelos percentuais de cada cor na Figura 17.

5.2.2.1.a A Comunidade 1

Esta comunidade, identificada com a cor verde na Figura 17, é composta por trinta e três áreas do conhecimento (o que representa 46,48% do total das áreas), sendo, portanto, a maior do período 2001-2003. A Tabela 13 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente da medida de grau. É possível perceber que nesta comunidade predominam áreas das humanidades e das ciências sociais.

Tabela 13: Relação de áreas e medidas de centralidade da Comunidade 1, identificada em 2001-2003

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● EDUCAÇÃO	26	0,133	0,583	0,874
● FÍSICA	18	0,076	0,543	0,64
● ANTROPOLOGIA	17	0,059	0,511	0,599
● CIÊNCIA POLÍTICA	15	0,061	0,504	0,502
● FILOSOFIA	15	0,03	0,515	0,567
● MULTIDISCIPLINAR	15	0,027	0,515	0,654
● HISTÓRIA	14	0,047	0,496	0,538
● COMUNICAÇÃO	14	0,02	0,473	0,49
● PSICOLOGIA	14	0,029	0,507	0,556
● CIÊNCIA DA COMPUTAÇÃO	13	0,044	0,511	0,532
● ECONOMIA	12	0,028	0,467	0,394
● PLANEJAMENTO URBANO E REGIONAL	11	0,037	0,424	0,309
● MATEMÁTICA	11	0,017	0,504	0,473
● ENGENHARIA DE PRODUÇÃO	11	0,018	0,496	0,426
● SOCIOLOGIA	9	0,004	0,446	0,356
● ARTES	7	0,001	0,424	0,316
● LETRAS	7	0,001	0,419	0,285
● ENFERMAGEM	6	0	0,446	0,348
● GEOGRAFIA	6	0,008	0,429	0,181
● DIREITO	6	0,005	0,405	0,198
● ARQUITETURA E URBANISMO	5	0,001	0,389	0,156
● LINGÜÍSTICA	5	0	0,414	0,227
● ADMINISTRAÇÃO	5	0,003	0,414	0,188
● ENSINO	5	0,002	0,429	0,229
● SERVIÇO SOCIAL	4	0	0,429	0,228
● ENGENHARIA ELÉTRICA	3	0	0,378	0,108
● EDUCAÇÃO FÍSICA	3	0	0,402	0,147
● ASTRONOMIA	2	0	0,359	0,088
● DEMOGRAFIA	2	0	0,34	0,057
● ARQUEOLOGIA	2	0	0,368	0,077
● CIÊNCIA DA INFORMAÇÃO	2	0	0,409	0,146
● TURISMO	1	0	0,299	0,025
● TEOLOGIA	1	0	0,333	0,042

Fonte: Dados da pesquisa

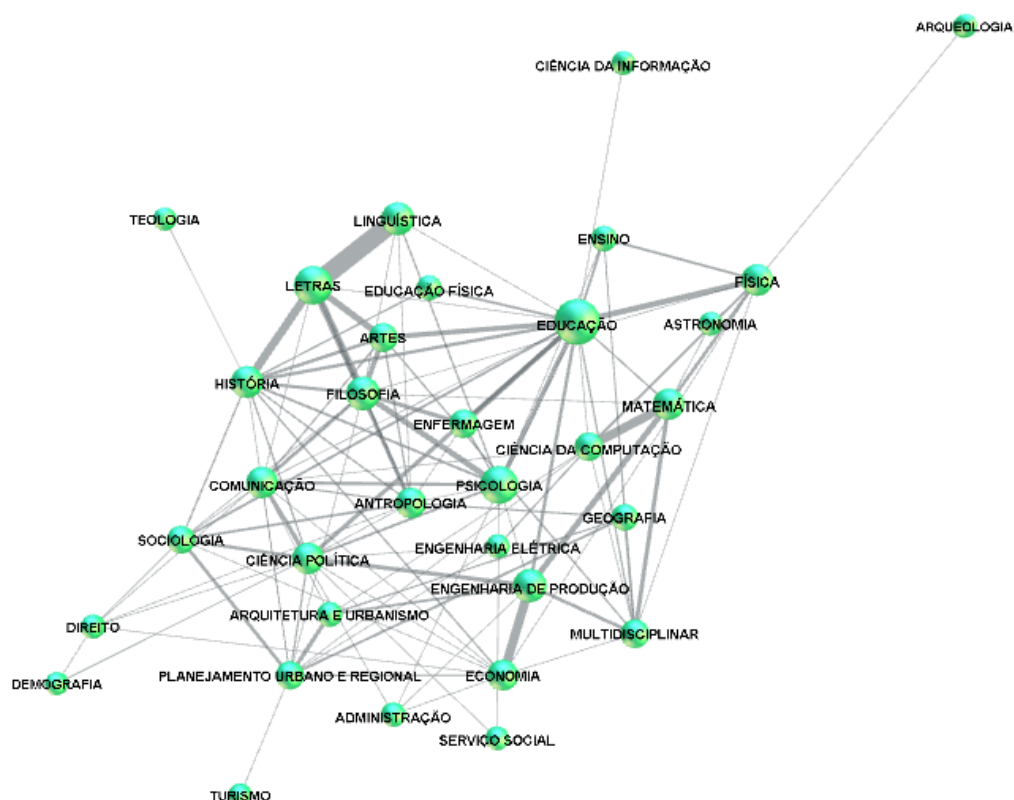
Como mostra a Tabela 13, todas as medidas indicam que Educação é a área mais central. Física e Antropologia apresentam o segundo e terceiro maior alto grau em relação às demais. Isso significa que tais áreas são mais ativas e possuem maior número de artigos com mais de uma área. Física está estruturalmente bem posicionada na rede, segundo a centralidade de proximidade (*Closeness Centrality*), possuindo, depois da área de Educação, mais chances que as demais de aumentar o número de artigos com mais de uma área. Outros resultados merecem destaque.

Podemos observar, por exemplo, que Ciência Política possui uma centralidade de grau menor que Antropologia, no entanto, em relação à centralidade de intermediação (*Betweenness Centrality*), é considerada a terceira área mais relevante, ou seja, possui melhor posição estratégica (de intermediador) entre áreas do que Antropologia.

Verifica-se, por outro lado, que a área Multidisciplinar possui uma centralidade de grau menor que as áreas acima mencionadas e, no entanto, no que se refere à centralidade do autovetor (*Eigenvector Centrality*), é considerada a segunda mais relevante, ou seja, tem mais ligações com áreas que possuem maior grau de centralidade (o que pode indicar a importância de um determinado nó).

A Figura 18 permite visualizar a rede que dá forma à Comunidade 1, onde os símbolos maiores sinalizam áreas mais centrais ou de maior peso.

Figura 18: Grafo que destaca a Comunidade 1, identificada em 2001-2003



Fonte: Elaborado pelo autor

Nota-se que as interações mais fortes dentro dessa comunidade, pela largura dos vínculos (arestas), acontecem entre as áreas de Letras e Linguística; Matemática e Ciências da Computação; Engenharia de Produção e Economia.

5.2.2.1.b A Comunidade 2

Esta comunidade, identificada com a cor vermelha na Figura 17, é composta por doze áreas (o que representa 16,9% do total das áreas), sendo, portanto, a segunda maior do triênio 2001-2003. A Tabela 14 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente de grau. É fácil observar que esta comunidade tem alta relação com as Engenharias.

Tabela 14: Relação de áreas e medidas de centralidade da Comunidade 2 identificada em 2001-2003

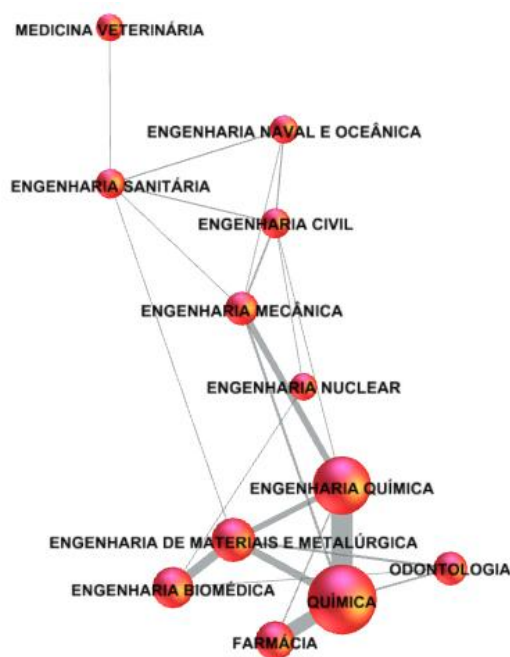
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● QUÍMICA	17	0,038	0,464	0,594
● ENGENHARIA QUÍMICA	14	0,034	0,476	0,443
● ENGENHARIA MECÂNICA	9	0,015	0,429	0,271
● ENGENHARIA BIOMÉDICA	9	0,01	0,455	0,378
● ENGENHARIA CIVIL	8	0,012	0,432	0,212
● ODONTOLOGIA	8	0,005	0,443	0,368
● ENGENHARIA SANITÁRIA	8	0,041	0,414	0,167
● FARMÁCIA	6	0,002	0,446	0,297
● ENGENHARIA DE MATERIAIS E METALÚRGICA	6	0,009	0,419	0,187
● ENGENHARIA NAVAL E OCEÂNICA	5	0,003	0,376	0,09
● ENGENHARIA NUCLEAR	4	0,001	0,389	0,149
● MEDICINA VETERINÁRIA	1	0	0,294	0,015

Fonte: Dados da pesquisa

Nota-se na Tabela 14 que Química e Engenharia Química são as áreas que apresentam maior centralidade de grau, ou seja, que possuem o maior número de ligações com outras áreas, com destaque para a primeira. Engenharia Química, entretanto, apresenta maior centralidade de proximidade (*Closeness Centrality*), o que revela mais chances de interagir com as outras áreas.

Observa-se, por outro lado, que Engenharia Sanitária, embora apresente um grau menor que Química e Engenharia Química, mostra a maior centralidade de intermediação, o que indica que tem maior posição estratégica, sendo a única responsável, por exemplo, pela troca/controla de informações entre Medicina Veterinária e as demais áreas do grafo em que está representada a Comunidade 2 (Figura 19).

Figura 19: Grafo que destaca a Comunidade 2 identificada no período 2001-2003



Fonte: Elaborado pelo autor

A Figura 19 mostra, também, que os vínculos mais fortes (observado pela largura das arestas) nessa comunidade se dão entre as áreas de Química, Engenharia Química e Farmácia; Engenharia Biomédica e Engenharia de Materiais e Metalurgia; assim como também entre Engenharia Química e Engenharia Mecânica.

5.2.2.1.c A Comunidade 3

Esta comunidade, identificada com a cor azul na Figura 17, é composta por onze áreas (o que representa 15,49% do total das áreas), sendo, portanto, a terceira maior do triênio 2001-2003. A Tabela 15 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente de grau. Observa-se que nesta comunidade prevalecem áreas relacionadas à saúde.

Tabela 15: Relação de áreas e medidas de centralidade da Comunidade 3 identificada em 2001-2003

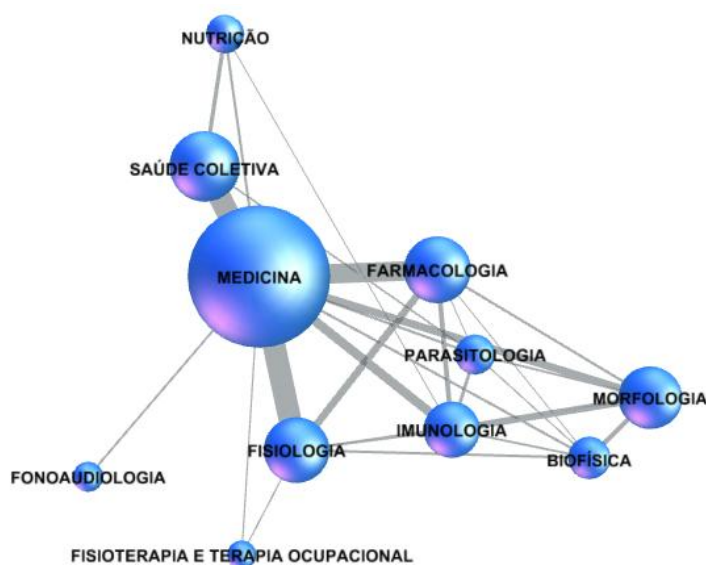
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● MEDICINA	27	0,135	0,583	1
● SAÚDE COLETIVA	22	0,104	0,579	0,854
● FARMACOLOGIA	15	0,02	0,496	0,63
● BIOFÍSICA	15	0,021	0,493	0,652
● IMUNOLOGIA	13	0,016	0,479	0,584
● FISILOGIA	12	0,016	0,483	0,538
● MORFOLOGIA	10	0,002	0,44	0,45
● PARASITOLOGIA	10	0,005	0,446	0,487
● NUTRIÇÃO	10	0,016	0,476	0,456
● FISIOTERAPIA E TERAPIA OCUPACIONAL	2	0	0,374	0,119
● FONOAUDIOLOGIA	2	0	0,409	0,146

Fonte: Dados da pesquisa

Verifica-se que todas as métricas da Tabela 15 apontam Medicina, destacadamente, como a área mais central, seguida de Saúde Coletiva.

Biofísica possui a terceira maior centralidade de intermediação e do autovetor (*Eigenvector Centrality*), ou seja, tem papel destacado como ponte do fluxo de informações entre áreas e está ligada a áreas mais representativas (em termos de maior conexão), além de possuir também, junto com Farmacologia, a terceira maior centralidade de grau.

Figura 20: Grafo que destaca a Comunidade 3 identificada no período 2001-2003



Fonte: Elaborado pelo autor

Na Figura 20 vemos a representação da Comunidade 3. Nota-se que as interações mais fortes nessa comunidade acontecem entre as áreas de Medicina, Fisiologia, Farmacologia e Saúde Coletiva.

5.2.2.1.d A Comunidade 4

A comunidade 4, identificada com a cor amarela na Figura 17, é composta por oito áreas (o que representa 15,49% do total das áreas do período 2001-2003). A Tabela 16 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente de grau. Esta comunidade mostra tendência temática em áreas da biologia e biomedicina.

Tabela 16: Relação de áreas e medidas de centralidade da Comunidade 4 identificada em 2001-2003

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● BIOQUÍMICA	20	0,054	0,526	0,783
● GENÉTICA	12	0,05	0,47	0,457
● MICROBIOLOGIA	12	0,011	0,452	0,477
● BOTÂNICA	6	0,006	0,419	0,225
● CIÊNCIA E TECNOLOGIA DE ALIMENTOS	4	0	0,387	0,169
● AGRONOMIA	3	0	0,366	0,12
● BIOLOGIA GERAL	2	0	0,352	0,106
● RECURSOS PESQUEIROS E ENGENHARIA DE PESCA	1	0	0,321	0,036

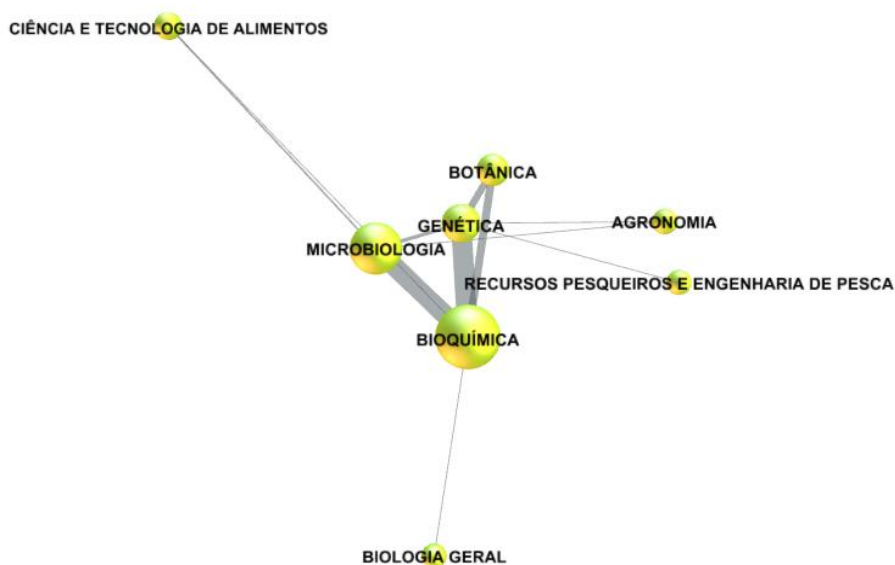
Fonte: Dados da pesquisa

Podemos observar que todas as medidas apresentadas na Tabela 16 situam Bioquímica como a área mais central dessa comunidade. Genética é a segunda com maior grau, proximidade e intermediação.

Entretanto, no que se refere à centralidade do autovetor (*Eigenvector Centrality*), Microbiologia é a que apresenta a segunda maior medida, colocando-se, depois de Bioquímica, como a área mais importante dessa comunidade, em termos de interação com áreas de maior grau (com maiores relações).

A Figura 21 mostra a representação da Comunidade 4. Podemos verificar que as ligações mais fortes acontecem entre Bioquímica e as áreas anteriormente mencionadas, e também com a área de Botânica.

Figura 21: Grafo que destaca a Comunidade 4 identificada no período 2001-2003



Fonte: Elaborado pelo autor

5.2.2.1.e A Comunidade 5

A última comunidade identificada, caracterizada pela cor roxa na Figura 17, é composta por sete áreas (o que representa 9,86% do total das áreas do período 2001-2003). A Tabela 17 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente de grau. Esta comunidade mostra múltipla tendência temática.

Tabela 17: Relação de áreas e medidas de centralidade da Comunidade 5, identificada em 2001-2003

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● GEOCIÊNCIAS	13	0,05	0,486	0,308
● ECOLOGIA	8	0,032	0,419	0,244
● ZOOLOGIA	8	0,01	0,417	0,285
● OCEANOGRAFIA	4	0,004	0,424	0,137
● PROBABILIDADE E ESTATÍSTICA	3	0,001	0,419	0,17
● ENGENHARIA DE TRANSPORTES	3	0,002	0,395	0,101
● RECURSOS FLORESTAIS E ENGENHARIA FLORESTAL	1	0	0,297	0,02

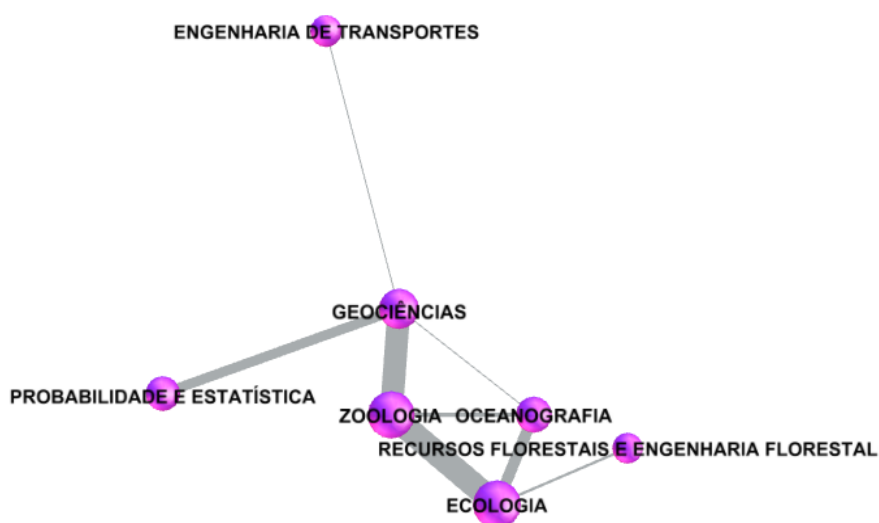
Fonte: Dados da pesquisa

Destaca-se, nesse grupo, a área de Geociências, que apresenta os maiores valores de todas as medidas de centralidade. Ecologia e Zoologia possuem a segunda e terceira maior centralidade de intermediação.

Entretanto, é Oceanografia que possui a segunda maior centralidade de proximidade, o que indica que esta área, depois de Geociência, é a que tem mais possibilidades de aumentar o número de trabalhos em colaboração com outras áreas.

Na Figura 22 vemos a representação da Comunidade 5. Podemos notar, por outro lado, que os vínculos mais fortes (observados pela largura das arestas), nessa comunidade, acontecem entre as áreas de Zoologia e Ecologia e, Geociências e Zoologia.

Figura 22: Grafo que destaca a Comunidade 5 identificada no período 2001-2003



Fonte: Elaborado pelo autor

5.2.2.1.f *Interação entre as Comunidades*

Para identificar as áreas que vinculam as comunidades identificadas, isolamos as áreas que possuem apenas arestas entre elas (comunidades). Dessa forma, identificamos vinte seis áreas como principais responsáveis por conectar as cinco comunidades observadas no período de 2001-2003.

A Tabela 18, que mostra a relação dessas áreas e suas respectivas medidas de centralidade, destaca Medicina e Educação como principais áreas, tanto em termos de conexão (grau), de intermediação (troca de informações com áreas pouco conectadas no grafo), proximidade, tendo maiores chances de aumentar os trabalhos com mais de uma área, e vínculo com os nós mais representativos em termos de grau (o que indica importância do nó).

Tabela 18: Relação de áreas e medidas de centralidade responsáveis por conectar as comunidades identificadas em 2001-2003

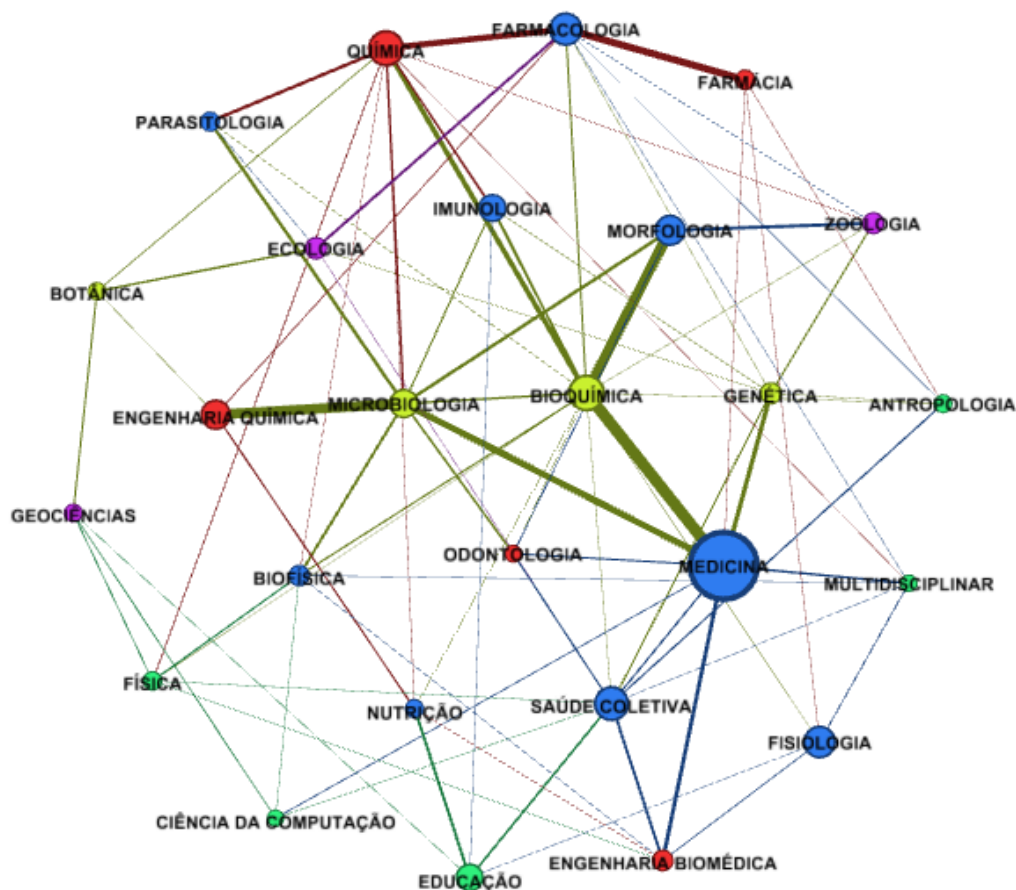
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● MEDICINA	27	0,135	0,583	1
● EDUCAÇÃO	26	0,133	0,583	0,874
● SAÚDE COLETIVA	22	0,104	0,579	0,854
● BIOQUÍMICA	20	0,054	0,526	0,783
● FÍSICA	18	0,076	0,543	0,64
● QUÍMICA	17	0,038	0,464	0,594
● ANTROPOLOGIA	17	0,059	0,511	0,599
● FARMACOLOGIA	15	0,02	0,496	0,63
● BIOFÍSICA	15	0,021	0,493	0,652
● MULTIDISCIPLINAR	15	0,027	0,515	0,654
● ENGENHARIA QUÍMICA	14	0,034	0,476	0,443
● IMUNOLOGIA	13	0,016	0,479	0,584
● CIÊNCIA DA COMPUTAÇÃO	13	0,044	0,511	0,532
● GEOCIÊNCIAS	13	0,05	0,486	0,308
● GENÉTICA	12	0,05	0,47	0,457
● MICROBIOLOGIA	12	0,011	0,452	0,477
● FISILOGIA	12	0,016	0,483	0,538
● MORFOLOGIA	10	0,002	0,44	0,45
● PARASITOLOGIA	10	0,005	0,446	0,487
● NUTRIÇÃO	10	0,016	0,476	0,456
● ENGENHARIA BIOMÉDICA	9	0,01	0,455	0,378
● ECOLOGIA	8	0,032	0,419	0,244
● ZOOLOGIA	8	0,01	0,417	0,285
● ODONTOLOGIA	8	0,005	0,443	0,368
● BOTÂNICA	6	0,006	0,419	0,225
● FARMÁCIA	6	0,002	0,446	0,297

Fonte: Dados da pesquisa

Por outro lado, cabe notar, por exemplo, que as áreas Multidisciplinar e Ciências da Computação possuem centralidade de grau que as situam em posição intermediária e, no entanto, em relação à centralidade de proximidade, colocam-se entre o grupo mais destacado, possuindo chances de aumento de maior interação (aumento do número de trabalhos em colaboração) com outras áreas, similares às de Antropologia, por exemplo.

A Figura 23 mostra o grafo das principais áreas que conectam as cinco comunidades identificadas no período de 2001-2003.

Figura 23: Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas no período 2001-2003



Fonte: Elaborado pelo autor

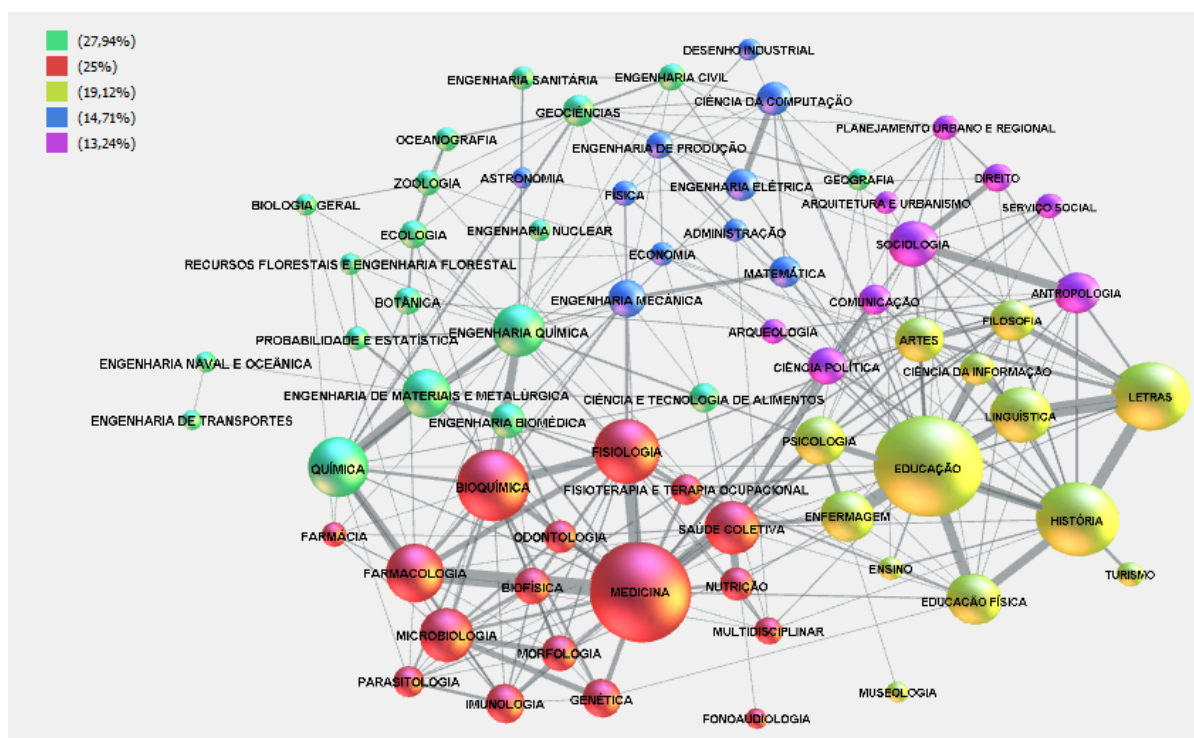
Pode-se notar na Figura 23 que, a comunidade 3 (cor azul), com nove áreas e que tem forte tendência temática no campo da saúde é a que mais contribui para a interação com as outras comunidades/áreas na UFRJ no triênio 2001-2003.

5.2.2.2 Comunidades e Principais Áreas no Triênio 2010-2012

A rede de colaboração entre áreas do conhecimento, no período de 2010 a 2012, contém 68 nós (68 áreas), três áreas a menos que no triênio anterior (2001-2003). Essas áreas estabeleceram 283 relacionamentos, arestas.

A Figura 24 mostra a visualização da rede, destacando as cinco comunidades identificadas em 2010-2012, mesmo número de comunidades identificadas no triênio anterior, as quais são caracterizadas por uma determinada cor.

Figura 24: Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos, UFRJ (2010-2012)



Fonte: Elaborado pelo autor

A seguir, analisamos cada uma das comunidades identificadas, considerando as medidas de centralidade, a fim de identificar as áreas mais representativas. Partimos das comunidades maiores (que contém maior número de áreas) para as comunidades menores, conforme indicado pelos percentuais de cada cor na Figura 24.

5.2.2.2.a A Comunidade 1

Identificada com a cor verde na Figura 24, esta comunidade é composta por dezenove áreas do conhecimento (o que representa 27,94% do total das áreas), sendo, portanto, a maior do período 2010-2012. A Tabela 19 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por

ordem decrescente de grau. Esta comunidade tem forte influência temática nas Engenharias.

Verifica-se que Engenharia Química apresenta o maior número de conexões, maior centralidade de grau e é também a área que possui a maior centralidade de intermediação, sendo, portanto, a que tem papel mais relevante em relação a sua posição de intermediador entre áreas menos conectadas na rede.

Tabela 19: Relação de áreas e medidas de centralidade da Comunidade 1, identificada em 2010-2012

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● ENGENHARIA QUÍMICA	19	0,101	0,496	0,537
● QUÍMICA	14	0,07	0,528	0,603
● GEOCIÊNCIAS	13	0,063	0,489	0,428
● ENGENHARIA DE MATERIAIS E METALÚRGICA	10	0,071	0,438	0,27
● ENGENHARIA BIOMÉDICA	9	0,028	0,459	0,342
● ECOLOGIA	8	0,022	0,456	0,282
● ENGENHARIA CIVIL	7	0,024	0,456	0,246
● CIÊNCIA E TECNOLOGIA DE ALIMENTOS	6	0,014	0,441	0,242
● BOTÂNICA	6	0,011	0,416	0,253
● ZOOLOGIA	6	0,012	0,424	0,204
● RECURSOS FLORESTAIS E ENGENHARIA FLORESTAL	5	0,007	0,409	0,139
● GEOGRAFIA	5	0,004	0,404	0,2
● PROBABILIDADE E ESTATÍSTICA	4	0,003	0,419	0,172
● BIOLOGIA GERAL	4	0,003	0,394	0,152
● ENGENHARIA NUCLEAR	3	0,001	0,406	0,117
● OCEANOGRAFIA	3	0	0,362	0,082
● ENGENHARIA SANITÁRIA	2	0	0,37	0,071
● ENGENHARIA NAVAL E OCEÂNICA	2	0,03	0,309	0,027
● ENGENHARIA DE TRANSPORTES	1	0	0,237	0,004

Fonte: Dados da pesquisa

A Química tem a segunda maior medida de grau e, no entanto, é a área com maior centralidade de proximidade e centralidade do autovetor, o que indica estar estruturalmente bem posicionada na rede (com maiores chances de interação) e ligada a áreas mais representativas, em termos de grau.

A área Engenharia de Matérias e Metalurgia, embora tenha menor medida de grau que as áreas anteriormente mencionadas, possui a segunda maior centralidade de intermediação, o que aponta ter papel relevante como intermediadora de informações entre áreas mais isoladas da rede.

A Figura 25 mostra a representação da Comunidade 1. Pode-se notar na Figura 25, que as áreas os vínculos mais fortes nessa Comunidade, observado pela

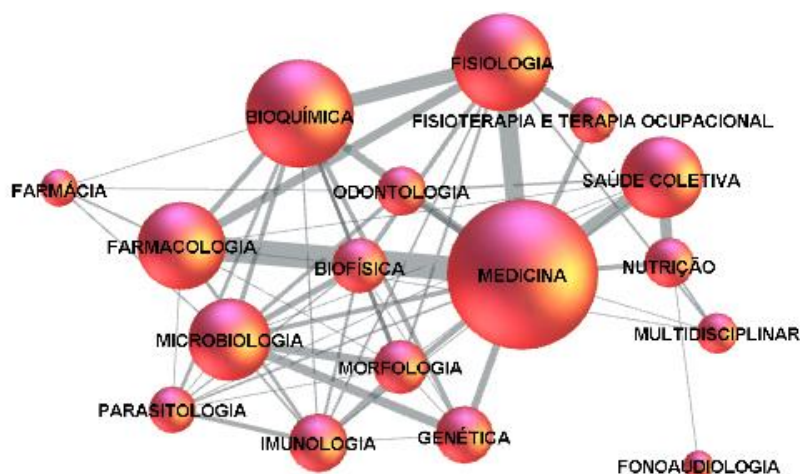
Tabela 20: Relação de áreas e medidas de centralidade da Comunidade 2, identificada em 2010-2012

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● MEDICINA	23	0,097	0,545	1
● SAÚDE COLETIVA	20	0,104	0,523	0,885
● BIOQUÍMICA	15	0,021	0,475	0,662
● FARMACOLOGIA	15	0,03	0,496	0,682
● FISILOGIA	13	0,02	0,479	0,582
● MICROBIOLOGIA	12	0,014	0,475	0,573
● IMUNOLOGIA	11	0,004	0,45	0,58
● MORFOLOGIA	10	0,004	0,421	0,51
● BIOFÍSICA	10	0,007	0,424	0,453
● PARASITOLOGIA	9	0,002	0,444	0,505
● MULTIDISCIPLINAR	9	0,021	0,493	0,441
● FARMÁCIA	7	0,003	0,409	0,304
● GENÉTICA	7	0,001	0,414	0,375
● ODONTOLOGIA	6	0,011	0,435	0,29
● NUTRIÇÃO	5	0,003	0,419	0,247
● FISIOTERAPIA E TERAPIA OCUPACIONAL	3	0,001	0,421	0,208
● FONOAUDIOLOGIA	1	0	0,345	0,074

Fonte: Dados da pesquisa

Farmacologia é a terceira área mais central dessa comunidade, de acordo com todas as medidas de centralidade (grau, intermediação, proximidade e centralidade do autovetor), mostradas na Tabela 20.

Figura 26: Grafo que destaca a Comunidade 2, identificada no período 2010-2012



Fonte: Elaborado pelo autor

A Figura 26 mostra a representação da Comunidade 2. Verifica-se que, no que concerne às interações mais fortes entre áreas dessa comunidade, estas acontecem entre Medicina e Farmacologia; Medicina e Fisiologia e; Medicina e Fisiologia.

5.2.2.2.c A Comunidade 3

Identificada com a cor amarela na Figura 24, esta comunidade é composta por treze áreas (representando 19,12% do total de áreas), sendo a terceira maior do período 2010-2012. A relação de áreas e respectivas medidas de centralidade que compõe esta comunidade, ordenadas por ordem decrescente de grau, é mostrada na Tabela 21. Como é possível notar, nesta comunidade prevalecem áreas das ciências humanas e sociais.

Tabela 21: Relação de áreas e medidas de centralidade da Comunidade 3, identificada em 2010-2012

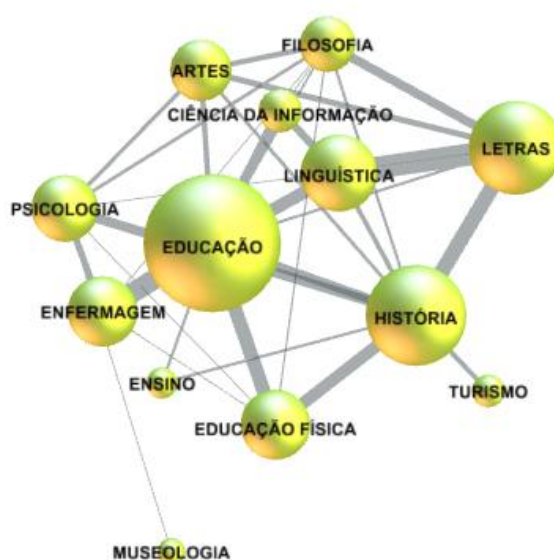
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● EDUCAÇÃO	24	0,147	0,549	0,859
● HISTÓRIA	19	0,092	0,523	0,779
● FILOSOFIA	14	0,02	0,493	0,736
● PSICOLOGIA	13	0,049	0,482	0,623
● EDUCAÇÃO FÍSICA	12	0,025	0,5	0,662
● LETRAS	11	0,011	0,469	0,607
● ARTES	10	0,02	0,469	0,515
● ENFERMAGEM	6	0,003	0,432	0,339
● LINGUÍSTICA	6	0,003	0,441	0,348
● CIÊNCIA DA INFORMAÇÃO	6	0,003	0,406	0,264
● ENSINO	3	0,001	0,39	0,153
● TURISMO	1	0	0,345	0,066
● MUSEOLOGIA	1	0	0,327	0,052

Fonte: Dados da pesquisa

Educação destaca-se como área com as mais altas medidas de centralidade, seguida por História. Filosofia dispõe da terceira melhor medida de grau, proximidade e centralidade do autovetor (*Eigenvector Centrality*), o que a coloca bem posicionada estruturalmente na rede. No entanto, em relação à medida de intermediação (*Betweenness Centrality*), Psicologia apresenta o terceiro maior valor, o que a situa entre as principais áreas da comunidade em termos de intermediador de áreas menos conectadas na rede (capacidade de ponte do fluxo de informação).

A Figura 27 mostra a representação da Comunidade 3. Observa-se que as ligações mais fortes (maior largura das arestas) acontecem entre as áreas de Letras e Linguística e Letras e História. Por outro lado, chama atenção também os fortes vínculos entre as áreas de Educação e Enfermagem, assim como também entre Educação Física e História.

Figura 27: Grafo que destaca a Comunidade 3, identificada no período 2010-2012



Fonte: Elaborado pelo autor

5.2.2.2.d A Comunidade 4

Esta comunidade, identificada com a cor azul na Figura 24, é constituída por dez áreas do conhecimento (o que representa 14,71% do total das áreas), sendo, portanto, a quarta do período 2010-2012. A Tabela 22 mostra a relação de áreas que compõem essa comunidade e as respectivas medidas de centralidade, ordenadas por ordem decrescente de grau. É possível identificar uma forte tendência de áreas das ciências exatas nesta comunidade.

Tabela 22: Relação de áreas e medidas de centralidade da Comunidade 4, identificada em 2010-2012

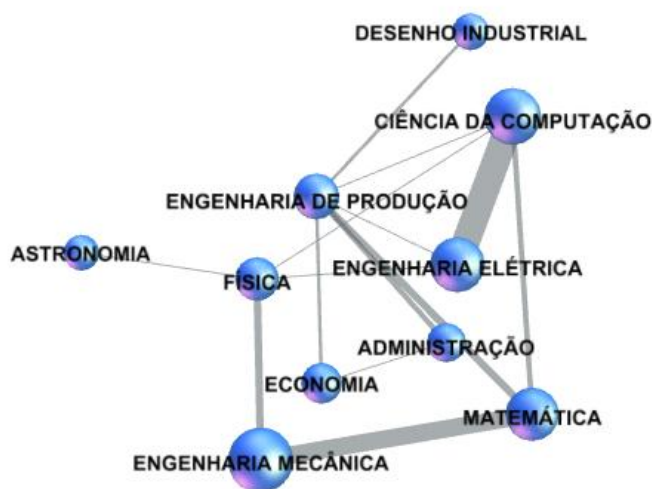
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● ENGENHARIA MECÂNICA	12	0,034	0,5	0,473
● CIÊNCIA DA COMPUTAÇÃO	9	0,018	0,441	0,248
● ECONOMIA	8	0,021	0,459	0,27
● ENGENHARIA DE PRODUÇÃO	7	0,016	0,409	0,151
● MATEMÁTICA	6	0,008	0,441	0,226
● FÍSICA	6	0,021	0,435	0,182
● ENGENHARIA ELÉTRICA	5	0,008	0,387	0,105
● ADMINISTRAÇÃO	4	0,002	0,394	0,152
● ASTRONOMIA	2	0,001	0,36	0,069
● DESENHO INDUSTRIAL	2	0,001	0,374	0,089

Fonte: Dados da pesquisa

Todas as medidas, mostradas na Tabela 22, apontam Engenharia Mecânica como área mais central. A área de Economia, embora tenha o terceiro maior grau, apresenta a segunda maior medida de proximidade e de centralidade do autovetor (*Eigenvector Centrality*).

No que concerne à medida de intermediação (*Betweenness Centrality*), Economia divide com Física a segunda maior medida, embora esta última possua centralidade de grau menor que a primeira, o que destaca a posição destas áreas de conectar outras ao restante da rede.

Figura 28: Grafo que destaca a Comunidade 4, identificada no período 2010-2012



Fonte: Elaborado pelo autor

A Figura 28 mostra a representação da Comunidade 4. Observa-se que os vínculos mais fortes dessa comunidade se dão entre as áreas de Ciência da Computação e Engenharia Elétrica e entre as áreas de Matemática e Engenharia Mecânica.

5.2.2.2.e A Comunidade 5

Identificada com a cor roxa na Figura 24, esta comunidade é constituída por nove áreas do conhecimento (o que representa 13,24% do total das áreas), sendo, portanto, a menor do período 2010-2012. A Tabela 23 mostra a relação de áreas dessa comunidade e as principais medidas de centralidade, ordenadas por ordem decrescente de grau. A Tabela 23 aponta também que esta comunidade é formada essencialmente por áreas das ciências sociais e humanas.

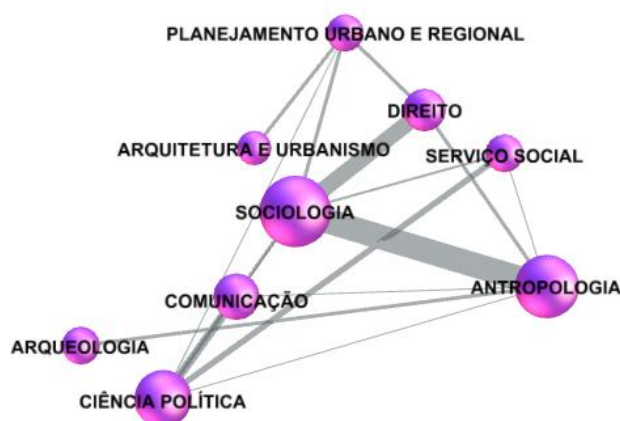
Tabela 23: Relação de áreas e medidas de centralidade da Comunidade 5, identificada em 2010-2012

Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● SOCIOLOGIA	14	0,017	0,489	0,704
● ANTROPOLOGIA	13	0,019	0,462	0,579
● COMUNICAÇÃO	12	0,014	0,475	0,641
● CIÊNCIA POLÍTICA	11	0,012	0,479	0,605
● PLANEJAMENTO URBANO E REGIONAL	9	0,011	0,435	0,361
● DIREITO	5	0,005	0,381	0,213
● SERVIÇO SOCIAL	5	0,001	0,421	0,308
● ARQUITETURA E URBANISMO	4	0,002	0,383	0,135
● ARQUEOLOGIA	3	0,004	0,387	0,138

Fonte: Dados da pesquisa

Sociologia apresenta as maiores medidas de grau, proximidade (*Closeness Centrality*) e centralidade do autovetor (*Eigenvector Centrality*). Antropologia possui a segunda maior medida de grau, no entanto, é a que tem a maior centralidade de intermediação, ou seja, tem posicionamento destacado em termos de ligação (ponte) entre áreas. Comunicação, por outro lado, embora tenha a terceira maior medida de grau, apresenta a segunda mais alta centralidade de proximidade e centralidade do autovetor, o que indica estar estruturalmente bem posicionada na rede, ligada a áreas mais representativas em termos de grau, e com maiores chances de aumentar o nível de interação.

Figura 29: Grafo que destaca a comunidade 5, identificada no período 2010-2012



Fonte: Elaborado pelo autor

A Figura 29 mostra a representação da Comunidade 5. Destacam-se como interações mais fortes dessa comunidade as estabelecidas entre as áreas de Sociologia e Antropologia e entre Sociologia e Direito.

5.2.2.2.f Interação entre Comunidades

Para identificar as áreas que vinculam as comunidades identificadas, isolamos as áreas que possuem apenas arestas entre elas (comunidades). Dessa forma, identificamos dezesseis áreas como principais responsáveis por conectar as comunidades observadas em 2010-2012. Essas áreas, e suas respectivas medidas de centralidade, são mostradas na Tabela 24.

Curioso observar que, em relação ao triênio 2001-2003, Educação aparece agora como área mais central em termos de grau (número de interações), intermediação (capacidade de conexão, ponte, entre áreas menos conectadas na rede) e proximidade (maiores chances de aumentar os trabalhos em colaboração, por estar estruturalmente bem posicionada na rede). Destaca-se, sobretudo, o papel de intermediador dessa área neste último triênio (2010-2012), conforme apontado pela métrica de *Betweenness Centrality*.

Tabela 24: Relação de áreas e medidas de centralidade responsáveis por conectar as comunidades identificadas em 2010-2012

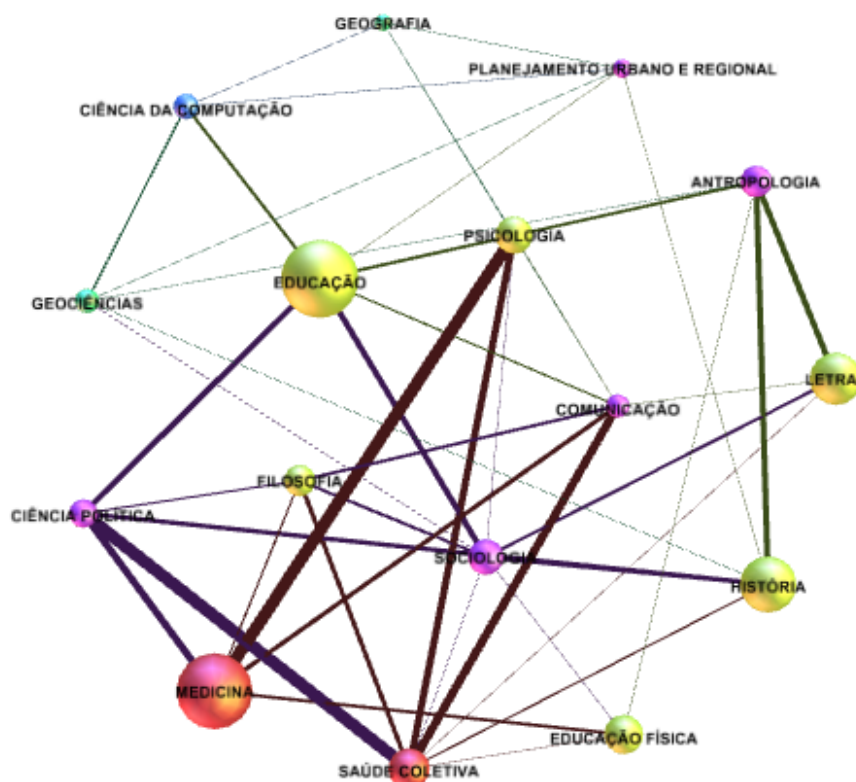
Nodes	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● EDUCAÇÃO	24	0,147	0,549	0,859
● MEDICINA	23	0,097	0,545	1
● SAÚDE COLETIVA	20	0,104	0,523	0,885
● HISTÓRIA	19	0,092	0,523	0,779
● SOCIOLOGIA	14	0,017	0,489	0,704
● FILOSOFIA	14	0,02	0,493	0,736
● GEOCIÊNCIAS	13	0,063	0,489	0,428
● PSICOLOGIA	13	0,049	0,482	0,623
● ANTRPOLOGIA	13	0,019	0,462	0,579
● EDUCAÇÃO FÍSICA	12	0,025	0,5	0,662
● COMUNICAÇÃO	12	0,014	0,475	0,641
● CIÊNCIA POLÍTICA	11	0,012	0,479	0,605
● LETRAS	11	0,011	0,469	0,607
● CIÊNCIA DA COMPUTAÇÃO	9	0,018	0,441	0,248
● PLANEJAMENTO URBANO E REGIONAL	9	0,011	0,435	0,361
● GEOGRAFIA	5	0,004	0,404	0,2

Fonte: Dados da pesquisa

Medicina possui a segunda maior centralidade de grau e de proximidade. No entanto, é a área que, neste período, tem mais ligações com áreas mais representativas, em termos de grau, de acordo com a medida de centralidade do autovetor (*Eigenvector Centrality*).

Saúde Coletiva, por outro lado, embora possua menor grau que Medicina e Educação, apresenta a segunda maior medida de intermediação e centralidade do autovetor. A Figura 30 mostra as principais áreas responsáveis pela interação entre as comunidades identificadas.

Figura 30: Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas no período 2010-2012



Fonte: Elaborado pelo autor

Verifica-se que a comunidade 3 (cor amarela), com seis áreas - Educação, História, Letras, Psicologia, Filosofia e Educação Física -, é a que mais contribui para a interação com as outras comunidades/áreas na UFRJ no triênio 2010-2012. Isto revela que as áreas de humanas e letras apresentam maior participação, na interação entre áreas, em relação ao período 2001-2003.

Não obstante, podemos notar também que Medicina e Saúde Coletiva têm presença significativa na interação entre comunidades/áreas neste último triênio e destacam-se por estabelecer vínculos fortes (observado pela largura das arestas na Figura 30) com áreas de outros campos, o que não acontece tanto com Educação, por exemplo.

5.3 PRODUÇÃO CIENTÍFICA E AUTORIA

Até agora, a identificação e visualização dos domínios institucionais da UFRJ guardavam relação com as grandes áreas e áreas do conhecimento devidamente registradas nos artigos em periódicos cadastrados na base EspaçoSIGMA.UFRJ. A partir deste ponto, no entanto, os domínios são analisados do ponto de vista da relação produção-autoria.

No período de 2001 a 2012, foram identificados 44.233 artigos publicados em periódicos, que contavam com 56.178 autores, dos quais 27.930 são autores externos, isto é, autores sem vínculo formal com a UFRJ. As análises a seguir focam inicialmente na totalidade da publicação em periódicos e, posteriormente, somente na produção de autores com algum tipo de vínculo formal com a UFRJ.

A análise da produção considerou o perfil da atividade colaborativa e a produtividade do corpo social da instituição, ou seja, os autores com diferentes tipos de vínculos formais com a UFRJ. Com base nas informações originais da base EspaçoSIGMA.UFRJ, os autores foram manualmente reclassificados em seis grupos, de acordo com os respectivos vínculos formais com a instituição, da seguinte forma: Docente (D); Discente de Pós-graduação (P); Egresso Discente de Pós-graduação (exP); Discente de Graduação (G); Técnico (T) e; Outros (O).

As seções a seguir apresentam, assim, uma análise da produção de artigos da UFRJ publicados em periódicos, focando na autoria e no tipo de vínculo institucional dos autores, bem como no nível de colaboração e produtividade. Por fim, identificamos a formação de comunidades ou domínios a partir dos autores, assim como as especificações das comunidades mais representativas, em termos de tamanho, em cada período (triênios).

5.3.1 Contexto geral: autoria, produção e vínculo institucional

A soma de todos os artigos completos da UFRJ publicados em periódicos é 44.233 para os quatro triênios. Mas se considerarmos as publicações com um único autor esse número cai para 14.019, enquanto que publicações com dois ou mais autores, isto é, artigos em colaboração, totalizam 30.214.

A Tabela 25 mostra, em termos percentuais, a relação entre o número de autoria e número de artigos. Vemos que o número total de artigos da UFRJ em cada triênio mantém-se praticamente constante em todo o período estudado. No entanto, observa-se uma diminuição da fração de artigos com 1, 2, 3 e 4 autores nos quatro triênios e um aumento do número de artigos com 5 ou mais autores. Esse quadro sugere uma tendência crescente da atividade colaborativa na equipe da UFRJ, fenômeno também observado para publicações científicas brasileiras nos anos 2000 (VANZ; STUMPF, 2012; MENA-CHALCO *et al.*, 2014).

Tabela 25: Distribuição (%) da totalidade de artigos em periódicos da UFRJ de acordo com o número de autores e o período de publicação

Autorias	Número de Artigos			
	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%
1 autor	35,78	33,15	29,60	28,30
2 autores	20,87	17,69	16,73	16,67
3 autores	17,74	16,59	16,32	14,60
4 autores	11,09	12,10	12,48	11,44
5 autores	6,13	7,44	8,25	8,96
6 autores	3,47	5,05	6,10	6,32
7 autores	2,48	3,24	3,85	4,28
8 autores	1,08	2,12	2,53	2,80
9 autores	0,58	1,24	1,57	2,10
10 autores	0,29	0,56	0,93	1,50
Mais de 10	0,49	0,82	1,62	3,04
TOTAL %	100	100	100	100
TOTAL Nº	11.003	10.840	11.696	10.694

Fonte: Dados da pesquisa

Muitos fatores estão associados à colaboração na ciência: dividir custos com equipamentos, facilitar a interação interdisciplinar, qualificar pessoal para determinado procedimento, a busca por visibilidade e reconhecimento, entre outros (KATZ; MARTIN, 1997). Neste trabalho, introduzimos um outro fator que pode ter relação direta com o ato de colaborar na ciência: o vínculo com a instituição, que é

uma informação registrada na base EspaçoSIGMA.UFRJ, tal como mencionado na metodologia.

A Tabela 26 apresenta, então, a distribuição das autorias da produção em periódicos cadastrado na base EspaçoSIGMA.UFRJ pelo tipo de vínculo institucional do/a autor/a. Os totais, portanto, referem-se aos autores que apresentam vínculo formal com a UFRJ, cuja soma foi 28.248¹⁹.

Tabela 26: Distribuição (%) dos autores dos artigos em periódicos da UFRJ de acordo com o vínculo institucional e o triênio da publicação

Vínculo institucional	Número de Artigos			
	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%
Docente	35,33	28,46	27,05	27,70
Discente Pós-Graduação	31,33	32,98	31,75	32,17
ex Discente Pós-Graduação	11,64	17,07	21,54	21,89
Discente Graduação	10,76	10,01	7,23	6,81
Técnico	4,95	4,34	3,87	3,08
Outros*	6,00	7,14	8,56	8,34
TOTAL %	100	100	100	100
TOTAL Nº	6.004	6.680	7.771	7.793

Fonte: Dados da pesquisa

* Inclui ex-professores, ex-técnicos e ex-alunos de graduação

Nota-se no triênio 2001-2003 que predomina a autoria de docentes (35,33%). Muito próximo da autoria de docente está a autoria de alunos de pós-graduação (31,33%); mais distantes estão os egressos da pós-graduação (11,64%) e alunos de graduação (10,76%).

A participação de técnicos e de autores com outros tipos de vínculos (incluem outros egressos - docentes, técnicos e alunos de graduação) como autores de artigos em periódicos representa 4,95% e 6,00%, respectivamente.

A partir do triênio 2004-2006, percebe-se que a maior parte das autorias é de alunos de pós-graduação (32,98%), enquanto os docentes aparecem como o

¹⁹ Autores externos à instituição, ou seja, sem vínculo formal com a UFRJ, somaram 27.930, sendo 4.535, 6.174, 8.314 e 8.907 em 2001-2003, 2004-2006, 2007-2009 e 2010-2012, respectivamente.

segundo tipo de vínculo mais presente entre os autores (28,46%). Esta situação se mantém até o último triênio estudado.

Nos períodos seguintes, chama atenção o aumento apresentado pelos egressos da pós-graduação, que passam a representar mais de 21% dos autores. Uma possível explicação para esta situação é o fato de que, nos últimos anos, cresceu o número de vagas para universidades e institutos de pesquisa do setor público (BRASIL, 2012), abrindo, assim, novas oportunidades para os jovens recém formados na UFRJ. Estes, por sua vez, devem estar buscando apoio em seus grupos de pesquisa de origem para o desenvolvimento de pesquisa em suas atuais instituições.

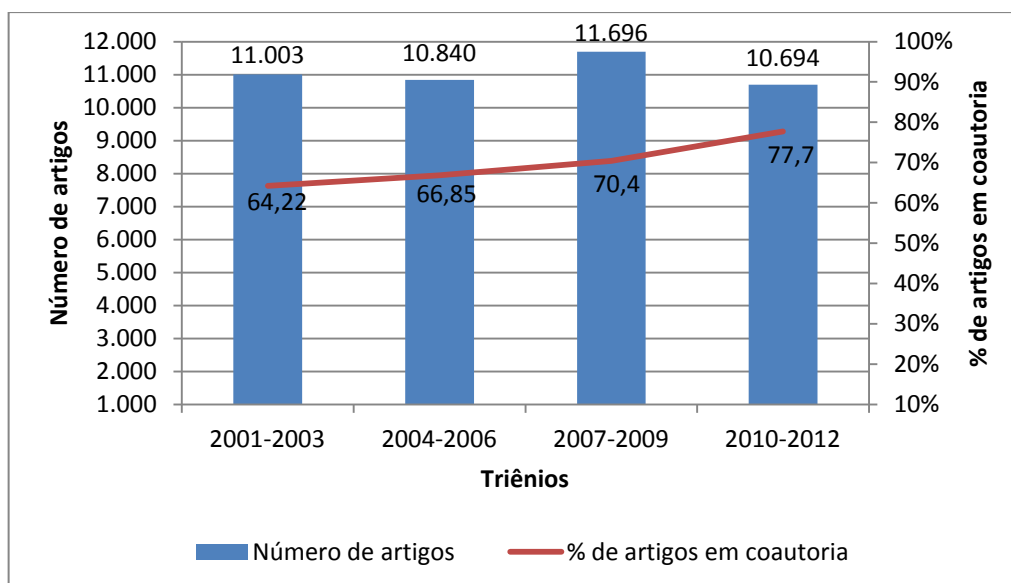
Uma vez conhecidos os autores do ponto de vista do vínculo com a UFRJ, passamos, então, a detalhar se (e como) estes tipos de vínculo estabelecem colaborações.

5.3.2 Coautoria e vínculo institucional

Artigos com ao menos uma coautoria representam 68,31% do total de artigos publicados pela UFRJ no período (44.233) e possuem 95% do total de autores (56.178), dos quais 27.930 são externos à instituição. Esse quadro, por si só, indica um elevado nível de trabalho colaborativo na UFRJ, inclusive com pares externos.

Na Figura 31, que mostra a distribuição dos artigos por triênios e o percentual de artigos em coautoria, verifica-se que no período 2001-2003, a quantidade total de artigos publicados em periódicos foi de 11.003, dos quais 7.066 (64,22%) eram artigos em coautoria. Em 2004-2006, foram 7.246 (66,85%) artigos assinados por mais de um autor. Nos triênios seguintes, esse grupo de artigos somou 8.234 (70,40%) e 7.668 (77,70%), respectivamente.

Figura 31: UFRJ – Artigos publicados em periódicos e percentuais de artigos em coautoria por triênios



Fonte: Elaborado pelo autor

Assim, como já apontado anteriormente, embora a produção total de artigos em periódicos da UFRJ se mantenha praticamente constante em todo o período estudado, os artigos em coautoria crescem de um triênio ao outro.

Considerando a coautoria como um *proxy* de colaboração na ciência, decidimos investigar mais detalhadamente a relação entre o tipo de vínculo e trabalho colaborativo, que é apresentada na Tabela 27. A análise considerou artigos “sem colaboração” aqueles que têm somente um autor e que somam, no período 5.118. Já os artigos “com colaboração” são aqueles com duas ou mais coautorias.

A Tabela 27 mostra a distribuição de autores com vínculos institucionais em relação à participação deles em artigos com colaboração e sem colaboração da UFRJ. A soma dos autores em artigos de ambos os tipos (com e sem colaboração) supera os totais apresentados na Tabela 26 para cada período. A diferença é explicada pelo fato de que um autor poder ser encontrado em ambos os tipos de publicações.

Tabela 27: Distribuição (%) de autores em artigos com e sem colaboração da UFRJ de acordo com o vínculo institucional e período da publicação

Vínculo Institucional	Número de Artigos							
	Artigos sem colaboração				Artigos com colaboração			
	2001-2003	2004-2006	2007-2009	2010-2012	2001-2003	2004-2006	2007-2009	2010-2012
	%	%	%	%	%	%	%	%
Docente	76,54	57,31	59,34	53,63	30,24	26,28	25,69	26,87
Discente PG	17,62	30,38	29,67	36,23	31,12	30,64	29,69	29,16
ex Discente PG	1,35	3,08	4,00	2,95	15,19	20,73	24,40	25,12
Discente Grad	0,90	2,77	0,92	1,52	11,19	10,15	7,20	6,88
Técnico	1,42	2,08	1,69	1,76	5,23	4,44	3,90	3,09
Outros*	2,17	4,38	4,38	3,91	7,04	7,76	9,11	8,88
TOTAL %	100	100	100	100	100	100	100	100
TOTAL Nº	1.334	1.300	1.301	1.253	5.695	6.325	7.484	7.503

Fonte: Dados da pesquisa

* Inclui ex-professores, ex-técnicos e ex-alunos de graduação
 Nota: Artigos sem colaboração referem-se a artigos de autoria única. Artigos com colaboração referem-se a artigos assinados por 2 ou mais autores.

Verifica-se que a fração de autores de artigos em periódicos com colaboração aumentou de 94,85% (n = 5.695), no primeiro triênio, para 96,28% (n = 7.503), no último. Uma característica importante é a participação de autores externos (dados não mostrados). No período, este conjunto de autores quase duplicou, passando de 4.535 para 8.907.

Pode-se notar que professores e alunos de pós-graduação se constituem como os principais tipos de vínculos nos artigos com colaboração em todos os períodos. Na verdade, 56-61% dos artigos da UFRJ em colaboração são de autoria de um professor ou de um estudante de pós-graduação. No entanto, estes dois tipos de vínculos institucionais são ainda mais frequentes nos artigos não colaborativos, 88-94%.

Curiosamente, os alunos egressos de pós-graduação apresentam crescente número de artigos em colaboração, sendo, portanto, este o terceiro grupo mais representativo da UFRJ em artigos deste tipo. Uma possível explicação, como dito anteriormente, é que os egressos da pós-graduação são agora pesquisadores

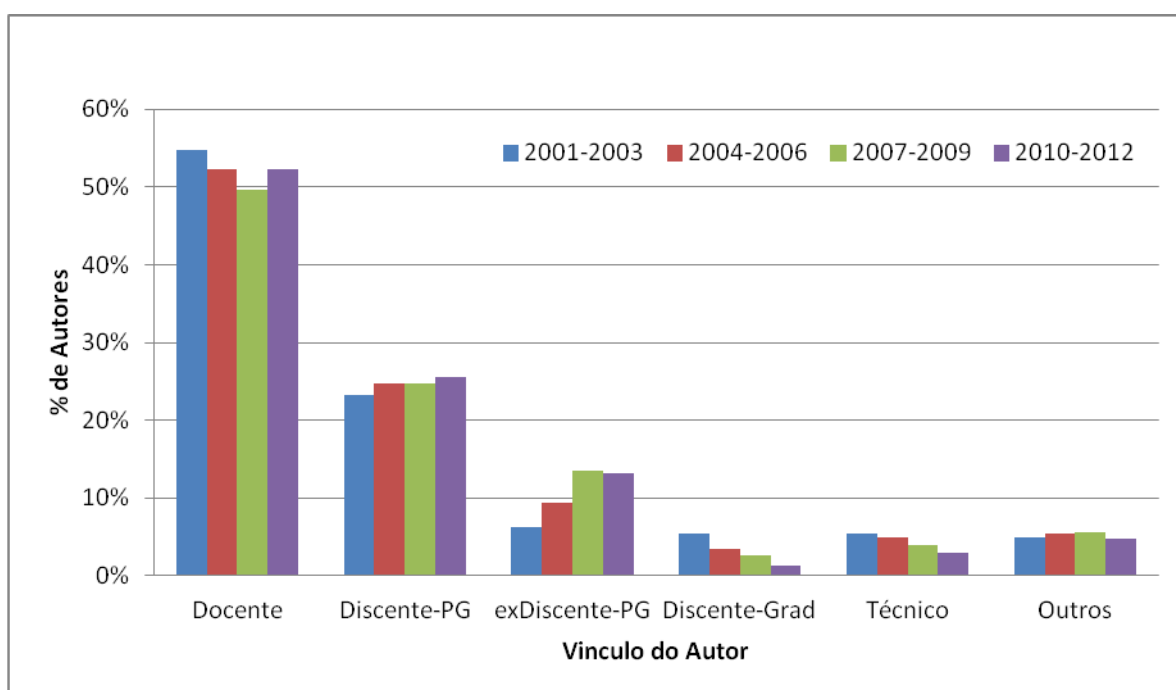
independentes ligados a outras instituições, mas ainda mantendo projetos de colaboração com seus institutos de origem ou orientadores.

5.3.3 Coautoria, produtividade e vínculo institucional

No sentido de verificar a relação entre produtividade, colaboração (estimada através da coautoria) e vínculo do autor, foram selecionados autores com 3 ou mais artigos publicados em colaboração para cada triênio.

Em 2001-2003, foram identificados 2.472 autores (o que representa 41,17% da população de autores desse triênio), dos quais 613 são externos à UFRJ. Em 2004-2006, foram identificados 2.645 autores (39,60%), dos quais 723 são externos à instituição. Já no período de 2007-2009, foram identificados 3.163 autores (40,70%), dos quais 867 são externos à UFRJ. Por fim, no triênio 2010-2012, foram identificados 2.946 autores (37,80%), dos quais 839 são externos à instituição. Percebe-se, assim, que o grupo de autores com 3 ou mais publicações oscila em torno de 40% do total de autores.

Figura 32: Distribuição (%) de autores com três ou mais artigos publicados em colaboração, de acordo com o vínculo institucional com a UFRJ



Fonte: Elaborado pelo autor

A Figura 32 mostra os autores mais produtivos (com três ou mais artigos de periódicos) e que publicaram em coautoria, em termos percentuais, por tipo de vínculo com a UFRJ para cada triênio estudado.

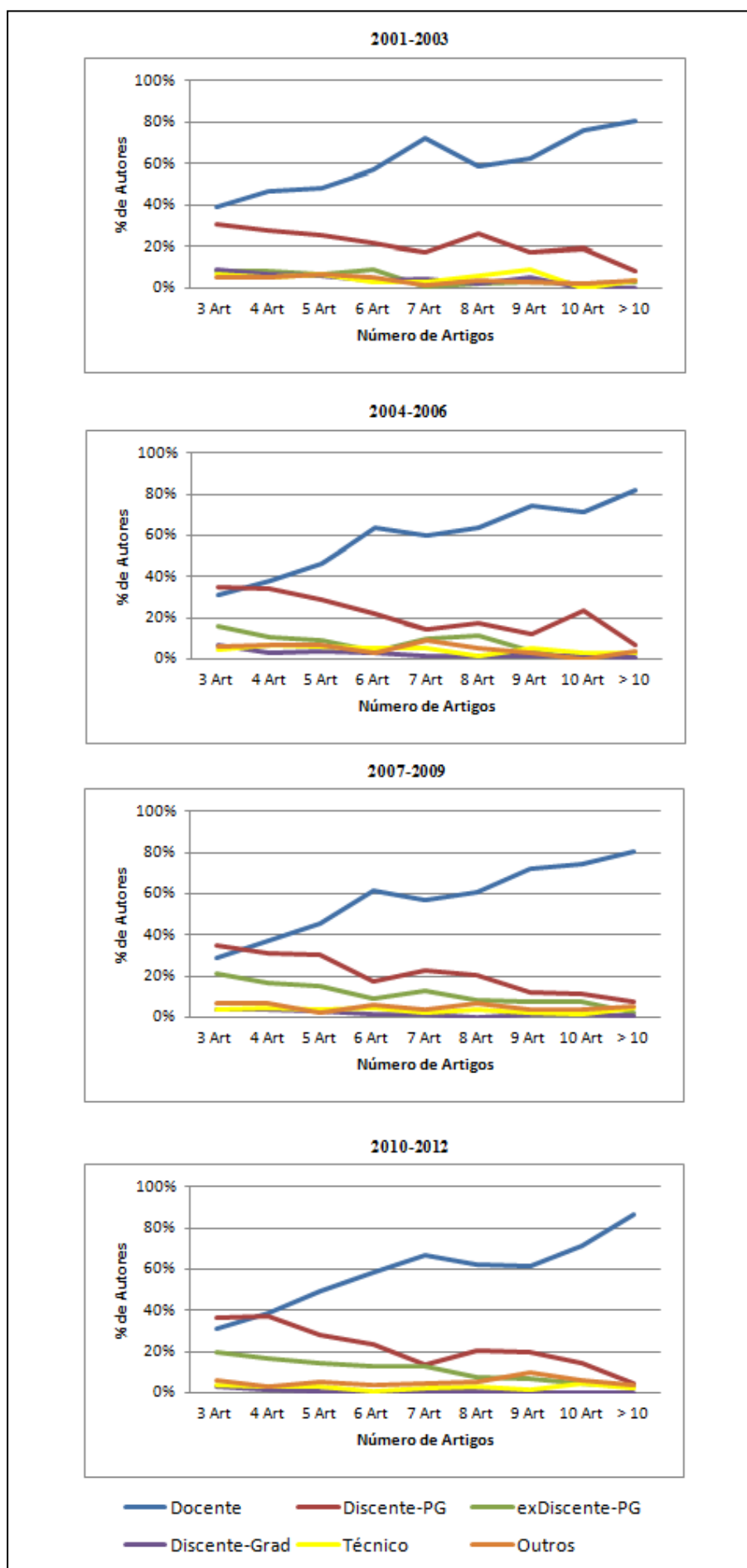
Verifica-se que, em média, 52% dos autores com três ou mais artigos em colaboração são docentes e 24% são alunos de pós-graduação, os quais tem apresentado pequeno aumento de produtividade de um período ao outro. Os alunos egressos da pós-graduação têm o aumento mais destacado de participação: de 6,19% em 2001-2003, para 13,15% em 2010-2012.

Considerando a diversidade em termos de produtividade no conjunto de autores, verificamos o vínculo institucional dos autores de acordo com o número de artigos em colaboração (Figura 33).

Nota-se que a participação de docentes eleva-se conforme aumenta a quantidade de artigos produzidos em coautoria. Este grupo representa mais de 40% entre os autores que publicaram cinco ou mais artigos em colaboração em todos os períodos, e 80% ou mais entre aqueles com mais de 10.

Os outros tipos de vínculos institucionais mostram tendências opostas: a participação se reduz à medida que o número de artigos em colaboração aumenta. Por exemplo, o percentual de estudantes de pós-graduação autores de artigos com colaboração representa, em média, de 30-40% entre os autores com 3, 4 e 5 artigos, e menos de 10% entre aqueles com mais de 10 artigos.

Figura 33: Distribuição (%) de autores por número de artigos da UFRJ publicados em colaboração e vínculo institucional



Fonte: Elaborado pelo autor

A seguir, são apresentadas as análises de rede para identificação das comunidades e/ou domínios institucionais da UFRJ a partir da autoria das publicações em periódico.

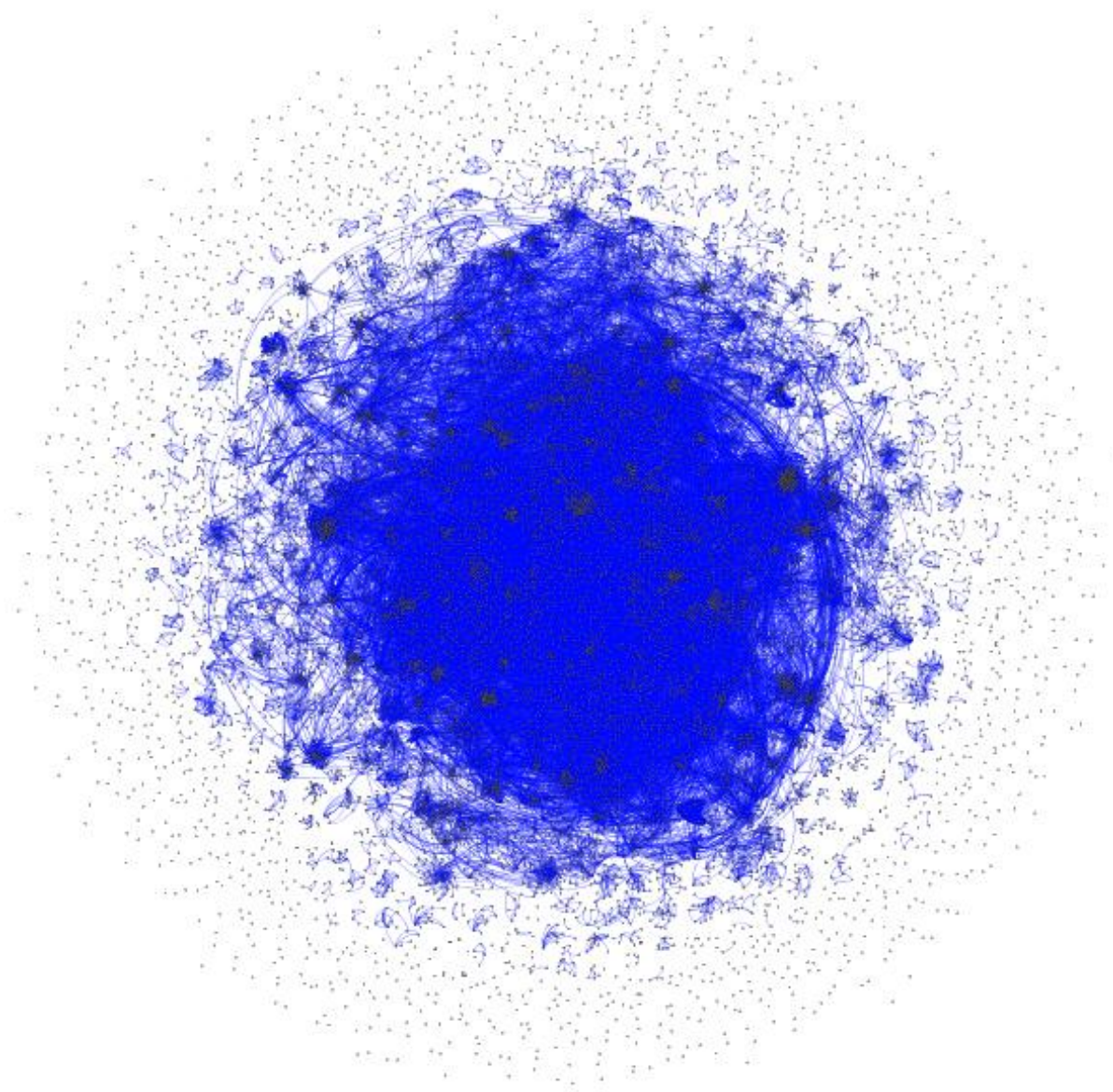
5.3.4 Redes de Interação entre Autores: Identificação de Comunidades

Para iniciar a análise de redes, elaboramos a Figura 34, que mostra o grafo que representa o total de autores com vínculo formal com a UFRJ ($n = 28.248$), responsáveis pelos 44.233 artigos publicados em periódicos no período de 2001 a 2012. Cada ponto do grafo representa um autor. Uma aresta na cor azul indica pelo menos uma publicação realizada em colaboração entre dois pesquisadores.

Os nós que aparecem visualmente isolados, mais externamente no grafo, correspondem a autores de artigos sem colaboração ou subconjunto de autores (geralmente formado por poucos pesquisadores) desligados da grande rede interligada.

Nas redes de interação de coautoria, em média, cada autor da UFRJ mantém colaboração com 6,73 pesquisadores. Observa-se que no grafo existem vários subgrupos, ou seja, conjuntos/agrupamentos de nós bem próximos uns dos outros, de autores que formam comunidades (*clusters*). Podemos notar, visualmente, que autores desses subgrupos colaboram mais entre si.

Figura 34: Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos em periódicos no período 2001-2012

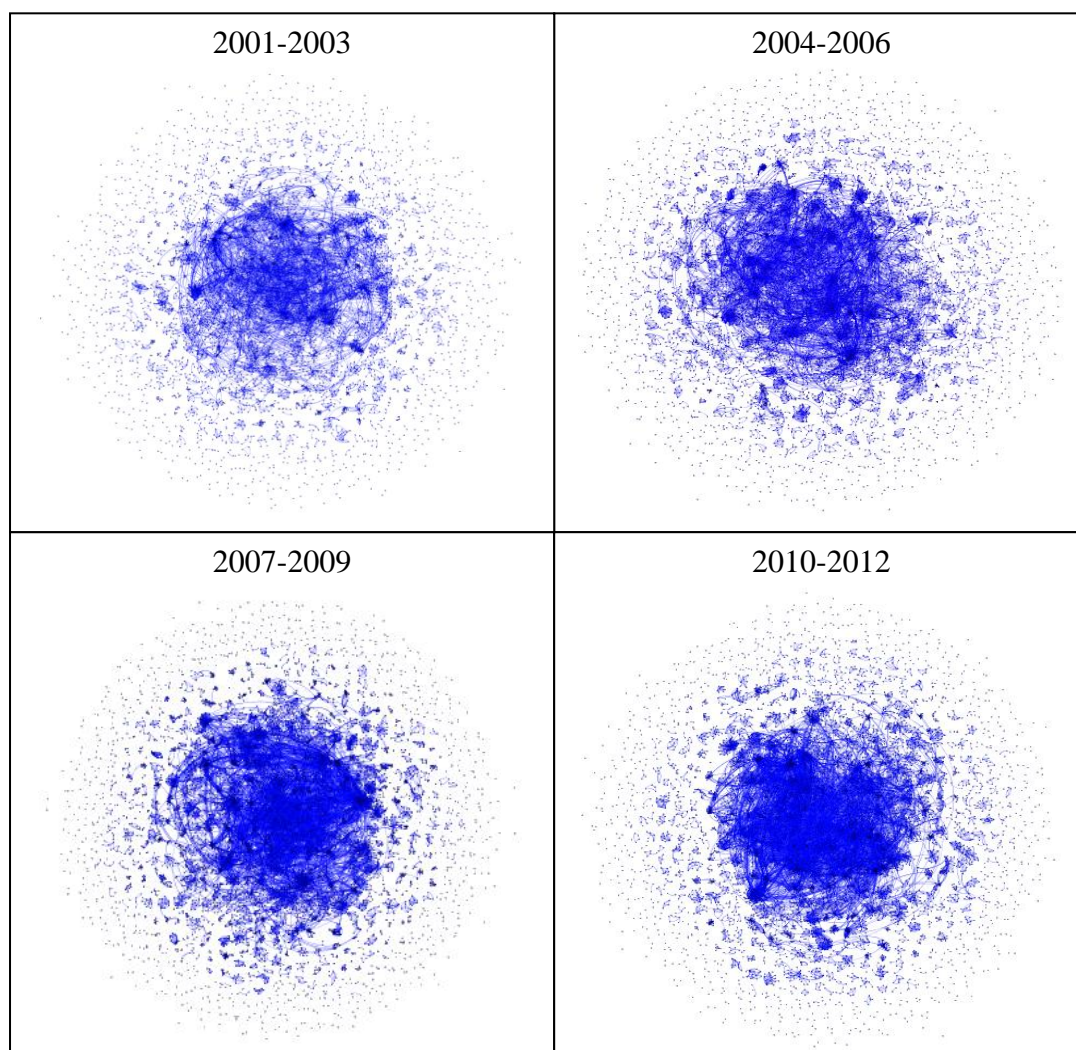


Fonte: Elaborado pelo autor

Nota: O total de autores é 28.248; e as coautorias estão representadas por arestas na cor azul.

A Figura 35 mostra o grafo de autores por triênios. Observa-se que, ao longo dos triênios, a densidade do grafo aumenta. Essa simples representação visual possibilita observar o incremento da atividade colaborativa na UFRJ ao longo do tempo.

Figura 35: Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos em periódicos por triênios



Fonte: Elaborado pelo autor

Nota: O total de autores é 28.248, distribuídos por triênios; e as coautorias estão representadas por arestas na cor azul

No triênio 2001-2003, o grau médio foi de 4,33, ou seja, cada autor colaborou, em média, com pouco mais de quatro autores, e a transitividade, ou seja, a tendência que os autores têm de se agruparem (coeficiente de *clustering*) foi de 0,76. Em 2004-2006, o grau médio foi de 4,87 e o coeficiente de *clustering* foi de 0,77. No período seguinte, 2007-2009, o grau médio foi de 5,12 e a transitividade foi de 0,78. Já o último triênio apresentou grau médio de 5,57 e transitividade de 0,79.

Essas medidas indicam que, ao longo dos períodos, novas colaborações foram estabelecidas, formando novos agrupamentos/comunidades, tornando, dessa

maneira, maior a transitividade no grafo. Em outras palavras, ao logo dos triênios, a tendência de todos os nós se agruparem aumenta.

A fim de verificar a rede de autores com participação mais intensa em termos de colaboração na instituição, identificamos a formação de diferentes subgrupos de pesquisadores que formam comunidades (*clusters*) para cada triênio, focalizando, assim, nossa atenção na maior comunidade de cada período.

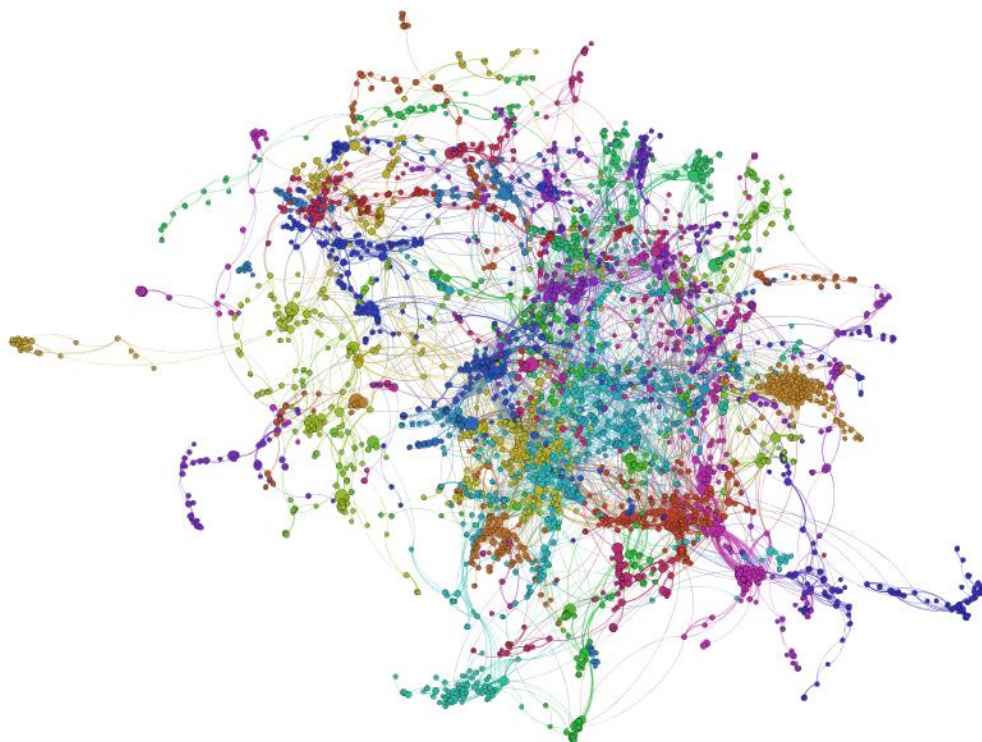
5.3.4.1 Comunidades de Autores no Triênio 2001-2003

A rede de coautoria da UFRJ, no período 2001-2003, é composta por 5.695 autores (totalidade de autores que publicaram em artigos de periódicos com pelo menos uma coautoria, ou em colaboração), conforme mostrado na Tabela 27.

A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel *et al.* (2008), identificou um total de 40 comunidades, envolvendo 3.736 nós (autores), identificadas na Figura 36 através das cores. Os nós estabeleceram entre si 11.304 relações (colaborações), representadas na rede por arestas (Figura 36). O grau médio foi de 6,05, ou seja, em média, estes autores participaram de publicações com pouco mais de seis autores e a transitividade (coeficiente de *clustering*) de 0,76.

A rede de autores que compõem as comunidades é inferior ao total de autores em artigos com colaboração devido à exclusão de autores (nós) menos representativos, ou grupos isolados.

Figura 36: Grafo das comunidades de autores da UFRJ de publicações em periódicos identificadas no período 2001-2003

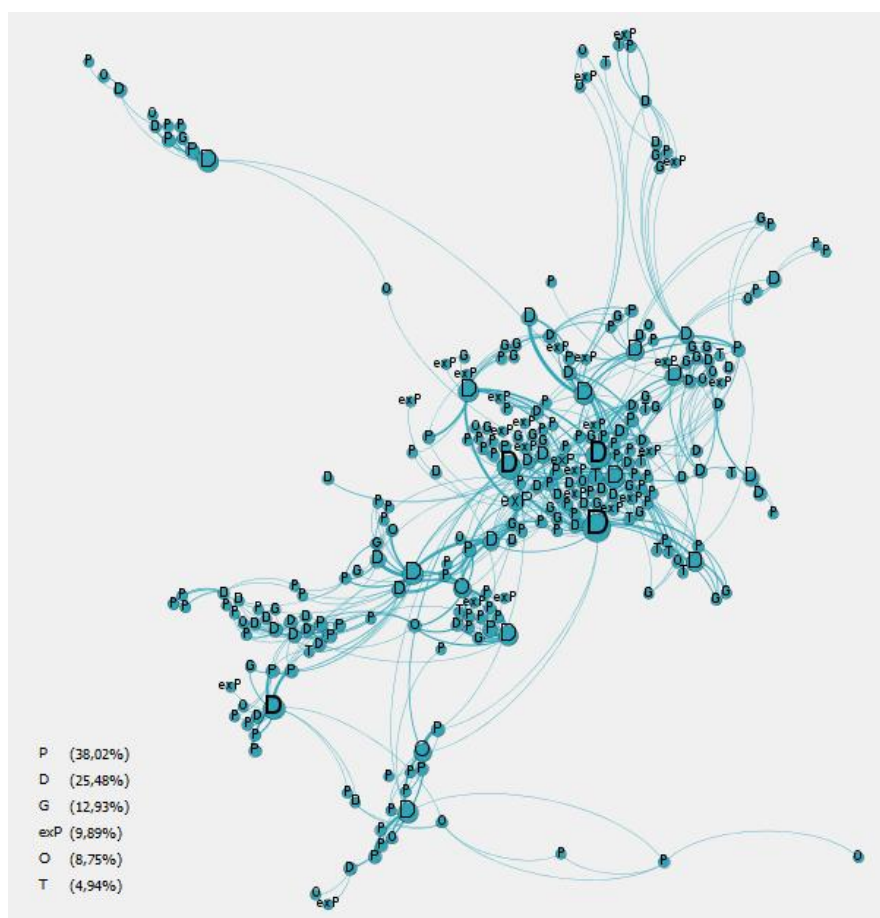


Fonte: Elaborado pelo autor

O grafo da Figura 36 mostra um emaranhado de relações envolvendo as 40 principais comunidades de autores da UFRJ, agrupadas por cores, identificadas no período 2001-2003. A seguir, procedemos à identificação da maior delas, no sentido de verificar especificações de seus integrantes.

A maior comunidade de autores identificada no período de 2001-2003 é constituída por 236 autores (o que representa 7,04% do total de autores das 40 comunidades) e, 689 relacionamentos, arestas (Figura 37). O grau médio é de 5,24, ou seja, cada autor colabora, em média, com mais de cinco pesquisadores. O coeficiente de transitividade apresentado é de 0,79.

Figura 37: Grafo da maior comunidade de autores de artigos em periódicos da UFRJ no período 2001-2003



Fonte: Elaborado pelo autor

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação (38,02%), seguida por docentes (25,48%) e alunos de graduação (12,93%). Alunos egressos da pós-graduação representam 9,89% e os autores com outros tipos de vínculo 8,75%. Já os técnicos representam 4,94%. Esses dados corroboram, de maneira geral, a distribuição da participação de atores da UFRJ em artigos com colaboração apresentada na Tabela 27, com a diferença de que nesta comunidade a proporção de alunos de graduação é superior ao número de alunos egressos de pós-graduação.

Do total de autores dessa comunidade, identificamos os vinte que apresentam as maiores medidas de centralidade de grau, ou seja, são autores que mais interagem, tem maior número de colaborações (Tabela 28).

Tabela 28: Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2001-2003

Id	Vínculo*	Grau	Áreas de Atuação
330	D	46	Ciências Biológicas; Ciências da Saúde
2139	D	34	Ciências Biológicas, Ciências da Saúde
342	D	26	Ciências Biológicas
339	exP	24	Ciências Biológicas, Ciências da Saúde
343	O	20	Ciências Biológicas
344	D	19	Ciências Biológicas
2140	D	19	Ciências da Saúde
268	D	17	Ciências Biológicas
406	D	16	Ciências Biológicas
2042	D	15	Ciências Agrárias, Ciências da Saúde
2320	D	15	Ciências da Saúde, Ciências Biológicas
445	D	14	Ciências Biológicas
540	D	14	Ciências Biológicas
2156	T	14	Ciências da Saúde
223	D	13	Ciências Biológicas
412	D	12	Ciências Biológicas, Ciências Humanas
1924	D	12	Ciências da Saúde, Ciências Agrárias
338	P	11	Ciências Humanas
341	D	11	Ciências Biológicas
541	D	11	Ciências Biológicas

Fonte: Dados da pesquisa

* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação); exP (Egresso Discente de Pós-graduação); T (Técnico); O (Outros)

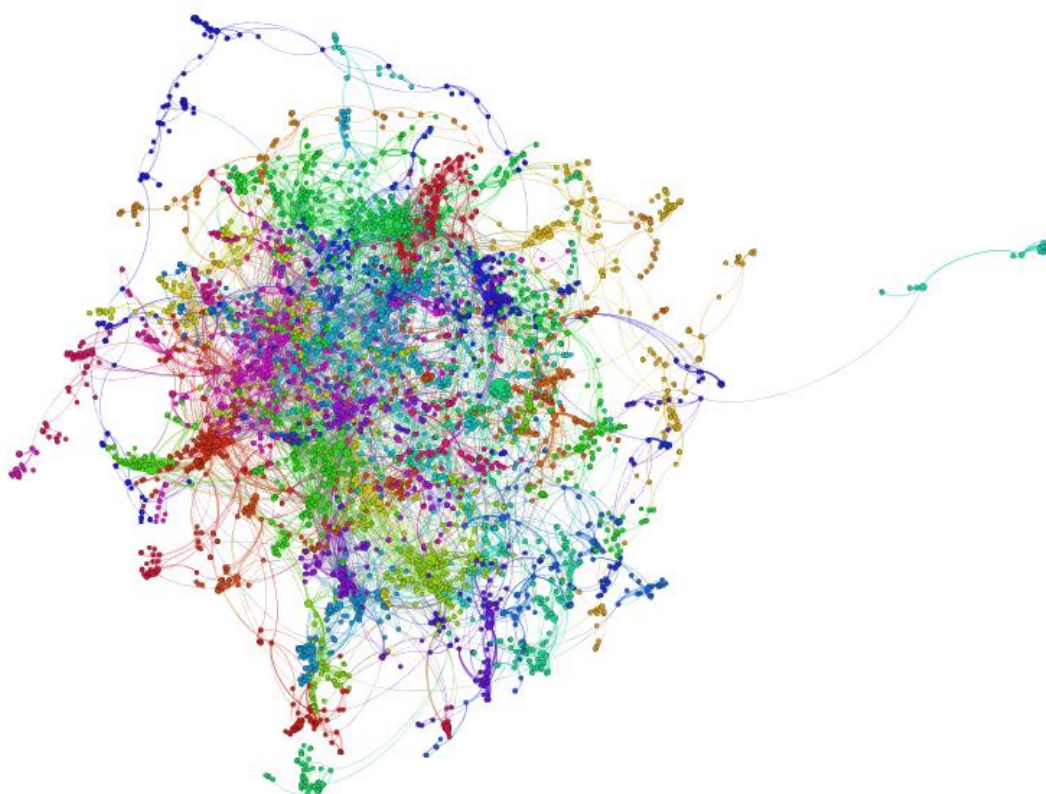
Verifica-se que dos vinte autores identificados, dezesseis são docentes. Os outros quatro autores têm os seguintes tipos de vínculo: egresso da pós-graduação, aluno de pós-graduação, técnico e um que possui outros tipos de vínculo (podendo ser professor egresso, técnico egresso ou egresso da graduação).

No que se refere às áreas dos autores, verificamos que dez deles atuam na grande área de Ciências Biológicas, dois nas Ciências da Saúde e um nas Ciências Humanas. Os outros autores identificados atuam em duas áreas do conhecimento, assim distribuídos: quatro deles em Ciências Biológicas e Ciências da Saúde; dois em Ciências Agrárias e Ciências da Saúde e; um autor em Ciências Biológicas e Ciências Humanas.

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações importantes (Figura 38). O tamanho do nó (e do Id) é

arestas (Figura 39). O grau médio da rede foi de 6,45 e a transitividade (coeficiente de *clustering*) de 0,77. Esses dados mostram uma expansão do número de subgrupos (agrupamentos), resultante do maior nível de interação entre autores em relação ao triênio anterior.

Figura 39: Grafo que mostra as 45 comunidades de autores da UFRJ identificadas no período 2004-2006



Fonte: Elaborado pelo autor

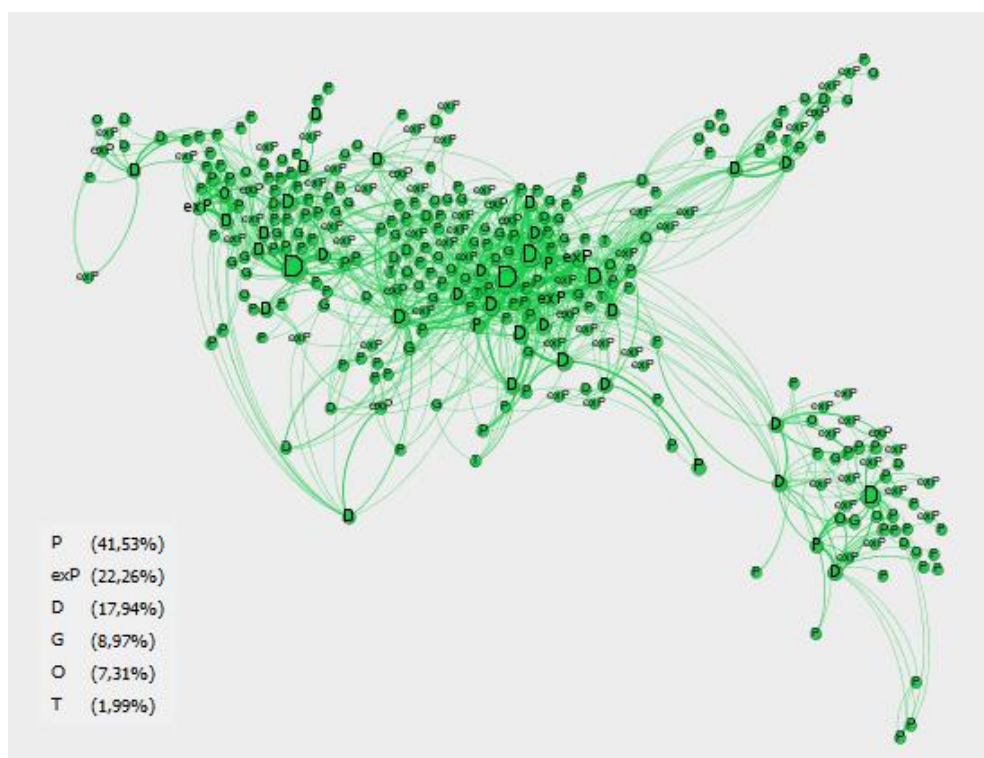
O grafo da Figura 39 mostra a complexidade das relações envolvendo as 45 comunidades de autores da UFRJ, agrupadas por cores, identificadas no período 2004-2006. Podemos observar pequenos subgrupos quase isolados da rede.

A seguir, destacamos a maior comunidade de autores, no sentido de verificar especificações de seus integrantes.

A maior comunidade de autores, identificada no período de 2004-2006, é constituída por 301 autores (o que representa 6,81% do total de autores das 45 comunidades), 65 autores a mais que no triênio anterior. Estes autores

estabeleceram 1.127 relacionamentos, arestas (Figura 40). O grau médio é de 7,48, ou seja, cada autor colabora, em media, com mais de sete pesquisadores. O coeficiente de transitividade apresentado é de 0,80.

Figura 40: Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2004-2006



Fonte: Elaborado pelo autor

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação (41,53%), seguida por alunos egressos da pós-graduação (22,26%) e docentes (17,94%). Alunos de graduação representam 8,97% e autores com outros tipos de vínculo 7,31%. Já o número de técnicos participantes nessa comunidade representam 1,99%.

Estes dados mostram, de maneira geral, coerência com a distribuição da participação de atores da UFRJ em artigos com colaboração apresentada na Tabela 27, com a diferença de que nesta comunidade a proporção de egressos da pós-graduação supera o número de docentes.

Os vinte autores que apresentam as maiores medidas de centralidade de grau são mostrados na Tabela 29. Observa-se que dos vinte autores identificados,

dezessete são docentes, dois são egressos da pós-graduação e um é aluno de pós-graduação.

Tabela 29: Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2004-2006

Id	Vínculo*	Grau	Áreas de Atuação
56	D	70	Ciências Biológicas
233	D	58	Ciências Biológicas
54	D	45	Ciências Biológicas
593	D	44	Ciências Biológicas
55	D	39	Ciências Biológicas
57	D	38	Ciências Biológicas
182	D	33	Ciências Biológicas
600	D	31	Ciências Biológicas
558	D	30	Ciências Biológicas
442	D	25	Ciências Biológicas
3389	D	25	Ciências da Saúde
53	exP	24	Ciências Biológicas
337	D	24	Engenharias, Ciências Biológicas
395	D	23	Ciências Biológicas
244	D	22	Ciências Biológicas
397	D	22	Ciências Biológicas
833	P	21	Ciências Biológicas
4797	D	21	Ciências da Saúde
1047	D	20	Ciências Biológicas
6096	exP	20	Ciências Biológicas

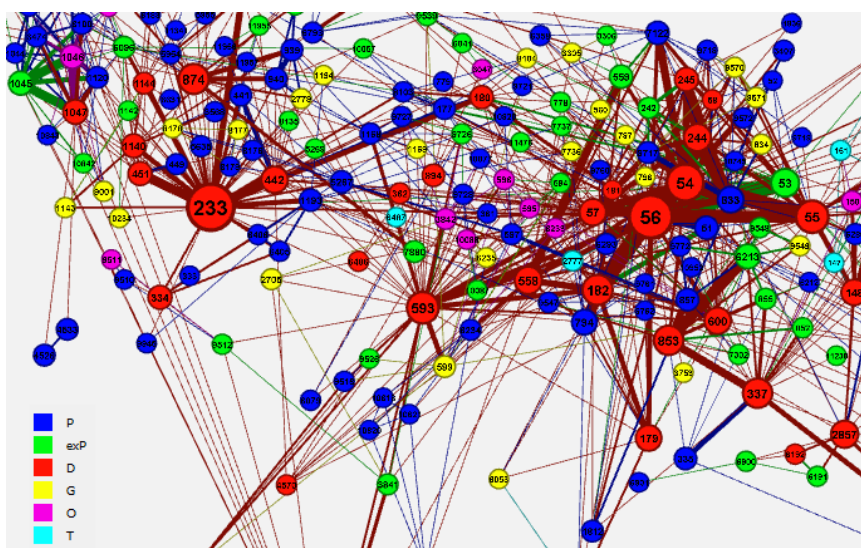
Fonte: Dados da pesquisa

* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação);
exP (Egresso Discente de Pós-graduação)

No que tocante às grandes áreas dos autores, verificamos que dezessete deles atuam nas Ciências Biológicas, dois nas Ciências da Saúde. Apenas um autor atua em duas áreas do conhecimento: Engenharias e Ciências Biológicas.

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações importantes (Figura 41). Podemos notar dois grandes polos de colaboração. Um envolvendo os docentes Id = 56 e 54, os quais tem fortes ligações com outros docentes e, em seu entorno, alunos e egressos da pós-graduação.

Figura 41: Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2004-2006



Fonte: Elaborado pelo autor

Por outro lado, temos o docente Id= 233, que concentra sua colaboração com alunos de pós-graduação.

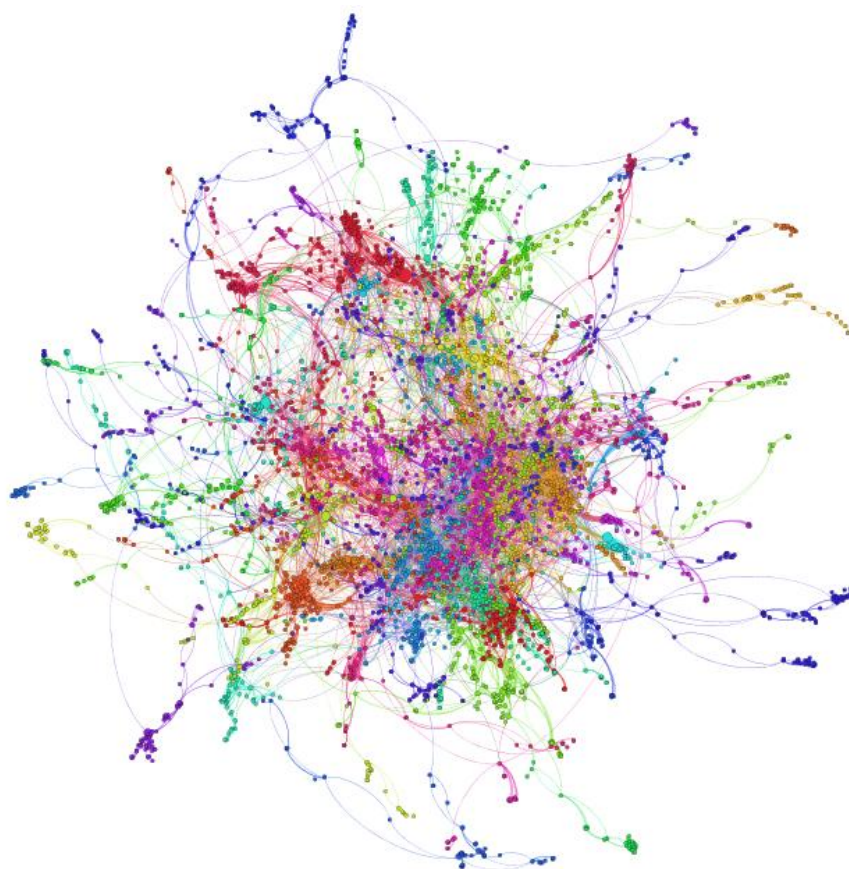
5.3.4.3 Comunidades de Autores no Triênio 2007-2009

No período 2007-2009, a rede de autores da UFRJ é composta por 7.484 autores (total de autores em artigos com colaboração, Tabela 27).

A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel *et al.* (2008), identificou um total de 50 comunidades, envolvendo 5.268 autores, os quais mantiveram entre si 17.548 ligações, representadas na rede por arestas (Figura 42). O grau médio foi de 7,48 e a transitividade (coeficiente de *clustering*) de 0,78. Há, portanto, relativo aumento na formação de agrupamentos em relação ao período anterior.

O grafo da Figura 42 mostra as diversas relações envolvendo as 50 comunidades de autores da UFRJ, agrupadas por uma determinada cor, identificadas no período 2007-2009.

Figura 42: Grafo que mostra as 50 comunidades de autores da UFRJ identificadas no período 2007-2009

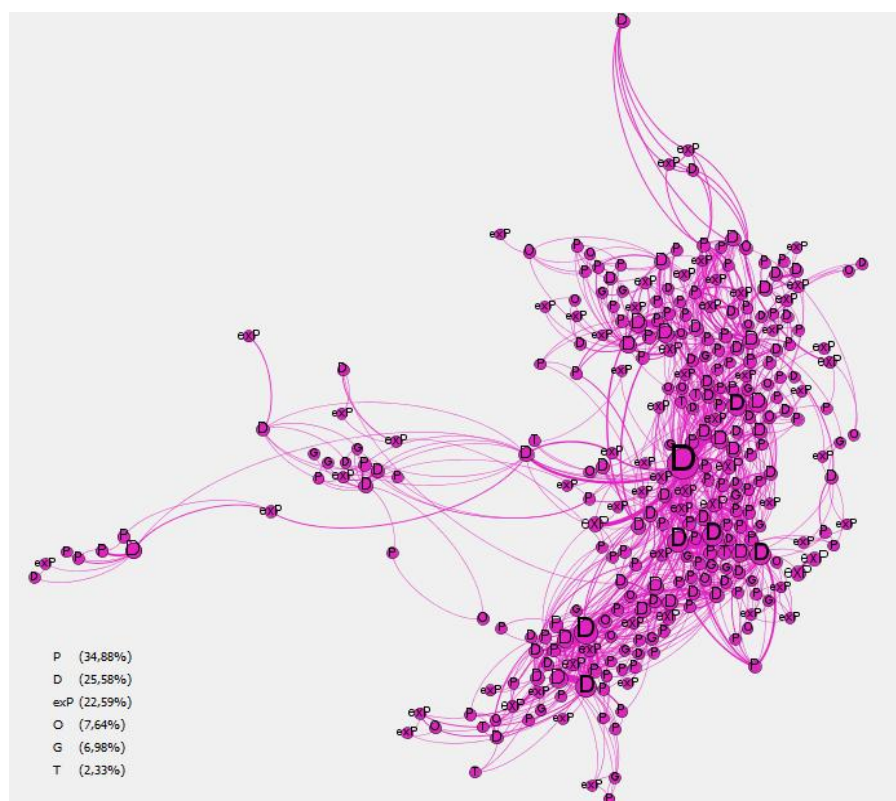


Fonte: Elaborado pelo autor

Destacamos a seguir, a maior delas, com o intuito de observar especificações de seus integrantes.

A maior comunidade de autores identificada no período de 2007-2009 é constituída por 301 autores (o que representa 5,71% do total de autores das 50 comunidades), número similar ao triênio anterior. Estes autores estabeleceram 1.133 relacionamentos, arestas (Figura 43). O grau médio é 7,52, isto é, cada autor colabora, em media, com mais de sete pesquisadores, e o coeficiente de *clustering* é de 0,79.

Figura 43: Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2007-2009



Fonte: Elaborado pelo autor

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação (34,88%), seguida por docentes (25,58%) e egressos da pós-graduação (22,59%). Autores com outros tipos de vínculo representam 7,64%, alunos de graduação 6,98% e técnicos 2,33%.

Estes dados mostram total coerência com a distribuição da participação de autores da UFRJ em artigos com colaboração apresentada na Tabela 27.

Os vinte autores com as maiores medidas de centralidade de grau são mostrados na Tabela 30. Podemos notar que dezoito dos autores são docentes, um é aluno de pós-graduação e o outro é egresso da pós-graduação.

Em relação às áreas dos autores, verificamos que todos atuam na área de Ciências Biológicas, sendo que um deles também participa das Ciências Humanas.

Tabela 30: Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2007-2009

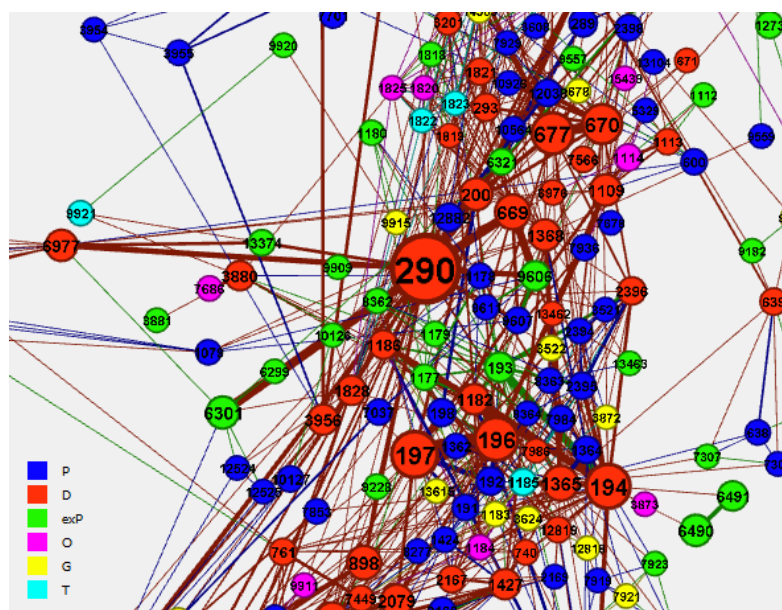
Id	Vínculo*	Grau	Áreas de Atuação
290	D	54	Ciências Biológicas
194	D	41	Ciências Biológicas
1095	D	38	Ciências Biológicas
1365	D	34	Ciências Biológicas
1368	D	33	Ciências Biológicas
193	exP	31	Ciências Biológicas
1828	D	29	Ciências Biológicas
2078	D	29	Ciências Biológicas
677	D	27	Ciências Biológicas
19	D	26	Ciências Biológicas
293	D	26	Ciências Biológicas
1427	D	26	Ciências Biológicas
294	D	25	Ciências Biológicas
1829	D	25	Ciências Biológicas, Ciências Humanas
3957	D	25	Ciências Biológicas
602	D	24	Ciências Biológicas
1827	D	24	Ciências Biológicas
670	D	23	Ciências Biológicas
1826	P	23	Ciências Biológicas
669	D	22	Ciências Biológicas

Fonte: Dados da pesquisa

* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação);
exP (Egresso Discente de Pós-graduação)

Uma visualização detalhada da rede em torno do grupo de autores mais centrais (Figura 44) mostra maior concentração de colaboração entre docentes. Chama a atenção, por outro lado, a significativa participação, além dos alunos de pós-graduação, dos egressos da pós-graduação em torno dos autores mais representativos em termos de grau (Id = 290 e 194).

Figura 44: Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2007-2009



Fonte: Elaborado pelo autor

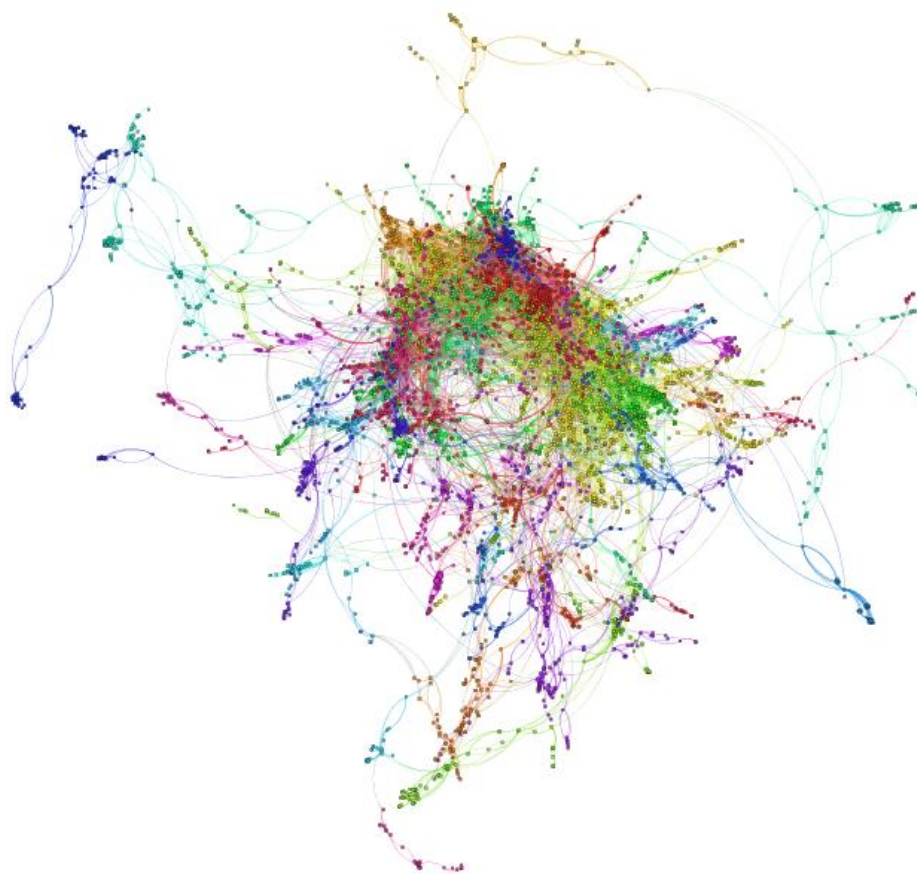
5.3.4.4. Comunidades de Autores no Triênio 2010-2012

A rede de autores da UFRJ, no período 2010-2012, é composta por 7.503 autores (total de autores em artigos com colaboração, Tabela 27). A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel *et al.* (2008), apontou um total de 42 comunidades, envolvendo 5.169 autores, os quais mantiveram entre si 19.328 conexões, representadas na rede por arestas (Figura 45). O grau médio foi de 7,47 e a transitividade (coeficiente de *clustering*) de 0,78.

Note-se que o número de autores e de comunidades diminuiu em relação ao período anterior, 2007-2009, entretanto, o número de conexões aumentou. Isto indica que o aumento da colaboração ocorreu pelo incremento dos agrupamentos já existentes em períodos anteriores.

O grafo da Figura 45 mostra as diversas relações envolvendo as 40 comunidades de autores da UFRJ, agrupadas por cores, identificadas no período 2010-2012. A seguir, destacamos a maior delas.

Figura 45: Grafo que mostra as 40 comunidades de autores da UFRJ identificadas no período 2010-2012



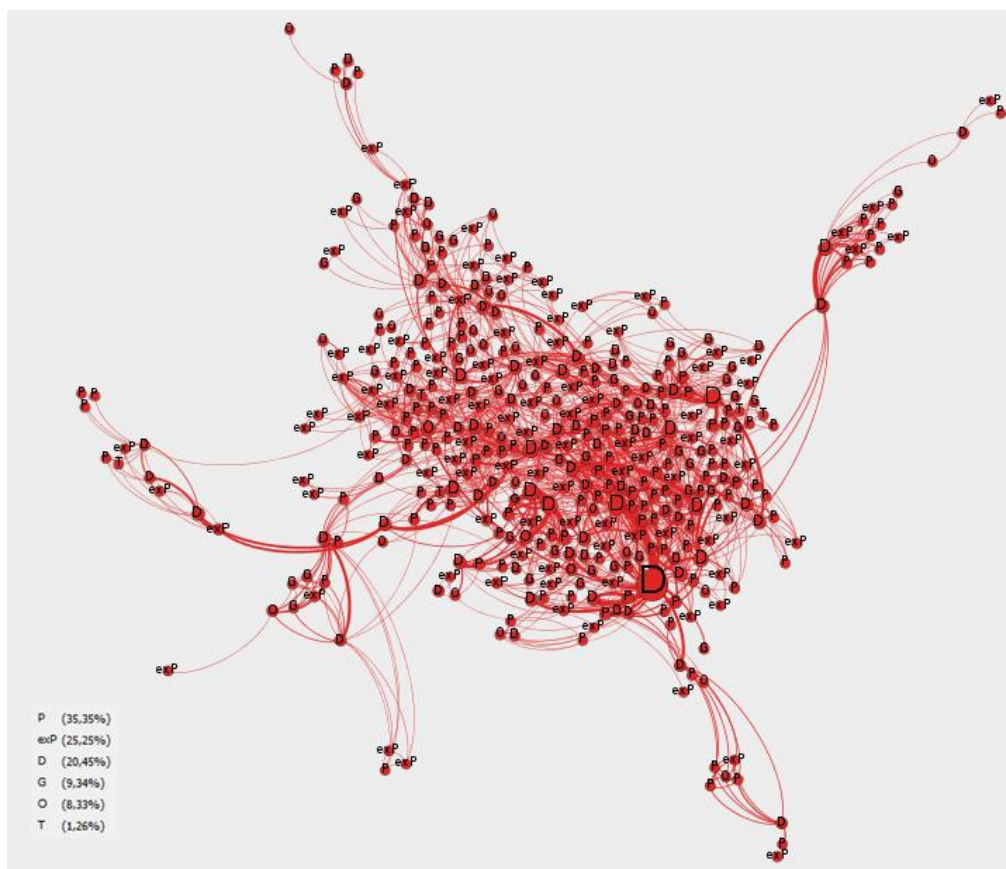
Fonte: Elaborado pelo autor

A maior comunidade de autores identificada no período de 2010-2012 é constituída por 396 autores (o que representa 7,66% do total de autores das 40 comunidades), 95 autores a mais que no triênio anterior. Estes autores estabeleceram 1.607 relacionamentos, arestas (Figura 46).

O grau médio é 8,12, isto é, cada autor colabora, em média, com pouco mais de oito pesquisadores, e o coeficiente de *clustering* é de 0,81.

Ao verificar os autores dessa comunidade de acordo com seus respectivos tipos de vínculo, vemos que os grupos mais representativos são os alunos de pós-graduação (35,35%), seguido pelos egressos da pós-graduação (25,25%) e docentes (20,45%). Os estudantes de graduação representam 9,34%, os autores com outros tipos de vínculo 8,33% e os técnicos 1,26%.

Figura 46: Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2010-2012



Fonte: Elaborado pelo autor

Desse total, identificamos os vinte autores que apresentam as maiores medidas de centralidade (Tabela 31).

Verifica-se que dezenove deles são docentes e um aluno de pós-graduação. Em relação às áreas dos autores, constatamos que todos atuam na área de Ciências Biológicas, sendo que três deles atuam também nas seguintes áreas: Ciências da Saúde; Ciências Exatas e da Terra e; Ciências Humanas.

Tabela 31: Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2010-2012

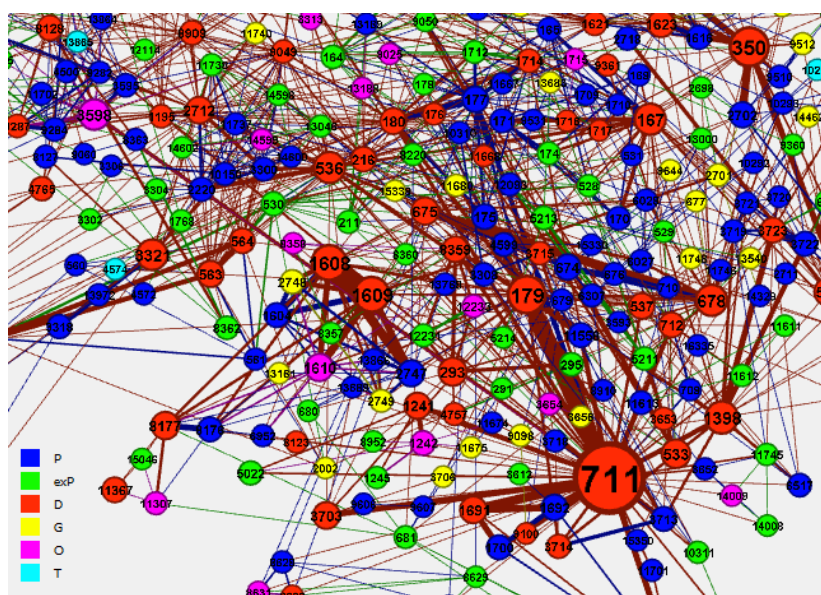
Id	Vínculo*	Grau	Áreas de Atuação
711	D	80	Ciências Biológicas
536	D	47	Ciências Biológicas
2712	D	44	Ciências Biológicas
180	D	43	Ciências Biológicas
179	D	37	Ciências Biológicas
678	D	37	Ciências Biológicas
350	D	36	Ciências Biológicas
9049	D	36	Ciências Biológicas
167	D	34	Ciências Biológicas
168	D	34	Ciências Biológicas
334	D	33	Ciências Biológicas
177	P	32	Ciências Biológicas
1436	D	27	Ciências Biológicas
1623	D	27	Ciências Biológicas
563	D	25	Ciências Biológicas
564	D	25	Ciências Biológicas
1398	D	25	Ciências Biológicas, Ciências Exatas e da Terra
1714	D	25	Ciências Biológicas, Ciências Humanas
1608	D	23	Ciências Biológicas
1609	D	23	Ciências Biológicas, Ciências da Saúde

Fonte: Dados da pesquisa

* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação)

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações importantes (Figura 47).

Figura 47: Grafo que destaca os autores mais centrais da maior comunidade da UFRJ identificada em 2010-2012



Fonte: Elaborado pelo autor

É possível notar forte interação entre os docentes identificados com os números 711 e 179. O primeiro mantendo, em seu entorno, colaborações com outros docentes e, o último, colaborando mais intensamente com alunos de pós-graduação.

6 DISCUSSÃO E CONCLUSÃO

Esta tese foi motivada pelos seguintes questionamentos: de que maneira estão configurados os domínios científicos da UFRJ? Eles são estáticos ou mudam ao longo do tempo? Os diversos domínios interagem entre si? A partir desses questionamentos, o principal objetivo deste trabalho foi mapear os domínios científicos da UFRJ, com o intuito de verificar conhecimentos e competências da instituição como reflexo das interações entre os diferentes autores de seu corpo social, assim como as trajetórias de sua comunidade científica.

Pretendeu-se alcançar esse objetivo geral, por meio de objetivos mais específicos, analisando o corpo social, particularmente o corpo docente, da UFRJ em relação às atividades em colaboração e à produtividade; observar algumas características da produção intelectual, em relação ao número de autores por item de produção e o tipo e nível de colaboração; verificar a existência de núcleos de conhecimentos e o nível de interação entre eles ao longo do tempo.

Por conseguinte, a consecução desses objetivos demandou a identificação e caracterização da produção científica da UFRJ realizando uma análise bibliométrica de toda a produção de artigos publicados em periódicos, recuperada da base de dados institucional EspaçoSIGMA.UFRJ, envolvendo o período compreendido de 2001 a 2012.

Um dos problemas que apresenta a delimitação temática dos domínios científicos é a diversidade de esquemas de classificação que existem, tanto para a mensuração de insumos (*inputs*) como de resultados (*outputs*), dificultando as possibilidades de comparação (GÓMEZ *et al.*, 2006). Neste trabalho, optamos pela classificação de áreas do conhecimento utilizada pelo CNPq (MCTI), que está configurado em nove grandes áreas do conhecimento, que por sua vez se desdobram em setenta e cinco áreas. O uso dessa classificação aproxima a análise da instituição à realidade de pesquisa no País.

A partir da análise das informações de áreas do conhecimento das produções, registradas espontaneamente pelos autores dessas produções – artigos publicados em periódicos, os resultados indicam que as grandes áreas temáticas com maior número de publicações científicas, na UFRJ, são: Ciências da Saúde, Ciências Biológicas e Ciências Exatas e da Terra. Muito embora sejam as de maior representatividade, estas grandes áreas são também as que apresentam decréscimo na fração de artigos publicados entre o 1º e 4º triênio do estudo.

Para as demais áreas, encontramos perfis distintos. A grande área Engenharias revela produção constante de artigos publicados em periódicos, representando 13% do total em todos os períodos considerados. Já a grande área Ciências Humanas é a que registra o maior crescimento, em todo o período, na fração de artigos publicados: de 9,95%, em 2001-2003, para 21,61%, em 2010-2012.

A área de Linguística, Letras e Artes mostra oscilação no número de artigos publicados ao longo dos quatro triênios, respondendo, em média, por 8% do total. Já Ciências Sociais Aplicadas registra aumento na fração de artigos nos três primeiros períodos, de 5% para 8%, e queda no último triênio, respondendo por 5% do total de artigos em 2010-2012. Outras grandes áreas, Ciências Agrárias e Outros, representam menos de 1% do total de artigos publicados em todos os períodos considerados.

Assim, nenhuma das áreas que tradicionalmente tem o artigo em periódico como principal veículo para comunicar seus resultados de pesquisa mostrou crescimento real na produção científica da UFRJ, diferentemente do apresentado pela área de Humanas, que duplicou o percentual de artigos de 2001-2003 para 2010-2012 (rever Tabela 1). Esse quadro nos leva a pensar em pelo menos três possíveis explicações, que podem não ser exclusivas: (a) as grandes áreas tradicionais estão saturadas, ou seja, já chegaram ao seu limite de produção, (b) as ciências humanas estão aderindo com maior velocidade à dinâmica de produção em periódicos, provavelmente como resultado do sistema de avaliação das pós-graduações e (c) as limitações do registro de informações na base de dados.

Em nível mais específico da análise, áreas de conhecimento, os resultados reforçam a forte presença das Ciências da Saúde, Biológicas e Engenharias: áreas

com maior quantidade de registros são Medicina, Química, Bioquímica, Zoologia, Microbiologia, Engenharia Química, Engenharia de Materiais e Metalúrgica. Além destas, encontramos uma destacada presença de registros nas áreas de Letras, História, Educação, Psicologia (rever Figura 15).

Outro aspecto observado foi o nível de interação entre as áreas. Verificou-se que em torno de 30% dos artigos publicados são vinculados a mais de uma área do conhecimento.

Ao longo do período estudado, há uma dinâmica da interação entre as áreas, na qual se destacam, como áreas centrais, as Ciências da Saúde, Engenharias e Ciências Humanas. Ou seja, são estas as grandes áreas que mais interagem entre si e com as outras grandes áreas. Importante destacar que a interação aqui está na compreensão do autor/a dos trabalhos que entendem que os mesmos têm abordagem múltipla, interdisciplinar, isto é, que ultrapassam os limites de suas próprias áreas.

A análise mais detalhada relativamente ao desdobramento das grandes áreas nas respectivas áreas, totalizando 75 áreas, permitiu verificar aspectos mais específicos de sua interação, possibilitando o mapeamento de comunidades de áreas do conhecimento. Os respectivos mapas ou visualizações de domínios permitem identificar relações entre documentos, detectar autores mais representativos numa disciplina, ou analisar a estrutura de determinada área de conhecimento e sua dinâmica através de representações em sucessivos espaços temporais (TUKEY, 1972).

Dessa forma, nossa investigação identificou, no período de 2001-2003, cinco comunidade de áreas, envolvendo um total de 71 áreas do conhecimento. Verificamos que a Comunidade 1 é composta, predominantemente, por áreas das humanidades e ciências sociais. A Comunidade 2 tem alta relação com as engenharias. Na Comunidade 3, prevalecem áreas relacionadas à saúde. Já a Comunidade 4 mostra tendência temática em áreas da biologia e biomedicina. A Comunidade 5, por outro lado, mostra múltipla temática.

A análise dessas comunidades permitiu identificar 26 áreas responsáveis por conectar as cinco comunidades identificadas, entre as quais se destaca Medicina, seguida por Educação, como as principais áreas, tanto em termos de conexão

(grau), de intermediação (troca de informações com áreas pouco conectadas na rede), proximidade, tendo maiores chances de aumentar os trabalhos com mais de uma área, e vínculo com os nós mais representativos em termos de grau (o que indica importância da área). Sobressaem, também, as áreas Multidisciplinar e Ciências da Computação em relação à centralidade de proximidade, colocando-se entre o grupo mais destacado, possuindo chances de aumento de maior interação (aumento do número de trabalhos em colaboração) com outras áreas.

Além disso, verificou-se que a Comunidade 3, com forte tendência temática no campo da saúde, é a que mais contribuiu para a interação com as outras comunidades/áreas na UFRJ no triênio 2001-2003.

No período 2010-2012, identificamos, tal como no período anterior, também cinco comunidades de áreas, envolvendo, no entanto, um número menor de áreas, um total de 68. Verificamos que a Comunidade 1 tem forte influência temática nas engenharias. A Comunidade 2, por outro lado, revelou forte relação com as áreas da saúde e biomédicas. Na Comunidade 3 prevalecem áreas das ciências humanas e sociais. A Comunidade 4 mostra forte tendência de áreas das ciências exatas. Já a Comunidade 5 mostrou também tendência por áreas das ciências sociais e humanas.

A análise dessas comunidades permitiu identificar 16 áreas responsáveis por conectar as comunidades observadas em 2010-2012. Diferentemente do triênio anterior (2001-2003), Educação é apontada como área mais central em termos de grau (número de interações), intermediação (capacidade de conexão, ponte, entre áreas menos conectadas na rede) e proximidade (maiores chances de aumentar os trabalhos em colaboração, por estar estruturalmente bem posicionada na rede). Medicina revelou ter a segunda maior centralidade de grau e de proximidade. No entanto, é área que, em 2010-2012, tem mais ligações com áreas mais representativas, em termos de grau, de acordo com a medida de centralidade do autovetor (*Eigenvector Centrality*).

Por outro lado, verificou-se que a Comunidade 3, com forte tendência de áreas das ciências humanas e sociais, é a que mais contribuiu para a interação com as outras comunidades/áreas na UFRJ no triênio 2010-2012. Este fato mostra, portanto, uma maior representação da área de humanas no domínio científico da

UFRJ nos últimos anos do período estudado. Isto pode indicar novas formas de configuração da produção científica na instituição, em que as pesquisas de docentes de diversas áreas têm assumido maior peso no universo das ciências sociais ou humanas.

Um último bloco de resultados diz respeito aos autores das publicações da UFRJ e interações que estabelecem a partir da medida de coautoria.

Embora haja limitações das medidas de coautoria, este indicador tem se utilizado com sucesso por muitos pesquisadores da área da Bibliometria/Cientometria para estudar a colaboração entre pessoas, instituições e países. Dentre as vantagens do uso da técnica para medir a colaboração científica, podemos mencionar a possibilidade de verificação dos dados por outros autores, visto que eles estão disponíveis publicamente (KATZ; MARTIN, 1997). Além disso, é um método fácil e prático que possibilita a análise de grandes amostras, permitindo, em nosso caso, observar aspectos da atividade científica colaborativa do corpo social da UFRJ.

Verificamos que, em termos absolutos, cresce o número de artigos em coautoria na UFRJ. Entretanto, em termos relativos relacionados à produção de cada triênio analisado, observamos uma tendência de queda no número de artigos contendo entre 1 e 4 autores, enquanto que tende a crescer o número de artigos contendo 5 ou mais autores (rever Tabela 25).

Este quadro constitui traço característico da ciência contemporânea, sendo cada vez mais uma atividade coletiva, multidisciplinar e na qual participam diversos atores (CALERO *et al.*, 2006). Certamente, o salto qualitativo que supõe a colaboração científica reflete a passagem da pequena ciência para a grande ciência, conforme descrito por Solla Price (1963).

A complexidade cada vez maior dos problemas de pesquisa supõe que pesquisadores isolados são capazes de tratar uma quantidade limitada de informações. Assim sendo, pesquisadores procuram se agrupar e interagir através de colégios invisíveis (SOLLA PRICE, 1965; CALLON *et al.*, 1995), ou materializam essa relação em publicações conjuntas, aproveitando os conhecimentos e competências de um conjunto de pesquisadores (DURDEN; PERRI, 1995; ENDERSBY, 1996). Além disso, considerando que conteúdos (e metodologias)

envolvidos nas modernas pesquisas são cada vez mais complexos, a tendência de realizar trabalhos em colaboração deve continuar aumentando (BORDONS *et al.*, 2004).

Nesta tese, introduzimos outro fator que pode ter relação direta com o ato de colaborar na ciência: o vínculo institucional. Assim, inicialmente, observamos a distribuição das autorias da produção científica (artigos em periódicos) da UFRJ pelo tipo de vínculo do/a autor/a.

Verificamos que, no período de 2001-2003, predomina a autoria de docentes (35,33%). Muito próxima da autoria docente esta a autoria de alunos de pós-graduação (31,33%); mais distantes estão os egressos da pós-graduação (11,64%) e alunos de graduação (10,76%). A participação de técnicos e autores com outros tipos de vínculos (que incluem outros egressos - docentes, técnicos e alunos de graduação) como autores de artigos em periódicos representa 4,95% e 6,00%, respectivamente (rever Tabela 26).

A partir do período 2004-2006, constatamos que a maior parte das autorias é de alunos de pós-graduação (32,98%), enquanto os docentes aparecem como o segundo tipo de vínculo mais presente entre os autores (28,46%). Esta situação se mantém até o último triênio estudado, chamando nossa atenção, também, o aumento apresentando pelos egressos da pós-graduação, que passam a representar 21% dos autores.

A análise de como estes tipos de vínculo estabelecem colaborações mostrou que, como era esperado, professores e alunos de pós-graduação são os mais frequentes nos artigos em colaboração (coautoria) em todos os triênios estudados. Na verdade, 56-61% dos artigos da UFRJ em colaboração são de autoria de um professor ou de um estudante de pós-graduação.

Chama atenção a participação dos egressos da pós-graduação, que constituem o terceiro grupo mais representativo da UFRJ em artigos em colaboração. Este fato pode estar relacionado com a inserção dos egressos da pós-graduação da UFRJ em universidades e institutos de pesquisa do setor público, resultado da política de expansão de vagas nessas instituições dos últimos anos (Cf. BRASIL, 2012), mas ainda mantendo projetos de colaboração com seus institutos de origem ou orientadores.

Muito embora a autoria conjunta de trabalhos não seja uma tarefa fácil, alguns autores encontram uma associação direta entre colaboração e produtividade (DURDEN; PERRI, 1995; ENDERSBY, 1996). A necessidade de pesquisadores por publicar pode estar relacionada a diversos motivos, como a busca por prestígio e competitividade com a finalidade de melhorar suas trajetórias científicas (BARNES, 1985), ou a obtenção de credibilidade em relação a seus trabalhos (LATOURE; WOOLGAR, 1997). Adicionalmente, no caso de pesquisadores que ainda não consolidaram sua carreira científica, publicar constitui um requisito fundamental para superar com êxito os processos de avaliação e credenciamento docente.

No Brasil, as avaliações constantes das agências de fomento podem constituir um fator que impulsiona o trabalho colaborativo. O pesquisador precisa publicar para poder ser avaliado (HERMES-LIMA, 2005), fazendo da coautoria um meio para aumentar o número total de publicações de cada pesquisador (VANZ; STUMPF, 2010).

Assim, ao selecionar autores com três ou mais artigos publicados em colaboração para cada triênio na UFRJ, verificamos que, em média, 52% dos autores são docentes e 24% são alunos de pós-graduação, os quais tem apresentado pequeno aumento de produtividade de um período ao outro. Os alunos egressos da pós-graduação têm o aumento mais destacado de participação: de 6,19% em 2001-2003, para 13,15% em 2010-2012.

Por outro lado, a relação entre a produtividade e a colaboração (coautoria) e ligação institucional mostrou uma tendência do aumento da participação de docentes conforme o número de artigos aumenta: professores representam cerca de 80% de autorias entre autores com mais de 10 artigos em colaboração. Este perfil, embora esperado, sugere que professores da UFRJ buscam fortemente associações e colaborações como estratégias para publicar seus estudos.

Sabe-se, contudo, que muitas destas associações são, de fato, com os alunos, seja da graduação ou da pós-graduação, que estão sob a sua orientação. Isso explica a grande presença destes tipos de ligações institucionais como coautores em publicações da UFRJ. A este respeito, é importante mencionar que, como resultado do processo de avaliação dos programas de pós-graduação

conduzida pela CAPES, muitos programas de pós-graduação da instituição exigem dos alunos publicar um (ou mais) artigo como requisito para a defesa de tese.

A maior parte das redes do mundo real caracteriza-se por conter comunidades muito entrelaçadas e ocultas (PALLA *et al.*, 2005). Além disso, cabe considerar que os membros dessas comunidades formam subgrupos, configurando uma imbricada rede entre todos eles.

A identificação de comunidades possibilita definir e melhor conhecer os subgrupos em função da estrutura de escolha de seus enlaces, isto é, com base na premissa de que os membros de cada subgrupo tendem a escolher os mesmos colaboradores e serem escolhidos pelos mesmos autores (PALLA *et al.*, 2005). Dessa forma, a permanência em um subgrupo estabelece-se em função das similaridades de escolha, dadas e recebidas, por cada autor.

Nesse sentido, as métricas de grau e de transitividade mostraram que, ao longo do período analisado, novas colaborações entre autores da UFRJ foram estabelecidas, formando novos agrupamentos/comunidades, tornando, dessa maneira, maior a transitividade no grafo. Em outras palavras, ao longo dos triênios, a tendência de todos os nós (autores) se agruparem aumenta.

Assim, no período de 2001-2003, identificamos um total de 40 comunidades de autores, envolvendo 3.736 autores. Em média, esses autores participaram de publicações com pouco mais de seis autores (o grau médio foi de 6,05), e a transitividade foi de 0,76. Ao analisar a maior comunidade identificada nesse período, verificamos que ela é composta por 236 autores, e que cada autor colabora, em média, com cinco pesquisadores. O coeficiente de transitividade apresentado foi de 0,79. A classificação dos autores por tipo de vínculo institucional mostrou que a maioria é composta por alunos de pós-graduação (38,02%), seguida por docentes (25,48%) e alunos de graduação (12,93%). Alunos egressos da pós-graduação representam 9,89% e os autores com outros tipos de vínculo 8,75%. Já os técnicos representam 4,94%.

No período 2004-2006, houve um aumento do nível de colaboração e expansão do número de agrupamentos/comunidades ($n = 45$) em relação ao período anterior. No triênio seguinte (2007-2009), também se observou aumento do nível de

colaboração e formação de agrupamentos/comunidades ($n = 50$) em relação ao período anterior.

Por fim, no último triênio, identificamos um total de 42 comunidades, envolvendo 5.169 autores (o grau médio foi de 7,47, e a transitividade foi de 0,78), ou seja, diminuiu o número de autores e de comunidades em relação ao período anterior. Isto indica que houve uma intensificação das interações nos agrupamentos já existentes em períodos anteriores. Ao analisar a maior comunidade identificada nesse período (2010-2012), verificamos que ela é composta por 396 autores, 95 autores a mais que no triênio anterior, e que cada autor colabora, em média, com pouco mais de oito pesquisadores (grau médio 8,12). A transitividade apresentada foi de 0,81. A classificação de autores dessa comunidade por tipo de vínculo mostrou que a maioria é composta por alunos de pós-graduação (35,35%), seguido pelos ex-alunos de pós-graduação (25,25%) e docentes (20,45%). Os estudantes de graduação representam 9,34%, os autores com outros tipos de vínculo 8,33% e os técnicos 1,26%.

Observar o desempenho científico do corpo social de uma instituição é de particular interesse para obter informações sobre o grau de eficiência e envolvimento dos diferentes recursos humanos na ciência. Muitos estudos têm buscado esse relacionamento por meio de diferentes abordagens. Larivière (2012), por exemplo, investigou o desempenho de todos os alunos de doutorado matriculados em cursos de pós-graduação em universidades de Quebec, que, segundo o autor, é o primeiro estudo em grande escala com esse foco. Entre os resultados, o autor destaca a alta contribuição de estudantes de doutorado: participando da autoria de cerca de 30% dos artigos de toda a província.

No caso da UFRJ, nosso foco não foi apenas no desempenho dos alunos de doutorado, mas em todos os tipos de relações institucionais. Esta análise só foi possível graças à principal fonte de informação deste estudo, o EspaçoSIGMA.UFRJ. Trata-se, portanto, de um estudo singular, de ampla abrangência, uma vez que se debruça para todo o corpo social de uma das principais instituições de ensino e pesquisa do Brasil, mas cujos resultados não podem ser comparados.

Cabe ressaltar que uma das contribuições destacáveis deste estudo é a particular visão que proporciona sobre a produção científica da UFRJ, já que, a diferença de outros estudos bibliométricos, a abordagem conduzida não se restringe às limitações de realizar estudos deste tipo utilizando as tradicionais bases de dados internacionais – *Web of Science* ou *Scopus*. Bases como estas não permitem observar determinadas especificações da produção científica como, por exemplo, o tipo de vínculo institucional dos autores como aqui mostrado.

Além disso, cumpre mencionar que as bases internacionais apresentam viés linguístico e de cobertura, por catalogarem principalmente periódicos de língua inglesa e mais preferencialmente periódicos publicados em países do hemisfério norte, especialmente Europa e Estados Unidos. Por outro lado, é bastante baixa a cobertura da produção das áreas de ciências sociais e humanidades (HOEKMAN, 2010), não permitindo, portanto, uma visão mais abrangente das áreas no total da produção científica de instituições, em especial nos países em desenvolvimento, já que nestes países grande parte do novo conhecimento é publicado e divulgado por periódicos locais, muitos dos quais não são incorporados às bases mencionadas por não possuírem circulação internacional (UNESCO, 2010).

Da mesma forma, esta tese caracteriza-se por apresentar uma abordagem alternativa aos trabalhos conduzidos a partir da principal base de dados brasileira para este tipo de estudos, a Plataforma Lattes do CNPq. Apesar da indiscutível abrangência e relevância das informações contidas nessa Plataforma, existem aspectos que devem ser considerados para sua utilização (CAÑIBANO; BOZEMAN, 2009), entre as quais: o fato das informações não serem validadas, já que os dados são inseridos pelos próprios usuários sem posterior validação, e a questão da duplicação do registro das produções, já que autores de um mesmo trabalho (coautores) podem registrar a mesma produção mais de uma vez, entre outros.

Por outro lado, cabe considerar que a maioria dos sistemas de informação das universidades e dos instrumentos externos de coleta de dados cria uma representação artificial e forçosamente desarticulada do todo institucional. O EspaçoSIGMA.UFRJ (principal fonte de dados primários deste estudo), ao contrário, adota um modelo de inferência dos níveis de integração das atividades-fim que compreende o registro de múltiplas caracterizações, mediante o qual é possível

localizar e determinar a intensidade das diversas relações entre instâncias da organização funcional e estrutural da UFRJ (FIGUEIREDO, 2006).

A utilização de tais dados torna este estudo, talvez, o primeiro com uma visão mais completa do desempenho científico de todo o corpo social de uma instituição. A despeito de todos os resultados, o trabalho apresentado nesta tese não pretendeu ser exaustivo e, ao longo de seu desenvolvimento, algumas lacunas e novos questionamentos surgiram e poderão ser alvo de futuras investigações.

Seria, por exemplo, interessante descer a níveis de agregação temática mais específica, para conhecer com maior precisão o perfil do domínio científico ao nível de subáreas e especialidades.

Outros aspectos que ficaram pendentes para futuros análises são os relativos à identificação de relações de colaboração entre programas de pós-graduação da instituição, que nos permita detectar a existência de redes de programas. Se bem que já iniciamos este caminho com a identificação de redes de áreas do conhecimento e de autores, achamos que há muita informação por revelar nesse sentido.

Por outro lado, acreditamos que seria oportuno complementar os resultados deste estudo com uma análise qualitativa, de modo a ter uma visão mais completa dos domínios científicos da UFRJ.

REFERÊNCIAS

ARCHAMBAULT, E.; VIGNOLA-GAGNE, E.; CÔTÉ, G.; LARIVIERE, V.; GINGRASB, Y. Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. **Scientometrics**, v. 68, n.3, p.329-342, 2006.

BAILÓN-MORENO, R. **Ingeniería del conocimiento y vigilancia tecnológica aplicada a la investigación en el campo de los tensioactivos**. Desarrollo de un Modelo Cienciométrico Unificado [tesis doctoral]. Granada: Universidad de Granada, 2003.

BARABÁSI, A.; JEONG, H; NÉDA, Z.; RAVASZ, E; SCHUBERT, A; VICSEK, T. Evolution of the social network of scientific collaborations. **Physica A**, v. 311, p. 590-614, 2002.

BARBOSA, D.; AVELINO, L; SOUZA, R.; OLIVEIRA, C.; JUSTEL, C. Medidas de centralidade e detecção de comunidades em rede de co-autoria. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 43, Ubatuba. **Anais...** Ubatuba, 2011. p. 1-10.

BARNES, J. A. Class and committees in a Norwegian Island Parish. **Human Relations**, v. 7, n. 1, p. 39-58, 1954.

BARNES, B. **Sobre ciencia**. Barcelona: Labor, 1985.

BAYER, A.; SMART, J.; MCLAUGHLIN, G. Mapping Intellectual structure of a scientific subfield through authors cocitations. **Journal of the American Society for Information Science**, v. 41, n. 6, p. 444-452, 1990.

BECK, M.; GASPAR, V. Scientometric evaluation of scientific performance at the Faculty of Natural Sciences, kossth Lajos University, Debrecen, Hungary. **Scientometrics**, v. 20, p.27-54, 1991.

BEGHTOL, G. Domain analysis, literary warrant, and consensus: the case of fiction. **Journal of the American Society for Information Science**, v.46, n.1, 1995.

BEN-DAVID, J.; ZLOCZOWER, A. El desarrollo de la ciencia institucionalizada en Alemania. In: BARNES, B. (Ed.). **Estudio sobre sociología de la ciencia**. Madrid: Alianza Universidad, 1980.

BERARD, E. **Essays in object-oriented software engineering**. Prentice Hall, 1992.

BLONDEL, V.; GUILLAUME, J.; LAMBIOTTE, R.; LEFEBVRE, E. Fast unfolding of communities in large networks. **J. Stat. Mech**, p.100-108, 2008.

BORDONS, M.; MORILLO, F.; GÓMEZ, I. Analysis of cross-disciplinary research through bibliometric tools. In: MOED, H; GLANZEL, W; SCHMOCH, U. (Ed.). **Handbook of quantitative science and technology research**. The use of publication and patent statistics in studies of S&T systems. Dordrecht, the Netherlands: Kluwer Academic, 2004. p. 437-456.

BORGATTI, S.P.; MEHRA, A.; BRASS, D.J.; LABIANCA, G. Network analysis in the social sciences. **Science**, v. 323, n. 5916, p. 892-895, 2009.

BORNER, K.; CHEN, C.; BOYACK, K. Visualizing knowledge domains. **Annual Review of Information Science and Technology**, v. 37, p. 179–255, 2003.

BRAAM, R.; MOED, H.; VAN RAAN, A. Mapping of science by combined co-citation and word Analysis: structural aspects. **Journal of the American Society for Information Science**, v.42, n.4, p. 233-251, 1991.

BRADFORD, S. **Documentation**. London: Crosby Lockwood & Sons, 1948.

BRADFORD, S. Sources of information on specific subject. **Engineering: an Illustrated Weekly Journal**, v.137, n.3550, p. 85-6, 1934.

BRASIL. Ministério da Educação. **Análise sobre a expansão das universidades federais 2003 a 2012**; Relatório da Comissão Constituída pela Portaria nº 126/2012. Brasília, 2012.

BRASIL, Ministério da Ciência e Tecnologia. **Indicadores**: apresentação/histórico. Brasília, 2002. Disponível em: <<http://www.mct.gov.br/index.php/content/view/740.html>>. Acesso em: 20 jun. 2013

BRAUN, T. Bibliometric indicators for the evaluation of universities - intelligence from the quantitation of the scientific literature. **Scientometrics**, v. 45, n. 3, p. 425-432, 1999.

BUSH, V. **Science, the endless frontier**: a report to the president. Washington D.C.: U.S. Government Printing Office, 1945a.

BUSH, V. As we may Think. **The Atlantic Montly**. v. 176, p. 101-108, 1945b.

CABAL, A. C. **Ciencia, tecnología, educación superior, gerencia ambiental e integración**: reflexiones. Santa Fe de Bogotá, D.C., Colombia: Solórzano Editores, 2000.

CALERO, C.; BUTER, R.; CABEYO VALDÉS, C.; NOYONS, E. How to identify research groups using publication analysis: an example in the field of nanotechnology. **Scientometrics**, v. 66, n. 2, p. 365-76, 2006.

CALLON, M.; COURTIAL, J.; PENAN, H. **Cienciometría. El estudio cuantitativo de la actividad científica**: de la bibliometría a la vigilancia tecnología. Gijón: Trea, 1995.

CAMÍ, J.; SUÑEN, E; MÉMDEZ-VÁSQUEZ, R. Mapa bibliométrico de España 1994-2002: biomedicina y ciencias de la salud. **Medicina Clínica**, v. 124, n. 3, p. 93-101, 2005.

CAÑIBANO, C.; BOZEMAN, B. Curriculum vitae method in science policy and research evaluation: the state-of-the-art. **Research Evaluation**, Guildford, v. 18, n. 2, p. 86-94, 2009.

- CHEN, C. Information visualization. **Information Visualization**, n. 1, p. 1-4, 2002.
- CHEN, C. Visualizing semantic spaces and author co-citation networks in digital libraries. **Information Processing and Management**, n. 35, p. 401-420, 1999.
- CHEN, C.; CARR, L. Visualizing the evolution of a subject domain: a case study. **IEEE Visualization**, San Francisco: IEEE Computer Society, p. 449-452, 1999.
- CHEN, C.; HICKS, D. Tracing knowledge diffusion. **Scientometrics**, n. 59, p. 199-211, 2004.
- CHEN, C.; IBEKWE, F.; HOU, J. The structure and dynamics of co-citation cluster: a multiple-perspective co-citation analysis. **Journal of the American Society for Information Science and Technology**, v. 61, n. 7, p. 1386-1409, 2010.
- CHEN, C.; PAUL, R.; O'KEEFE, B. Fitting the jigsaw of citation: information visualization in domain analysis. **Journal of the American Society for Information Science and Technology**, v. 53, n. 4, p. 315-330, 2001.
- CLAUSET, A.; NEWMAN, J.; MOORE, C. Finding community structure in very large networks. **Physical Review**, vol. 79, n. 6, p. 66-111, 2004.
- COLE, F. J.; EALES, N. B. The history of comparative anatomy. Part I: A statistical analysis of the literature. **Science Progress**, v.11, n.44, p.578-596, 1917.
- COSTA, J. **La esquemática**: visualizar la información. Barcelona: Paidós, 1998.
- CRANE, D. Social structure in a group of scientist: a test of the "invisible college" hypothesis. **American Sociological Review**, v. 34, n.3, p. 335-53, 1969.
- CROSBY, A. W. **The measure of reality**: quantification and western Society 1250-1600. London: Cambridge University Press, 1997.
- DE ARENAS, J.; CRONIN, B. The contribution of higher education institutions to the development of the Mexican health sciences base. **Journal of Information Science**, vol. 15, n. 6, p. 333-338, 1989.
- DELGADO, H.; RUSSELL, J. Impact of studies published in the international literature by scientists at the National University of Mexico. **Scientometrics**, v. 23, n. 1, p. 75-90, 1992.
- DOYLE, B. Semantic roadmaps for literature searchers. **Journal of the Association for Computing Machinery**, v. 8, n. 4, 1961.
- DURDEN, G. C.; PERRI, T. J. Co-authorship and publication efficiency. **Atlantic Economic Journal**, v. 23, n. 1, p. 69-76, 1995.
- EATON, J.; WARD J.; KUMAR, A.; REINGEN, P. Structural analysis of co-author relationships and author productivity in selected outlets for consumer behavior research. **Journal of Consumer Psychology**, v. 8, n. 1, p. 39-59, 1999.

EDWARDS, Paul N. **The closed world**: computers and the politics of discourse in cold war America. Massachusetts: MIT Press, 1997.

EGGHE, L; ROUSSEAU, R. **Introduction to informetrics**: quantitative methods in Library, Documentation and Information Science. Amsterdam: Elsevier, 1990.

ELSEVIER. **Sciverse.Scopus Facts and Figures**. 2010. Disponível em: <<http://www.info.sciverse.com/scopus>>. Acesso em: set. 2011.

ENDERSBY, J. W. Collaborative research in the social sciences: multiple authorship and publication credit. **Social Science Quarterly**, v. 77, n. 2, p. 375-92, 1996.

FIGUEIREDO, A.M.D. **Espaço SIGMA. Uma visão integrada da atividade acadêmica**: descrição geral. Rio de Janeiro: Scire-COPPE/UFRJ, 35p, 2006. Disponível em <http://projetos.scire.coppe.ufrj.br/adjutorium/images/documents/ies_espacosigma_descricaogeral_v61.pdf> Acesso em: abr. 2010.

FIGUEIREDO, A.M.D. **EspaçoSIGMA.UFRJ. SIGMA.Súmula. UFRJ**: produção intelectual (1998-2009). Rio de Janeiro: Scire-COPPE/UFRJ, 2009. Disponível em <http://download.scire.coppe.ufrj.br/Download/Sumula/2009_05_SIGMA_Sumula_UFRJ_ProdIntelectual_1998_2009.pdf> Acesso em: abr. 2010.

FORTUNATO, S. Community detection in graphs. **Physics Reports**, n. 486, p. 75-174, 2010.

FRAME, J.; CARPENTER, M. International research collaboration. **Social Studies of Science**, v. 9, n. 4, p. 481-497, 1979.

FREEMAN, L. C. **The development of social network analysis**: a study in the sociology of science. Vancouver, BC: Empirical, 2004.

GALCERÁN, H; DOMÍNGUEZ, S. **Innovación tecnológica y sociedad de masas**. Madrid: Síntesis, 1997.

GARFIELD, E. Citation indexes to science: a new dimension in documentation through association of ideas. **Science**, v. 122, n. 3159, p 108-11, 1955.

GARFIELD, E. **From bibliographic coupling to co-citation analysis via algorithmic historio-bibliography**. A citationist's tribute to Belver C. Drexel Filadélfia, EUA, 2001. Disponível em: <<http://garfield.library.upenn.edu/papers/drexelbelvergriffith92001.pdf>> Acesso em: jun. 2010.

GARFIELD, E. The history and meaning of the Journal Impact Factor. **Journal of the American Medical Association**, v. 295, n. 1, p 90-3, 2006.

GARFIELD, E. Mapping the world of science. **150 Anniversary Meeting of the AAAS**, Philadelphia, PA, 1998. Disponível em: <<http://www.garfield.library.upenn.edu/papers/mapsciworld.html>> Acesso em: abr. 2013.

GARFIELD, E.; SHER, I. **Genetics citation index**. Philadelphia: Institute for Scientific Information, 1963.

GARFIELD, E.; SHER, I.; TORPIE, R. **The use of citation data in writing the history of science**. Philadelphia: Institute for Scientific Information, 1964.

GIBBONS, M.; GEORGHIOU, L. **Evaluation of research**: a selection of current practices. Paris: OCDE, 1987.

GLÄNZEL, W. National characteristics in international scientific co-authorship relations. **Scientometrics**, v. 51, n. 1, p. 69-115, 2001.

GLÄNZEL, W.; SHUBERT, A. Analyzing scientific networks through co-authorship. In: MOED, H.; GLÄNZEL, W.; SCHMOCH, U. (Eds) **Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems**. Dordrecht, the Netherlands: Kluwer Academic Publishers, 2004, p. 257-276.

GODIN, B. **Measurement and statistics on science and technology**: 1920 to present. New York: Routledge, 2005.

GODIN, B. Outline for a history of science measurement. **Science, Technology, & Human Values**, v. 27, n.1, p. 3-27, 2002.

GOLDANI, M. Z.; SILVA, C. H. da; NASCIMENTO, L. F. M. do; BLANK, D. A Questão da produção do conhecimento: desafios na gestão dos programas de Pós-Graduação. **Revista Brasileira de Pós-Graduação**, Brasília, v. 7, n. 12, p. 104-116, jul., 2010.

GÓMEZ, I.; BORDONS, M. Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. **Política Científica**, v. 46, p. 21-26, 2006.

GONZÁLES, G.; LÓPEZ, C.; LUJAN, J. **Ciencia, tecnología y sociedad. Una introducción al estudio social de la ciencia**. Madrid: Editora Tecnos, 1996.

GRANOVSKY, Y. V. Is possible to measure science? V.V. Nalimov's research in scientometrics. **Scientometrics**, v. 52, n. 2, p. 127-150, 2001.

GRIFFITH, B.; SMALL, H.; STONEHILL, J.; DEY, S. The structure of scientific literatures II: toward a macro and microstructure for science. **Science Studies**, p. 339-365, 1974.

GROSS, P. L. K.; GROSS, E. M. College libraries and chemical education. **Science**, v. 66, n.1713, p. 385-389, 1927.

GUERREIRO-BOTE. Methods for the analysis of the uses of scientific information: the case of the University of Extremadura (1996-1997). **Libri**, v. 52, n.2, p. 99-109, 2002.

GUTIÉRREZ, P. D. Indicadores científicos: evaluaciones negativas proposiciones positivas. **Investigación Bibliotecológica: Archivonomía, Bibliotecología, e Información**, v. 12, n. 25, p. 64-118, 1998.

HERMES-LIMA, M. Publicar e perecer? **Ciência Hoje**, São Paulo, p. 76-77, jan./fev, 2005.

HERRERO-SOLANA V. Producción científica de la Universidad Nacional de Mar del Plata: análisis de dominio. **Nexos**, v. 8, n. 14, p. 4-10, 2001.

HJORLAND, B. Domain analysis in information science: eleven approaches-traditional as well as innovative. **JDOC**, v. 58, n. 4, p. 422-462, 2002.

HJORLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain analysis. **Journal of the American Society for Information Science**, v. 46, n. 6, p. 400-425, 1995.

HOEKMAN, J. Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe. **Research Policy**, v. 39, p. 662-673, 2010.

HOOD, W. W.; WILSON, C. S. The literature of bibliometrics, scientometrics, and informetrics. **Scientometrics**, v. 52, n. 5, 2001.

HOOK, P. Domain maps: purpose, history, parallels with cartography and applications. In: INTERNATIONAL CONFERENCE INFORMATION VISUALIZATION, 11. **Proceedings....** Zurich, July 4-6, 2007, p. 442-446.

HORTA, J. S. B.; MORAES, M. C. M. O sistema CAPES de avaliação da Pós-Graduação: da área de educação à grande área de Ciências Humanas. **Revista Brasileira de Educação**, Rio de Janeiro, n. 30, set./ dez., 2005.

HUANG, M; CHANG, Y. Characteristics of research output in social sciences and humanities: from a research evaluation perspective. **Journal of the American Society for Information Science and Technology**, v. 59, n. 11, p.1819-1828, 2008.

INGWERSEN, P.; BJONEBORN, L. Methodological issues of webometric studies. In: MOED, H; GLANZEL, W; SCHMOCH, U. (Ed.). **Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems**. Dordrecht, the Netherlands: Kluwer Academic Publishers, 2004, p. 339-69.

IRIBARREN-MAESTRO, I. **Producción científica y visibilidad de los investigadores de la Universidad Carlos III de Madrid en las bases ISI, 1997-2003**. Tesis (Doctorado). Madrid: UC3 de Madrid, 2006. Disponível em: <<http://e-archivo.uc3m.es/bitstream/10016/1088/1/TESIS%20IIM.pdf>> Acesso em: nov. 2011.

JAPIASSÚ, H.; MARCONDES, D. **Dicionário de filosofia**. Rio de Janeiro: Jorge Zahar, 1996.

JIMÉNEZ-CONTRERAS, E. **Universidad de Granada: 1975-1987. La transición científica**. Granada: Universidad de Granada, 1996.

KATZ, J.; MARTIN, B. What is research collaboration? **Research Policy**, n. 26, p. 1-18, 1997.

KLOVDAHL, A. A note on images of networks. **Social Networks**, n. 3, p. 197-214,

1981.

KNOKE, D.; KUKLINSKI, J. **Network analysis**. Beverly Hills: Sage Publications, 1982.

KRAUSKOPF, M. Scientometric indicators as a means to assess the performance of state supported universities in developing countries: the Chilean case. **Scientometrics**, vol. 23 (1), 105-121, 1992.

KRAUSKOPF, M.; VERA, M. I.; ALBERTINI, R. Assessment of a university's scientific capabilities and profile: the case of the Faculty of Biological Sciences of the Pontificia Universidad Católica de Chile. **Scientometrics**, v. 34, n. 1, p. 87-100, 1995.

KUHN, T. S. **The structure of scientific revolutions**. Chicago: University of Chicago Press, 1962.

LARIVIERE, Vincent. On the shoulders of students? The contribution of PhD students to the advancement of knowledge. **Scientometrics**, v. 90, n. 2, p. 463-481, 2012.

LATOURE, B.; WOOLGAR, S. **A vida de laboratório: a produção de fatos científicos**. Rio de Janeiro: Relume Dumará, 1997.

LETA, J. Indicadores de desempenho, ciência brasileira e a cobertura das bases informacionais. **Revista USP**, n. 89, p. 62-77, 2011.

LEYDESDORFF, L. Evaluation of research and evolution of science indicators. **Curr. Sci.**, v. 89, n. 9, p. 1510-7, 2005.

LEYDESDORFF, L.; VAUGHAN, L. Co-occurrence matrices and their applications in Information science: extending ACA to the Web environment. **Journal of the American Society for Information Science and Technology**, v. 57, n. 12, p. 1616-1628, 2006.

LIN, X.; SOERGEL, D.; MARCHIONINI, G. A Self-organizing semantic map for information retrieval. In: ANNUAL INTERNATIONAL ACM/SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 14. **Proceedings...** Chicago, 1991, p. 262-269.

LÓPEZ-YEPES J. La publicación periódica de carácter científico como medio de información documental. Origen y evolución histórica. In: LÓPEZ-YEPES J. (Comp.). **Fundamentos de información y documentación**. Madrid: Eudema, 1989, p. 101-33.

LOTKA, A. J. The frequency distribution of scientific productivity. **Journal of the Washington Academy of Science**, v. 16, n. 2, p. 317-325, 1926.

McCAIN, K. Cocited author mapping as a valid representation of intellectual structure. **Journal of the American Society for Information Science**, v. 37, n. 3, p. 111-122, 1996.

McCain, K. Mapping authors in intellectual space: a technical overview. **Journal of the American Society for Information Science**, v. 41, n. 6, p. 433-443, 1990a.

McCain, K. Mapping authors in intellectual space: population genetics in the 1980s. In: BORGMAN, CL, **Scholarly communication and bibliometrics**. London: Sage Ed., 1990b, p. 194-216.

McORMICK, B.; DEFANTI, T.; BROWN, M. Visualization in scientific computing. **Computer Graphics**, v. 21, n. 6, p. 129-147, 1987.

McCain, K. W.; VERNER, J. M.; HISLOP, G. W.; EVANCO, W.; COLE, V. **Combining bibliometric and knowledge elicitation techniques to map a knowledge domain**. 2006. Disponível em: <<http://vw.indiana.edu/sackler03/ppts/McCain.ppt#1>>. Acesso em: fev. 2010.

MARICONDA, P. R. O controle da natureza e as origens da dicotomia entre fato e valor. **Scientle Studia**, v. 4, n. 3, p. 453-72, 2006.

MARICONDA, P. R.; VASCONCELOS, J. **Galileu e a nova física**. São Paulo: Odysseus, 2006. (Coleção Imortais da Ciência).

MARTIN, B. The use of multiple indicators in the assessment of basic research. **Scientometrics**, v. 36, n. 3, p. 343-362, 1996.

MARTIN, B.R.; IRVINE, J. Some partial indicators of scientific progress in radio-astronomy. **Res Policy**, v. 12, n. 2, p. 61-90, 1983.

MEADOWS, A. J. **Communication in Science**. London: Butterworths, 1974.

MENA-CHALCO, J.; DIGIAMPIETRI, L.; LOPES, F.; CESAR-JR., R. Brazilian bibliometric coauthorship networks. **Journal of the Association for Information Science and Technology**, v. 65, n. 7, p. 1424-1445, 2014.

MENEGHINI, R. (Coord.). Produção científica. In: FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO – FAPESP. **Indicadores de ciência, tecnologia e inovação em São Paulo – 2001**. São Paulo, 2005, cap. 6, p. 6.1-6.22. Disponível em: <<http://www.fapesp.br/indct/cap06/cap06.htm>>. Acesso em: maio 2009.

MIGUEL S.; MOYA-ANEGÓN, F.; HERRERO-SOLANA, V. A New approach to institutional domain analysis: multilevel research fronts structure. **Scientometrics**, v. 74, n. 3, p. 331-344, 2008.

MOED, H.F.; DE BRUIN, R.; VAN LEEUWEN, T. New bibliometrics tools for the assessment of national research performance: database description, overview of indicators and first applications. **Scientometrics**, v. 33, p. 381-442, 1995.

MOED, H.; VAN RAAN, A. Indicators of research performance: applications in university research policy. In: VAN RAAN, A. (Ed.). **Handbook of quantitative studies of Science and Technology**. Dordrecht: Elsevier, 1998, p. 177-192.

MOLINA, J.; DOMÉNECH, M. Redes de publicaciones científicas: un análisis de la estructura de coautorías. **Redes - Revista Hispana para el Análisis de Redes**

Sociales, v. 1, n. 3, 2002. Disponible em: <http://revista-redes.rediris.es/html-vol1/vol1_3.htm> Acceso em: nov. 2011.

MOLINA, J. L. **El análisis de redes sociales**: una introducción. Barcelona: Ediciones Bellaterra, 2001.

MORENO, J. **Fundamentos de la sociometría**. Buenos Aires: Paidós, 1962.

MOYA-ANEGÓN, F. Visualización y análisis de la estructura científica española: ISI Web of Science 1990-2005. **El Profesional de la Información**, v.15, n.4, jul./ago. 2006.

MOYA-ANEGÓN F.; HERRERO SOLANA, V. Análisis de dominio de la Revista Mexicana Investigación Bibliotecológica. **Información, Cultura y Sociedad**, n. 5, p. 10-29, 2001.

MOYA-ANEGÓN, F.; RODRÍGUEZ, Z. **Resultados de investigación científica con visibilidad internacional de la Universidad de Granada (ISI-WOS, 1990-2003)**. Granada: UGR, 2006.

MOYA-ANEGÓN, F.; VARGAS-QUESADA, B.; HERRERO-SOLANA, V.; CHINCHILLA, Z.; CORERA-ALVAREZ E.; MUÑOS, F. A new technique for building maps of large scientific domains based on the cocitation of classes and categories. **Scientometrics**, v. 61, n. 1, p. 129-145, 2004.

NEDERHOF, A.; ZWAAN, R.; DE BRUIN, R.; DEKKER, P. Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural sciences: a comparative study. **Scientometrics**, v. 15, n. 5, p. 423-435, 1989.

NEIGHBORDS, J. **Software construction using components**. Thesis (PhD). Department of Information and Computer Science, University of California, Irvine, 1981.

NEWMAN, M. E. J. The Structure and Function of Complex Networks. **SIAM Review**, v. 45, n. 2, p. 167-256, 2003.

NOTESS, G. Scholarly web searching: Google Scholar and Scirus. **Online**, v. 29, n. 4, 2005.

NOYONS, E. Bibliometric mapping of science in a science policy context. **Scientometrics**, v. 50, n. 1, p. 83-98, 2001.

NOYONS, E. Science maps within a science policy context. In: MOED, H.; GLANZEL, W.; SCHMOCH, U. (Ed.). **Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p. 237-256.

NOYONS, E.; MOED, H.; LUWEL, M. Combining mapping and citation analysis for evaluative bibliometric purposes: a bibliometric study. **Journal of the American Society for Information Science**, v. 50, n. 2, p. 115-131, 1999.

OCDE. **Main science and technology indicators**. Paris, 2014.

OCDE. **Manual Frascati**: medição de atividades científicas e tecnológicas. João Pessoa: CNPq – IBICT, 1978. (Cadernos de Informação em Ciência e Tecnologia, n. 2).

OCDE. **Principios básicos propuestos para la recogida e interpretación de datos de innovación tecnológica**; Manual de Oslo. 2. ed. París, 1996.

OKUBO, Y. **Bibliometric indicators and analysis of research systems**: methods and examples. Paris: OCDE/ GD, 1997.

OSCA-LLUCH, J.; HABA, J. Dissemination of spanish social sciences and humanities. **Journal of Information Science**, n. 31, p. 230-237, 2005.

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

PAISLEY, W. An oasis where many trails cross: the improbable cocitation networks of a multidiscipline. **Journal of the American Society for Information Science**, v. 41, n. 6, p. 459-468, 1990.

PALLA, G.; DERÉNYI, I.; FARKAS, I.; VICSEK, T. Uncovering the overlapping community structure of complex networks in nature and society. **Nature**, n. 435, p. 814-818, 2005.

PERSSON, O. The intellectual base and research fronts of JASIS 1986-1990. **Journal of the American Society for Information Science and Technology**. v. 45, n. 1, p. 31-38, 1994.

PERSSON, O.; MELIN, G. Studying research collaboration using co-authorships. **Scientometrics**, v. 36, n. 3, p. 363-376, 1996.

PERSSON, O.; MELIN, G.; DANELL, R.; KALAOUDIS, A. Research collaboration at Nordic universities. **Scientometrics**, v. 39, n. 2, p. 209-223, 1997.

PIETRO-DIAZ. Domain analysis: an introduction. **Software Engineering Notes**, v. 15, n. 2, p. 47-54, 1990.

PRITCHARD, A. Statistical bibliography or bibliometrics. **Journal of Documentation**, v. 25, n. 4, p. 348-349, 1969.

PROVAN, K.; FISH, A.; SYDOW, J. Interorganizational networks at the network level: a review of the empirical literature on whole networks. **Journal of Management**, v. 33, n. 3, p. 479-516, 2007.

RICYT - Red Iberoamérica de Indicadores de Ciencia y Tecnología. **El estado de la ciencia**. Principales indicadores Iberoamericanos – Interamericanos. Buenos Aires, 2001.

RICYT - Red Iberoamérica de Indicadores de Ciencia y Tecnología. **El Estado de la Ciencia**. Principales indicadores de Ciencia y Tecnología. Buenos Aires, 2009.

RIP, A. Societal challenges for R&D evaluation. In: SHAPIRA, P.; KUHLMANN, S. (Ed.). **Learning from science and technology policy evaluation: experiences from the United States and Europe**. Cheltenham: Esward Elgar Publishing Limited, 2003, p.32-53.

RIVELLINI, G.; RIZZI, E.; ZACCARIN, S. The science network in Italian population research: an analysis according to the social network perspective. **Scientometrics**, v. 67, n. 3, p. 407-418, 2006.

RODRIGUEZ, A. P.; GOMEZ, C. G.; MOYA-ANEGON, F. Detecting, identifying and visualizing research groups in co-authorship networks. **Scientometrics**, v. 82, n. 2, p. 307–319, 2010.

RODRIGUEZ, J. **Análisis estructural y de redes**. Madrid: Centro de Investigaciones Sociológicas, 1995. ISBN: 84-7476-224-3.

RUIZ—BAÑOS, R. **Ciencimetría: análisis de la investigación internacional sobre Arqueología mediante el método de las palabras asociadas (1980-1993)** [tesis doctoral]. Granada: Universidad de Granada, 1997.

RUSSELL, J. M.; DELGADO, H.; ROSAS, A. M.; BLANCAS, G. Estudio bibliométrico de la producción biomédica internacional de los investigadores de la Universidad Nacional Autónoma de México. **Revista Española de Documentación Científica**, v. 15, n. 2, p. 129-139, 1992.

SALVÁ, M. N. R. Reflexos da avaliação Capes no trabalho docente de pós-graduação *stricto-sensu* no campo da Saúde Coletiva. **Revista Cadernos de Estudos Sociais e Políticos**, v. 4, n. 7, jan./jun., 2015.

SANCHO, R. Medición de las actividades de Ciencia y Tecnología, estadísticas e indicadores empleados. **Rev. Esp. Doc. Cient.**, v. 24, n. 4, p. 382-404, 2001.

SCHLUETER, R.; BORLUND, P. Introduction to bibliometrics for construction and maintenance of thesauri: methodical considerations. **Journal of Documentation**, v. 60, n. 5, p. 524-549, 2004.

SCHWARTZMAN, S.; KRIEGER, E.; GALEMBECK, F.; GUIMARÃES, E. A.; BERTERO, C. O. Ciência e tecnologia no Brasil: uma nova política para um mundo global. In: SCHWARTZMAN, S. (Coord.). **Ciência e tecnologia no Brasil: política industrial, mercado de trabalho e instituições de apoio**. Rio de Janeiro: Editora da Fundação Getúlio Vargas, 1995, p.1-59.

SCOTT, J. **Social network analysis**. London: Sage, 2000.

SMALL, H. Co-citation in the scientific literature: a new measure of the relationship between two documents. **Journal of the American Society for Information Science**, v. 24, pp. 265-269, 1973.

SMALL, H. Paradigms, citations, and maps of science: a personal history. **Journal of the American Society for Information Science**, v. 54, n. 5, p. 394-399, 2003.

SMALL, H. Visualizing science by citation mapping. **Journal of the American Society for Information Science and Technology**, v. 50, n. 9, p. 799-813, 1999.

SMALL, H.; GRIFFITH, B. The structure of scientific literature, I: identifying and graphing specialties. **Science Studies**, v. 4, p. 17-40, 1974.

SMITH, M. The trend toward multiple authorship in psychology. **American Psychologist**, n. 13, p. 596-599, 1958.

SOLANA, V. H. Producción científica de la Universidad Nacional de Mar del Plata: análisis de dominio. **Nexos**, v. 8, n. 14, 4-10, 2001.

SOLLA PRICE, J. D. **Little science, big science**. New York: Columbia University Press, 1963.

SOLLA PRICE, J. D. Networks of scientific papers. **Science**, v. 149, n. 3683, p. 510-515, 1965.

SPINAK, E. **Diccionario enciclopédico de bibliometría, ciencia de la información e infometría**. Caracas: UNESCO, 1996. ISBN: 92-9143-007-2.

SPINAK, E. Indicadores Cienciométricos. **Acimed**, n. 9 (Suppl.), p. 42-49, 2001.

STORER N. W. The internationality of science and the nationality of scientists. **International Science Journal**, n. 22, p. 87-104, 1970.

STUMPF, I. R. C. Passado e futuro das revistas científicas. **Ciência da Informação**, v. 25, n. 3, p. 125-128, 1996.

TAGUE-SUTCLIFFE, J. An introduction to Informetrics. **Information Processing & Management**, v. 28, n. 1, p. 1-3, 1992.

TARGINO, M. G.; GARCIA, J. C. R. Ciência brasileira na base de dados do Institute for Scientific Information (ISI). **Ciência da Informação**, Brasília, v. 29, n. 1, p. 1-20, jan./abr. 2000.

TUKEY, J. W. Some graphic and semigraphic displays. In: T. A. BANCROFT (Ed.), **Statistical papers in honor of George W. Snedecor**. Ames: The Iowa State University Press, 1972, p. 293-316.

UNESCO - UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. **Science report 2010**. Paris, 2010.

UNESCO - UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. **Recommendation Concerning the International Standardization of Statistics on Science and Technology**. Paris, 1978.

VAN LEEUWEN, T. N. Descriptive versus evaluative bibliometrics. In: MOED, H.F.; GLANZEL, W.; SCHMOCH, U. (Eds.). **Handbooks of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems**. Dordrecht: Kluwer Academics Publishers, 2004, p. 373-388.

VAN RAAN, A. F. Measuring science. In: MOED, H.; GLANZEL, W.; SCHMOCH, U. (Ed.). **Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p.19-50.

VAN RAAN, A. F. Scientometrics: state-of-the-art. **Scientometrics**, v. 38, n. 1, p. 205-218, 1997.

VAN RAAN A. F.; PETERS H. Structuring scientific activities by co-authors analysis. **Scientometrics**, v. 20, n. 1, p. 235-255, 1991.

VANZ, S.; STUMPF, I. Colaboração científica: revisão teórico-conceitual. **Perspectivas em Ciência da Informação**, v. 15, n. 2, 2010.

VANZ, S.; STUMPF, I. Scientific output indicators and scientific collaboration network mapping in Brazil. **COLLNET Journal of Scientometrics and Information Management**, 2012.

VARGAS QUESADA, B. **Visualización y análisis de grandes dominio científicos mediante redes pathfinder (PFNET)**. Tesis (Doctorado), Universidad de Granada, Granada, 2005.

VELHO, L. Conceitos de ciência e a política científica, tecnológica e de inovação. **Sociologias**, v. 13, n. 26, p.128-153, 2011.

VESSURI, H. El proceso de institucionalización. In: SALOMÓN, J.; SACHS, C. (Comp.). **Una búsqueda incierta. Ciencia, tecnología y desarrollo**. México: Fondo de Cultura Económica/UNU, 1995, p. 199-234.

VIOTTI, E. B. Brasil: de política de C&T para política de inovação? Evolução das políticas brasileiras de ciência, tecnologia e inovação. In: VELHO, L.; SOUZA PAULA, M. C. (Org.). **Avaliação de políticas de ciência, tecnologia e inovação: diálogo entre experiências internacionais e brasileiras**. Brasília: Centro de Gestão e Estudos Estratégicos, 2008.

WAGNER, C.; LEYDESDORFF, L. Network structure, self-organization, and the growth of international collaboration in science. **Research Policy**, n. 34, p. 1608-1618, 2005.

WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

WENGER, E. **Communities of practice: learning, meaning, and identity**. Cambridge: Cambridge University Press, 1998.

WHITE, H. Pathfinder networks and author cocitation analysis: a remapping of paradigmatic information scientist. **Journal of the American Society for Information Science and Technology**, v. 5485, p. 423-434, 2003.

WHITE, H.; McCAIN, K. Visualization a discipline: an author co-citation analysis of information science, 1972-1995. **Journal of the American Society for Information Science and Technology**, v. 49, n. 4, p. 327-355, 1998.

WHITE, H.; McCAIN, K. Visualization of literatures. **Annual Review of Information Systems and Technology**, n. 32, p. 99-168, 1997.

WHITE, H; GRIFFITH, B. Author cocitation: a literature measure of intellectual structure. **Journal of the American Society for Information Science**, v. 32, n. 3, p. 163-172, 1981.

ZIMAN, J. **A força do conhecimento**: a dimensão científica da sociedade. Belo Horizonte: Itatiaia [1977], 1981.

ZIMAN, J. **¿Qué es la ciencia?** Madrid, Cambridge University Press, 2003, 384p.

ZIPF, G.K. **Psycho-biology of languages**. Houghton-Mfflin, 1935.

ZUCCALA, A. Modeling the invisible college. **Journal of the American Society for Information Science and Technology**, v. 57, n. 2, p. 152-168, 2006.

Fontes Eletrônicas:

<http://www.sigma-foco.scire.coppe.ufrj.br/site/foco/index.htm>

<http://www.sigma-sumula.scire.coppe.ufrj.br/>

http://www.sigma-ajuda.scire.coppe.ufrj.br/index.php?option=com_frontpage&Itemid=1

<http://pajek.imfm.si/doku.php?id=pajek>

<https://gephi.org/>

<http://www.vosviewer.com>