

DOMÍNIOS CIENTÍFICOS: mapeamento de áreas do conhecimento da Universidade Federal do Rio de Janeiro

Roberto Mario Lovón Canchumani*
Jacqueline Leta**
Antonio MacDowell de Figueiredo***

RESUMO: Este artigo tem como principal objetivo mapear os domínios científicos da Universidade Federal do Rio de Janeiro (UFRJ), identificando as principais áreas de conhecimento da instituição. Trata-se de um estudo de natureza quantitativa, envolvendo informações de áreas vinculadas à produção científica dos programas de pós-graduação da UFRJ, que utiliza técnicas bibliométricas e de análise de redes a partir das medidas de co-ocorrência. O trabalho investigou 30.669 registros recuperados da base de dados institucional, Espaço SIGMA.UFRJ, de artigos publicados em periódicos no período de 2001 a 2012. 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, 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.

Palavras-chave: Domínios Científicos. Áreas do Conhecimento. Bibliometria. UFRJ. Mapas da Ciência.

* Doutor em Ciência da Informação pela Universidade Federal do Rio de Janeiro, Brasil.
E-mail: rlovon@gmail.com.

** Doutora em Química Biológica Gestão Educação e Difusão Em Ciência e Tecnologia pela Universidade Federal do Rio de Janeiro, Brasil. Professora Associada da Universidade Federal do Rio de Janeiro, Brasil.
E-mail: jleta@bioqmed.ufrj.br.

*** Doutor em Thermische Verfahrenstechnik pelo Universitaet Stuttgart, Alemanha. Professor Adjunto na Universidade Federal do Rio de Janeiro, Brasil.
E-mail: amdf@scire.coppe.ufrj.br.

I INTRODUÇÃO

No Brasil, 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 (LETA, 2011). 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 (OKUBO, 1997; VAN RAAN, 2004).

Considerando que a produção científica de uma dada instituição é representada pelo conjunto das produções científicas geradas em suas dependências (BRAUN, 1999), é possível imaginar que o perfil e os padrões científicos deste conjunto são determinados por cada

um dos setores ou domínios que compõem a instituição, 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 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 (MOYA-ANEGÓN *et al.*, 2004).

Não obstante, para este tipo de estudo, requer-se a disponibilidade de uma base de informações abrangente, preferencialmente de acesso público, que seja capaz de identificar, ao longo do tempo, a configuração e as disponibilidades de conhecimentos e competências que a instituição possui.

A Universidade Federal do Rio de Janeiro (UFRJ) contou por mais de uma década com uma base abrangente de dados que representava sua estrutura organizacional e o registro de atividades de ensino e pesquisa assim como os resultados alcançados (FIGUEIREDO, 2006). Esta base era denominada EspaçoSIGMA.UFRJ¹ e, a partir das informações contidas foram levantadas evidências que permitiram responder algumas questões de pesquisa, como: de que maneira estavam configurados os domínios científicos da UFRJ? Eles são estáticos ou mudam ao longo do tempo? Os diversos domínios interagiam entre si?

O presente trabalho, que tem como campo empírico uma das maiores e mais importantes instituições de ensino e pesquisa do país, a Universidade Federal do Rio de Janeiro (UFRJ), pretende investigar empiricamente a configuração de seus domínios científicos, um conceito incorporado na Ciência da Informação nos anos de 1990, tal como apresentado na próxima seção.

2 DOMÍNIOS CIENTÍFICOS E ANÁLISE DE DOMÍNIOS

Estudos sobre análise de domínio são mais predominantes na engenharia de *software*. Nesta área, os trabalhos nessa 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” (NEIGHBORS, 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).

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 e ALBRECHTSEN, 1995).

Domínio, nesse contexto, pode ser entendido como a designação dada a uma comunidade de discurso, vinculada a um âmbito em que se desenvolve uma determinada 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

¹ <http://www.sigma.ufrj.br/site/espaco/index.htm>

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

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 (HJORLAND e ALBRECHTSEN, 1995). Isto é, entre os níveis social e individual.

Ações e elementos que envolvem os atores de um determinado domínio (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. Também é um discurso 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).

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), os estudos bibliométricos, que se constituem 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 e HERRERO SOLANA, 2001; VARGAS QUESADA, 2005), é o foco do presente trabalho.

2.2 Bibliometria e análise de domínio: algumas aplicações

Considerando que um domínio pode ser entendido sob as diferentes perspectivas de onde se desenvolve uma atividade, seja ela no

nível temático ou institucional, a diferenciação e separação das partes do domínio pode proporcionar um melhor conhecimento de 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, visualizar/ampliar a capacidade de delimitar a estrutura de relações existentes em uma determinada disciplina (MOYA-ANEGÓN e 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 uma série 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, que é 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.

Hjørland (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, na Suécia, 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, nos EUA, que apresentam uma revisão das técnicas para análise, visualização e posicionamento espacial da informação; e, mais recentemente, Moya-Anegón *et al.* (2006) na Universidade de Granada, Espanha, que desenvolveram uma análise da estrutura científica espanhola baseada na análise de cocitação de autores e de revistas; entre outros.

Nos países latino-americanos, o desenvolvimento deste tipo de estudo 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 desta 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 Leta *et al.* (2006) que destacam o forte papel das universidades públicas como principais instituições de pesquisa no Brasil; 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 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 MATERIAS E METODOS

Para atingir o objetivo proposto, o trabalho utilizou a informação, envolvendo o período de 2001 a 2012, recompilada e armazenada na base de dados EspaçoSIGMA.UFRJ. Esta base 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. Em 2013, por decisões da Pró-Reitora de Pós-graduação da UFRJ, essa coleta deixou de ser realizada para fins avaliativos. A partir de então, a base se encontra inativa.

A escolha da base EspaçoSIGMA.UFRJ 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 UFRJ, que ainda não foi explorada.

3.1 Unidades de análise

Eliminamos a produção científica publicada em periódico como unidade alvo de análise. Para identificar os domínios temáticos, foram consideradas, então, as informações de área vinculadas à produção dos programas de pós-graduação das instâncias/centros da UFRJ.

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).

3.2 Coleta e análise de dados

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

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

A partir destes critérios, foi possível extrair 30.669 registros da produção da UFRJ de artigos publicados em periódicos no período de 2001 a 2012. Todos estes registros possuíam registro de pelo menos uma *área do conhecimento vinculada* a eles, que foi atribuído pelo docente, aluno ou técnico no momento da inclusão da produção na base EspaçoSIGMA.UFRJ.

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

Para análise das áreas do conhecimento da produção, utilizou-se a classificação do CNPq para consolidar as publicações em grandes áreas.

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. Para este estudo, optou-se pela análise de redes a partir de medidas de co-ocorrência.

Para a geração dos mapas para visualização dos domínios da UFRJ, com base nas análises de redes sociais, foi utilizado o *software Gephi*.

4 RESULTADOS

A seguir mostra-se empiricamente a aplicação do método apresentando a fim de atingir os objetivos propostos.

4.1 Análise da produção por grandes áreas de conhecimento

A produção científica da UFRJ foi analisada a partir dos artigos, catalogados na base EspaçoSIGMA.UFRJ, que apresentavam o registro da área do conhecimento (um registro espontaneamente realizado pelos autores). Os 30.669 artigos completos publicados em periódicos, no período de 2001 a 2012, com a informação da área de conhecimento foram, então, agrupados em nove grandes áreas de conhecimento, conforme a classificação da tabela de áreas do conhecimento do CNPq.

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 as maiores frações de artigos publicados. São elas, respectivamente, Ciências da Saúde, Ciências Biológicas e Ciências Exatas e da Terra.

Tabela 1 – UFRJ – Distribuição percentual de artigos publicados em periódicos na base EspaçoSIGMA.UFRJ 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

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

Fonte: Dados da pesquisa

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. A grande área Engenharias apresenta produção constante de artigos publicados em periódicos, representando 13% do total em todos os períodos. Já a grande área Ciências Humanas é a que registra o maior crescimento na fração de artigos publicados.

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 a área de Ciências Sociais Aplicadas registra aumento na fração de artigos nos três primeiros períodos e queda no último triênio. 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, sugerem 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).

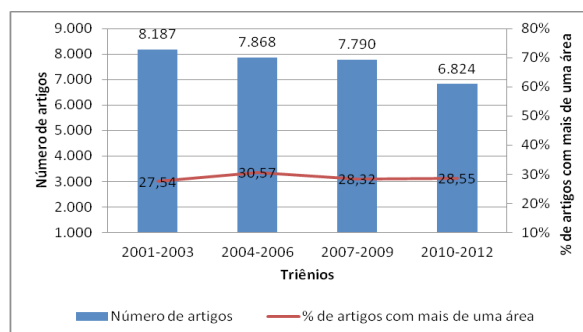
4.2 Análise da produção por grandes áreas de conhecimento

Do total de 30.669 artigos publicados em periódicos no período de 2001 a 2012, 8.806 artigos (o que representa 28,71% do total) apresentavam 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.

A Figura 1 mostra a distribuição dos artigos por triênios e o percentual de artigos associados a mais de uma área do conhecimento em quatro triênios ao longo do período 2001-2012. 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 percentual de artigos que registram mais de uma área do conhecimento é quase constante no período estudado, em torno de 28%.

Figura 1 - UFRJ - Total e percentual de artigos publicados em periódicos registrados na base EspaçoSIGMA.UFRJ vinculados a mais de uma área do conhecimento por triênios

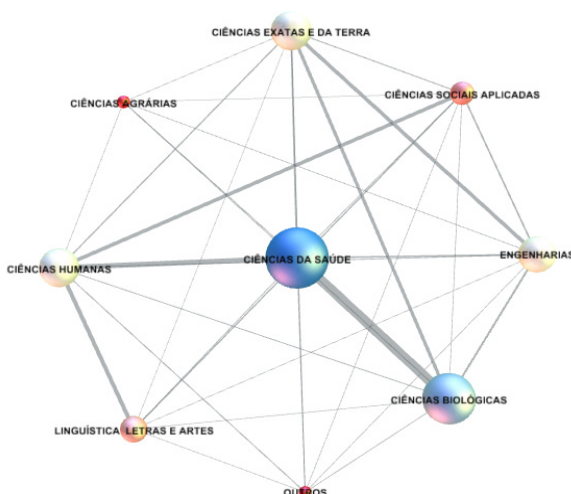


Fonte: Elaboração própria

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

A fim de visualizar esta interação, foi elaborada a Figura 2, 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.

Figura 2 - Mapa da interação entre grandes áreas do conhecimento de artigos publicados em periódicos da UFRJ, cadastrados na base EspaçoSIGMA.UFRJ (2001-2012)



Fonte: Elaboração própria

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. *É possível* observar, por exemplo, 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.

Em outro ponto do mapa, observa-se um 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 há *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)*.

Nota-se, assim, que o compartilhamento de áreas gera uma medida de colaboração entre áreas do conhecimento, mesmo que, neste caso, a intenção de estabelecer uma colaboração venha de um ato espontâneo daquele/a que

registrou a produção na base. Informações como esta, 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, assim, visualizar dinâmicas das áreas do conhecimento e avaliar o grau de inter-relação entre especialidades. Dessa forma, a elaboração de séries históricas permite observar as mudanças nas tendências de pesquisa, 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, elaborou-se 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.

A fração de artigos com uma área nas Ciências Biológicas é praticamente 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 um decréscimo, para 4%, no último. A área de 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). Já as 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%).

Tabela 2 - UFRJ - Total e percentual de artigos cadastrados na base EspaçoSIGMA.UFRJ 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

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.

4.3 A centralidade das grandes áreas de conhecimento

Para melhor interpretação da interação entre as grandes áreas, calculou-se a centralidade de grau para cada período analisado. Para esta análise, considera-se que a 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.

Tabela 3 – UFRJ – Centralidade de grau de grandes áreas do conhecimento dos artigos em periódicos cadastrados na base EspaçoSIGMA.UFRJ

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
LINGUÍSTICA, 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

Pode-se observar que a área das 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.

Isso significa que as 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.

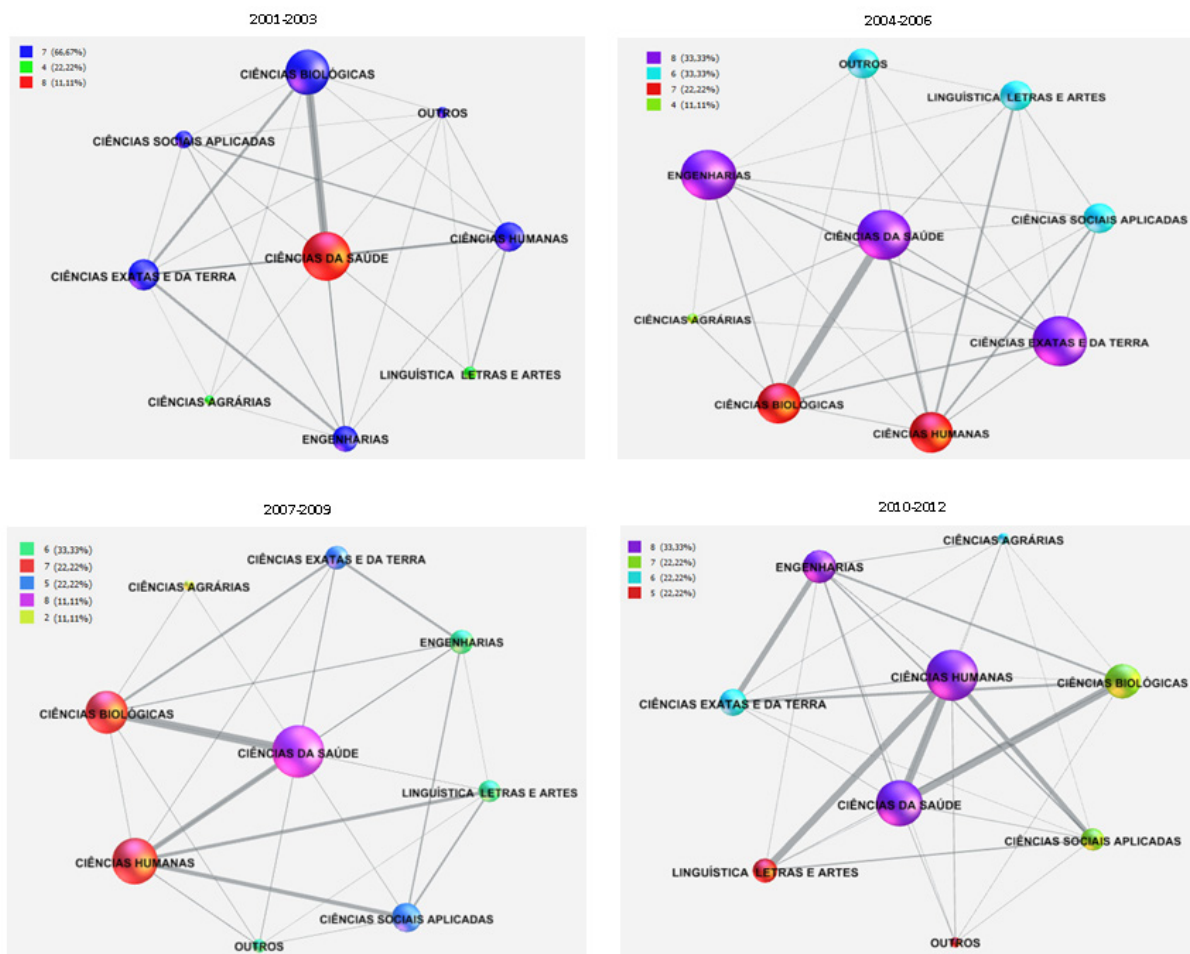
Numa posição mais intermediária encontram-se Linguística, Letras e Artes e as Ciências Sociais Aplicadas. Por outro lado, a grande área mais periférica e, conseqüentemente,

a que menos 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 3, onde são agrupadas cada uma das grandes áreas por grau (representado por uma determinada cor) e período de tempo (triênios).

No triênio 2004-2005, as grandes áreas de Engenharia e Ciências Exatas e da Terra tinham posição central junto com Ciências da Saúde, as três áreas apresentavam grau 8. Já no último triênio (2010-2012), a área das Ciências Humanas toma a posição de Ciências Exatas e da Terra, assumindo, junto com Engenharia e Ciências da Saúde, como grande área central da instituição.

Figura 3 - Mapa da interação entre grandes áreas do conhecimento indicadas nos artigos publicados em periódicos cadastrados na base EspaçoSIGMA.UFRJ por triênios



Fonte: Elaboração própria

4.4 Rede de interação entre as áreas do conhecimento: identificação de comunidades

No sentido de ter uma visão mais detalhada do domínio por áreas do conhecimento, foi realizada uma aproximação a nível microestrutural das áreas. Isto permite observar mais detalhadamente o comportamento das diversas áreas que compõem cada uma das grandes áreas, descritas/analísadas até aqui, assim como verificar o grau de contribuição de

cada uma delas para o domínio institucional da UFRJ ao longo do período estudado.

Cabe mencionar que o desenvolvimento de pesquisas no campo de redes complexas tem apresentado, nas últimas décadas, significativos avanços na descrição e análise dos traços topológicos e dinâmicos de redes de tipo social, biológica ou tecnológica (BARABÁSI *et al.*, 2002). Nesse cenário, um dos aspectos que tem concentrado particular atenção tem sido a existência de subconjuntos de nós fortemente entrelaçados, às vezes conectados e outras vezes desconectados

da rede. Neste último caso, formam-se agrupamentos ou comunidades (*clusters*) que se encontram ligadas ou não à rede (FORTUNATO, 2010).

Essas comunidades cumprem papel relevante nas propriedades de estruturas complexas. Identificar e analisar a natureza das comunidades aí ocultas é uma importante tarefa para revelar a organização informal e a natureza dos fluxos de informação que acontecem dentro desses sistemas complexos.

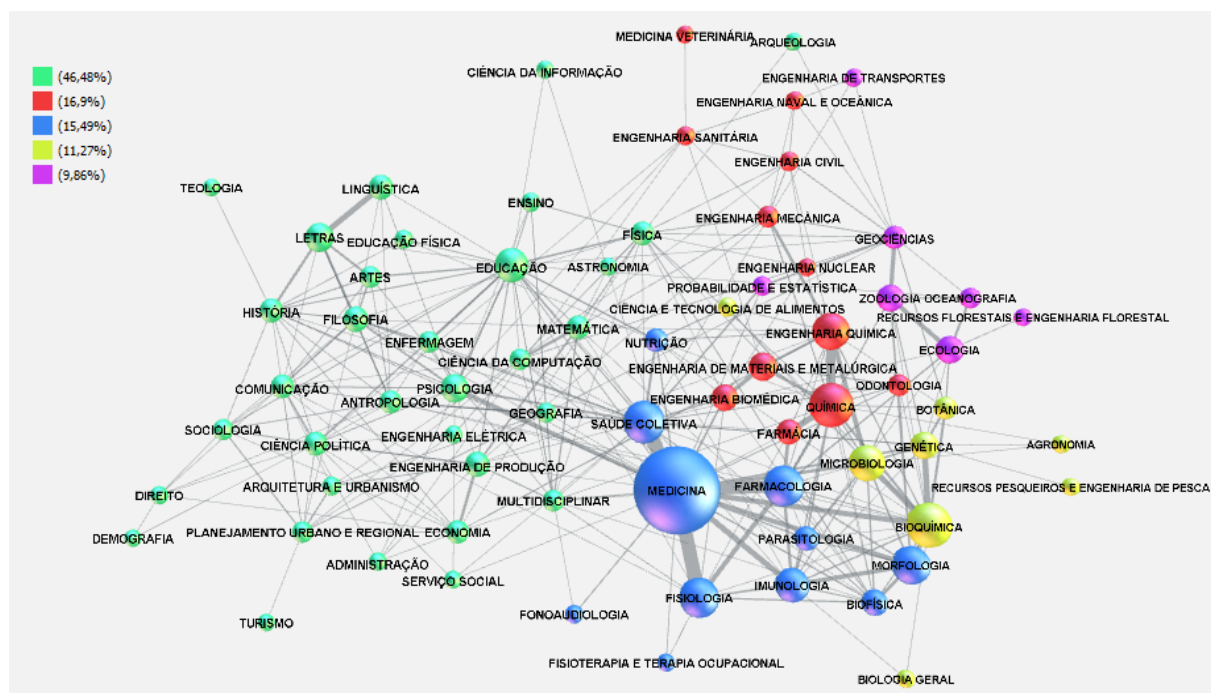
Importante destacar que tais comunidades são compostas por diferentes subconjuntos de nós, os quais se caracterizam por estarem densamente ligados mais entre si do que em relação ao resto da rede. A identificação dessas

comunidades facilita a compreensão das redes. Como nossa intenção é *observar eventuais mudanças destas comunidades*, os dados a seguir são apresentados em dois triênios: 2001-2003 e 2010-2012.

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 4 mostra a rede, destacando as cinco comunidades identificadas, as quais são caracterizadas por cores distintas.

Figura 4 – Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos cadastrados na base EspaçoSIGMA.UFRJ (2001-2003)



Fonte: Elaboração própria

Verifica-se que a Comunidade 1 (cor verde) é composta, predominantemente, por produções originárias nas **áreas das humanidades e ciências sociais. A Comunidade 2** (cor vermelha) tem alta relação com as engenharias. Na Comunidade 3 (cor azul), prevalecem áreas relacionadas à saúde. Já a Comunidade 4 (cor amarela) mostra tendência temática em áreas da biologia e biomedicina. A Comunidade 5 (cor roxa), por outro lado, mostra múltipla temática.

Para identificar as áreas que vinculam as comunidades identificadas, foram isoladas as áreas que possuem apenas arestas entre elas (comunidades). Dessa forma, identificaram-se vinte seis áreas como principais responsáveis por

conectar as cinco comunidades observadas no período de 2001-2003.

A Tabela 4, que mostra a relação dessas áreas e suas respectivas medidas de centralidade, destaca Medicina e Educação como principais áreas, quando se consideram as seguintes medidas: grau (número de conexões); intermediação ou *Betweenness Centrality* (que mede a troca de informações com áreas pouco conectadas no grafo); proximidade ou *Closeness Centrality* (que mede as chances de aumentar os trabalhos com mais de uma área e; *Eigenvector Centrality* (que mede o vínculo com os nós mais representativos em termos de grau, indicando a importância do nó).

Tabela 4 - Relação de áreas e medidas de centralidade responsáveis por conectar as comunidades identificadas em artigos publicados em periódicos e cadastrados na base EspaçoSIGMA.UFRJ em 2001-2003

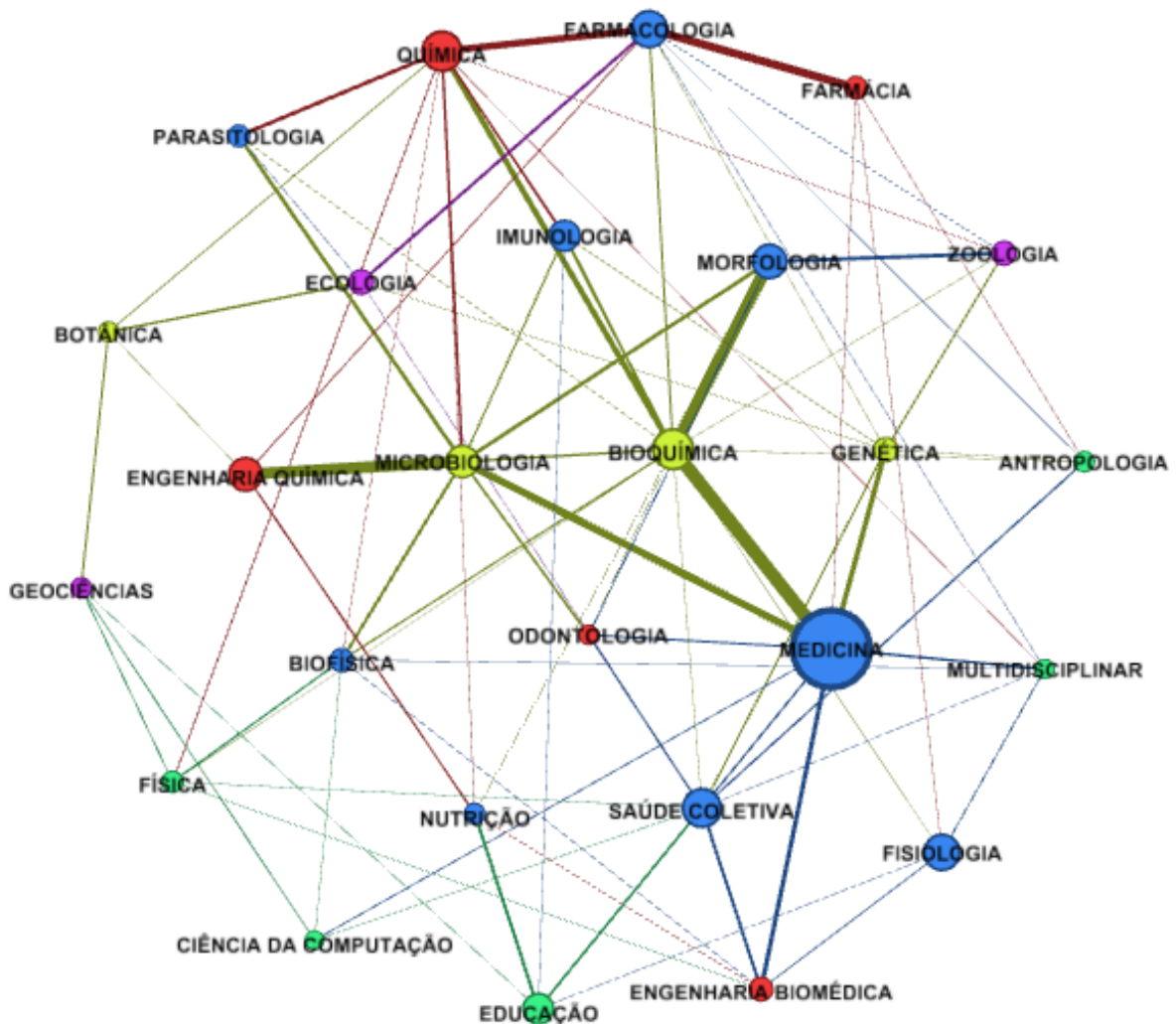
Nodos	▼ Grau	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
● MEDICINA	27	0,135	0,583	1
● EDUCAÇÃO	26	0,133	0,583	0,874
● SAÚDE COLETIVA	22	0,104	0,579	0,854
● BIOQUÍMICA	20	0,054	0,526	0,783
● FÍSICA	18	0,076	0,543	0,64
● QUÍMICA	17	0,038	0,464	0,594
● ANTROPOLOGIA	17	0,059	0,511	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 5 mostra o grafo das principais áreas que conectam as cinco comunidades identificadas no período de 2001-2003.

Figura 5 - Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas em artigos publicados em periódicos e cadastrados na base EspaçoSIGMA.UFRJ no período 2001-2003



Fonte: Elaboração própria

Pode-se notar na Figura 5 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.

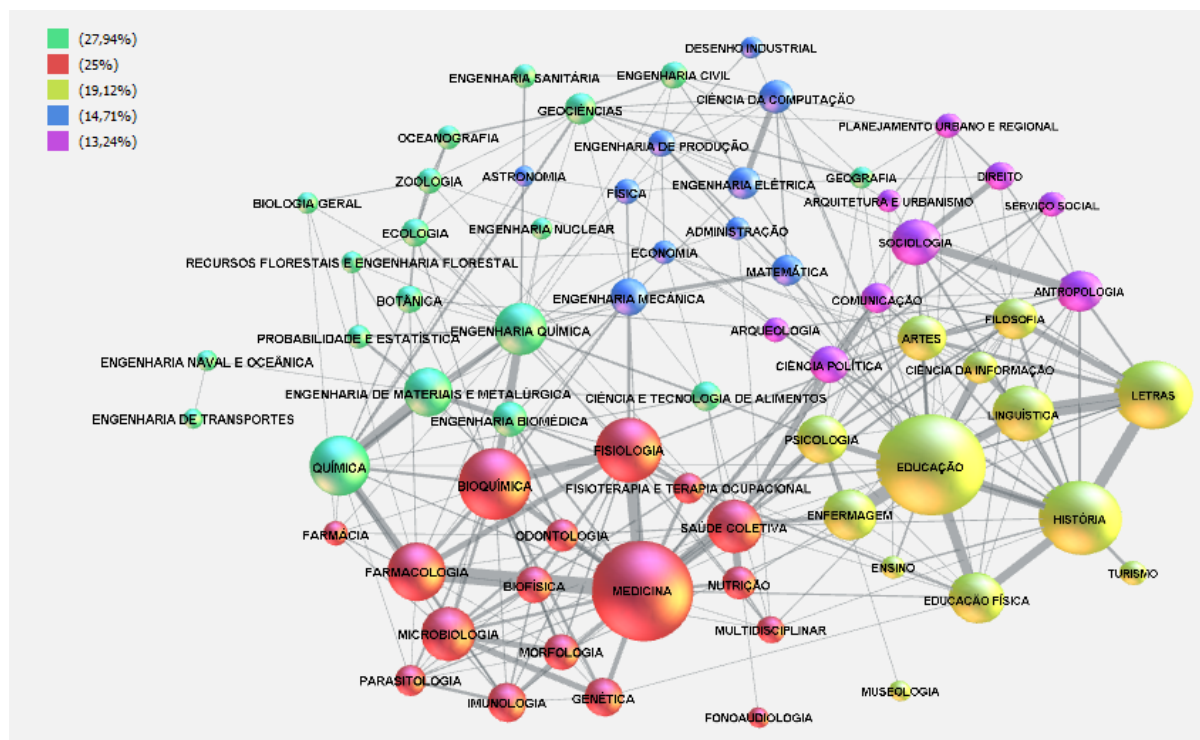
Comunidades e Principais Áreas Identificadas no Triênio 2010-2012

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 6 mostra a rede, destacando as cinco comunidades identificadas, as quais são caracterizadas por cores distintas.

Verifica-se que a Comunidade 1 (cor verde) tem forte influência temática nas engenharias. A Comunidade 2 (cor vermelha), por outro lado, revelou forte relação com as áreas da saúde e biomédicas. Na Comunidade 3 (cor amarela) prevalecem áreas das ciências humanas e sociais. A Comunidade 4 (cor azul) mostra forte tendência de áreas das ciências exatas. Já a Comunidade 5 (cor roxa) mostrou também tendência por áreas das ciências sociais e humanas.

Figura 6 - Mapa das 5 comunidades de áreas identificadas a partir de artigos publicados em periódicos cadastrados na base EspaçoSIGMA.UFRJ (2010-2012)



Fonte: Elaboração própria

Assim como no período anterior, para identificar as áreas que vinculam as comunidades identificadas, isolamos as áreas que possuem apenas arestas entre elas (comunidades). Dessa forma, identificamos

dezesesseis áreas como principais responsáveis por conectar as comunidades observadas em 2010-2012. Essas áreas, e suas respectivas medidas de centralidade, são mostradas na Tabela 5.

Tabela 5 - Relação de áreas e medidas de centralidade responsáveis por conectar as comunidades identificadas em artigos publicados em periódicos e cadastrados na base EspaçoSIGMA.UFRJ 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
● ANTROPOLOGIA	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

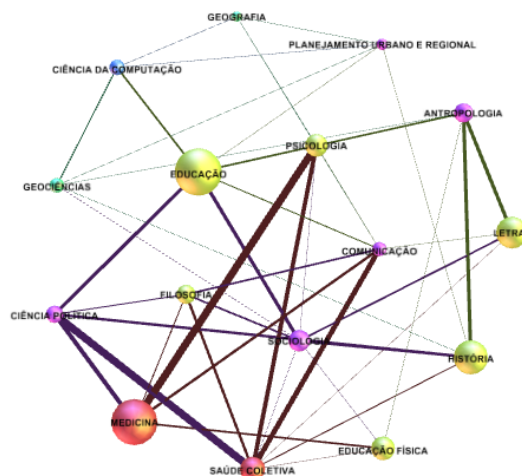
Fonte: Dados da pesquisa

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*.

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 7 mostra as principais áreas responsáveis pela interação entre as comunidades identificadas.

Figura 7 - Grafo que mostra as principais áreas responsáveis por conectar as comunidades identificadas em artigos publicados em periódicos e cadastrados na base EspaçoSIGMA.UFRJ no período 2010-2012



Fonte: Elaboração própria

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 7) com áreas de outros campos, o que não acontece tanto com Educação, por exemplo.

5 DISCUSSÃO E CONCLUSÃO

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 (GOMEZ *et al.*, 2006). Neste trabalho, optamos pela classificação de áreas do conhecimento utilizada pelo CNPq, 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.

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 (Tabela 1). Esse quadro sugere 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, encontrou-se uma destacada presença de registros nas áreas de Letras, História, Educação, Psicologia.

Por outro lado, verificou-se que em torno de 30% dos artigos publicados são vinculados a mais de uma área do conhecimento e que, 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, a presente investigação identificou, no período de 2001-2003, cinco comunidades de áreas, envolvendo um total de 71 áreas

do conhecimento. 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, foram identificadas, tal como no período anterior, também cinco comunidades de áreas, envolvendo, no entanto, um número menor de áreas, um total de 68. A análise dessas comunidades permitiu identificar 16 áreas responsáveis

por conectar as comunidades observadas em 2010-2012. Diferentemente do triênio anterior (2001-2003), Educação aparece como a 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 a 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.

Artigo recebido em 10/03/2017 e aceito para publicação em 02/04/2017

SCIENTIFIC DOMAINS: mapping of fields of knowledge of Federal University of Rio de Janeiro

ABSTRACT: *This paper aims mapping the scientific domains of the Federal University of Rio de Janeiro (UFRJ), identifying the its main fields of expertise. It is a quantitative study that encompasses information related to the fields of knowledge from the scientific production of UFRJ's graduate programs, which uses the network analysis technique based on the co-occurrence measures. The study investigated 30,669 records retrieved from the institutional database, EspaçoSIGMA.UFRJ, articles published in journals from 2001 to 2012. 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.*

Key-words: *Scientific Domains; Fields of Knowledge; Bibliometrics; UFRJ; Maps of Science.*

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*, 68(3), 329-342, 2006.
- BARABÁSI, A.; JEONG, H.; NÉDA, Z.; RAVASZ, E.; SCHUBERT, A.; VICSEK, T. Evolution of the social network of scientific collaborations. *Physica A*, 311, p. 590-614, 2002.
- BEGHTOL, G. Domain Analysis, Literary Warrant, and Consensus: the case of fiction. *Journal of the American Society for Information Science*, 46(1), 1995.
- BERARD, E. *Essays in Object-Oriented Software Engineering*. Prentice Hall, 1992.
- BORNER, K.; CHEN, C.; BOYACK, K. Visualizing Knowledge Domains. *Annual Review of Information Science and Technology*, v. 37, p. 179-255, 2003.
- BRAUN, T. Bibliometric indicators for the evaluation of universities - intelligence from the quantitation of the scientific literature. *Scientometrics*, vol. 45, No. 3, 425-432, 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.; 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*, 53(4): 315-330, 2001.
- 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 (6), 333-338, 1989.
- DELGADO, H.; RUSSELL, J. Impact of studies published in the international literature by scientists at the National University of Mexico. *Scientometrics*, vol. 23 (1), 75-90, 1992.
- 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_descricao geral_v61.pdf>
- FORTUNATO, S. Community detection in graphs. *Physics Reports*, 486, pp. 75-174, 2010.
- GÓMEZ CARIDAD, I.; BORDONS GANGAS, M. Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. *Política Científica*, vol. 46, p. 21-26, 2006.
- GRIFFITH, B.; SMALL, H.; STONEHILL, J.; DEY, S. (1974). The structure of scientific literatures II: toward a macro and microstructure for science. *Science Studies*, p. 339-365, 1974.
- GUTIÉRREZ, P. D. Indicadores científicos: evaluaciones negativas proposiciones positivas. *Investigación Bibliotecológica: Archivonomía, Bibliotecología, e Información*, vol. 12 (25), 64-118, 1998.
- HJORLAND, B. Domain Analysis in Information Science: eleven approaches-traditional as well as innovative. *JDOC*. 58(4): 422-462, 2002.
- HJORLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain analysis. *Journal of the American Society for Information Science*, 46(6), 400-425, 1995.
- 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*, 59(11), 1819-1828, 2008.
- 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*, vol. 34 (1), 87-100, 1995.
- LETA, J. Indicadores de desempenho, ciência brasileira e a cobertura das bases informacionais. *Revista USP*. n 89, p. 62-77, 2011.
- LETA, J; GLANZEL, W; THUS, B. Science in Brazil. Part 2: sectoral and institutional research profiles. *Scientometrics*. vol. 67(1), 87-105, 2006.

- 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: 20 fev. 2010.
- MOYA-ANEGÓN F.; HERRERO SOLANA, V. Análisis de Dominio de la Revista Mexicana Investigación Bibliotecológica. **Información, Cultura y Sociedad**. 5: 10-29, 2001.
- 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.; 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**, 61(1), 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 behavioral sciences: A comparative study. **Scientometrics**, 15(5), 423-435, 1989.
- NEIGHBORS, J. (1981). **Software Construction Using Components**. Thesis (PhD). Department of Information and Computer Science, University of California, Irvine, 1981.
- NOYONS, E. Bibliometric mapping of science in a science policy context. **Scientometrics**, 50 (1), p. 83-98, 2001.
- OKUBO, Y. **Bibliometric Indicators and Analysis of Research Systems: Methods and Examples**. Paris: OCDE/ GD (97) 41, 1997.
- OSCA-LLUCH, J.; HABA, J. Dissemination of Spanish Social Sciences and Humanities. **Journal of Information Science**, nº 31, p. 230-237, 2005.
- PERSSON, O. The Intellectual Base and Research Fronts of *JASIS* 1986-1990. **Journal of the American Society for Information Science and Technology**. 45(1), pp. 31-38, 1994.
- PIETRO-DIAZ. Domain Analysis: an introduction. **Software Engineering Notes**, 15(2), p. 47-54, 1990.
- RICYT (Red Iberoamérica de Indicadores de Ciencia y Tecnología). **El Estado de la Ciencia**. Principales indicadores de Ciencia y Tecnología. Buenos Aires: RICyT, 2009.
- 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**, vol. 15(2), 129-139, 1992.
- SCHLUETER, R.; BORLUND, P. Introduction to Bibliometrics for Construction and Maintenance of Thesauri: methodical considerations. **Journal of Documentation**, 60(5), pp. 524-549, 2004.
- SMALL, H.; GRIFFITH, B. The structure of scientific literature, I: identifying and graphing specialties. **Science Studies**, 4, pp. 17-40, 1974.
- SOLANA, V. H. Producción científica de la Universidad Nacional de Mar del Plata: Análisis de dominio. **Nexos**, vol. 8 (14), 4-10, 2001.
- SPINAK, E. **Diccionario Enciclopédico de Bibliometría, Cienciometría e Infometría**. Caracas: UNESCO. ISBN: 92-9143-007-2, 1996.
- 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, p.293-316, 1972.
- VAN RAAN, A. F. Measuring Science. In: MOED, H; GLANZEL, 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, p.19-50, 2004.
- VARGAS QUESADA, B. **Visualización y análisis de grandes dominio científicos mediante redes pathfinder (PFNET)**. Tesis (Doctorado), Universidad de Granada, Granada, 2005.
- 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, p. 199-234, 1995.
- WENGER, E. **Communities of Practice: Learning, Meaning, and Identity**, Cambridge University Press, 1998.
- 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**. 49(4), pp. 327-355, 1998.