

Institucionalização da pesquisa científica no Brasil: cartografia temática e de redes sociais por meio de técnicas bibliométricas¹

Scientific research institutionalization in Brazil: thematic and social network cartography by means of bibliometric techniques

Nair Yumiko KOBASHI²

Raimundo Nonato Macedo dos SANTOS³

RESUMO

Indicadores de produção científica produzidos por metodologias tradicionais são questionados quando utilizados como parâmetros exclusivos de avaliação da pesquisa científica de países, regiões ou mesmo áreas de conhecimento. Diante do fato, é necessário buscar alternativas teóricas e metodológicas para mapear globalmente e de forma confiável a pesquisa científica nos vários contextos. Para fazer face ao desafio, apresenta-se como hipótese de trabalho a idéia de que os repositórios de dissertações e teses podem ser fontes confiáveis para conhecer a ciência produzida no país. Por meio de fundamentação teórica que associa os Estudos Sociais da Ciência, a Organização e Representação do Conhecimento e Métodos bibliométricos avançados, podem ser vislumbradas novas abordagens sobre as formas de institucionalização social e cognitiva da pesquisa científica no Brasil. A visualização gráfica dessas informações, por meio de cartografias dinâmicas, é útil por sua funcionalidade para oferecer uma visão global de conjuntos de informações e evidenciar relações e estruturas entre elas. Da mesma forma, configura-se como interface poderosa para uso em sistemas de recuperação de informação.

Palavras-chave: institucionalização da pesquisa científica; organização do conhecimento; cartografia temática; indicadores bibliométricos; dissertações; teses.

ABSTRACT

Pointers of scientific production yielded by traditional methodologies are questioned when used as exclusive parameters for evaluation of scientific research done in countries, regions or even knowledge areas. For this reason, it is necessary to seek theoretical and methodological alternatives to map scientific research in various contexts fully and reliably. To face this challenge, a working hypothesis is presented, based on the idea that the repositories of dissertations and theses can be reliable sources of the science produced in Brazil. By means of theoretical grounds that associate Science Social Studies, the Organization and advanced Representation of

¹ Projeto de pesquisa desenvolvido em colaboração entre o Programa de Pós-Graduação em Ciência da Informação da Escola de Comunicação e Artes da Universidade de São Paulo e do Programa de Pós-Graduação em Ciência da Informação da Pontifícia Universidade Católica de Campinas.

² Professora Doutora, Escola de Comunicações e Artes, Departamento de Biblioteconomia e Documentação, Universidade de São Paulo. Av. Prof. Lúcio Martins Rodrigues, 443, Butantã, 05508-900, São Paulo, SP, Brasil. E-mail: <nykobash@usp.br>.

³ Professor Doutor, Programa de Pós-Graduação em Ciência da Informação, Pontifícia Universidade Católica de Campinas. Rua Marechal Deodoro, 1099, Centro, 13045-901, Campinas, SP, Brasil. Correspondência para/Correspondence to: R.N.M. SANTOS. E-mail: <rnsantos@puc-campinas.edu.br>.

Recebido e aceito para publicação em 8/5/2006.

the Knowledge and bibliometric Methods, new approaches can be glimpsed as social and cognitive institutionalization of the scientific research in Brazil. The graphical visualization of this information by means of dynamic cartography is useful due to its functionality, providing a global vision of sets of information and evidencing relations and structures among them. In addition, it is a powerful interface, which is used in information recovery systems.

Key words: *institutionalization of scientific research; organization of knowledge; thematic cartography; bibliometric pointers; dissertation; theses.*

INTRODUÇÃO

A pesquisa em Ciência da informação requer mudanças epistemológicas profundas. Nessa perspectiva, em um texto seminal, García Gutiérrez (2002, 2004, 2005), afirma que pode ela ser refundada por meio da superação das características de uma disciplina

Historicamente acostumbrada a resolver esos problemas acumulando escaso bagaje teórico, ante unos profesionales que no han conseguido ser denominados mediante una etiqueta de amplio consenso, ante unas especializaciones - dentro de la propia profesión genérica - frguadas em el desencuentro mutuo, ante unas operaciones hasta ahora desarraigadas de los procesos de cognición y, portanto, de los procesos de mediación. (GARCÍA GUTIÉRREZ, 2002, p.13).

Para o autor andaluz, a pesquisa sobre os dispositivos de memória institucionalizados

pueden ser calificadas ...como mediaciones. Y estas mediaciones, sobre las inscripciones del pasado, contribuyen, em gran medida, a la reinterpretación de la memoria y, por tanto, a la construcción de la cultura, de la conciencia y las identidades del presente. (GARCÍA GUTIÉRREZ, 2002, p.13).

Os gestos de refundação do campo, acima enunciados, não são tarefas triviais pois requerem abordagens que privilegiem a abertura transdisciplinar em “*nuevos espacios de regociación teórica*” (GARCÍA GUTIÉRREZ, 2002, p.21). Semelhante visão vem sendo desenvolvida por pesquisadores brasileiros, dos quais destacamos aqui Aldo Barreto, Maria Nélide González de Gomes e Suzana Müller, cujas contribuições têm sido decisivas para reconfigurar o campo da Ciência da Informação no contexto brasileiro. A tomada de posição desses pesquisadores tem repercutido, embora não na profundidade requerida, em estudos que procuram aliar pertinência ao campo, profundidade teórica e rigor metodológico, infor-

mados por visão filosófica. Esses estudos desenvolvem-se, por outro lado, evitando aderir aos modismos que periodicamente surgem na área, os quais revelam, via de regra, ausência de visão crítica e escassa bagagem teórica.

A perspectiva de pesquisa ora exposta espera ser uma contribuição para fortalecer as melhores tradições da pesquisa em Ciência da Informação no país. O projeto, que tem como objeto empírico bases de dados de dissertações e teses brasileiras, justifica-se fundamentalmente porque explora dispositivos com funções de perpetuação da memória da atividade intelectual desenvolvida no âmbito acadêmico. Tais memórias, inscrições armazenadas em meio eletrônico, são representações cujo fluxo depende da significação. Sob essa ótica, cabe a pergunta, as memórias atualmente disponíveis têm a garantia de reverberar, ou mesmo sobreviver, nos espaços de circulação hoje privilegiados - as redes eletrônicas -, espaços simultaneamente indeterminados e desprovidos de subjetividade (BHABHA, 1998)? Pode-se afirmar, com segurança, que muitos dispositivos de memória estão submersos. Portanto, resgatá-los e analisá-los são modos de promover a reapropriação do conhecimento produzido. Dito de outro modo, explorar as bases de dados de dissertações e teses produzidas no país, para produzir indicadores, significa rememorar e reavaliar a atividade científica desenvolvida na universidade. É a reflexão sobre o processo de pesquisa sobre esse objeto que apresentamos a seguir, cujo percurso, de abordagem interdisciplinar, entrecruza três campos: os Estudos Sociais da Ciência, a Organização e Representação do Conhecimento e Métodos Bibliométricos avançados. A abordagem proposta parte da leitura crítica de conhecimentos anteriormente produzidos, já que, como afirma Granger, a descontinuidade radical dos paradigmas é um mito, porque os modelos científicos sempre se estabelecem “*a partir de um momento anterior do pensamento.*” (GRANGER, 1989, p.25).

DESAFIOS DOS ESTUDOS DE INSTITUCIONALIZAÇÃO DA CIÊNCIA

Uma primeira questão a ser problematizada nos estudos sobre a institucionalização da ciência diz respeito à natureza e função dos indicadores quantitativos e qualitativos produzidos pela correlação e análise de elementos bibliográficos disponibilizados em bases de dados referenciais. Tais bases, elaboradas inicialmente para armazenar e dar acesso a informações referenciais, configuram-se hoje como dispositivos plurifuncionais amplamente utilizados para avaliar o estado da arte da ciência e da tecnologia.

São três os indicadores clássicos gerados a partir das bases referenciais: os indicadores de produção científica, os indicadores de citação e os indicadores de cooperação (COURTIAL, 1990; MACIAS-CHAPULA, 1998). Apesar das restrições sobre sua representatividade, as prioridades da pesquisa e seu financiamento, em praticamente todos os países, apóiam-se largamente nesses indicadores (PRAT, 1998; TREZNIAK, 1998). Mais precisamente, a produção de indicadores, em nível internacional, baseia-se fundamentalmente nos dados do *Institut for Scientific Information* (ISI), que indexa em torno de 8.500 periódicos científicos dos cerca de um milhão que circulam hoje pelo mundo.

Os dados do ISI estão organizados em três bases de dados. A mais antiga (1961) é a *Science Citation Index* (SCI), multidisciplinar, indexa mais de 5.700 títulos das áreas de ciências exatas e biológicas, tais como agricultura, neurociência, astronomia, bioquímica, biologia, biotecnologia, física, química, ciência da computação e matemática. A *Social Sciences Citation Index* (SSCI), lançada em 1972, indexa mais de 1.725 periódicos das áreas de Ciências Sociais. Entre as disciplinas cobertas por esta base figuram história, direito, sociologia, lingüística, psicologia, antropologia, ciência política, saúde pública e estudos urbanos. A terceira base é a *Arts & Humanities Citation Index* (AHCI), criada em 1978, que indexa 1.144 periódicos da área de artes e humanidades.

A representatividade da base SCI para a produção internacional, especialmente em certos domínios (como Física e Biologia, por exemplo), é pouco contestada pela comunidade científica. O

mesmo não ocorre com a base SSCI porque cobre, via de regra, campos do conhecimento menos internacionalizados. Com efeito, muitas publicações conceituadas da Europa, que não têm o inglês como língua mãe, não são por ela indexadas. Em razão do baixo grau de representatividade, a SSCI é encarada com certa reserva pela comunidade acadêmica internacional, que recomenda cautela na utilização dos indicadores por ela produzidos.

Mesmo em relação às ciências físicas e biológicas, argumenta-se o viés do SCI, em benefício dos periódicos e publicações de língua inglesa, mais especificamente dos norte-americanos, que predominam amplamente na base, comprometem a representatividade da ciência européia e a dos países periféricos. Em algumas áreas científicas, especialmente naquelas que têm um caráter mais aplicado, como a agricultura e saúde, essa distorção é bastante acentuada (LANDI; FURTADO, 2002).

O viés de cobertura utilizado por essa instituição é o principal fator que tem levado os pesquisadores, tanto de países com produção científica de impacto internacional quanto de países periféricos, a discutirem a validade desses indicadores como parâmetros exclusivos de julgamento da performance de pesquisa científica de países, regiões ou mesmo áreas de conhecimento. Observa-se que os dados de partida preferencialmente utilizados nas análises - artigos científicos - não contemplam as peculiaridades de publicação dos resultados de pesquisa das diversas áreas do conhecimento. Com efeito, certas áreas privilegiam a comunicação da ciência por meio de livros, capítulos de livros ou trabalhos publicados em anais de eventos, sendo secundária a publicação de artigos em periódicos científicos.

A metodologia de avaliação de produção científica por meio da mensuração de artigos publicados em revistas internacionais de alto impacto, ou a análise de citações de artigos dessas revistas, parece, portanto, não ser suficientemente potente para dar conta da produção científica efetiva de cada país. Diante do fato, surge a pergunta: que tipos de publicações podem servir de base para mapear de forma confiável a pesquisa científica dos países? Além disso, quais são os parâmetros necessários para analisá-los de modo a obter resultados significativos?

A tentativa de refundação da pesquisa sobre a produção científica requer, desse modo, a adoção

de novos parâmetros teóricos e metodológicos para mapear e interpretar as formas de institucionalização da ciência. Estudos críticos produzidos em países da América Latina, como o México e Cuba, que já têm alguma tradição na produção de indicadores, como também os dados e procedimentos metodológicos desenvolvidos em países da Europa com histórico recente de produção de indicadores, especificamente Espanha e Portugal são fontes preciosas a serem estudadas. Por não serem regiões ou países líderes na produção de pesquisa, esses países têm desenvolvido esforços para criar parâmetros e metodologias adequadas de avaliação da ciência em seus contextos.

Além de metodologias de produção de indicadores adequados, o país necessita constituir infra-estruturas efetivas de informação em Ciência e Tecnologia. Experiências internacionais inovadoras podem subsidiar a institucionalização de estruturas de armazenamento e recuperação aptas a promover estudos que contemplem a realidade brasileira. Abordagens novas vêm sendo feitas em Observatórios de Ciência e Tecnologia, que associam, muitas vezes, a análise de dados à constituição de repositórios (bases de dados bibliográficos), motivadas pela ausência de sistemas de informação de nível nacional nesses países (GUSMÃO, 2000, 2002; OBSERVATOIRE..., 2003).

Vale a pena observar, por último, que, no caso dos países que não fazem parte do grupo de liderança em nível internacional, atribui-se a maior parte dos problemas relacionados à produção de indicadores à precariedade ou mesmo inexistência de bases de dados exaustivas que cubram com confiabilidade a produção científica nacional. A proposição de parâmetros para a construção dessas bases de dados coloca-se, também, como um problema que deve ser pensado no contexto da pesquisa brasileira. Sabe-se que as técnicas bibliométricas avançadas requerem bases de dados com metadados estruturados de forma adequada para a produção de indicadores. Para construir tais bases é necessário, por outro lado, preparar quadros que compreendam a importância do rigor na constituição desses repositórios. É urgente, nesse sentido, criar e fortalecer Grupos de Pesquisa aptos a mapear a ciência brasileira e preparar profissionais aptos a garantir a qualidade das bases de dados. São esses os requisitos para o desenvolvimento de estudos sobre a produção científica que sirvam efetivamente para subsidiar a

tomada de decisões relativas à política de pesquisa, tanto em nível nacional (estudos macro), quanto nas instâncias diretamente envolvidas na pesquisa: grupos de pesquisa, universidades, institutos de pesquisa e programas de pós-graduação (estudos micro).

A precariedade das bases de dados sobre a produção científica brasileira ficou bastante evidenciada na pesquisa "Temas e linhas de pesquisa em Ciência da Informação no Brasil" (KOBASHI, 2004) desenvolvido em associação com o projeto Desenvolvimento de indicadores da produção científica da Ciência da Informação no Brasil (SANTOS, 2005a, b). É alarmante o grau de desorganização, dispersão e falta de padronização dos repositórios de produção científica nacionais. Esses fatos indicam que a realização de análises rigorosas, qualquer que seja a área de conhecimento considerada, encontra barreiras muitas vezes intransponíveis.

Os estudos acima referidos indicaram que, no caso brasileiro, a análise temática de dissertações e teses produzidas em programas de pós-graduação pode fornecer dados preciosos sobre a pesquisa feita no país. Com efeito, os programas de pós-graduação são instâncias altamente institucionalizadas e avaliadas de forma sistemática (COORDENAÇÃO..., 2004). As instituições que abrigam tais programas criam repositórios com fins administrativos e para responder à instância de avaliação competente - a Capes. Conta-se, portanto, com uma massa significativa de dados referenciais sobre essa produção. Pode-se vislumbrar a magnitude da massa de dados disponíveis, por um exemplo específico: a Base de Dissertações e Teses defendidas na Universidade de São Paulo, com cobertura do período 1934-2005, contém atualmente cerca de 70 mil registros. Sabe-se, por outro lado, que partes de teses e dissertações são publicadas sob a forma de artigos ou livros. Desse modo, as dissertações e teses são bases de informações bastante confiáveis para mapear temas de pesquisa, orientadores e linhas de pesquisa e transformá-los em indicadores.

Estudo de natureza semelhante, que tem como objeto empírico as dissertações e teses produzidas nos últimos 30 anos no Programa de pós-graduação em Energia Nuclear do IPEN - São Paulo (KOBASHI; SANTOS, 2006) tem sido fundamental para o desenvolvimento e refinamento de aspectos metodológicos específicos de análise bibliométrica para a formulação de indicadores e a geração e

interpretação de cartografias temáticas e de relações sociais.

A utilidade da visualização de dados por meio de mapas tem respaldado em estudos sobre a percepção, que mostram que o ser humano tem primeiro uma percepção global de uma cena antes de atentar para os detalhes. Tufte (1983) e Bertin (1977) abriram caminhos para explorar essas características da percepção global. A visualização da informação, portanto, baseia-se nessas hipóteses de trabalho sobre a percepção e procura aplicá-las aos sistemas de informação, particularmente nos aspectos relacionados à recuperação de informação, tais como:

1. Exploração rápida de conjuntos de informações desconhecidas;
2. Evidenciação de relações e estruturas nas informações;
3. Fornecimento de alternativas de acesso a informações pertinentes;
4. Classificação interativa de informação.

O OBJETO DE ESTUDO

De forma global, a pesquisa em curso tem vista propor parâmetros metodológicos para mapear tematicamente a produção de ciência no Brasil, descrever sua institucionalização cognitiva e social e produzir indicadores, tendo como objetos empíricos as dissertações e teses produzidas pelo sistema nacional de pós-graduação. Esse objetivo global desdobra-se nos seguintes objetivos específicos:

- caracterizar as áreas de estudo selecionadas quanto aos aspectos de institucionalização cognitiva e social;
- levantar parâmetros de análise e selecionar os mais adequados para os objetivos globais da pesquisa;
- definir a cadeia de análise e os elementos bibliográficos a serem considerados no estudo (processos e técnicas);
- socializar e disseminar os resultados obtidos.

O desenvolvimento desta pesquisa justifica-se, primeiramente, pela presença ainda tímida de estudos desta natureza no país. Como colocado

acima, as fontes e métodos utilizados convencionalmente nos estudos de institucionalização da ciência e produção de indicadores têm recebido críticas por parte dos pesquisadores, tanto de países com produção científica de impacto internacional quanto de países periféricos. Não se considera, atualmente, que os dados produzidos pelo ISI sejam adequados para, isoladamente, julgar a performance de pesquisa científica de países, regiões ou mesmo áreas de conhecimento. É fato amplamente conhecido que diversas áreas privilegiam a comunicação da ciência por meio de livros, capítulos de livros ou trabalhos publicados em anais de eventos e não por meio da publicação de artigos em periódicos científicos (VELHO, 2004). Desse modo, esta pesquisa procura identificar caminhos alternativos de estudo da institucionalização da ciência e sua representação, com base na hipótese de que as teses e dissertações defendidas em uma disciplina são reflexos importantes da atividade de pesquisa institucional. Com efeito, de todas as atividades acadêmicas, tais como publicação de artigos em revistas, organização de colóquios e financiamento de projetos de pesquisa, é ela a mais organizada, certificada por pares e controlada pelas instituições universitárias; é ela também que participa ativamente da reprodução do corpo de pesquisadores da disciplina, sob a supervisão de um orientador previamente reconhecido por autoridades legitimadas. Analisar as teses e dissertações permitirá, portanto, ter um olhar privilegiado sobre a produção de saber de uma dada disciplina. No caso brasileiro, isto é ainda mais verdadeiro, tendo em vista os vieses dos indicadores ISI, expostos acima.

O conhecimento científico e técnico produzido no país não conta, até o presente momento, com estudos suficientes de institucionalização cognitiva e social do conhecimento produzido nas universidades e institutos de pesquisa. Não se dispõem, dessa forma, de estudos abrangentes sobre as temáticas de pesquisa dos programas de pós-graduação, as relações sociais estabelecidas (cooperação) e seu mapeamento. Semelhante situação é responsável, em larga medida, pelas dificuldades enfrentadas pelas universidades e institutos de pesquisa para avaliar sua própria produção e produzir indicadores confiáveis. Compromete-se, dessa forma, a formulação de políticas de pesquisa, incluídas aí a definição de áreas ou projetos prioritários para destinação de recursos.

Esse fato afeta também a formulação de políticas nacionais de pesquisa porque a elaboração de indicadores de produção de âmbito nacional depende da captação e consolidação de dados produzidos localmente por cada instituição de ensino e/ou pesquisa. A solução para esse problema é urgente, na medida em que a apropriação do conhecimento científico pela sociedade, na atual etapa de desenvolvimento, é crucial tanto para o desenvolvimento econômico quanto para a formulação de políticas públicas.

CONSTRUÇÃO DO REFERENCIAL TEÓRICO E DA OPÇÃO METODOLÓGICA

A construção de indicadores rigorosos e significativos é uma atividade complexa. Responder aos desafios do empreendimento implica investigar a institucionalização cognitiva e social da pesquisa, tarefa igualmente complexa porque envolve o estudo de um amplo e heterogêneo espectro de atividades, com resultados e exigências distintos, envolvendo múltiplos agentes e instituições e políticas de ciência. O mapeamento da ciência e a produção de indicadores são antecedidas por uma série de ações: identificação de repositórios de informações, análise de sua estrutura e qualidade para verificar sua pertinência para o objetivo proposto.

As especificidades de cada país quanto aos estágios de institucionalização de cada campo da ciência apontam para a necessidade de redefinir pressupostos antes de produzir indicadores. Produzir indicadores não significa, portanto, apenas manipular dados estatísticos e representá-los por meio de tabelas e gráficos. É preciso, ademais, conhecer em profundidade as características e hábitos de cada setor para elaborar hipóteses de trabalho pertinentes e utilizar metodologias de coleta e análise que produzam resultados confiáveis. No caso brasileiro, o desafio é maior dada a precariedade das fontes de informação disponíveis e, em alguns casos, sua ausência. Deve-se ressaltar que os indicadores de produção científica vêm ganhando importância crescente como instrumentos para análise da atividade científica e das suas relações com o desenvolvimento econômico e social. Sua construção tem sido incentivada pelos órgãos de fomento à pesquisa como meio para se obter uma

visão acurada da produção de ciência, de modo a subsidiar a política científica e avaliar seus resultados.

Os indicadores podem ser definidos, em uma primeira abordagem, como dados estatísticos que representam aspectos da realidade. No caso da atividade científica, há diversos métodos de abordagem, conhecidos por duas denominações, ao menos: Cientometria e Bibliometria. Guardadas as diferenças terminológicas, alguns aspectos operacionais e campos de aplicação, elas têm em comum o fato de se ocuparem do desenvolvimento de metodologias para análise e construção de indicadores, com base em abordagem interdisciplinar. Baseiam-se em indicadores construídos a partir de documentos publicados, que são observados quanto à quantidade, autoria, citações e temas. Esses parâmetros são empregados como medidas indiretas da atividade da pesquisa científica e contribuem para a compreensão dos objetivos da pesquisa, das estruturas da comunidade científica, do seu impacto social, político e econômico (SPINAK, 1996, 1998; ASTON & KLAVANS, 1997; OKUBO, 1997; TRZESNIAK, 1998). Os indicadores bibliométricos ou cientométricos são caracterizados como aproximações da realidade ou uma expressão incompleta dela. Apesar de sua importância, não se dispõe de procedimentos de construção que gozem de suficiente reconhecimento e uso internacional (SPINAK 1998; ORGANISATION..., 2002).

Há um conjunto expressivo de indicadores bibliométricos empregados na análise da produção científica. Eles podem ser divididos em indicadores de produção, indicadores de citação e indicadores de ligação (COURTIAL, 1990; CALLON et al., 1993; NARIN *et al.*, 1994; OKUBO, 1997; MACIAS-CHAPULA, 1998; SPINAK, 1998).

- Os Indicadores de produção científica são construídos pela contagem do número de publicações por tipo de documento (livros, artigos, publicações científicas, relatórios etc.), por instituição, área de conhecimento, país, etc.

- Os indicadores de citação são construídos pela contagem do número de citações recebidas por uma publicação de artigo de periódico. É o meio mais reconhecido de atribuir crédito ao autor.

- Os indicadores de ligação são construídos pela co-ocorrências de autoria, citações e palavras, sendo aplicados na elaboração de mapas de estruturas de conhecimento e de redes de relacionamento

entre pesquisadores, instituições e países. Emprega técnicas de análise estatística de agrupamentos.

Segundo Macias-Chapula (1998, p.136) “[...] a ciência necessita ser considerada como um amplo sistema social, no qual uma de suas funções é disseminar conhecimentos. Sua segunda função é assegurar a preservação de padrões e, a terceira, é atribuir crédito e reconhecimento para aqueles cujos trabalhos têm contribuído para o desenvolvimento das idéias em diferentes campos”. Dadas as peculiaridades da ciência, a comunidade científica de cada área ou subárea adota diferentes processos de publicação e socialização do conhecimento produzido. É fato conhecido que as áreas de ciências exatas e biológicas não têm a mesma cultura de publicação das ciências sociais e humanas. Enquanto as primeiras tendem a privilegiar a publicação de artigos em revistas, nas ciências humanas e sociais, privilegia-se a publicação de livros (MACIAS-CHAPULA, 1998; PRAT, 1998; SPINAK, 1998; TARGINO & GARCIA, 2000). É, portanto, inadequada a universalização do critério de avaliação da produção científica baseada tão somente em artigos de periódicos.

A bibliometria tem evoluído para uma disciplina científica com vários subcampos e estruturas de comunicação científica correspondentes desde o começo da década de 80. Diversos são os conceitos dos subcampos da *bibliometria*, *informetria*, *cienciometria* e *tecnometria* não existindo, no entanto, uma fixação terminológica na área. O campo está se constituindo como disciplina científica, incluindo todos os aspectos estatísticos e matemáticos relacionados aos problemas da biblioteconomia, da documentação e da informação, com fortes vínculos com os aspectos teóricos da recuperação da informação (WORMELL, 1998).

Atualmente, o campo da bibliometria como um todo inclui todos os aspectos quantitativos e os modelos da comunicação científica e do armazenamento, disseminação e recuperação da informação científica. Esse conceito de bibliometria é muito mais amplo do que as definições usuais do termo e objetiva incorporar todas as orientações correntes, como também suas aplicações à política científica e à recuperação da informação (WORMELL, 1998).

Os métodos bibliométricos, por sua vez, vêm sendo aplicados não somente aos estudos cientométricos e à avaliação da pesquisa em ciência e

tecnologia (C&T), mas também à análise de suas relações sociais e econômicas. Com efeito, a ciência, por ser uma força vigorosa na sociedade contemporânea, seu mapeamento e avaliação são questões consideradas cruciais. Uma forma de avaliação da pesquisa científica é aquela feita por pares. Essa avaliação é de natureza qualitativa, usada para julgar propostas de pesquisa, avaliar grupos de pesquisa, entre outros aspectos. Os indicadores bibliométricos, por sua vez, representam o aspecto quantitativo da atividade de avaliação.

Deve-se considerar, no entanto, que os aspectos quantitativos também estão presentes na avaliação por pares, já que se levam em consideração o número de publicações e as atividades acadêmicas próprias de áreas institucionalizadas (Van RAAN, 2003). Esse método se configura, ao nível dos programas de pesquisa, como um instrumento indispensável para a tomada de decisão na política científica, particularmente em relação às prioridades a serem estabelecidas e a fazer balanços retrospectivos.

A análise bibliométrica pode ser feita em nível macro ou micro. Ou seja, podem-se fazer avaliações dos campos como um todo, por exemplo, avaliar a performance do país em termos de áreas científicas. A atividade de pesquisa pode ser analisada sistematicamente, no nível intermediário, em termos de grandes instituições, tais como universidades e institutos de pesquisa. Além disso, a análise pode ser restringida para um nível micro, isto é, ao nível da prática concreta de pesquisa: departamentos, grupos de pesquisa e programas entre universidades e grandes instituições (Van RAAN, 2003).

O princípio básico subjacente às abordagens bibliométricas é a idéia de que a comunicação dos resultados da pesquisa é um aspecto central da ciência. Embora as publicações não sejam os únicos indicadores da atividade científica elas são certamente elementos muito importantes do processo de troca de conhecimento.

O uso de indicadores bibliométricos deve ser avaliado de forma equilibrada e objetiva, particularmente nas Ciências sociais e Humanas, considerando-se que os métodos clássicos utilizados nas ciências exatas e biológicas não podem ser universalizados, tendo em vista os hábitos e valores presentes em cada

comunidade. No entanto, como afirma Van Raan, “devemos nos acautelar contra a aceitação fácil e a persistente caracterização das Ciências sociais e Humanidades como sendo “bibliometricamente inacessíveis.” (Van RAAN, 2003, p.8).

Estudos rigorosos devem, no entanto, qualquer que seja o campo considerado, estabelecer padrões de comparação com grupos de áreas similares. A pesquisa não poderá avançar sem dispor de padrões amplamente aceitos. Outro aspecto a ser considerado é o tamanho do corpus e a dimensão temporal porque a análise de tendências deve estar fincada em parâmetros adequados para cada campo do conhecimento, tendo em vista o seu grau de institucionalização. Segundo Van Raan (2003), “fotos instantâneas” de atividades científicas não são úteis, já que grupos de pesquisa requerem tempo para estabelecer suas posições. Desse modo, seria incorreto julgar a performance de pesquisa com base em número restrito de anos. Quanto ao tipo de documento a ser considerado, por exemplo, afirmam que livros e capítulos de livros constituem em torno de um terço de todas as publicações holandesas em psicologia social. Além disso, um campo de pesquisa pode ser definido por várias abordagens: com base em códigos de classificação e/ou seleção de palavras-chave em uma base de dados específica, conjunto selecionado de revistas, uma base de dados de publicações de campo específico, ou alguma combinação dessas abordagens. Deve ser revisto, no entanto, o método tradicional de tratamento lingüístico computacional aplicado a títulos e resumos para derivar temas (MEERTENS *et al.*, 1992).

Outra alternativa seria trabalhar com informações constantes dos campos de descritores de bases de dados. Parte-se aqui do princípio de que a indexação realizada com base em tesouros fornece informações temáticas mais padronizadas do que as ocorrências estatísticas de palavras de um texto integral. Deve-se lembrar que a organização da informação por meio de tesouros, taxonomias e ontologias vem sendo reconhecida como imprescindível para tratar grandes massas de informações, mesmo no ambiente da Web.

A padronização dos descritores e palavras-chave do corpus considerado, parametrizadas por

vocabulários controlados adotados pela própria fonte de dados é abordagem apropriada para mapear a ciência do ponto de vista temático, já, via de regra, os termos produzidos em cada área são uma das bases de referência dos tesouros, taxonomias e ontologias. Os mapas gerados a partir desses dados e métodos são representações da produção científica da área expressa por meio de conceitos produzidos e utilizados pela própria área. Tem-se, portanto, neste caso, a garantia das terminologias das áreas para a geração das representações cartográficas da institucionalização cognitiva de um dado campo do conhecimento. É necessário dedicar especial atenção à padronização dos descritores para evitar dispersões que possam comprometer a análise e a validade dos resultados. Fica claro, portanto, que estudos de mapeamento da produção científica e produção de indicadores requer, a constituição de bases de dados *ad hoc*.

Em síntese, o marco teórico da pesquisa, ora apresentado, pretendeu dar uma visão panorâmica dos pressupostos e métodos adotados neste projeto. Pretende-se, com isso alcançar resultados objetivados de análise da atividade de pesquisa e sua representação cartográfica. Intuímos, além disso, que os aspectos interdisciplinares da pesquisa aparecerão com força no mapeamento, tendo em vista as características da pesquisa na contemporaneidade.

O P E R C U R S O M E T O D O L Ó G I C O

Para uma primeira aproximação do objeto de estudo, o *corpus* empírico da pesquisa foi constituído de informações bibliográficas de teses e dissertações produzidas em campos específicos do conhecimento: Energia nuclear e Ciência da Informação. A área de Cronobiologia será o próximo campo a ser analisado. A seleção dessas áreas foi estabelecida porque integram campos distintos do conhecimento: Ciências Exatas, Ciências da Vida e Ciências Sociais. Esse fato permitirá identificar as especificidades metodológicas necessárias para cada campo.

No desenvolvimento da pesquisa, foram realizados os seguintes procedimentos:

- Identificação das fontes de informação de dissertações e teses dessas áreas.

- Acesso e recuperação de referências bibliográficas.

- Tratamento de dados bibliográficos para fins de análise bibliométrica;

- Análises e validação de resultados.
- Representação gráfica dos dados obtidos;
- Relatórios intermediários e relatório final.

São fontes de dados:

- as bases de dados de teses e dissertações produzidos por universidades e instituições de ensino e pesquisa do país;

- base de coleta de dados da Capes.

A pesquisa requer, por outro lado, a utilização de técnicas e tecnologias da informação imprescindíveis para assegurar o rigor da pesquisa. Para isso foram utilizados os seguintes softwares:

- para análise bibliométrica: programa INFOTRANS, da empresa alemã luK, empregado para a preparação dos dados com fins bibliométricos. Esse programa permite transformar registros bibliográficos em registros bibliométricos, padronizá-los, complementá-los com dados externos e reorganizá-los, de modo a eliminar registros duplicados.

- para produção de listas de frequência e matrizes de co-ocorrência: programa *VantagePoint*, desenvolvido pelo grupo de pesquisa *Technology Policy and Assessment Center* do *Georgia Institute of Technology*, empregado para a produção de listas de frequência e matrizes que servem de base para gerar gráficos e tabelas. É utilizado para grandes massas de dados. Para análises de nível micro utilizamos o *Dataview*, programa elaborado pelo *Centre de Recherche Rétrospective de Marseille* (CRRM) da *Universidade Aix-Marseille III, Centre de St. Jérôme*, Marselha, França.

- para geração automática de gráficos e clusters: programa *MATRISME*, desenvolvido pelo *Centre de Recherche Rétrospective de Marseille* (CRRM) da *Universidade Aix-Marseille III, Centre de St. Jérôme*, Marselha, França. que permite a geração automática de gráficos e clusters. Pode ser aplicado em associação com o *Dataview*, construindo gráficos a partir das matrizes geradas naquele programa.

Os dados estão sendo analisados por meio de métodos bibliométricos e cientométricos. A análise

foi antecedida de procedimentos de reformatação e padronização de elementos bibliográficos e temáticos. Quanto aos aspectos temáticos, procedeu-se à análise e consolidação dos descritores, levando-se em conta a linguagem de indexação da base de dados e a sua política de indexação. Esses procedimentos foram parametrizados por métodos estatísticos específicos. Os resultados obtidos devem ser submetidos à avaliação crítica de especialistas dos campos de conhecimento abordados, para validação.

CONSIDERAÇÕES FINAIS

Vislumbramos, em termos de resultados concretos deste projeto, além da análise concreta de três campos de conhecimento (Ciência da Informação, Energia Nuclear e Cronobiologia): a) a constituição de um referencial teórico de apoio à pesquisa e à formação de recursos humanos; b) o desenvolvimento de competências para estabelecer adequações, avaliações, análises, exploração em geral dos indicadores; c) a disseminação da produção de conhecimentos sobre métodos e técnicas de análise de informação científica e de suas aplicações na produção de indicadores, em apoio ao processo de planejamento e de gestão e planejamento dos recursos destinados à pesquisa de organizações pertinentes. Espera-se que esses resultados permitam:

- a médio e a longo prazo, que a área da bibliometria, infometria e cientometria se desenvolvam no país;

- que a mensuração das atividades científicas e tecnológicas tenham, por meio da implementação deste projeto, diagnósticos mais consistentes e, portanto, confiáveis;

- formar e capacitar recursos humanos para a pesquisa e atuação na área de análise de informação científica e produção de indicadores;

- promover a disseminação de informação científica e tecnológica pertinente e validada aos usuários dos diferentes setores;

- produzir material para fins didáticos abordando os principais métodos quantitativos utilizados na análise de informação para fins de gestão.

REFERÊNCIAS

- ASTON, W.B.; KLAUVANS, R.A. *Keeping abreast of science and technology: technical intelligence for business*. Columbus, OH: Batelle Press, 1997. 560p.
- BERTIN J. *La graphique et le traitement graphique de l'information*. Paris: Flammarion, 1977.
- BHABHA, H.K. *O local da cultura*. Belo Horizonte: UFMG, 1998.
- CALLON, M.; COURTIAL, J.P.; PENAN, H. *La scientométrie*. Paris: Presses Universitaires de France, 1993. 126p. (Collection "Que sais-je?", v.2727).
- COORDENAÇÃO E APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. *Relatório de Gestão Institucional*. Brasília, 2002. Disponível em: <<http://www.capes.gov.br>>. Acesso em: mar. 2004.
- COURTIAL, J.P. *Introduction à la scientométrie: de la bibliométrie à la veille technologique*. Paris: Anthropos, 1990.
- GARCÍA GUTIÉRREZ, A. *La memoria subrogada: mediación, cultura y consciencia en la red digital*. Granada: Universidad de Granada, 2002.
- GARCÍA GUTIÉRREZ, A. *Otra memoria es posible: estrategias descolonizadoras de la archivo mundial*. Buenos Ayres: La Cruzia, 2004.
- GARCÍA GUTIÉRREZ, A. *Fijaciones*. Madrid: Ed. Biblioteca Nueva, 2005.
- GRANGER, G. *Por um conhecimento filosófico*. São Paulo: Contexto, 1989.
- GUSMÃO, R. Developing and using indicators of multilateral S&T cooperation policy making: the experience from European Research Programme. *Scientometrics*, v.4, n.3, p.493-514, 2000.
- GUSMÃO, R. Nuevas estructuras de producción y difusión de indicadores de C&T: um panorama internacional. *Cuadernos del CENDES*, n.51, p.19-41, 2002.
- KOBASHI, N.Y. *Temas e linhas de pesquisa em Ciência da Informação*. Projeto de pesquisa desenvolvido na Universidade de São Paulo, 2004.
- KOBASHI, N.Y.; SANTOS, R.N.M. *Análise de dissertações e teses do Programa de Pós-graduação em Energia Nuclear do IPEN-SP*. Projeto de pesquisa desenvolvido na Universidade de São Paulo e PUC-Campinas, 2006.
- LANDI, F.R.; FURTADO, J. (Org.). *Indicadores de Ciência, Tecnologia e Inovação*: 2001. São Paulo: FAPESP, 2002.
- MACIAS-CHAPULA, C.A. O papel da informetria e da cienciométrica e sua perspectiva nacional e internacional. *Ciência da Informação*, Brasília, v.27, n.2, p.134-140, 1998.
- MEERTENS, R.W.; NEDERHOF, A.J.; WILKE, H.A.M. Social Psychological Research in The Netherlands, 1980-1988. *European Journal of Social Psychology*, 22, p.93-100, 1992.
- NARIN, F.; OLIVASTRO, D.; STEVENS, K.S. Bibliometric theory, practice and problem. *Evaluation Review*, v.18, n.1, 1994.
- OBSERVATOIRE DES SCIENCES ET DES TECHNIQUES - OST. *Les chiffres clés de la science et de la technologie*. Paris: Economica, 2003.
- ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. *STI outlook 2002: country response to policy questionnaire*. Paris, 2002. Available from: <<http://www.oecd.org/dataoecd/60/32/2762738.pdf>>. Access apr. 2004.
- OKUBO, Y. *Bibliometric indicators and analysis of research systems: methods and examples*. Paris: OECD, 1997, 69p. (STI Working Papers, 1997/1).
- PRAT, A.M. Avaliação da produção científica como instrumento para o desenvolvimento da ciência e da tecnologia: relatos de experiências. *Ciência da Informação*, Brasília, v.27, n.2, p.206-209, 1998.
- SANTOS, R.N.M.; KOBASHI, N.Y. Aspectos metodológicos da produção de indicadores em ciência e tecnologia. Apresentado no VI CINFORM. Salvador, 17 de junho de 2005a.
- SANTOS, R.N.M. *Desenvolvimento de indicadores da produção científica da Ciência da Informação no Brasil*. Projeto de pesquisa desenvolvido na PUC-Campinas, 2005b.
- SPINAK, E. *Dicionário enciclopédico de bibliometria, cienciométrica e informetria*. Caracas: UNESCO, 1996.
- SPINAK, E. Indicadores cienciométricos. *Ciência da Informação*, Brasília, v.27, n.2, p.141-148, 1998.
- TARGINO, M.G.; GARCIA, J.C.R. Ciência brasileira na base de dados do Instituto for Scientific information (ISI). *Ciência da Informação*, Brasília: DF, v.29, n.1, p.103-107, 2000.
- TRZESNIAK, P. Indicadores quantitativos: reflexões que antecedem seu estabelecimento. *Ciência da Informação*, Brasília, v.27, n.2, p.159-64, 1998.
- TUFTE E.R. *The visual display of quantitative information*. Connecticut: Graphics Press, 1983.
- VAN RAAN, A.F.J. *The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments*. Pre-print, 2003.
- VELHO, L. *Cuidado com os rankings científicos*. Disponível em: <<http://www.prometeu.com.br/bb-lea.asp>>. Acesso em: 2004.
- WORMELL, I. Informetria: explorando bases de dados como instrumentos de análise. *Ciência da Informação*, v. 27, n. 2, 1998, p. 210-216.

EQUIPE DE PESQUISA

Coordenação: Nair Yumiko Kobashi e Raimundo Nonato Macedo dos Santos.

Participantes: Marivalde Moacir Francelin e Mônica de Fátima Loureiro (alunos de doutorado do Programa de Pós-graduação em Ciência da Informação da Escola de Comunicação e Artes da Universidade de São Paulo); César Antonio Pereira, Regiane Alcântara Eriel e Oscar Eliel (alunos de mestrado do Programa de Pós-graduação em Ciência da Informação da Pontifícia Universidade Católica de Campinas); Mery Igami (Mestre em Ciências da Comunicação e Diretora da Biblioteca do Instituto de Pesquisas Energéticas e Nucleares); Evelin Costa (aluna de graduação do Curso de Ciência da Informação da Pontifícia Universidade Católica de Campinas) e Priscila Nozaki (aluna de graduação do Curso de Biblioteconomia da Escola de Comunicação e Artes da Universidade de São Paulo).