

**VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação**  
**28 a 31 de outubro de 2007 • Salvador • Bahia • Brasil**

GT 7 – Produção e Comunicação da Informação em CT&I  
Comunicação oral

**A CONSTRUÇÃO DO CONHECIMENTO CIENTÍFICO:  
a Web Semântica como objeto de estudo**

Cíntia Braga Ferreira Pinheiro (USP/Ribeirão Preto, cintiabfp@ffclrp.usp.br)  
Edberto Ferneda (USP/Ribeirão Preto, ferneda@ffclrp.usp.br)

**Resumo:** A ciência é uma atividade social crítica e dinâmica cujo objetivo é a produção de conhecimento sobre diferentes aspectos da natureza. Um conceito importante para o desenvolvimento deste trabalho é o de comunidade científica caracterizada pela prática de uma especialidade, por uma formação teórica comum, pela circulação abundante de informação no interior do grupo. Este estudo tem por objetivo caracterizar a pesquisa brasileira em Web Semântica, considerado um tema na fronteira de pesquisa de duas matrizes disciplinares: a Ciência da Computação e a Ciência da Informação. Optou-se por um estudo cienciométrico, tendo como fonte para coleta dos dados o Curriculum Lattes de pesquisadores doutores. A integração entre comunidades científicas formadas em matrizes disciplinares diferentes encontra um ambiente favorável nos espaços de fronteira que demarcam as matrizes disciplinares. Comprova-se isso ao observar-se a comunidade pesquisadora sobre Web Semântica no Brasil, composta por pesquisadores de diversas matrizes como a Linguística, a Ciência da Informação e a Ciência da Computação.

**Palavras-chave:** Web Semântica. Cienciometria. Comunidade Científica

**Abstract:** Science is a critical and dynamic social activity whose objective is the knowledge production on different aspects of the nature. An important concept for the development of this work is of scholarly community characterized by the practical one of a specialty, for a common theoretical formation, the abundant circulation of information in the interior of the group. This study it has for objective to characterize the Brazilian research in Semantic Web, considered a subject in the border of research of two matrices discipline matrices the Computer Science and the Information Science. It was opted to a scientometrics study, having as source of the data the Curriculum Lattes of doctors. The integration between scholarly communities formed in matrices of different disciplines finds a favorable environment in the scientific boundaries that demarcate the discipline matrices. Observing that researching community on Semantic Web in Brazil, composed for researchers of diverse matrices as Linguistics, Information Science and Computer Science.

**Keywords:** Semantic Web. Scientometrics. Scientific Community.

## Introdução

A ciência é uma atividade social crítica e dinâmica cujo objetivo é a produção de conhecimento sobre diferentes aspectos da natureza. Essa atividade social possui objetivos diversos, tais como a descrição, o controle, a predição e a explicação dos aspectos naturais e sociais formadores da natureza.

As relações sociais e a estruturação das disciplinas científicas em relação às suas práticas metodológicas e o desenvolvimento científico são objetos de estudo da história/sociologia da ciência. Kuhn provocou um impacto na Sociologia da Ciência ou Estudos Sociais da Ciência ao privilegiar os aspectos históricos e sociológicos na análise da prática científica.

Num sentido amplo, o paradigma kuhniano refere-se àquilo que é compartilhado por uma comunidade científica, é uma forma de fazer ciência, uma matriz disciplinar. Uma comunidade científica caracteriza-se pela prática de uma especialidade científica, por uma formação teórica comum, pela circulação abundante de informação no interior do grupo e pela unanimidade de juízo em assuntos profissionais.

Este estudo tem por objetivo caracterizar a pesquisa brasileira em Web Semântica, considerado um tema na fronteira de pesquisa de duas matrizes disciplinares: a Ciência da Computação e a Ciência da Informação. Procura também identificar o papel de cada matriz na institucionalização social do tema. Para isso fundamenta-se na perspectiva cienciométrica de análise das publicações científicas produzidas pelos pesquisadores brasileiros.

## A Ciência da Informação e a Ciência da Computação como Matrizes Disciplinares

A Sociologia da Ciência ou os Estudos Sociais da Ciência têm procurado analisar a estrutura das disciplinas científicas em relação às suas práticas sociais, procurando oferecer novas perspectivas sobre a construção do saber e o desenvolvimento científico e tecnológico.

Como afirmam Machado e Teixeira (2007) Kuhn estabelece uma postura relativista no estudo da ciência, o pesquisador passa a ser considerado um negociador que produz diversas formas de interpretação do real que precisam ser legitimadas pelo maior número de pessoas, consolidando-se assim o paradigma defendido. A postura relativista rejeita a idéia de critérios únicos, estáveis e atemporais da ciência “o acordo é puramente social, ele é fruto das interações e das negociações entre os pesquisadores que compartilham de um sistema de crença que, em sua época, lhes parece objetivo, ao passo que esse mesmo sistema foi objeto de negociação.” (MACHADO; TEIXEIRA, 2007)

Um conceito importante na teoria kuhniana e na Sociologia da Ciência para o desenvolvimento deste trabalho, é o de comunidade científica, pois a socialização entre pesquisadores permite uma adesão aos valores e crenças do grupo. Uma comunidade científica caracteriza-se pela prática de uma especialidade científica, por uma formação teórica comum, **pela circulação abundante de informação no interior do grupo** e pela unanimidade de juízo em assuntos profissionais.

Dencker (2001) concorda que “Kuhn identificou, na ciência, uma estrutura comunitária em que os participantes são submetidos a uma iniciação profissional e educação similares, o que leva a compartilhar dos mesmos paradigmas.”

Uma matriz disciplinar forma seus pesquisadores principalmente através do ensino superior, pois nele se consolida e dissemina paradigmas através de generalizações simbólicas, fortalecimento de crenças em determinados modelos e compartilhamento de valores. Estes elementos fornecem a uma determinada comunidade científica analogias ou metáforas preferidas, embora possam ocorrer divergências quanto ao seu uso. A estrutura da comunidade científica se estabelece através do compartilhamento de exemplos e possíveis soluções. Para Kuhn (2003) o pesquisador é formado através do treinamento dentro de

determinada matriz disciplinar, consolidando um conhecimento tácito gerado através do ‘aprender fazendo’ ao identificar problemas semelhantes e aplicar as soluções aceitas por aquela comunidade específica. Portanto o saber científico é adquirido através deste treinamento recebido dentro de uma comunidade científica no âmbito de sistemas de convenções que têm origem nos processos cognitivos que incluem os juízos e acordos socialmente aceitos pelo grupo.

Uma comunidade científica pode ser definida como um grupo produtor e legitimador do conhecimento científico e apresenta algumas características: seus integrantes são treinados de forma semelhante; compartilham e absorvem a mesma literatura técnica/científica; aprendem as mesmas lições; o grupo possui um objeto científico próprio, embora este possa ser abordado por pontos de vistas de diferentes escolas; o grupo possui um sistema de comunicação amplo com avaliação pelos pares; e constituem uma única audiência para essa comunicação. De maneira resumida pode-se definir comunidade científica como grupo de pesquisadores praticantes de determinada especialidade científica.

Fleck afirma que os objetos científicos são construídos por uma comunidade científica ou ‘coletivos de pensamento’ e cada coletivo elabora seu próprio ‘estilo de pensamento’ ou matriz disciplinar. (LOWY, 1994)

Diante dos fundamentos teóricos apresentados, pretende-se caracterizar duas ciências que tiveram origem na revolução científica e técnica provocada pela Segunda Guerra Mundial: a Ciência da Informação e a Ciência da Computação, como matrizes disciplinares distintas que compartilham objetos de estudo como a recuperação da informação, a Web e suas estruturas. Ambas as matrizes consideradas campos interdisciplinares e que desenvolvem pesquisa em ciência aplicada.

### **Ciência da Computação**

A Ciência da Computação é um campo interdisciplinar desenvolvido num núcleo teórico das ciências exatas ou ciências duras. Para Denning et al (1989) esta disciplina trata apenas de processos que possam ser executados através de algoritmos. Fonseca Filho (1999, p.13) define a Ciência da Computação ”como um corpo de conhecimento formado por uma infra-estrutura conceitual e um edifício tecnológico onde se materializam o hardware e o software. A primeira fundamenta a segunda e a precedeu.” Os paradigmas desta matriz disciplinar, direta ou indiretamente, estão materializados em programas (software) ou dispositivos (hardware) comprovando seu caráter de ciência aplicada.

A Segunda Guerra Mundial é um marco para a Ciência da Computação quando efetivamente foram construídos os primeiros computadores digitais e ocorre um desenvolvimento científico e tecnológico exponencial com o surgimento de diversas sub-especialidades. Estudar o estado da arte na Ciência da Computação atualmente exige a seleção de perspectivas, campos ou áreas específicas como: linguagem de programação, teoria da computação, robótica, inteligência artificial, computação gráfica, etc.

A Matemática e a Engenharia são bases fundamentais para o desenvolvimento da Ciência da Computação a partir da segunda metade do século passado. É a união de matemáticos e engenheiros através deste interesse em comum que forma uma nova comunidade científica preocupada em pesquisar e desenvolver dispositivos e programas capazes de executar operações computacionais que substituem a noção de intuitivo na execução de procedimentos por uma execução formalizada através de uma estrutura matemática: os algoritmos.

Um fator para o surgimento de uma nova comunidade científica e a consolidação da Ciência da Computação como matriz disciplinar foi o interesse comum entre matemáticos e engenheiros em desenvolver máquinas (hardware) e sistemas (software) capazes de executar raciocínios presentes no mundo real. A partir de 1960 a Ciência da Computação se consolida

definitivamente como uma matriz disciplinar com a formação não só de grupos de pesquisa como de pesquisadores doutores na área. Vale lembrar a teoria kuhniana que considera a educação como um fator primordial na formação de seguidores de uma matriz disciplinar.

Desde a década de 50 do século passado vários pesquisadores buscavam soluções para o tratamento computacional da informação, com destaque a Hans Peter Luhn, desenvolvedor de vários projetos que modificaram radicalmente os métodos tradicionais de armazenamento, tratamento e recuperação da informação. (FERNEDA, 2003).

A Recuperação da Informação consolidou-se como área de pesquisa de duas matrizes disciplinares: a Ciência da Computação e a Ciência da Informação, matriz disciplinar que também tem origem na revolução científica e técnica provocada pela Segunda Guerra Mundial.

### **Ciência da Informação**

Assim como a Ciência da Computação tem seu desenvolvimento histórico estruturado no conceito abstrato de número e no desenvolvimento no campo da Matemática, a Ciência da Informação está vinculada historicamente ao desenvolvimento da escrita e a possibilidade de se registrar fisicamente o conhecimento através dela. O desenvolvimento das técnicas de escrita até o surgimento da prensa de tipo móvel permitiu o registro, estoque e a recuperação do conhecimento.

Num rápido percurso histórico da Ciência da Informação podemos destacar a figura de Paul Otlet, que no início do século XX cunhou do termo Documentação sistematizou e previu tecnologias para sua operacionalização. Outro destaque histórico dentro da Ciência da Informação é Vannevar Bush que no período pós Segunda Guerra Mundial publicou um artigo considerado um marco na busca de soluções para os problemas da explosão informacional gerada no período pós-guerra. No artigo “As We may think” Bush propõe como solução para os problemas de gerenciamento de informações o desenvolvimento de uma máquina capaz de agregar diversas tecnologias de informação da época.

Sob o ponto de vista de desenvolvimento paradigmático da Ciência da Informação, diversos autores têm procurado estabelecer fronteiras, objetos de estudo, definições, identificar uma estrutura disciplinar ou colaborações disciplinares. (SARACEVIC, 1996; LE COADIC, 2004; NEHMY et al., 1996; EUGÊNIO, FRANÇA E PEREZ, 1996; ARAÚJO, 2003; PINHEIRO, LOURENÇO, 1995; SARACEVIC, 1995; SAYÃO, 2001)

Para Saracevic (1996) a Ciência da Informação tem a Biblioteconomia, a própria Ciência da Computação, a Comunicação e a Ciência Cognitiva como disciplinas que estabeleceram relações mais profundas dentro de seu escopo. Le Coadic (2004, p.20) salienta também outras disciplinas como principais colaboradoras da Ciência da Informação como a Psicologia, Linguística, Sociologia, Informática, Matemática, Lógica, Estatística, Eletrônica, Economia, Direito, Filosofia, Política e Telecomunicações. O destaque é que a interdisciplinaridade da Ciência da Informação está ligada a um núcleo teórico onde predominam as ciências sociais.

Assim como a Ciência da Computação, considera-se a década de 60 do século XX o início da consolidação da Ciência da Informação como matriz disciplinar com a formação de uma comunidade científica preocupada em solucionar problemas e estudar o tratamento documental e o uso de tecnologias nesse processo. A constituição de uma comunidade científica é que determina a cientificidade de uma matriz disciplinar pois pressupõe a existência de mecanismos eficientes para a socialização. Oliveira, Mota e Alvarado (2004) destacam alguns elementos que caracterizam a consolidação de uma matriz disciplinar: existência de instituições capazes de abrigar os grupos de pesquisadores, a existência de pesquisadores capacitados para o desenvolvimento e treinamento da atividade científica, canais de comunicação que permitam o fluxo de informação dentro da comunidade. Desta forma caracterizam a comunidade científica brasileira em Ciência da Informação através da

identificação das instituições de pesquisa, dos canais de comunicação (periódicos e eventos) e das associações profissionais confirmando a cientificidade da Ciência da Informação.

### **Ciência da Informação, Ciência da Computação e o compartilhamento de objetos de estudo**

Apresentadas as duas matrizes disciplinares, cumpre estabelecer qual o objeto de estudo comum a elas. Considerado o componente tecnológico da Ciência da Informação, a Ciência da Computação permite o compartilhamento de diversos temas de pesquisa como a Recuperação da Informação, a ambiência digital da informação, a estrutura hipertextual dos documentos digitais, a estruturação de metadados representativos da informação, etc.

A Recuperação da Informação foi um dos primeiros objetos de estudo destas duas matrizes: a Ciência da Informação e a Ciência da Computação. Saracevic (1996) afirma que a Recuperação de Informação pode ser considerada a vertente tecnológica da Ciência da Informação e é resultado da relação desta com a Ciência da Computação.

O advento da Internet como grande repositório de documentos e das informações neles contidas favoreceu o surgimento de pesquisas sobre recuperação de informação na Web. Mecanismos de busca foram desenvolvidos especificamente para este novo ambiente informacional e atualmente com o objetivo de melhorar a recuperação de informação neste grande repositório, as pesquisas buscam encontrar formas de agregar valor semântico às páginas da Web.

Em 2001, Berners-Lee, Hendler e Lassila publicam no Scientific American uma proposta para solucionar problemas na recuperação de informações na Web: o estabelecimento de padrões ou metadados que permitam ao computador compreender a semântica dos textos digitais, disponíveis na Internet. Os autores denominaram esta nova metodologia para conteúdos digitais de 'Web Semântica'.

As pesquisas envolvendo maior padronização das páginas da Web através de metadados podem ser consideradas um paralelo na agenda das duas matrizes disciplinares. Portanto o paradigma da organização da informação digital através de normas e padrões como solução para o aumento do grau de precisão na recuperação da informação passa a ser compartilhado por duas ciências que constroem, analisam e comunicam suas pesquisas de forma diferenciada e para públicos específicos e distintos.

### **Cienciometria**

A Cienciometria é um segmento da Sociologia da Ciência que tem como objetos de estudo as disciplinas, assuntos, áreas e campos científicos, analisando os fatores que diferenciam as sub-disciplinas (revistas, autores, documentos, como os cientistas se comunicam) utilizando métodos de análise de conjuntos e de correspondência, objetivando identificar os domínios de interesses, concentração de assuntos e compreender a comunicação entre os cientistas (MACIAS-CHAPULA, 1998). É o estudo dos aspectos quantitativos da ciência como disciplina ou atividade econômica.

O termo Cienciometria surgiu na antiga União Soviética sendo utilizado especialmente na Hungria. Dobrov e Karenol, em 1969, estão entre os primeiros pesquisadores a utilizar o termo. (VANTI, 2002) O termo originalmente referia-se a aplicação de métodos quantitativos e matemáticos para o estudo da história da ciência e do progresso tecnológico. Mas é com o surgimento da revista Scientometrics, em 1977, que o termo consolida-se dentro da área.

Para Tague-Sutcliffe (1992) apud Macias-Chapula (1998) a Cienciometria, também chamada de Cientometria, estuda os aspectos quantitativos da ciência como disciplina ou atividade econômica. É de interesse da Cienciometria o crescimento quantitativo das ciências, o desenvolvimento das disciplinas ou sub-disciplinas, a obsolescência de paradigmas científicos, etc.

Lara (2006) define a Cienciometria como uma área que objetiva estudos dos aspectos quantificáveis da ciência como atividade humana e do conhecimento científico, seja através do uso de instrumentos para mensuração da produção e atividade científica de um país ou de uma comunidade científica específica.

A Cienciometria tem como instrumentos ferramentas capazes de mapear o intercâmbio da pesquisa entre países ou grupos de pesquisadores, identificar a evolução e desenvolvimento de paradigmas, construir indicadores para suporte nas decisões sobre política científica e tecnológica, estabelecer relações entre a pesquisa científica com inovação e desenvolvimentos econômico. (LARA, 2006)

A Cienciometria procura oferecer parâmetros para estudos historiográficos da ciência, para análise do desenvolvimento científico e tecnológico, para análise de políticas científicas, etc. Pode-se concluir que a Cienciometria está relacionada à análise da ciência através de dados coletados de publicações científicas e trabalhados com o uso das técnicas bibliométricas.

São fontes de informação para a Cienciometria as publicações científicas e as patentes como indicadores efetivos dos resultados das pesquisas e do desenvolvimento tecnológico ao explicitar o conhecimento gerado nos campos da ciência e da tecnologia.

Como destaca Santos (2003, p.134) são premissa da Cienciometria alguns postulados:

uma obra científica é o produto objetivo da atividade intelectual criativa. Num contexto científico, uma publicação é uma representação da atividade de pesquisa de seu autor. O maior esforço deste autor é de persuadir os pares de que suas descobertas, seus métodos e técnicas são particularmente pertinentes. O modo de comunicação escrita fornecerá, portanto, todos os elementos técnicos, conceituais, sociais e econômicos que o autor busca afirmar ao longo de sua argumentação. A atividade de publicação científica é uma eterna confrontação entre as reflexões intrínsecas do autor e os conhecimentos que ele adquiriu pela leitura dos trabalhos originários dos outros autores.

As metodologias desenvolvidas dentro do escopo dos estudos métricos da informação têm oferecido instrumentos para análise de diversos estudos nos campos da Filosofia da Ciência; da Sociologia da Ciência; da Psicologia da Ciência; da Política Científica; da Ciência da Informação, etc. O uso destes estudos dentro de cada matriz disciplinar pode gerar indicadores científicos e tecnológicos, avaliações da produção científica, mapeamentos disciplinares e outros produtos que recebem análise conforme os objetivos de cada pesquisa. Especificamente a Cienciometria, considerada a Ciência da Ciência (PRICE, 1963 apud MUGNAINI; CARVALHO; CAMPANATTI-OSTIZ, 2006) têm contribuído para geração de panoramas de áreas de estudo, disciplinas, periódicos, pesquisadores, instituições, países ou qualquer conjunto de elementos mensuráveis para a ciência.

A Cienciometria serviu de base para o estudo da Web Semântica como campo de pesquisa transdisciplinar da Ciência da Informação e da Ciência da Computação.

### **Pesquisa Brasileira em Web Semântica**

Pesquisas buscando estabelecer estruturas de metadados que agreguem valor semântico às páginas Web, tecnologia denominada Web Semântica, têm se consolidado na agenda de pesquisas de diversas matrizes disciplinares, com destaque para a Ciência da Computação e Ciência da Informação. Procura-se caracterizar a pesquisa de doutores brasileiros sobre o tema num esforço em identificar a participação destas e de outras matrizes disciplinares na construção do conhecimento científico no âmbito nacional.

### **O conhecimento científico e sua comunicação**

A construção e/ou desenvolvimento do conhecimento científico envolve um processo de socialização de idéias pelos canais formais ou informais de comunicação, pela atividade

científica nos laboratórios ou grupo de pesquisa, pela formação acadêmica, pelo aliciamento de seguidores, todos os fatores que caracterizam uma comunidade científica. Podemos, portanto afirmar que o desenvolvimento do conhecimento científico está fortemente ligado ao processo de comunicação dentro da comunidade de pesquisadores.

O sujeito acessa o conhecimento cumulativo, construído por outros e, com base na própria capacidade de apreensão, análise e reflexão, gera novo conhecimento. No entanto, acredita-se que o conhecimento somente será de fato construído, com sua socialização aos outros. (VALENTIM, 2005, p. 11)

TARGINO (2000) destaca algumas funções da comunicação dentro do campo da ciência: fornecer respostas a perguntas específicas; promover a atualização profissional dos pesquisadores; introduzir os pesquisadores nos campos científicos; estimular descobertas científicas; divulgar áreas científicas emergentes; legitimar novos conhecimentos; redirecionar ou ampliar as áreas de interesses dos pesquisadores; fornecer *feedback* para o aprimoramento da produção científica.

É premissa da Cienciometria que a dinâmica da atividade científica possa ser observada através das publicações científicas, pois “se o documento é a expressão de uma pessoa ou grupo trabalhando em uma frente de pesquisa, pode-se dizer alguma coisa sobre as relações entre as pessoas a partir dos próprios documentos”. (MACIAS-CHAPULA, 1998, p. 134).

### **A Web Semântica na Plataforma Lattes**

Procura-se caracterizar através da identificação de Instituições de pesquisa, matrizes disciplinares, tipologia das publicações da comunidade científica responsável pela construção e desenvolvimento da pesquisa em Web Semântica considerando-se como representantes desta comunidade os cientistas doutores. A opção por um estudo por amostragem fundamenta-se no conceito kuhniano da necessidade de *training* dos pesquisadores dentro de suas matrizes disciplinares para amadurecimento e consolidação dos paradigmas dominantes daquela ciência. Visto que a pós-graduação no nível de doutorado permite uma extensa socialização dentro da matriz disciplinar, considera-se os pesquisadores doutores amadurecidos e efetivamente praticantes dos paradigmas consolidados em sua formação. Acredita-se que o estudo deste segmento da comunidade científica pesquisadora sobre Web Semântica refletirá adequadamente sua atual estrutura.

A abordagem cienciométrica teve como fonte para a coleta dos dados a Plataforma Lattes desenvolvida pelo Conselho Nacional de Pesquisa (CNPq) através do Currículo Lattes dos pesquisadores doutores. A Plataforma Lattes “representa a experiência do CNPq na integração de bases de dados de currículos e de instituições da área de ciência e tecnologia em um único Sistema de Informações.” (CONSELHO) Uma dessas bases de dados tem como função registrar as atividades profissionais de pesquisadores independente do grau de formação: graduação, especialização, mestrado, doutorado, pós-doutorado, inclusive estudantes ainda não graduados ou licenciados no ensino superior, trata-se do Currículo Lattes. “O Currículo Lattes registra a vida pregressa e atual dos pesquisadores sendo elemento indispensável à análise de mérito e competência dos pleitos apresentados às agências de fomento.” (CONSELHO)

Foram recuperados, em abril de 2007, os Currículos Lattes dos pesquisadores doutores que satisfizeram a seguinte estratégia de busca: através do formulário de Busca Avançada por Assunto foi preenchido o campo ‘Construa uma consulta com... esta frase exata’ com o termo Web Semântica. Segundo as instruções da Base de Dados através da estratégia adotada foram recuperados os Currículos Lattes dos pesquisadores doutores que tivessem no título ou no campo assunto de suas produções bibliográficas o termo ‘Web Semântica’.

Foram recuperados 141 Currículos Lattes que satisfizeram a estratégia de busca, conforme as orientações disponíveis na Base de Dados. Porém após a análise destes Currículos foram excluídos 38 pesquisadores doutores que não eram autores ou co-autores de publicações bibliográficas sobre Web Semântica totalizando um universo de pesquisa de 103 doutores responsáveis pela pesquisa científica brasileira em Web Semântica. Optou-se pela exclusão de 38 pesquisadores que tiveram seus perfis recuperados através da estratégia de busca adotada que no entanto não tiveram nenhuma publicação científica de sua autoria sobre o tema.

Foram analisados os Currículos Lattes destes doutores buscando coletar dados que permitissem o mapeamento das Instituições de pesquisa nas quais eles atuam, mapeamento de sua distribuição geográfica, identificação da matriz disciplinar a que pertencem, caracterização da produção científica dentro da área Web Semântica e identificação dos canais de comunicação predominantes e seus vínculos com as duas matrizes disciplinares Ciência da Computação e Ciência da Informação.

### **Quem contribui para a pesquisa em Web Semântica?**

Procurou-se coletar e analisar os dados presentes nos Currículos Lattes dos pesquisadores doutores cuja produção científica incluisse a pesquisa em Web Semântica com o intuito de identificar o grau de participação da Ciência da Informação e da Ciência da Computação na construção do conhecimento científico nesta área.

A pesquisa em Web Semântica pode ser considerada como uma das áreas de fronteira entre as essas duas matrizes disciplinares, pois faz uso de diversas estruturas em camadas de metadados que agregam valor semântico às páginas Web. Uma destas camadas forma-se uma árvore hierárquica de classes, de maneira que cada classe herda as características de uma ou mais classes superiores. (FERNEDA, 2003) Estes e outros conceitos se assemelham, em teoria, aos padrões de descrição bibliográfica e temática presentes na Ciência da Informação e têm sido estudados pelos pesquisadores da área. Sendo assim, a Recuperação de Informação na Web, principalmente o desenvolvimento da Web Semântica é um objeto de estudo comum às duas matrizes disciplinares.

Dos 103 doutores: 86 (83,49%) são da Ciência da Computação, 16 (15,54%) da Ciência da Informação, 1 (0,97%) da Linguística com interface com a Ciência da Computação (Tabela 1). Merece destaque a identificação de 4 (3,88%) pesquisadores da Ciência da Computação que possuem interface com a Ciência da Informação através do uso de seus canais de comunicação.

**Tabela 1. Matriz Disciplinar dos pesquisadores**

<b>Matriz Disciplinar</b>	<b>Pesquisadores</b>	<b>Porcentagem (%)</b>
Ciência da Computação	86	83,49
Ciência da Informação	16	15,54
Linguística	01	0,97

A presença de um maior número de pesquisadores dentro a Ciência da Computação era esperada visto que se trata de um paradigma gerado dentro desta matriz disciplinar e, portanto seus primeiros canais de comunicação foram os compartilhados por aquela comunidade. É esta matriz disciplinar que assume o papel central no desenvolvimento das pesquisas. A Ciência da Informação tem papel coadjuvante ao proporcionar novas perspectivas paradigmáticas.

A presença de pesquisadores da Ciência da Computação no sistema de comunicação científica da Ciência da Informação tem caráter positivo, porém espera-se que o inverso também aconteça para que a intersecção de paradigmas possa se concretizar no âmbito nacional. Prática importante para o crescimento científico e tecnológico, as pesquisas



transdisciplinares promovem a formação de redes de interação entre diferentes matrizes disciplinares que contribuem para o fortalecimento da área por elas compartilhada. Ao analisarmos posteriormente os canais de comunicação adotados pelos pesquisadores procuramos identificar a participação da Ciência da Informação na comunicação científica da Ciência da Computação .

A participação de Instituições particulares 20 (35,08%) de um total de 57 Instituições de pesquisa identificadas é animador na medida em que caracteriza um vínculo de pesquisadores graduados com doutorado na educação superior particular, sendo um possível indicativo de qualidade educacional, porém não se pode afirmar que estas Instituições têm investido na pesquisa em Web Semântica visto que a metodologia adotada não permite esta inferência. Podem estes pesquisadores serem co-participantes e co-autores em pesquisas realizadas pelas Instituições públicas de pesquisa.

Sob o ponto de vista da tipologia documental das produções bibliográficas encontradas nos Currículos Lattes, o destaque são os trabalhos apresentados em eventos (72,18%). São os eventos os canais de comunicação mais rápidos para a socialização de uma pesquisa e geralmente são usados na divulgação de pesquisas em andamento ou recém concluídas. (MEADOWS, 1999; MUELLER, 2005; SILVA et al. 2006) Portanto o canal de comunicação mais utilizado pelos pesquisadores em Web Semântica é compatível com o grau de desenvolvimento do tema visto que o artigo ‘The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities’ foi publicado em 2001. No entanto, 18,23% da produção bibliográfica científica brasileira está registrada em periódicos científicos, considerado o canal de comunicação científica mais respeitado pelos cientistas, principalmente nas ciências exatas, da saúde e sociais aplicadas. (MUELLER, 2005)

Meadows (1999) e Mueller (2005) destacam que diferentes matrizes disciplinares ou áreas do conhecimento possuem preferências por canais de comunicação distintos como consequência da natureza e especificidades de cada área que leva à adoção de práticas diferentes da atividade científica, e consequentemente a adoção de formas diferentes de comunicação e socialização dos conhecimentos gerados em suas pesquisas.

**Tabela 2. Canais de comunicação da pesquisa em Web Semântica**

<b>Tipo de Publicação</b>	<b>Porcentagem (%)</b>
Artigo de Periódico	18,23
Capítulo de Livro	6,00
Livro	2,39
Relatório Técnico	1,20
Trabalho de Evento	72,18

Os livros (2,39%) e capítulos de livros (6%) em geral são pouco utilizados como canal de comunicação dentro da Ciência da Computação, confirmando o baixo índice de adoção destes canais que também pode ser reflexo do estágio inicial do conhecimento e da pesquisa científica em Web Semântica. (MUELLER, 2005)

Dentro da matriz Ciência da Informação a opção principal como canal de comunicação também foram os trabalhos de eventos (48,58%) e os artigos de periódicos (40%), quase em igual proporção. Conforme Mueller (2005) existe também a preferência pela publicação de livros e capítulos de livros nas matrizes disciplinares integrantes das Ciências Sociais Aplicadas. Sendo compatível com esta preferência a identificação de duas publicações em cada um destes canais de comunicação dentro da Ciência da Informação, conforme a tabela 3.

**Tabela 3. Canais de comunicação da pesquisa em Web Semântica na Ciência da Informação**

<b>Tipo de Publicação</b>	<b>Porcentagem (%)</b>
Artigo de Periódico	40,00
Capítulo de Livro	5,71
Livro	5,71
Trabalho de Evento	48,58

Destaca-se a existência da participação da Ciência da Informação nos canais de comunicação da Ciência da Computação através dos trabalhos de eventos, prática a ser incentivada para a construção de um conhecimento científico transdisciplinar em Web Semântica.

A partir dos dados coletados considera-se a Web Semântica uma área de pesquisa na fronteira da Ciência da Computação com a Ciência da Informação da qual depende da integração das duas matrizes disciplinares para a construção e desenvolvimento do conhecimento científico de forma transdisciplinar.

### **Conclusão**

A ampliação dos espaços informacionais com o desenvolvimento da Web provocou um impacto nas pesquisas de diversas matrizes disciplinares dentre as quais podemos destacar a Ciência da Informação e a Ciência da Computação. Questões envolvendo os impactos tecnológicos na sociedade ou em comunidades específicas, a recuperação da informação na Web, a arquitetura de informação, usabilidade de sistemas de informação, têm sido objeto de estudo e análise destas e de outras matrizes disciplinares.

Pode-se considerar que pesquisas envolvendo a Internet e a Web tem um posicionamento de fronteira dentro dos limites das matrizes disciplinares. Estes temas fronteiriços podem ser considerados como transdisciplinares permitindo a adoção e contribuição de diversos paradigmas de diferentes áreas do conhecimento.

A Web Semântica está posicionada nesta fronteira de saberes e particularmente na pesquisa brasileira tem recebido contribuições tanto da Ciência da Computação quanto da Ciência da Informação na construção e desenvolvimento de sua tecnologia. Optou-se pela análise cienciométrica deste campo de pesquisa que permitiu através do estudo das publicações científicas identificar quais as matrizes disciplinares, canais de comunicação preferenciais de cada matriz, etc estabelecendo um painel do cenário atual da pesquisa sobre Web Semântica no Brasil.

A participação dos pesquisadores da Ciência da Computação em mais de 80 % da produção científica sobre o tema era um fato esperado neste estudo, no entanto revelou questões importantes como um tímido intercâmbio entre as duas matrizes que pode indicar o início de processo de compartilhamento de informações através dos canais de comunicação mais utilizados em cada matriz.

A integração entre comunidades científicas formadas em matrizes disciplinares diferentes que, no entanto compartilham o mesmo objeto de estudo é um processo difícil, nem sempre possível, mas que encontra um ambiente favorável nos espaços de fronteira que demarcam as matrizes disciplinares. Pode-se comprovar isso ao observarmos a comunidade pesquisadora sobre Web Semântica no Brasil, composta por pesquisadores de diversas matrizes como a Linguística, a Ciência da Informação e a Ciência da Computação que procuram compartilhar suas pesquisas em diversos canais de comunicação, principalmente os eventos científicos e periódicos de suas e de outras matrizes disciplinares.

Espera-se que este cenário permaneça e evolua para um maior intercâmbio de informações entre as matrizes disciplinares, proporcionando a cooperação multiparadigmática que a Web Semântica necessita.

## Referências Bibliográficas

ARAÚJO, C. A. A. A ciência da informação como ciência social. **Ci. Inf.**, Brasília, v. 32, n. 3, set./dez. 2003a. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652003000300003&lng=pt&nrm=iso&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652003000300003&lng=pt&nrm=iso&tlng=pt) Acesso em: 18 de agosto de 2006.

BERNERS-LEE, T.; HANDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities.

**Scientific American**, may 2001. Disponível

em: <http://www.scientificamerican.com/2001/0501issueberners-lee.html> Acesso em: 27 de novembro de 2004.

CONSELHO NACIONAL DE PESQUISA. **Plataforma Lattes**. Disponível em: [http://lattes.cnpq.br/conheca/con\\_apres.htm](http://lattes.cnpq.br/conheca/con_apres.htm) Acesso em: 09 de abril de 2007.

DENCKER, A. F. M. Paradigmas científicos e ensino superior. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 24., 2001, Campo Grande. **Anais...** São Paulo: Intercom, 2001. CD-ROM. Disponível em: <http://reposcom.portcom.intercom.org.br/handle/1904/5175> Acesso em: 08 de janeiro de 2007.

DENNING, P.J. et al. Computing as a discipline. **Communication of the ACM**, v. 32, n. 1, p. 9-23, 1989

DOBROV, GM; KARENNOI, A. A. The informational basis of scientometrics. In: MIKHAILOV, A. I. (ed.) **On theoretical problems of informatics**. Moscou: VINITI/FID, 1969. p. 165-191.

EUGÊNIO, M.; FRANÇA, R.O.; PEREZ, R. C. Ciência da informação sob a ótica paradigmática de Thomas Kuhn: elementos de reflexão. **Perspec. Ci. Inf.**, Belo Horizonte, v. 1, n. 1, p. 27-39, jan./jun. 1996.

FERNEDA, E. **Recuperação da informação**: análise sobre a contribuição da ciência da computação para a ciência da informação. 2003. 137 f. Tese (Doutorado em Ciência da Comunicação). Escola de Comunicações e Artes - USP, São Paulo, 2003.

FLECK, L. **Genesis and development of a scientific fact**. Chicago: University of Chicago Press, 1979.

FONSECA FILHO, C. **História da computação**: teoria e tecnologia. São Paulo: LTr, 1999.

KUHN, T. **A estrutura das revoluções científicas**. São Paulo: Perspectiva, 2003.

LARA, M. L. G. Termos e conceitos da área de comunicação e produção científica. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M (orgs.) **Comunicação & produção científica**: contexto, indicadores e avaliação. São Paulo: Angellara, 2006. p. 387-416.

LE COADIC, Y.-F. **A ciência da informação**. 2. ed. Brasília: Briquet de Lemos, 2004.

LOWY, I. Ludwick Fleck e a presente história das ciências. **Manguinhos**, v.I, n. 1, p. 7-18, 1994.

MACIAS-CHAPULA, C. A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ci. Inf.**, Brasília, v. 27, n. 2, p. 134-140, maio/ago. 1998.

MACHADO, C. J. S.; TEIXEIRA, M. O. Descoberta, invenção e inovação segundo os estudos sociais anglo-saxões e europeus das ciências. **DataGramaZero**, v. 8, n. 2, abr. 2007. Disponível em: [http://www.dgz.org.br/abr07/F\\_I\\_aut.htm](http://www.dgz.org.br/abr07/F_I_aut.htm) Acesso em: 06 de abril de 2007.

MEADOWS, A. J. **A comunicação científica**. Brasília: Briquet de Lemos, 1999.

- MUELLER, S. P. M. A publicação da ciência: áreas científicas e seus canais preferenciais. **DataGramaZero**, v.6, n. 1, fev. 2005. Disponível em: [http://www.datagramazero.org.br/fev05/F\\_I\\_art.htm](http://www.datagramazero.org.br/fev05/F_I_art.htm) Acesso em: 13 de março de 2005.
- MUGNAINI, R.; CARVALHO, T.; CAMPANATTI-OSTIZ, H. Indicadores de produção científica: uma discussão conceitual. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M (orgs.) **Comunicação & produção científica**: contexto, indicadores e avaliação. São Paulo: Angellara, 2006. p.313-340.
- NEHMY, R. M. Q. et al. A ciência da informação como disciplina científica. **Perspec. Ci. Inf.**, Belo Horizonte, v. 1, n.1, p. 9-25, jan./jun. 1996.
- OLIVEIRA, M.; MOTA, F. R. L.; ALVARADO, R. U. Comunidade científica e cientificidade da ciência da informação. In: CONGRESSO DA BAD, 8., 2004, Lisboa. Disponível em: <http://sapp.telepac.pt/apbad/congresso8/com27.pdf> Acesso em: 21 de janeiro de 2005.
- PINHEIRO, L. V. R.; LOURENÇO, J. M. M. Traçados e limites da ciência da informação. **Ci. Inf.**, Brasília, v. 24, n. 1, jan./abr. 1995. Disponível em: <http://www.ibict.br/cienciadainformacao/viewarticle.php?id=576&layout=abstract> Acesso em: 18 de agosto de 2006.
- SANTOS, R. N. M. Indicadores estratégicos em ciência e tecnologia: refletindo a sua prática como dispositivo de inclusão/exclusão. **Transinformação**, v. 15, Ed. Esp., p. 129-140, set./dez. 2003.
- SARACEVIC, T. A natureza interdisciplinar da ciência da informação. . **Ci. Inf.**, Brasília, v. 24, n. 1, jan./abr. 1995. Disponível em: <http://www.ibict.br/cienciadainformacao/viewarticle.php?id=575&layout=abstract> Acesso em 18 de agosto de 2006.
- SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspec. Ci. Inf.**, Belo Horizonte, v. 1, n.1, p. 41-62, jan./jun. 1996.
- SAYÃO, L. F. Modelos teóricos em ciência da informação – abstração e método científico. **Ci. Inf.**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.
- SILVA, E.L. et al. Panorama da pesquisa em ciência da informação no Brasil. **Inf. & Soc.: Est.**, v. 16, n. 1, p. 205-228, 2006.
- TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. **Inf. & Soc.: Est.**, v. 10, n. 2, p. 37-85, 2000.
- TAGUE-SUTCLIFFE, J. An introduction to informetrics. **Information Processing & Management**, v. 28,n. 1, p. 1-3, 1992
- VALENTIM, M. L. P. Construção do conhecimento científico. In: VALENTIM, M. L. P.(org.) **Métodos qualitativos de pesquisa em ciência da informação**. São Paulo: Polis, 2005. p. 07-28.
- VANTI, N.A.P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002