



XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)

**GT 07 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação**

**ANÁLISE BIBLIOMÉTRICA SOBRE O TEMA WEB SEMÂNTICA NO BRASIL:  
PRIMEIRAS APROXIMAÇÕES**

***BIBLIOMETRIC STUDY ON SEMANTIC WEB IN BRAZIL: FIRST APPROACHES***

**Cintia Braga Ferreira Pinheiro<sup>1</sup>, Ely Francina Tannuri de Oliveira<sup>2</sup>**

**Modalidade de apresentação:** Poster

**Resumo:** Esta pesquisa objetiva, de forma geral, analisar o conceito de Web Semântica como objeto de estudo, no Brasil, por meio de procedimentos bibliométricos, a partir da análise da produção científica no tema. De forma mais específica, busca identificar o grupo de pesquisadores brasileiros mais produtivos no tema; identificar abordagens da pesquisa sobre Web Semântica, por meio dos projetos de pesquisa vigentes dos pesquisadores brasileiros mais produtivos; e analisar a internacionalização da ciência brasileira no tema, por meio da rede colaborativa institucional, gerada entre o Brasil e demais países. Como procedimento de pesquisa, utilizou-se a base Scopus para o levantamento das referências sobre o tema Web Semântica, usando-se palavras-chave como termos de busca, que resultaram em um total de 1.384 documentos. Após a triagem, resultou um *corpus* de 1.049 artigos. Os dados foram agrupados, apresentados em tabelas e analisados no sentido de atender aos objetivos propostos. Gerou-se a rede colaborativa *two-mode* entre instituições brasileiras e estrangeiras, por meio do software *Ucinet*. Conclui-se que já se destacam, no Brasil, instituições e pesquisadores, muitos deles líderes de Grupos de Pesquisa no tema. A produção científica brasileira sobre *Web Semântica*, apesar da temática recente, é consistente, revelando sua inserção nas principais revistas e eventos nacionais e internacionais. A análise da rede mostra que as instituições nacionais, que somam 19 no total, cooperam com 59 universidades, laboratórios, empresas e instituições de pesquisa internacionais, revelando assim a internacionalização da pesquisa brasileira.

Palavras-chave: Análise Bibliométrica; Web Semântica; Colaboração científica; Redes colaborativas; Grupos de pesquisa.

<sup>1</sup> Doutoranda em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho. Mestre em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho. Especialista de Laboratório da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP.

<sup>2</sup> Livre\_Docente na área de Estudos Bibliométricos da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho. Graduada em Matemática e Pedagogia. Atualmente é professora doutora da Universidade Estadual Paulista Júlio de Mesquita Filho.

**Abstract:** *This research aims, in general, to analyze the concept of Semantic Web as an object of study in Brazil, through bibliometric procedures, from the analysis of scientific literature on the subject. More specifically, this study aims to identify the most productive group of Brazilian researchers working on the theme; to identify approaches to research on Semantic Web, through ongoing research projects from the most productive Brazilian researchers, and to analyze the internationalization of Brazilian science on the subject by analyzing the institutional collaborative network, generated between Brazil and other countries. As research procedure, we used Scopus database to collect references on the subject Semantic Web, through keywords, which resulted in a total of 1,384 documents. After screening to evaluate the relevance of the work to the subject, we obtained a corpus of 1,049 articles. Data were grouped, presented in tables and analyzed in order to meet the proposed objectives. We generated a two-mode collaborative network across Brazilian and foreign institutions, using Ucinet software. We concluded that Brazilian institutions and researchers stand out, and these researchers lead research groups in the field. The Brazilian scientific production on Semantic Web, despite recent, is consistent, revealing its inclusion in major journals and national and international events. The network analysis demonstrates the internationalization of Brazilian research and 19 national institutions cooperating with 59 universities, laboratories, companies and international research institutions.*

**Keywords:** *Bibliometric analysis; Semantic Web; Scientific Collaboration; Collaborative Networks; Research Groups.*

## 1 INTRODUÇÃO

Os avanços tecnológicos na área da informação têm levado a pesquisas interessadas em solucionar problemas na área de recuperação de informações, especialmente na *web*. Surge, assim, uma nova metodologia para conteúdos digitais denominada *Web Semântica*. Em 2001, Berners-Lee, Hendler e Lassila publicaram, no *Scientific American*, uma proposta com pesquisas envolvendo a *Web Semântica*.

A *Web Semântica* constitui um interdomínio, no qual diversas áreas do conhecimento contribuem ativamente para seu desenvolvimento e implantação: Inteligência Artificial; Desenvolvimento da Web; Linguística Computacional; Engenharia de Documento e Bibliotecas Digitais; Desenvolvimento de Interfaces; e Ciências Sociais e Humanidades (MORALES DEL CASTILLO, 2011). Também a Organização do Conhecimento (KO) em sua vertente mais humanística pode contribuir para o desenvolvimento de sistemas de recuperação da informação, com base na *Web Semântica* (MARTÍNEZ-ÁVILA; SAN SEGUNDO; ZURIAN, 2014; MARTÍNEZ-ÁVILA, 2015).

A presente pesquisa objetiva, de forma geral, analisar o conceito de *Web Semântica* como objeto de estudo, no Brasil, por meio da análise da produção científica no tema. De forma mais específica, busca identificar o grupo de pesquisadores brasileiros mais produtivos que trabalham com a temática *Web Semântica*; identificar as diferentes abordagens sobre o

tema; e analisar a internacionalização da ciência brasileira, por meio da análise da rede colaborativa institucional, gerada entre o Brasil e demais países.

Esta pesquisa justifica-se ao contribuir para o entendimento das particularidades no tema, para a formulação de agendas de pesquisa compartilhadas entre diversos domínios, além de permitir a inserção e fortalecimento do tema em âmbito nacional e internacional. Some-se a esta questão a necessidade de estudos por se tratar de temática recente, em que as análises e avaliações se tornam imprescindíveis.

## **2 PROCEDIMENTOS METODOLÓGICOS**

A fonte de informação para estudo foi a base *Scopus*, na medida em que traz as publicações dos pesquisadores brasileiros em periódicos indexados e eventos com maior relevância no tema *Web Semântica*.

Para o levantamento das referências sobre o tema Web Semântica, em 14/7/2016, procedeu-se a seguinte estratégia de busca: utilizaram-se como termos de busca “ *web semantics*” OR “*semantic web*” OR “*Ontology*” OR “*RDF*” OR “*OWL*” OR “*Sparql*” AND Brasil OR *Brazil* no campo *Affiliation country*, no período de 2011 a 2015, em todas as áreas do conhecimento, a saber: *Life Sciences*, *Health Sciences*, *Physical Sciences* e *Social Sciences & Humanities*.

A busca resultou em um total de 1.384 documentos. Realizou-se a triagem destes documentos por tipo de publicação, limitando-se aos artigos e trabalhos de eventos por serem as publicações mais representativas na área da *Web Semântica*. O *corpus* passou por nova triagem, resultando num total de 1.049 publicações. Foram excluídos trabalhos cujas temáticas tratavam de epistemologia, antropologia, sociologia, genética, biologia, medicina veterinária, educação e etnografia.

## **3 APRESENTAÇÃO DOS RESULTADOS E ANÁLISES DOS DADOS**

Apresenta-se a Tabela 1 com os dez autores mais produtivos, as instituições a que estão vinculados, número total de trabalhos encontrados na base de dados, e suas coautorias nacionais e internacionais. A análise da Tabela 1 destaca como autores mais produtivos, nos últimos cinco anos, três pesquisadores da Universidade Federal do Espírito Santo (UFES), a saber: GUIZZARDI, G., com 52 publicações, doutor em *Computer Science* pela *Universty of Twente* na Holanda e pós-doutorado em duas instituições italianas; ALMEIDA, J. P. A., com 27 publicações, com graduação em Engenharia de Computação pela UFES, mestrado em *Telematics* e doutorado em Ciência da Computação pela *Universty of Twente* na Holanda; e GUIZZARDI, R. S. S., com 19 publicações na *Scopus*, com graduação e mestrado na UFES em

Ciência da Computação e Informática, respectivamente, e doutora em *Computer Science* pela *Universty of Twente* na Holanda.

TABELA 1 - Autores mais produtivos na base *Scopus*

<b>Autor</b>	<b>Instituição</b>	<b>Trabalhos</b>	<b>Trabalhos em Coautoria</b>	<b>Coautoria Nacional</b>	<b>Coautoria Internacional</b>
Guizzardi, G.	UFES	53	49	22	27
Casanova, M. A.	PUC-RIO	42	42	24	18
Abel, M.	UFRGS	27	27	22	05
Almeida, J. P. A.	UFES	27	27	18	09
Almeida, M. B.	UFMG	24	22	16	06
Revoredo, K.	UNIRIO	22	22	19	03
Freitas, F.	UFPE	21	20	14	06
Baião, F.	UNIRIO	20	20	20	00
Guizzardi, R. S. S.	UFES	19	19	05	14
Carbonera, J. L.	UFRGS	19	19	14	05
<b>TOTAL</b>		<b>274</b>	<b>267</b>	<b>174</b>	<b>93</b>

Fonte: Elaboração dos autores

Em relação às publicações e ao tipo de autoria, evidencia-se a preferência por publicações em coautoria com 97,4% do *corpus* realizados em colaboração com instituições nacionais e internacionais, com predomínio de colaborações nacionais (65,2%). Considerando as temáticas de interesse dos pesquisadores, registrados no Currículo Lattes, buscou-se identificar seus atuais projetos de pesquisa envolvendo a Web Semântica, apresentados na Tabela 2.

TABELA 2 – Projetos de pesquisa\*

<b>Autor</b>	<b>Projeto</b>	<b>Início</b>
Guizzardi, G.	1. <i>Ontological foundations of service Science</i>	2013
	2. Contribuições teóricas, metodológicas e computacionais para a Engenharia de Ontologias de Referência	2012
	3. Interoperabilidade semântica de dados abertos governamentais com base em ontologias	2012
Casanova, M. A.	Web de dados: fundamentos, técnicas e aplicações	2014
Abel, M.	Interoperabilidade semântica na cadeia de exploração de petróleo	2015
Almeida, J. P. A.	Semântica dos aspectos estratégicos de arquiteturas corporativas baseada em ontologias	2015
Almeida, M. B.	1. Representação de processos organizacionais em domínios de conhecimento especializado intensivo no âmbito de saúde pública	2015
	2. Gestão da informação para fomentar a sustentabilidade da carteira de geração de energia elétrica com fontes renováveis alternativas	2015
Revoredo, K.	1. Aprendendo, adaptando e alinhando ontologias: metodologias e algoritmos	2016
	2. Evoluindo alinhamentos de ontologias: metodologias e algoritmos	2016
Freitas, F.	Um framework para integração de ontologias e banco de dados heterogêneos utilizando tecnologias da Web Semântica	2013
Baião, F.	1. <i>RISE BPM: propelling business process management by research and innovation staff Exchange activity</i>	2015
	2. Refinamento e alinhamento de ontologias: metodologias e algoritmos	2015
	3. Metodologias e ferramentas computacionais para suporte a aplicação de acesso intensivo a	2014

	dados em ambientes distribuídos.	
Carbonera, J. L.	Interoperabilidade semântica na cadeia de exploração de petróleo	2015

Fonte: Elaboração dos autores

\* Renata Silva Souza Guizzardi, nona autora mais produtiva, não tem cadastrado nenhum projeto em andamento envolvendo a Web Semântica.

Observa-se que os projetos de pesquisa envolvem aplicações variadas da *Web Semântica*, evidenciando seu caráter de ciência aplicada, como as pesquisas de GUIZZARDI, ABEL, ALMEIDA J. e ALMEIDA, M., BAIÃO e CARBONERA, que abrangem a aplicação da Web Semântica em temáticas como saúde, gestão, sustentabilidade, cadeia produtiva de petróleo e informação governamental. Destaca-se que alguns projetos envolvem o estudo *da Web Semântica* e sua aplicabilidade sem a definição específica num tema. GUIZZARDI, CASANOVA, REVOREDO, FREITAS e BAIÃO possuem projetos de pesquisa com este perfil de interesse.

A fim de identificar a participação destes autores em Grupos de Pesquisa cadastrados no CNPq e sua possível interdisciplinaridade, apresenta-se a Tabela 3. Pesquisadores como ALMEIDA, M. B. e FREITAS participam de Grupos de Pesquisa em instituições distintas que a sua.

Tabela 3 – Grupos de Pesquisa dos autores mais produtivos\*

<b>Pesquisador</b>	<b>Instituição</b>	<b>GP Atualizado</b>	<b>Área Predominante</b>	<b>Tema</b>
Guizzardi, G.	UFES	1(L)**	Computação	Ontologia
Abel, M.	UFRGS	1(L)	Computação	Ontologia
Almeida, J. P. A.	UFES	1(L)	Computação	Ontologia
Almeida, M. B.	UFMG	1(L) 1 (HEMOMINAS)***	Ciência da Informação Medicina	Ontologia Ontologia Aplicada à Biologia e Medicina
Revoredo, K.	UFRJ	1(L)	Computação	Ciência da Web/Web Semântica
Freitas, F.	UFPE	1(UFAL)***	Computação	Ontologia Aplicada EAD
Baião, F.	UNIRIO	1(L) 1	Computação	Web Semântica Ciência da Web/Web Semântica
Guizzardi, R. S. S.	UFES	1	Computação	Ontologia

Fonte: Elaboração dos autores

\* Marco Antonio Casanova, não está cadastrado em nenhum Grupo de Pesquisa atualizado; e Joel Luis Carbonera, doutorando não está cadastrado como pesquisador e sim como estudante; \*\* A identificação (L) corresponde à líder do GP; \*\*\* Indicação de GP vinculado à instituição distinta.

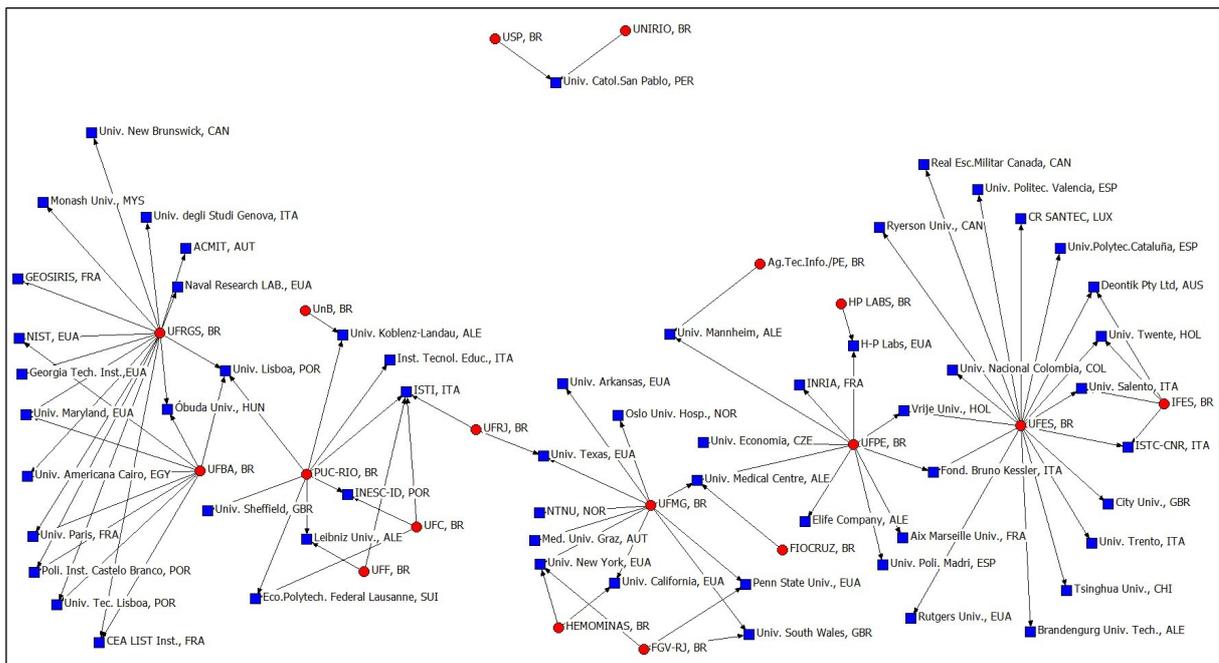
ALMEIDA, M. B. é o único pesquisador do grupo estudado, que está vinculado a GPs,

cuja área predominante difere da Computação: Ciência da Informação e Medicina. Estas temáticas podem ser explicadas pelo vínculo do pesquisador como docente na área da Ciência da Informação tanto na graduação quanto na pós-graduação. Destaca-se que 75% dos pesquisadores estudados estão cadastrados como líderes de seus respectivos grupos. O líder de um Grupo de Pesquisa pode ser considerado um pesquisador senior numa área do conhecimento, visto que suas atribuições devem ultrapassar questões técnicas e críticas da pesquisa, envolvendo habilidades ao mediar conflitos dentro e fora do grupo (SANTANA et al., 2014).

A rede colaborativa entre as instituições nacionais e internacionais é apresentada na Figura 1. A rede é formada por 19 instituições brasileiras que fizeram, no período, colaborações com 59 instituições estrangeiras na temática *Web Semântica*. As brasileiras estão representadas por círculos vermelhos, e os quadriláteros azuis representam as instituições estrangeiras, ligadas entre si por segmentos que indicam a intensidade de colaboração entre elas.

Todas as subredes são conectadas entre si, em sua maioria por meio de instituições estrangeiras. A Universidade de Lisboa é destaque por colaborar com três instituições brasileiras: UFRGS, UFBA e PUC-RIO, unindo essas subredes, e o *Istituto di Scienza e Tecnologie dell'Informazione "A.Fedo"*. O ISTI, instalado na cidade de Pisa (Itália), colaborou com a PUC-RIO, UFC e UFRJ, no entanto o Instituto não é ponto de conexão entre as subredes. O instituto de pesquisa italiano ISTI pertence a uma subrede que se conecta com as subredes, à direita da Figura 1, por meio de uma universidade brasileira, a UFRJ, única intermediadora entre as duas grandes subredes de colaboração científica identificadas.

FIGURA 1-Rede colaborativa *two-mode* entre instituições brasileiras e estrangeiras



Fonte: Elaboração dos autores e *software Ucinet*.

#### 4 CONSIDERAÇÕES FINAIS

Apesar da temática recente, a produção científica brasileira sobre Web Semântica tem sido consistente, com mais de mil trabalhos num período de cinco anos, indexados numa base de dados internacional, revelando sua inserção nas principais revistas e eventos nacionais e internacionais. Já se destacam algumas instituições e pesquisadores líderes de Grupos de Pesquisa no tema, coordenando outros pesquisadores e estudantes e gerindo as ações que possibilitam o desenvolvimento da Web Semântica, no Brasil.

Identificaram-se duas vertentes de estudo, a saber: aquelas que caracterizam a pesquisa brasileira em Web Semântica, sua aplicação em atividades e/ou domínios específicos como Gestão, Medicina, entre outros; e o estudo da Web Semântica em si, metodologias e requisitos para o próprio crescimento como área científica.

Agradecimentos ao apoio do CNPQ- Conselho Nacional de Desenvolvimento Científico e Tecnológico

#### REFERÊNCIAS

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: 20 de abril de 2016.

MARTÍNEZ-ÁVILA, D.; SAN SEGUNDO, R.; ZURIAN, F. A. Retos y oportunidades en organización del conocimiento em la intersección com las tecnologías de la información.

**Revista Española de Documentación Científica**, v. 37, n. 3, p. 1-13, 2014.

MARTÍNEZ-ÁVILA, D. Knowledge organization in the intersection with information technologies. **Knowledge Organization**, v. 42, n. 7, p. 786-498, 2015.

MORALES DEL CASTILLO, J. M. **Hacia la biblioteca digital semántica**. Gijón: Trea, 2011.

SANTANA, G. A. et al. Indicadores dos grupos de pesquisa da área de Gestão da Informação na Região Nordeste: um enfoque para a colaboração em artigos de periódicos. **Em Questão**, Porto Alegre, v. 20, n.3 – ed. esp., p. 229-252, 2014.