



COLABORAÇÃO ENTE ÁREAS DE PESQUISA EM ONTOLOGIAS NO BRASIL: Uma análise baseada na co-autoria no Evento ONTOBRAS

Eixo temático: Colaboração

Modalidade: Apresentação oral

1 INTRODUÇÃO

Ontologia é um tema que proporciona convergência entre áreas que pesquisam conceitos, classificações, modelagem e organização da informação e do conhecimento, como Ciência da Computação (CC), Ciência da Informação (CI) e Letras/Linguística (LET) e Filosofia, pois aborda a investigação de teorias e métodos para especificar formalmente conceitualizações compartilhadas em domínios específico. O desenvolvimento e o uso de ontologias também é interesse de pesquisa por qualquer área do conhecimento, quando buscam organizar e/ou integrar seus recursos e informações, e ontologias estão na base da Web Semântica, que compreende numa nova geração da Web que visa possibilitar que os recursos da web tenham seus significados compreendidos tanto por pessoas como por máquinas (Berners-Lee, Hendler e Lassina, 2001). Este trabalho investiga a colaboração entre áreas do conhecimento na pesquisa em ontologias no Brasil, através de um estudo bibliométrico focado na co-autoria de artigos publicados no evento ONTOBRAS, que é um encontro científico sobre ontologias, realizado no Brasil.

2 ONTOLOGIAS E O SEMINÁRIO ONTOBRAS

Derivada do grego *onto* (*ser*) + *logia* (*estudo*), pode-se dizer que Ontologia é a Ciência que estuda o ser. Desde então, as Ontologias se firmaram como uma área interdisciplinar, agregando pesquisadores de diversas ciências, como a Ciência da Informação, a Ciência da Computação, a Linguística, as Engenharias, entre outras.

De tempos em tempos, a necessidade de recuperação de um volume crescente de informação, faz com que sejam criados novos sistemas que atendam às demandas específicas da sociedade. De acordo com Café (2011, p.106) “a necessidade de agilizar buscas tem levado pesquisadores da área de Ciência da Computação, a desenvolver ontologias”. De acordo com Gruber (1993, apud SALES, 2009), uma ontologia é a

“representação de uma conceituação, um conjunto de conceitos estudados e especificados sobre uma determinada área do conhecimento”. Vital e Café (2011, p.117) defendem que as Ontologias:

[...] representam o conhecimento de um dado domínio em forma de uma rede relacional, intencional, onde as relações se sobrepõem aos possíveis ‘estados da coisa’. As ontologias podem ser consideradas mais complexas que as taxonomias, levando em consideração que não se limitam aos relacionamentos taxonômicos adotados na ordenação de classes e subclasses, estabelecendo outros tipos de relações semânticas, como as de associação, derivadas da explicitação das características dos conceitos.

O objetivo das Ontologias é melhorar a representação e recuperação da informação em meio digital, tornando-se uma ferramenta de referência para a representação de assuntos especializados. “No jargão dos pesquisadores em inteligência artificial, uma ontologia é um documento ou arquivo que define formalmente a relação entre termos” (SANTOS, 2011).

O Seminário ONTOBRAS tem como objetivo reunir a comunidade de pesquisadores e estudantes que investigam o tema Ontologia no Brasil, proporcionando uma oportunidade e um ambiente científico para trocar experiências a respeito de teorias, metodologias, linguagens e ferramentas relacionadas com o desenvolvimento de ontologias e aplicações. Desde 2008 ocorreram quatro edições do ONTOBRAS. A Primeira Edição¹ (2008) teve como objetivo discutir o estado da arte na área de Ontologias no Brasil e identificar grupos e projetos que estariam pesquisando sobre o tema. A Segunda Edição² (2009) teve como foco reunir os grupos de pesquisa brasileiros, para que houvesse o compartilhamento das experiências e conhecimentos acerca do tema ontologias, e a Terceira Edição³ (2010) objetivou consolidar o caráter interdisciplinar e de formação de comunidades em torno do tema ontologias. Quarta Edição⁴ (2011) teve como foco integrar os eventos sobre Ontologias já existentes no Brasil e incentivar a troca de experiências e formação de comunidades, consolidando o evento como um espaço multidisciplinar e internacional sobre o tema. Nesta edição, o ONTOBRAS juntou-se à Sexta Edição do MOST (International Workshop on Metamodels, Ontologies and Semantic Technologies). Na primeira, terceira e quarta

¹ Realizada pela UFF/Rio de Janeiro - <http://www.uff.br/ontologia/>

² Realizada pelo IME/Rio de Janeiro - <http://ontobra.comp.ime.eb.br/>

³ Realizada pelos grupos Engenharia de Conhecimento e CI da UFSC - <http://www.egc.ufsc.br/ontobras/>

⁴ Realizado pela UFRGS (Computação e Ciência da Informação) - <http://www.egc.ufsc.br/ontobras/>

edições ocorreram apresentação de trabalhos. Nos eventos de 2010 e 2011 foram apresentados, respectivamente 36 artigos (12 apresentados oralmente e 24 na forma de pôster) e 27 artigos (sete trabalhos longos, demonstrando resultados de pesquisa, e 20 trabalhos curtos, demonstrando trabalhos em andamento). Na edição de 2010, foram submetidos 65 (29 trabalhos longos, e 36 trabalhos curtos).

3 COLABORAÇÃO ENTRE ÁREAS NO EVENTO ONTOBRAS

Bibliometria é técnica quantitativa e estatística de medição dos índices de produção e disseminação do conhecimento científico (Araújo, 2006), e inclui a co-autoria. A co-autoria, conhecida também como autoria múltipla, autoria colaborativa, autoria em parceria, é a produção de um trabalho ou estudo científico com outrem (Ferreira, 1993). Dentre vários motivos para o estabelecimento da co-autoria apontados por Katz e Martin (1997, p.7) *apud* Vilan Filho (2010, p.33), destacamos: o incremento da especialização na ciência; o avanço das disciplinas científicas o que significa que pesquisadores requerem mais e mais conhecimento para fazer avanços significativos, demanda que frequentemente só pode ser atendida com esforço conjunto; a crescente profissionalização da ciência; o desejo de crescer e obter fertilização cruzada entre disciplinas.

Co-autoria é o instrumento usado neste trabalho para investigar a cooperação entre áreas de pesquisa no ONTOBRAS. Os dados bibliográficos (autores e artigos) formam extraídos dos anais dos eventos de 2010 e 2011. A definição da área de cada autor teve como critério a área do departamento em que o mesmo está vinculado, e o Currículo Lattes foi utilizado para resolver problemas de inconsistências de nomes e identificação de vínculos e áreas. Para autores que são estudantes, a instituição considerada foi aquela que determina o vínculo acadêmico, pois o trabalho tem como enfoque a colaboração entre áreas de pesquisa. Foram escolhidas para a análise as edições de 2010 e 2011 em função destas representarem o evento já consolidado.

A análise valeu-se de duas abordagens: investigar a cooperação entre áreas (através da verificação da co-autoria entre autores de áreas diferentes), e investigar a formação de grupos e de pesquisa e das áreas de pesquisa que compõem estes grupos (através a análise de grupos de co-autoria e as áreas de seus membros). O quadro 1 apresenta os dados gerais do evento (autores, artigos, autores por artigo e autores que produzem junto com autores de outras áreas). O evento apresenta uma média de 2,2

autores por artigo, sendo que somente 12% dos artigos tiveram a participação de um único autor. A cooperação entre áreas ocorre em 49% dos autores e em 42% dos artigos escritos, e aumentou na segunda edição, de 45% para 60% nos autores, de 38% para 48% nos artigos (quadro1). Estes índices atingem 49% dos artigos e 52% dos autores, quando excluídos os artigos/autores de autoria única. Conclui-se que há colaboração entre áreas no ONTOBRAS, alcançando praticamente a maioria simples em artigos e autores.

Quadro 1 – Totais do Evento

	2010	2011	Ambos
Artigos	36	27	63
Artigos de autoria única	5 (13%)	3(11%)	8 (12%)
Artigos feitos por autores de mais de uma área	14 (38%)	13 (48%)	27 (42%)
Autores	95	69	140
Autores por artigo	2,6	2,5	2,2
Autores que participaram em 2010 e 2011	-	-	24 (17%)
Autores com um único artigo sem co-autoria	5 (5%)	3 (4%)	-
Autores que escrevem com autores de outras áreas	43 (45%)	42 (60%)	69 (49%)

O quadro 2 e as redes da figura 1 demonstram a investigação da colaboração entre as áreas, destacando aquelas que mais colaboram, assim como a intensidade da colaboração. As áreas mais representativas são Ciência da Computação (CC), com 56,4% do total de autores (considerando os dois eventos), Ciência da Informação (CI), com 12%, e Engenharia da Produção (ENGP), com 17% dos autores. As seguintes características foram observadas com relação à colaboração de cada área com outras áreas. **Computação:** abrange a maioria dos autores (65%), em que quase a metade colabora com autores de outras áreas(43%). A colaboração ocorre com todas as áreas, com predominância para CI, Geologia (GEO) e ENGP. **Ciência da Informação:** participa com 12% de autores, que são os que mais colaboram (66%). A colaboração é considerável com CC e constante nas duas edições. Colaborações pequenas com Medicina (MED) e ENGP. **Engenharia da Produção:** participa com 17% dos autores, em que 37% colaboram. A colaboração é considerável com CC e a participação da área diminui em 2011. Colaborações pequenas com MED e ENGP; **Letras e Geologia:** participação dos autores inferior a 1%, restrita e considerável com CC (3 e 4 artigos respectivamente). **Engenharia e Educação:** participação dos autores inferior a 1%, exclusiva com CC, com dois artigos cada. ENG se mantém constante nas duas edições. **Medicina:** participação dos autores inferior a 1%, com ENGP em uma edição e CI em outra.

Conclui-se que há maior intensidade na colaboração entre CC e CI (9 artigos), seguida das colaborações entre CC e ENGP (cinco artigos), entre CC e GEOL (quatro artigos) e CC e Letras (LET, três artigos). Dentre estas áreas, os autores da CI são os que mais colaboram com outras áreas (66%), mas o índice é significativo para as demais (43% para CC e 37% para ENGP). CC, ENGP e CI destacam-se como áreas centrais à colaboração por combinarem características como a representatividade dos autores e a intensidade da colaboração de seus autores com outras área. Entretanto, há pouca colaboração entre ENGP e CI, e CC é a área central na colaboração. As demais áreas (GEOL, ENG, LET, MED e Educação - EDU) possuem representatividade de autores inferior a 1%, estando atreladas às áreas centrais (principalmente CC) e não cooperam entre si. Isso nos leva a formular a hipótese, a ser investigada, de que os autores destas áreas cooperam os autores das áreas centrais no campo da aplicação de ontologias.

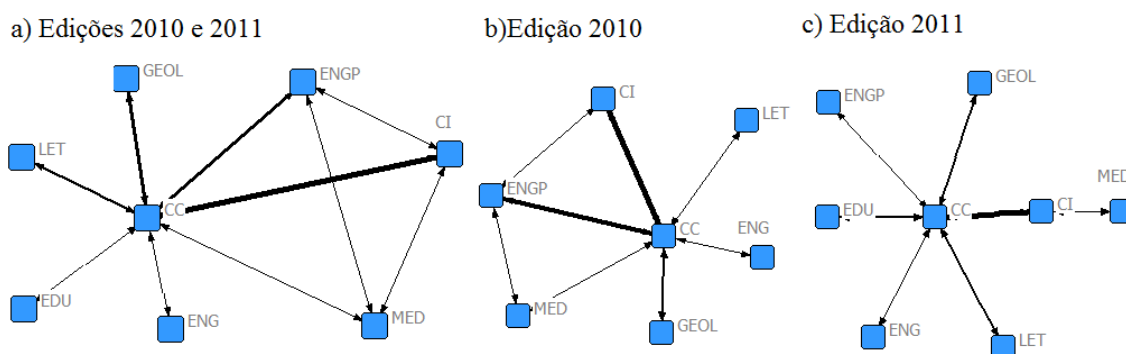


Figura 1 – Rede de Cooperação entre Áreas - em que os elos são determinados por artigos elaborados por autores de duas áreas

Quadro 2 – Totais por Área

Áreas	Autores 2010	Autores 2011	Total	Participação por área	Autores nas duas edições	Autores que escrevem c/ autores de outras áreas
CC	51	39	79	56,4%	11 (14%)	34 (43%)
CI	13	12	18	12,8%	7 (4%)	12 (66%)
ENG	3	2	4	0,28%	1	4 (100%)
ENGP	22	5	24	17,1%	3	9 (37%)
GEOL	2	3	3	0,21%	2	3(100%)
LET	3	6	9	0,64%	0	4 (44%)
MED	1	1	2	0,14%	0	2 (100%)
EDU	0	1	1	0,007%	0	1 (100%)

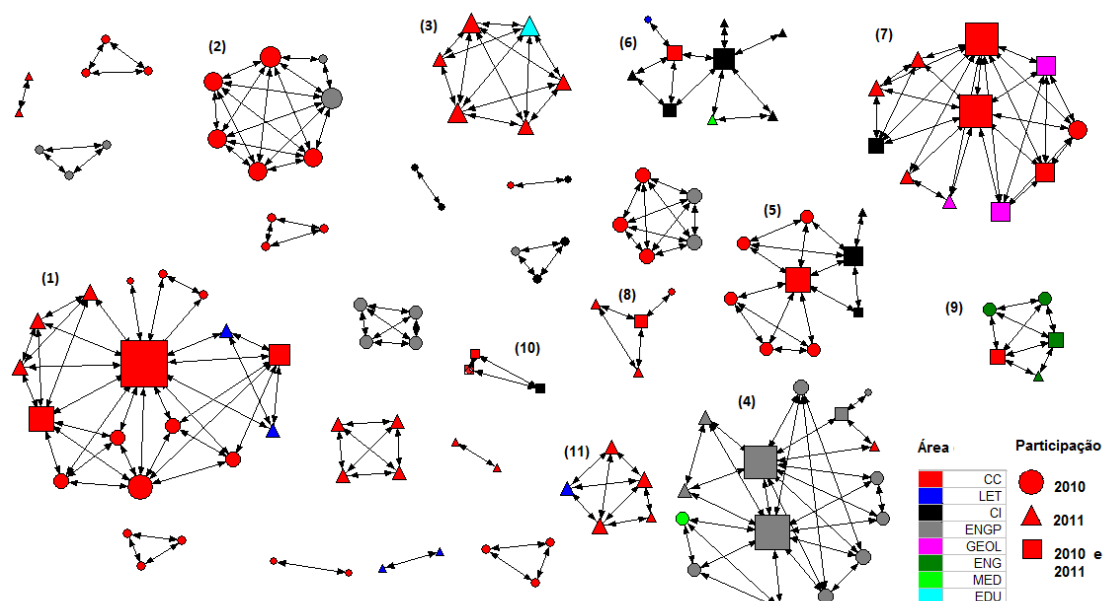


Figura 2 – Redes de colaboração entre autores (tamanho do nodo indica centralidade)

Quadro 3 – Categorias de Grupos

	Características	Grupos (figura 2)
A	Vários autores, artigos presentes nas duas edições, com centralidade em um ou dois autores, ambos de uma única área, com autores de outras áreas com participações menores (periféricas)	1-CC PUC-RS 4-ENGP UFSC 7-CC/UFRGS
B	Menores que os do tipo A, com autores e artigos presentes nas duas edições, com centralidades em autores de duas áreas.	5-CC/UFRJ e CI/UFF 6-CI/UFGM e CC/FGV
C	Pequenos, com a participação de um artigo em cada edição, autores centrais de duas áreas.	9- IME/ENG e LNCC/CC 10- CC/UFGM-CI/UFGM
D	Pequenos, com a participação de um artigo em cada edição, sem colaboração com outra área.	8-CC/USP
E	Pequenos, com a participação em um único evento, com autores centrais de duas áreas.	2- CC/UCSal 3- CC/UCSal
F	Formado por autores ligados apenas por um artigo, em uma única edição. Somente em um grupo ocorrem autores de áreas diferentes	Total de 14 grupos

A rede da figura 2 apresenta os grupos formados no ONOTBRAS através da co-autoria. Esta rede identifica também as áreas de cada participante (pelas cores dos nodos), o evento da participação de cada membro (através da forma do nodo) e os membros mais centrais (pelo tamanho do nodo). A partir da análise dos grupos, das áreas dos seus membros e das suas participações nos dois eventos, sete tipos de grupos foram identificados (quadro 3). Essa classificação de grupos nos leva a concluir que no ONTOBRAS existem grupos grandes e com presença marcante nos dois eventos (Tipo A, grupos 1, 4 e 7). Estes grupos, entretanto, caracterizam-se por estarem centrados em uma única área, em que a participação de autores de outras áreas se dá de forma periférica. Por outro lado, em grupos menores, mas com presença nos dois eventos (Tipos B e C grupos 5, 6, 9 e 10), observa-se que a cooperação entre as áreas ocorre em maior equilíbrio, com centralidade em autores de mais de uma área. Isso também ocorre com grupos que participaram de um único evento com mais de um artigo (Tipo D,

grupos 2 e 3). Nos grupos formados em torno de um único artigo/evento (Tipo F), observa-se que colaboração entre áreas é inexpressiva (1 grupo em 14), enquanto que nos demais tipos somente em um grupo (8) não há autores de várias áreas. Conclui-se que a formação dos grupos no ONTOBRAS ocorre em dois formatos favoráveis à colaboração entre áreas: pesquisadores buscam apoio às suas pesquisas juntando-se a grupos grandes, centrados em pesquisadores de uma única área (Tipo A); e pesquisadores de áreas distintas agrupam-se com a finalidade de crescer e obter fertilização cruzada, em que há equilíbrio de áreas no quesito centralidade (Tipos B, C e E).

4 CONSIDERAÇÕES FINAIS

Conclui-se que há colaboração entre áreas no ONTOBRAS com base principalmente nas seguintes constatações: que há grande colaboração de autores e a maioria absoluta destes autores colabora com autores de outras áreas; e que, excluindo-se grupos formados por autores que escreveram um único artigo, que possuem pouca representatividade no ONTOBRAS, os outros grupos consolidam-se e seguem modelos que incentivam a colaboração envolvendo autores de várias áreas. No sentido de completar este estudo, uma atividade está sendo desenvolvida, que envolve identificar elos de ligação não explícitos entre os grupos, através das citações, isto é, da identificação de grupos, autores e áreas que se relacionam por citar mesmos autores.

REFERÊNCIAS

- ARAÚJO, Carlos Alberto. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11-32, jan./jun. 2006.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American Magazine**, May 2001.
- BORGATTI, S.P. 2002. **Netdraw Network Visualization**. Analytic Technologies: Harvard, MA.
- CAFÉ, Lígia Maria Arruda; VITAL, Luciane Paula. Ontologias e taxonomias: diferenças. **Perspectivas em Ciência da Informação**, v.16, n.2, p.115-130, jun./ago. 2011.
- SALES, Rodrigo de; CAFÉ, Lígia. Diferenças entre tesauros e ontologias. **Perspectivas em Ciência da Informação**, v.14, n.1, p. 99-116, jan./abr. 2009.
- SANTOS, Livia Regina Nogueira dos. Ontologias aplicada a padronização dos currículos de pesquisadores: mapeamento do conhecimento. **DataGramaZero**. v.12 n.5 out/11.
- VILAN FILHO, Jaime leiro. **Autoria múltipla em artigos de periódicos científicos das áreas de informação no Brasil**.2010. (Tese de Doutorado) 215f.Faculdade de Ciência da Informação. Universidade Federal de Brasília. Brasília, DF: 2010.
- VITAL, Luciane Paula; CAFÉ, Lígia Maria Arruda. Ontologias e taxonomias: reflexões conceituais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 16, n. 2, p. 115-130, 2011.