



CIÊNCIA DA INFORMAÇÃO E WEB SEMÂNTICA: UMA RELAÇÃO EFETIVA OU APÓCRIFA?

Luís Machado¹, Maria da Graça Simões², Renato Rocha Souza³, Maurício Almeida⁴

¹Universidade de Coimbra, 0000-0003-3403-5618, luismachado@fe.uc.pt

²Universidade de Coimbra, gsimoes@darq.uc.pt

³Fundação Getúlio Vargas, 0000-0002-1895-3905, renato.souza@fgv.br

⁴Universidade Federal de Minas Gerais, mba@eci.ufmg.br

RESUMO A Web é o local privilegiado para procurar informação. As potencialidades de uma *web semântica* para recuperar informação afiguram-se essenciais. Na Ciência da Informação (CI), esta ideia surge com Otlet ao desenvolver uma “web semântica analógica”, justificando a ligação entre: Web Semântica (WS) e CI. Procura-se verificar e refletir sobre a presença da WS no *output* dos cursos de mestrado/doutoramento em CI, Portugal e Brasil (2005-2015). Para tanto: identificam-se os cursos, dissertações de mestrado (DM) e teses de doutoramento (TD) em CI associados à WS; identificam-se pontos de confluência entre as duas áreas nestes trabalhos; caracterizam-se o perfil académico dos autores e orientadores dos mesmos. Partiu-se do paradigma qualitativo: revisão bibliográfica e estudo empírico, centrado na análise de conteúdo dos trabalhos recuperados (RCAAP e OASISbr). Verificou-se 4 registos únicos no RCAAP; 40 no OASISbr. Identificaram-se 4 cursos de mestrado e 3 de doutoramento (Portugal), e 16 de mestrado e 9 de doutoramento (Brasil). Recuperaram-se 0 DM/TD (Portugal) e 16 DM, 5 TD (Brasil). Identificaram-se 3 pontos de confluência: a) contribuição mútua (62%), b) WS enquanto tema interdisciplinar (29%) e c) complementaridade de conhecimentos (9%). Contabilizaram-se 11 trabalhos (52%) onde os investigadores responsáveis apresentam, no conjunto, formação nas áreas CI/Informática. O estudo aponta para a WS como um campo de investigação da CI, embora essencialmente teórico, sendo o trabalho prático interdisciplinar, CI/Informática, incipiente e pouco representativo.

PALAVRAS-CHAVE *Web Semântica; Ciência da Informação; cursos de Mestrado/Doutoramento; Portugal; Brasil.*

ABSTRACT The Web is the privileged place for seeking information. The potential of the so-called *semantic web* to enhance retrieval of relevant information seems essential. In Information Science (IS), this idea was developed from the Otlet's “analogic semantic web”, justifying the connection between the Semantic Web (SW) and the field of IS. We aimed to exam and analyze the presence of subjects related to SW in the output of IS master and PhD courses in Portugal and Brazil (2005-2015). Specifically, we aimed to: identify IS courses, master's dissertations (MD) and PhD thesis (PD) associated to SW; identify the points of confluence between the IS and the SW in these works; characterize the academic profile of these works' authors and advisors. Based on a qualitative approach, we performed a literature review and an empirical study focused on content analysis of the retrieved works (RCAAP and OASISbr). As results, we have: verified 4 unique records in RCAAP and 40 in OASISbr; identified 4 Master's and 3 PhD courses (Portugal) and 16 masters and 9 PhD courses (Brazil); retrieved no MD/PD in Portugal, 16 MD and 5 PD in Brazil; identified 3 confluence points: a) mutual contribution (62%), b) WS as an interdisciplinary theme (29%) and c) complementary knowledge (9%); identified 11 works (52%) where the respective researchers have training in IS and Computer Science (CS). The study points to the WS as a field of investigation in IS, although essentially theoretical, insofar as practical interdisciplinary works in CI/CS are incipient and not representative.

KEYWORDS *Semantic Web; Information Science; Master's/PhD courses; Portugal; Brazil.*

COPYRIGHT Este trabalho está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional (<https://creativecommons.org/licenses/by/4.0/deed.pt>)

INTRODUÇÃO

Nas últimas décadas tem-se assistido a um aumento exponencial de informação produzida no espaço da *World Wide Web* (WWW ou simplesmente *Web*). Na transição da informação, do analógico para o digital, uma carência começou a evidenciar-se: a necessidade de uma recuperação mais eficaz da mesma, considerando-se a precisão e a revocação (Caro-castro, 2010; Patel-Schneider & Horrocks, 2007; Souza & Alvarenga, 2004). Neste contexto uma “nova” *Web*, denominada de *Web Semântica*, apresenta-se como solução para esta situação. Muitas esperanças foram colocadas nesta *Web Semântica* (WS), nomeadamente, como sendo capaz de criar um ambiente onde *agentes*, um tipo específico de *software* “inteligente”, processando e partilhando “conhecimento”, conseguem realizar tarefas complexas de acordo com as necessidades dos utilizadores (Berners-Lee, 2006; Berners-Lee & Fischetti, 2000; Berners-Lee, Hendler, & Lassila, 2001). A questão que se coloca neste cenário é a necessidade de tornar as máquinas capazes de raciocínios similares ao dos humanos decorrendo, daí, uma associação à chamada Inteligência Artificial, que os investigadores ligados à WS, logo nas publicações iniciais, procuraram afastar relacionando-a, antes, à modelação do conhecimento humano utilizando meios computacionais por forma a possibilitar inferências automáticas (Berners-Lee & Hendler, 2001; Shadbolt, Hall, & Berners-Lee, 2006) e, deste modo, garantir uma recuperação da informação eficaz.

A revisão da literatura relativa a este conceito (*Web Semântica*), revelou a necessidade de uma clarificação relativamente a outros conceitos associados, comumente utilizados como sinónimos, tais como: *Web de Dados* (*Web of Data*) ou *Web de Dados Ligados* (*Web of Linked Data* ou simplesmente *Linked Data*), conceitos esses que, apesar de intrinsecamente relacionados, são passíveis de distinção conceitual.

A WS apresenta-se, de acordo com a análise efetuada, como um constructo abrangente cujo objetivo final pode ser colocado como o de tornar a WWW uma base de dados global. Tornando possível, por um lado, uma pesquisa transversal (independente do tipo, formato e fonte dos dados), por outro, a obtenção de dados semanticamente interrelacionados, i.e., informação e não apenas uma listagem de documentos, muitas vezes sem ligação semântica entre si, como acontece na atual “web de documentos” (World Wide Web Consortium [W3C], 2013, 2015).

A esta *web*, resultante do desenvolvimento da WS, corresponde o conceito de *Web de Dados*, tal entendimento é reforçado pelo facto de se constatar que é esse o termo utilizado quando o objetivo da WS é mencionado (Berners-Lee, 1998; Bizer, Heath, & Berners-Lee, 2009; Hendler, Berners-Lee, & Miller, 2002; W3C, 2016).

Ao contrário da *Web de Dados*, ainda longe de se tornar efetiva, a *Web de Dados Ligados* está já em funcionamento (Hendler, 2014; Sletten, 2014), fazendo uso do conjunto de procedimentos e regras para a publicação de dados abertos (*Linked Open Data*), de especificações como RDF e SPARQL e das chamadas “ontologias leves”, tais como SKOS, FOAF ou Dublin Core (Bizer et al., 2009; Cambridge Semantics, 2016; W3C, 2016). A *Web de Dados Ligados* não é considerada como o objetivo da WS pois, ao contrário do que se pretende com a *Web de Dados*, a primeira mencionada apresenta grandes limitações na abrangência de ligações e, sobretudo, nas relações semânticas das mesmas (Hendler, 2011, 2014).

Pode-se, desta forma, entender a Web de Dados Ligados como uma parte da Web da Dados e esta como parte da WS que engloba, também, todas as especificações (padrões, linguagens, vocabulários, protocolos) desenvolvidas ou associadas à sua operacionalização, comumente designadas por “tecnologias da web semântica” (Bikakis, Tsinaraki, Gioldasis, Stavrakantonakis, & Christodoulakis, 2013; Parreiras, 2012).

A WS pretende, assim, potenciar a ideia original da WWW de disponibilizar o conhecimento humano, de forma abrangente, promovendo o trabalho colaborativo (Berners-Lee, 2009). Todavia, meio século antes desta *Web* digital, já o projeto da *Rede Universal de Documentação* de Paul Otlet, assenta nos mesmos ideais sendo considerada por vários autores como uma “web semântica” de madeira e cartão (Heuvel & Rayward, 2011; Laaff, 2011). Esta visão expandida da Documentação, de Otlet, é considerada, por muitos autores, como estando na gênese da área que, atualmente assume a nomenclatura de Ciência da Informação (CI) (Borges & Casado, 2009; A. M. da Silva & Ribeiro, 2008). Neste contexto, existe uma ligação entre o propósito da WS e as fundações teóricas da CI (Robredo, 2010; Souza & Alvarenga, 2004).

Concettualmente, a CI surge com frequência, na literatura, como uma área científica cujas demarcações, de conhecimento e de intervenção, se mostram complexas (Barreto, 2008; Saracevic, 1996; J. L. C. Silva & Freire, 2012; Zins, 2007a). Em termos epistemológicos a CI apresenta-se como uma área excessivamente flexível em termos de solidez disciplinar, a ponto de a comprometer (Galvão, 1998; Souza & Almeida, 2009). Parte desta problemática encontra-se relacionada com as intrincadas interações disciplinares presentes ao longo de todo o trajeto histórico da CI, em particular com a Biblioteconomia, a Arquivística e a Documentação (Borko, 1968; Le Coadic, 1996; Mikhailov & Giljarevskij, 1971; L. V. R. Pinheiro, 1997; A. M. da Silva & Ribeiro, 2008; Wersig & Neveling, 1975). Importa ressaltar que, relativamente ao conceito de disciplina, este é perspetivado segundo a visão epistemológica de Morin (2007), enquanto categoria autónoma de um corpo científico mais vasto.

Outra ligação disciplinar de relevo, que contribui para a complexidade conceitual da área da CI, é a que esta mantém com a Informática/Ciência da Computação (I/CC), a ponto da primeira ser “confundida” com a segunda (L. V. R. Pinheiro, 1997). Mesmo entre académicos relacionados com a área da CI é possível observar entendimentos similares, não só no que diz respeito à I/CC como à Biblioteconomia, à Comunicação ou à Matemática (Zins, 2007b).

Apesar destas díspares conceções da área, é possível, segundo Zins (2007a, 2007b, 2007c, 2007d), apresentar uma definição que reúne um razoável consenso entre especialistas contemporâneos, onde a CI é entendida como o estudo dos aspetos de mediação entre os utilizadores e os fenómenos dado, informação e conhecimento (enquanto artefactos humanos) e a sua implementação no domínio cultural/social (entendido como a forma da sociedade encarar a realidade).

A ligação entre a CI e a WS, embora concettualmente se afigura óbvia, em termos pragmáticos não se mostra tão evidente. O contexto teórico-prático que se consubstancia nas Dissertações de Mestrado (DM) e Teses de Doutoramento (TD) originárias de cursos de CI, cujo tema é a WS, é um campo de estudo profícuo, tanto mais que este tipo de estudos são indicadores importantes na aferição do desenvolvimento de uma área científica. Neste sentido, a reflexão que se propõe fornece indícios relevantes para a clarificação da relação CI-WS, em termos práticos e teóricos. A incidência do estudo em trabalhos produzidos em cursos de Portugal e do Brasil afigura-se pertinente, na medida em que estes países partilham afinidades históricas e culturais.

Assim, pretende-se: verificar e proceder a uma reflexão crítica no que respeita à presença da WS no *output* dos cursos de mestrado e doutoramento em Portugal e no Brasil em CI, considerados no período de 2005 a 2015. Para tal, propõe-se, caracterizar os repositórios agregadores de Portugal e do Brasil, respetivamente, RCAAP e OASISbr¹ dado serem a fonte de coleta de dados; identificar os cursos de mestrado e doutoramento em Portugal e no Brasil em CI certificados, respetivamente, pela A3ES e pela CAPES²; verificar o número de DMs e TDs em CI associados à WS, produzidas em Portugal e no Brasil (2005 a 2015), recuperadas de acordo com a metodologia desenhada para este estudo; identificar os pontos de confluência da CI com a WS, presentes nos trabalhos recuperados; caracterizar o perfil académico dos autores, orientadores e coorientadores dos trabalhos que constituem o corpus de estudo.

O desenho metodológico considerado para a efetivação destes propósitos assentou num paradigma qualitativo/pragmático, cuja descrição dos métodos e técnicas se apresentam no ponto que se segue.

METODOLOGIA

Desenhou-se uma metodologia de natureza qualitativa exploratória assente em duas fases. Na primeira recorreu-se a uma revisão bibliográfica para contextualizar os objetos de estudo: Web Semântica e Ciência da Informação, assim como para fundamentar teoricamente a parte empírica desenvolvida na segunda fase do estudo. A bibliografia recolhida foi analisada e cruzada. Relativamente à sua tipologia recorreu-se, essencialmente, a artigos científicos compreendidos, de um modo geral, entre 1990 e 2016.

No que se refere à segunda fase desenvolveu-se um estudo empírico. Iniciou-se por limitar a recolha dos dados aos repositórios RCAAP e OASISbr, dado que ambos se configuram como ponto de acesso à produção científica e académica nacional, multidisciplinar e de acesso aberto. Limitou-se, o estudo às DM e às TD, dado a, já referida, importância deste tipo de produção e sua ligação direta aos cursos superiores de uma área científica específica. Por fim, foi estabelecido como limite o período entre 2005 e 2015 pelo facto de, neste período, se ter assistido à evolução conceptual da Web Semântica, assim como da sua aplicabilidade em vários domínios do conhecimento, entre os quais na CI.

Seguiu-se a identificação dos cursos de mestrado e doutoramento através das estruturas oficiais associadas aos sistemas educativos dos respetivos países em estudo. Portugal: Direção Geral do Ensino Superior, Direção-Geral de Estatísticas da Educação e Ciência, Agência de Avaliação e Acreditação do Ensino Superior; Brasil: Conselho Nacional de Desenvolvimento Científico e Tecnológico, portal do Governo Federal do Brasil, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Para a pesquisa foi selecionado o termo *web semântica*, uma vez que, após o cruzamento entre os termos de quatro glossários sobre a WS e os programas, conteúdos e ementas dos cursos de CI identificados, se verificou ser esse o termo que reunia a pertinência e relevância para os objetivos do estudo.

Procedeu-se a testes prévios nos dois repositórios, RCAAP e OASISbr, no sentido de aferir e ajustar os procedimentos a aplicar, nomeadamente, no que concerne à forma de recuperar possíveis variações do

¹ RCAAP - Repositório Científico de Acesso Aberto de Portugal, OASISbr - Portal Brasileiro de Acesso Aberto à Informação Científica.

² A3ES - Agência de Avaliação e Acreditação do Ensino Superior, CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

termo eleito, como aos campos e filtros a utilizar nas respetivas plataformas digitais. A escolha dos campos *Título* e *Assunto* garantiu uma uniformização de critérios entre repositórios e direcionou a pesquisa a trabalhos onde se assume que a WS tenha um papel central, dados os trabalhos recuperados apresentarem o respetivo termo em um, ou nos dois elementos relevantes de identificação das DM e TD. Os critérios usados na recolha de dados foram estabelecidos de acordo com os objetivos específicos dado a articulação intrínseca destes com o processo de recolha (Quivy & Campenhoudt, 1992).

Num total de 519 registos recuperados eliminaram-se 297 (duplicados ou falsos positivos) constituindo a amostra 222 trabalhos. Desta amostra, foi apurada uma subamostra de 21, número relativo aos trabalhos que reuniam as condições para o estudo — ser proveniente de um curso de CI e apresentar o termo *Web Semântica* no título e/ou nas palavras-chave. O corpus deste estudo é constituído por estes 21 trabalhos.

Na análise dos resultados foi aplicado um conjunto de técnicas e procedimentos metodológicos denominado de análise de conteúdo (Bardin, 2011), adotando-se níveis analíticos diferentes, porque o cumprimento dos objetivos assim o exigiram. Para determinar os pontos de confluência entre a CI e a WS e o grau de proximidade temática na abordagem efetuada nos 21 trabalhos que constituíram o corpus do estudo, analisaram-se (nível mais exaustivo) os seus resumos e respetivos elementos internos³, e os objetivos, expressos na introdução dos documentos. Estabeleceram-se critérios para determinar os pontos de confluência grau de confluência e o grau de proximidade temática entre a CI e a WS (ver figura 1), expressos em uma escala de intensidade negativa (Bardin, 2011).

c1: Critérios usados na análise dos pontos de confluência da Web Semântica (WS) com a Ciência da Informação (CI), expressos nos resumos dos trabalhos constituintes da amostra.										
Presença de pontos de confluência no resumo	Frequência mínima								Grau de confluência atribuído	
	de ligações (entre os conceitos CI e WS)					dos termos			quantitativo	qualitativo
	diretas	e/ou	indiretas	e/ou	c/intermediário	CI	WS	tipo de relação		
Sim	2	e	indiferente			n.a	n.a	n.a	0	maior
	1	ou	3	e	indiferente	n.a	n.a	n.a	-1	↓
	0	e	1	ou	3	n.a	n.a	n.a	-2	↓
	0	e	0	e	1	n.a	n.a	n.a	-3	↓
Não	n.a	n.a	n.a	n.a	n.a	1	1	direta	-4	↓
	n.a	n.a	n.a	n.a	n.a	1 deles		indireta	-5	↓
	n.a	n.a	n.a	n.a	n.a	0	0	n.a	-6	menor
<i>n.a: não aplicável</i>										
c2: Critérios usados na análise da centralidade temática da Web Semântica (WS), expressa nos objetivos dos trabalhos constituintes da amostra.										
Função do conceito WS desempenhada nos objetivos	Presença/frequência do termo WS no(s) objetivo(s):						Grau de proximidade atribuído			
	Geral			Específicos			quantitativo	qualitativo		
Configura-se como o próprio objeto da ação	Apresenta						Em 50% ou mais		0	maior
Enquanto objeto da ação apenas se encontra nos objetivos específicos							Em menos de 50%		-1	↓
Surge como auxiliar do objeto da ação ou como elemento contextual	Surge no geral ou em um dos específicos e não nos dois em simultâneo						-2	↓		
Relação conceitual não explicitada	Não apresenta			Igual a 0%			-3	↓		
							-4	menor		

Figura 1. Critérios de análise dos resumos (c1) e objetivos (c2) dos trabalhos e respetiva escala gradativa

³ Introdução, objetivos, metodologia, resultados e conclusões.

Relativamente ao perfil académico dos autores, orientadores e coorientadores dos 21 trabalhos (nível analítico de grau moderado), a análise incidiu na formação académica dos mesmos. Consideraram-se como fontes para a recolha desses dados as plataformas digitais: *DeGóis*, para docentes portugueses e a *Lattes*, para docentes brasileiros.

Confrontou-se, posteriormente, esta última análise com a soma das duas anteriores (aos resumos e aos objetivos), no sentido de verificar a existência de uma correlação entre o grau de confluência/proximidade da relação WS-CI e a diferença quantitativa das formações em I/CC relativamente às formações em CI dos autores e orientadores de cada trabalho. Utilizou-se uma escala bipolar onde o polo de intensidade positiva máxima corresponde a um maior peso da área I/CC na formação e o polo inverso ao peso da área CI.

Por fim, com um nível analítico de menor grau, verificou-se a presença da Web Semântica enquanto tema dos trabalhos recuperados de acordo com as áreas científicas.

RESULTADOS

Verificou-se, quanto aos repositórios RCAAP e OASISbr, diferenças tanto no enquadramento normativo como nas funcionalidades (o segundo apresenta, e.g., mais opções de “pesquisa avançada”). Observou-se a existência de trabalhos recuperáveis apenas em um dos repositórios (quatro no RCAAP e 40 no OASISbr), ao contrário do espectável dado ambos agregarem os documentos existentes no outro.

Relativamente aos cursos de CI, ativos em 2016, identificaram-se quatro de mestrado e três de doutoramento em Portugal e, no Brasil, 16 de mestrado e 9 de doutoramento (ver tabela 1).

Tabela 1. Cursos de mestrado e doutoramento em CI, ativos em 2016, em Portugal e no Brasil certificados, respetivamente, pela A3ES e pela CAPES

Instituição de Ensino Superior	Ref. ^a	Designação do Curso	Tipologia ^b
Universidade de Coimbra	P.01	Ciência da Informação	M & D
Fundação Ensino e Cultura Fernando Pessoa	P.02	Ciência da Informação	D
Universidade do Porto	P.03	Ciência da Informação	M
Universidade do Porto & Universidade de Aveiro ^c	P.04	Informação e Comunicação em Plataformas Digitais	D
Universidade de Lisboa	P.05	Ciências da Documentação e Informação	M
Universidade do Algarve	P.06	Ciências Documentais	M
Universidade de São Paulo	B.01	Ciência da Informação	M & D
Universidade Estadual Paulista Júlio de Mesquita Filho/Marília	B.02	Ciência da Informação	M & D
Universidade Estadual de Londrina	B.03	Ciência da Informação	M
Universidade Federal da Bahia	B.04	Ciência da Informação	M & D
Universidade Federal da Paraíba/João Pessoa	B.05	Ciência da Informação	M & D
Universidade Federal de Pernambuco	B.06	Ciência da Informação	M
Universidade Federal Santa Catarina	B.07	Ciência da Informação	M & D
Universidade Federal Fluminense	B.08	Ciência da Informação	M & D
Universidade Federal do Rio de Janeiro	B.09	Ciência da Informação	M & D
Universidade de Brasília	B.10	Ciência da Informação	M & D
Universidade Federal de Minas Gerais	B.11	Ciência da Informação	M & D

Universidade do Estado de Santa Catarina	B.12	Gestão da Informação	M(P)
Universidade Federal do Rio Grande do Norte	B.13	Gestão da Informação e do Conhecimento	M(P)
Fundação Casa de Rui Barbosa	B.14	Memória e Acervos	M(P)
Universidade Federal do Estado do Rio de Janeiro	B.15	Biblioteconomia	M(P)
	B.16	Gestão de Documentos e Arquivos	M(P)

^a P.- Portugal; B.- Brasil. ^b Tipologia: M - mestrado; D - doutoramento; M(P) - mestrado profissional. ^c Curso lecionado em parceria pelas duas instituições.

Tendo-se verificado, nos cursos de Portugal, a continuação da utilização de nomenclaturas como *Ciência da Documentação* ou *Ciências Documentais* e a inexistência, nestas designações, de termos como *Biblioteconomia*, *Arquivística* ou *Gestão da Informação*. No Brasil, estes últimos termos, ou similares, surgem associados a mestrados profissionais, enquanto o termo *Ciência da Informação* é adotado por todos os cursos (de ambos os graus) classificados como académicos. Interessa, ainda, constatar que os dois cursos da Universidade Federal de Rio de Janeiro, *Biblioteconomia* e *Gestão de Documentos e Arquivos*, deixaram de ser incluídos na área do conhecimento da CAPES *Ciência da Informação*, em dezembro de 2015, passando estes cursos a estar integrados, respetivamente, nas novas áreas *Biblioteconomia* e *Arquivologia* (designação usada no Brasil), indo ao encontro da proposta de alteração da Tabela de Áreas do Conhecimento (TAC), efetuada em 2005 (ver tabela 2).

Tabela 2. Relação entre a Ciência da Informação, a Biblioteconomia e a Arquivística nas versões da TAC da CAPES

Ano da versão	Ciência da Informação (CI)	Biblioteconomia e Arquivística
1976	Subárea da <i>Comunicação</i>	Especialidades da área CI
1982	Uma área conjunta: <i>Ciência da Informação, Biblioteconomia e Arquivologia</i>	
1984	Área	Subáreas da área CI
2005 ^a	Três áreas autónomas: (i) <i>Ciência da Informação</i> ; (ii) <i>Biblioteconomia</i> ; (iii) <i>Arquivologia</i>	

^a Proposta preliminar da Comissão Especial de Estudos das entidades: CNPq (Conselho Nacional de Desenvolvimento Científico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e FINEP (Financiadora de Estudos e Projetos).

Nas DM e TD de CI cujo tema é a WS recuperaram-se, respetivamente, 16 e cinco, produzidas no Brasil. Estes 21 trabalhos (ver tabela 3) representam 12% de um total de 173 desenvolvidos neste país (cf. figura 2).

Tabela 3. Identificação dos 21 trabalhos que compõem a amostra do presente estudo

Ref. ^a	Ano ^b	Título do trabalho	Curso ^c
dm.01	2005	Web semântica: Uma análise focada no uso de metadados	B.02
dm.02	2006	Web semântica: aspetos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação	B.02
td.01	2006	Autoria de documentos para a Web semântica: Um ambiente de produção de conhecimento baseado em ontologias	B.10
dm.03	2008	A construção do conhecimento científico: a Web Semântica como objeto de estudo	B.02
dm.04	2008	Inserção da biblioteca digital de teses e dissertações no contexto da Web Semântica	B.12
dm.05	2008	Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspetiva da Ciência da Informação: uma abordagem do <i>marcont initiative</i> na era da Web Semântica	B.02
dm.06	2009	Perspetivas na Web Semântica para a Ciência da Informação	d.01 ^d
dm.07	2010	Web semântica e o governo eletrônico brasileiro: Revisão histórica e teórico-conceitual da representação descritiva e estudo acerca da taxonomia e adoção de metadados na representação da informação	B.04
td.02	2010	Representação iterativa: Um modelo para repositórios digitais	B.02
dm.08	2011	Estudo exploratório sobre compatibilização semântica de vocabulário utilizando ontologia: Fundamentação teórica-metodológica visando ao aperfeiçoamento do MHTX	B.11

dm.09	2011	Registros de marcas da junta comercial do estado de Pernambuco: Um olhar semântico para a organização de um acervo histórico	B.06
dm.10	2012	Artigos científicos digitais na web: Novas experiências para apresentação, acesso e leitura	B.08
dm.11	2012	As dimensões semânticas e pragmática da web e dos mecanismos de busca no ciberespaço	d.02 ^d
dm.12	2012	Organização em sistema caótico: Uso das <i>tags</i> para classificação da informação pelos usuários da rede social <i>twitter</i>	B.09
dm.13	2013	Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária	B.09
dm.14	2013	As tecnologias da Web Semântica no domínio bibliográfico	B.02
dm.15	2013	Representação da informação dinâmica em ambientes digitais	B.01
td.03	2013	Web Semântica e repositórios digitais educacionais na área de saúde: Uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca	B.10
dm.16	2014	Wikis semânticos: Da web para a Web Semântica	B.02
td.04	2015	Interoperabilidade e mapeamentos entre sistemas de organização do conhecimento na busca e recuperação de informações em saúde: Estudo de caso em ortopedia e traumatologia	B.01
td.05	2015	Representação e recuperação de informação em acervos digitais nos contextos da Web Semântica e web pragmática: Um estudo crítico	B.01

^a dm: dissertação de mestrado; td tese de doutoramento. ^b ano de defesa do trabalho. ^c referência que consta na Tabela 1. ^d cursos que em 2016 se encontravam desativados: d.01- Ciência da Informação (Pontifícia Universidade Católica de Campinas); d.02- Gestão da Informação (Universidade Estadual de Londrina).

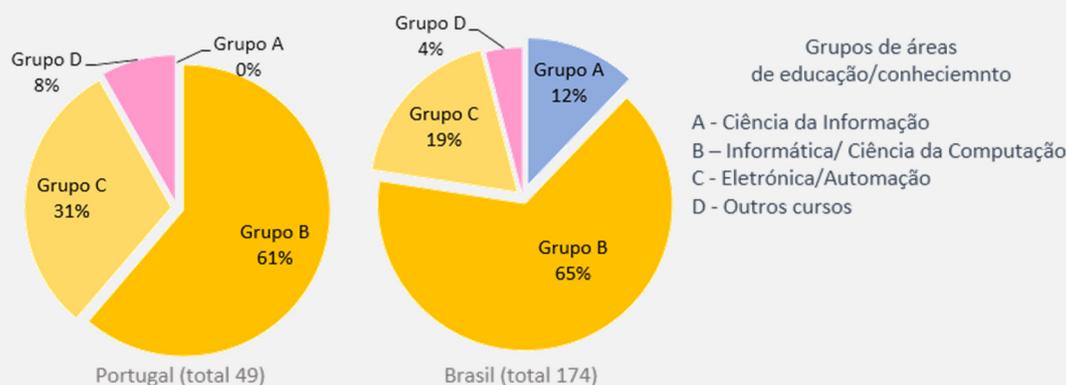


Figura 2. Distribuição percentual dos trabalhos recuperados (DM e TD), de acordo com o país de origem, pelos grupos de áreas de educação/conhecimento

Em Portugal, a inexistência de trabalhos cujo tema é a WS, produzidos em cursos considerados de CI, contrasta com os quatro (8% de 49) provenientes das áreas: Ciências da Educação; Ciências Dentárias; Gestão e Administração; e Tecnologia dos Processos Químicos, (incluídas no grupo D).

Identificaram-se três pontos de confluência entre a CI e a WS: (i) contribuição mútua entre as duas áreas (62%), (ii) a WS enquanto tema interdisciplinar entre a CI e a I/CC (29%) e (iii) a utilização conjunta de conhecimentos das duas áreas (9%). Na distribuição temporal destes pontos pelos elementos do resumo onde foram encontrados (ver figura 3) destaca-se o primeiro ponto de confluência pela presença ao longo de todo período analisado. Este ponto apresenta 54% das suas ocorrências no elemento *objetivos* (entre 2006 e 2013) e é o único a ser referido no elemento *introdução* (em 2014). Relativamente ao terceiro ponto, verifica-se que as duas únicas ocorrências surgem no último ano, 2015, ambas no elemento *metodologia* de teses de doutoramento.

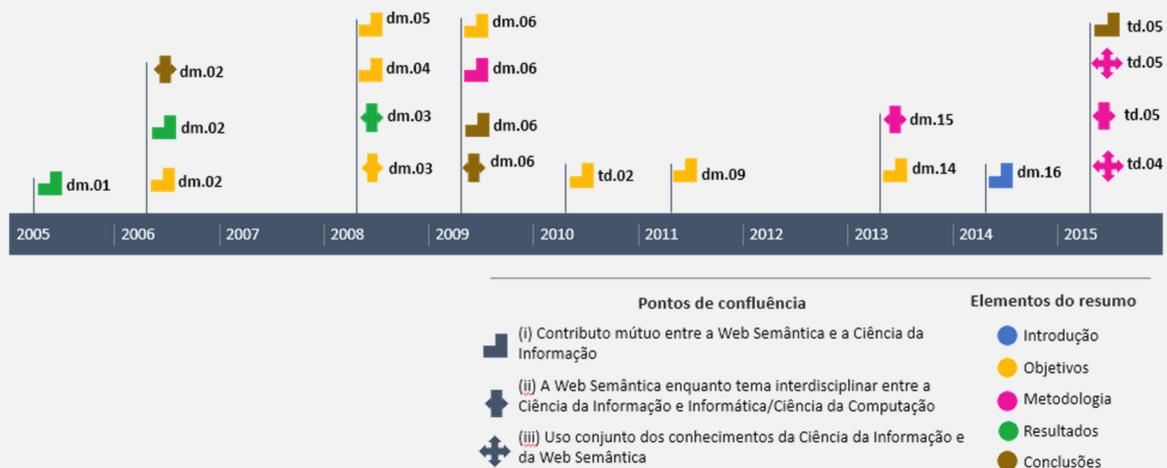


Figura 3. Distribuição temporal dos pontos de confluência pelos elementos do resumo onde foram identificados

Relativamente ao perfil académico dos 17 orientadores e dois coorientadores dos trabalhos considerados neste estudo, apurou-se que: dois (10%) apresentam formação só em CI; 11 (58%) em CI e outras áreas; e seis (32%) só em outras áreas. A formação na área da I/CC surge em sete (35%) docentes, sendo que destes apenas um acumula essa formação com a da área de CI. Nos 21 autores, contabilizaram-se: 10 (48%) que apresentam formação só em CI e 11 (52%) em CI e outras áreas. A formação na área da I/CC surge em nove (43%) autores, todas obtidas em graus anteriores ao mestrado.

Em termos de correlação, verificou-se a existência de uma bastante forte (expressa no coeficiente 0,781) entre as duas variáveis analisadas nos trabalhos: o grau de confluência encontrado nos resumos e o grau de proximidade determinado nos objetivos, ao contrário da correlação muito residual (coeficiente 0,145) entre a diferença quantitativa das formações em I/CC relativamente às de CI e o grau de confluência/proximidade da relação WS-CI (ver figura 4).

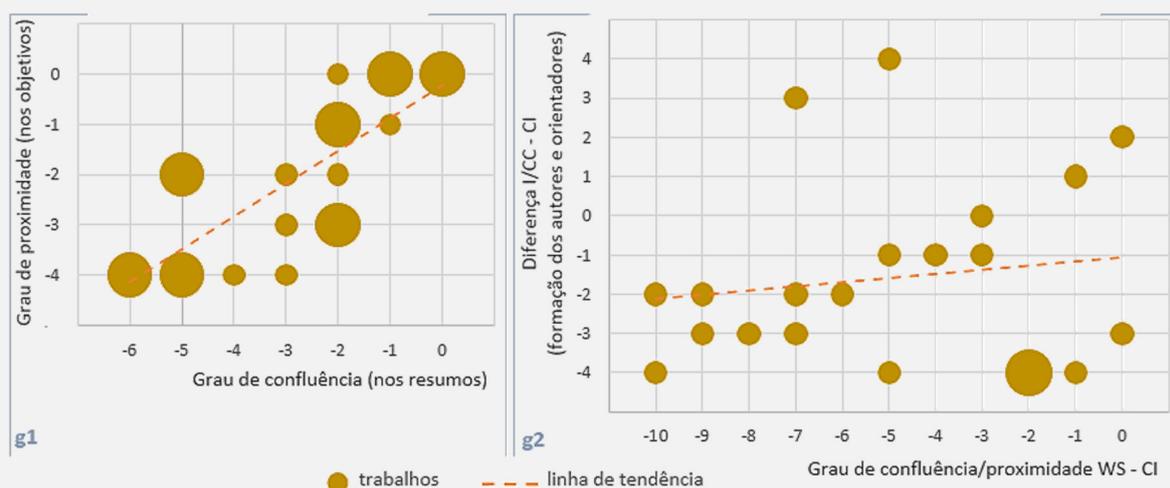


Figura 4. Gráficos de correlação entre pontos de confluência WS-CI e centralidade temática da abordagem à WS (g1) e entre relação WS-CI e diferença quantitativa das formações I/CC-CI (g2)

A confrontação efetuada entre a formação dos responsáveis pelos trabalhos e o grau de proximidade aferido, da relação WS-CI, permitiu também verificar uma correlação ainda mais afastada (coeficiente 0,092) se a ponderação for o equilíbrio entre os pesos de ambas as formações em vez da aproximação à I/CC.

CONCLUSÕES

Relativamente às fontes de obtenção de dados (RCAAP e OASISbr), conclui-se existir discrepâncias entre a informação que consta nos campos *assuntos* e *data* destes e os campos correspondentes nos repositórios das Instituições de Ensino Superior (RIES) de origem dos documentos. Discrepâncias com impacto nos trabalhos recuperados que, infere-se, resultam da existência, em vários RIES, de mais de um campo com informação passível de ser considerada “assunto” assim como com vários campos com diferentes datas registadas.

Conclui-se que em Portugal o panorama apresenta uma mudança de atitude no que concerne aos cursos da área de CI. A posição inicial pode ser descrita como uma associação entre a visão de Otlet (1934) e a posição de López Yepes (1995), i.e., uma visão da Documentação enquanto área que abarca as outras áreas relacionadas “elevando” o carácter científico desse conjunto às “Ciências Documentais”. A posição de chegada, pode ser apresentada como uma visão holística da área, materializando-se na CI enquanto entidade que incorpora aqueles saberes e outros, num todo maior que a soma das suas partes. No Brasil, conclui-se que existiu um percurso relativamente fluído da Documentação para a CI. Quanto ao relacionamento desta com a Biblioteconomia e a Arquivística, observa-se, o que se infere ser, uma procura de identidade da CI enquanto um todo e, simultaneamente, uma tentativa de afirmação por parte daquelas áreas disciplinares específicas.

No que respeita à distribuição da amostra pelos dois países, conclui-se existir um desfasamento maior que o esperado entre ambos, dado que todos os trabalhos recuperados considerados de CI, cujo tema é a WS, foram produzidos no Brasil. Assim, infere-se a existência de uma relação entre os contextos de origem da CI, europeu, para Portugal, e americano, no caso do Brasil, e a diferente ligação à I/CC, área dominante no que diz respeito ao tratamento do assunto em questão.

Relativamente aos três pontos de confluência entre as duas áreas (CI e WS) identificados (i: contribuição mútua; ii: a WS enquanto tema interdisciplinar; iii: utilização conjunta de conhecimentos), conclui-se, atendendo à relevância percentual dos mesmos, sua distribuição ao longo do período analisado e o elemento do resumo onde é referido, que a relação entre a CI e a WS apresenta-se como um campo, essencialmente, de investigação teórica. Dentro da amostra analisada, apenas em 2014 a contribuição recíproca entre as duas áreas é tomada como premissa deixando de ser uma hipótese a investigar e, em 2015, surgem trabalhos que se propõem fazer uso efetivo dos conhecimentos das duas áreas.

Em termos do perfil académico dos autores e orientadores dos trabalhos analisados, conclui-se que a amostra (21) é pouco representativa do carácter interdisciplinar da WS enquanto tema trabalhado pelas áreas da CI-I/CC, atendendo a que em apenas 11 trabalhos (52%) os investigadores responsáveis apresentam, no conjunto, formação nestas duas áreas (10% se apenas for atendido a formação dos orientadores) e a correlação quase inexistente desta com uma maior confluência/proximidade entre a WS e a CI.

Assim, da análise da amostra, conclui-se que, embora se possa entender a WS como um campo de estudo de direito próprio da CI, este configura-se, essencialmente, de investigação teórica sendo o trabalho prático interdisciplinar, CI-I/CC, incipiente e pouco representativo.

Considera-se ter contribuído com: uma clarificação de conceitos relativos à WS, indicadores sobre o seu estudo, na área CI, e relativos à relação CI-I/CC; subsídios para uma reflexão sobre as relações

disciplinares da CI no contexto dos cursos de ensino superior; alguns indicadores relacionados com os repositórios usados, nomeadamente quanto à ausência ou duplicação de registos. Questões que se julga merecedoras de futuros aprofundamentos pelas suas implicações tanto teóricas como práticas.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bardin, L. (2011). *Análise de conteúdo*. (L. A. R. A. Pinheiro, Trans.). São Paulo: Almedina.
- Barreto, A. de A. (2008). Uma quase história da ciência da informação. *DataGramaZero*, 9(2), 1–18. Retrieved from http://eprints.rclis.org/17637/1/DataGramaZero - Revista de Ciência da Informação - Artigo 01_Aldo.pdf
- Berners-Lee, T. (1998, October 14). A roadmap to the Semantic Web. Retrieved March 21, 2016, from <https://www.w3.org/DesignIssues/Semantic>
- Berners-Lee, T. (2006). The World Wide Web: Past, present and future (FAQ). Retrieved March 16, 2016, from <http://www.w3.org/People/Berners-Lee/FAQ.html>
- Berners-Lee, T. (2009). Web 2.0 Summit 09 discussion: A Conversation with Tim Berners-Lee interview by Tim O'Reilly [video]. San Francisco: O'Reilly Media. Retrieved from <https://youtu.be/KY5skobffk0>
- Berners-Lee, T., & Fischetti, M. (2000). *Weaving the Web*. HarperCollins.
- Berners-Lee, T., & Hendler, J. A. (2001). Publishing on the Semantic Web. *Nature*, 410(April), 1023–1024. Retrieved from <https://github.com/AKSW/semann/blob/master/pdf/T.B. Lee - Publishing on the semantic web.pdf>
- Berners-Lee, T., Hendler, J. A., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 35–43. Retrieved from http://ldc.usb.ve/~yudith/docencia/UCV/ScientificAmerican_FeatureArticle_TheSemanticWeb_May2001.pdf
- Bikakis, N., Tsinaraki, C., Gioldasis, N., Stavrakantonakis, I., & Christodoulakis, S. (2013). The XML and Semantic Web Worlds: Technologies, Interoperability and Integration. A survey of the state of the art. In I. E. Anagnostopoulos, M. Bieliková, P. Mylonas, & N. Tsapatsoulis (Eds.), *Semantic Hyper/Multimedia Adaptation* (pp. 319–360). Berlin: Springer. https://doi.org/10.1007/978-3-642-28977-4_12
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data - The story so far. University of Southampton, ePrints Soton. Retrieved from <http://eprints.soton.ac.uk/271285/>
- Borges, M. M., & Casado, E. S. (Eds.). (2009). *A Ciência da Informação criadora de conhecimento* (Vol. I). Coimbra: Imprensa da Universidade de Coimbra. <https://doi.org/http://dx.doi.org/10.14195/978-989-26-0319-3>
- Borko, H. (1968). Information Science: What Is It? *American Documentation*, 19(1), 3–5. Retrieved from <http://cdigital.uv.mx/bitstream/123456789/6699/2/Borko.pdf>

Cambridge Semantics. (2016). Many Names for the Semantic Web. Retrieved May 15, 2017, from <http://www.cambridgesemantics.com/semantic-university/many-names-semantic-web>

Caro-castro, C. (2010). Las clasificaciones bibliográficas: De los estantes a la web. *Tabula*, 13, 11–23. Retrieved from <http://hdl.handle.net/10366/121954>

Galvão, M. C. B. (1998). Construção de conceitos no campo da ciência da informação. *Ciência Da Informação*, 27(1), 46–52. <https://doi.org/10.1590/S0100-19651998000100006>

Hendler, J. A. (2011). “Why the Semantic Web will never work” (note the quotes) [Video]. Retrieved from <https://youtu.be/oKiXpO2rbJM>

Hendler, J. A. (2014). The Semantic Web: The Inside Story [video]. In *Summer School in cognitive Science: Web Science and the Mind*. Montréal: Institut des sciences cognitives, UQAM. Retrieved from <https://youtu.be/3Ap5FsxvjTQ>

Hendler, J. A., Berners-Lee, T., & Miller, E. (2002). Integrating Applications on the Semantic Web. *Journal of the Institute of Electrical Engineers of Japan*, 122(10), 676–680. Retrieved from <http://www.w3.org/2002/07/swint>

Heuvel, C. Van Den, & Rayward, W. B. (2011). Facing interfaces: Paul Otlet’s visualizations of data integration. *International Review of Research in Open and Distance Learning*, 62(12), 2313–2326. <https://doi.org/http://dx.doi.org/10.1002/asi.21607>

Laaff, M. (2011, July 22). Internet visionary Paul Otlet: Networked knowledge, decades before Google. (E. Ornstein, Trans.), *SPIEGEL ONLINE - International*. Retrieved from <http://www.spiegel.de/international/world/internet-visionary-paul-otlet-networked-knowledge-decades-before-google-a-775951.html>

Le Coadic, Y.-F. (1996). *A ciência da informação*. (M. Y. F. Gomes, Trans.). Brasília: Briquet: de Lemos/Livros.

López Yepes, J. (1995). *La Documentación como disciplina: Teoría e historia* (2nd ed.). Pamplona: Ediciones Universidade de Navarra.

Mikhailov, A. I., & Giljarevskij, R. S. (1971). An Introductory Course on Informatics/Documentation. *International Federation for Documentation*. Retrieved from <http://eric.ed.gov/?id=ED060875>

Morin, E. (2007). *Educação e complexidade: Os Sete Saberes e outros ensaios*. (M. da C. de Almeida & E. de A. Carvalho, Eds., E. de A. Carvalho, Trans.) (4th ed.). São Paulo, Brasil: Cortez Editora.

Otlet, P. (1934). *Traité de documentation: Le livre sur le livre: Théorie et pratique*. Bruxelles: Ediciones Mundaneum Palais Mondial. Retrieved from http://lib.ugent.be/fulltxt/handle/1854/5612/Traite_de_documentation_ocr.pdf

Parreiras, F. silva. (2012). *Semantic Web and Model-driven Engineering*. Piscataway: John Wiley & Sons, Inc.

Patel-Schneider, P. F., & Horrocks, I. (2007). A comparison of two modelling paradigms in the Semantic Web. *Web Semantics*, 5(4). <https://doi.org/http://dx.doi.org/10.1016/j.websem.2007.09.004>

Pinheiro, L. V. R. (1997). *A Ciência da Informação entre sombra e luz: Domínio epistemológico e campo interdisciplinar*. Universidade Federal do Rio de Janeiro. Retrieved from <http://ridi.ibict.br/handle/123456789/35>

Quivy, R., & Campenhoudt, L. Van. (1992). *Manual de investigação em ciências sociais*. (J. M. Marques, M. A. Mendes, & M. Carvalho, Trans.) (6th ed.). Lisboa: Gradiva.

Robredo, J. (2010). Ciência da informação e Web semântica: Linhas convergentes ou linhas paralelas? In *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento* (pp. 12–47). Brasília: IBICT. Retrieved from <http://repositorio.unb.br/handle/10482/7952>

Saracevic, T. (1996). Ciência da informação: origem, evolução e relações. *Perspectivas Em Ciência Da Informação*, 1(1), 41–62. Retrieved from <http://www.brapci.inf.br/article.php?dd0=0000003224&dd90=bc40e71f02>

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. *IEEE Intelligent Systems*, 21(3), 96–101. <https://doi.org/10.1109/MIS.2006.62>

Silva, J. L. C., & Freire, G. H. de A. (2012). Um olhar sobre a origem da ciência da informação: indícios embrionários para sua caracterização identitária. *Encontros Bibli*, 17(33), 1–29. <https://doi.org/10.5007/1518-2924.2012v17n33p1>

Silva, A. M. da, & Ribeiro, F. (2008). *Das “ciências” documentais à ciência da informação: Ensaio epistemológico para um novo modelo curricular* (2nd ed.). Porto: Edições Afrontamento.

Sletten, B. (2014, January 13). Keep On Keeping On. Dataversity, Smart Data. Retrieved from <http://www.dataversity.net/keep-on-keepin-on/>

Souza, R. R., & Almeida, M. B. (2009). Representação do conhecimento: identidade ou esvaziamento da Ciência da Informação? In *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 157–165). Coimbra: Imprensa da Universidade de Coimbra.

Souza, R. R., & Alvarenga, L. (2004). A Web Semântica e suas contribuições para a ciência da informação. *Ciência Da Informação*, 33(1), 132–141. Retrieved from <http://ref.scielo.org/mp2fh8>

Wersig, G., & Neveling, U. (1975). The phenomena of interest to Information Science. *Information Scientist*, 9(4). Retrieved from <http://sigir.org/files/museum/pub-13/18.pdf>

World Wide Web Consortium. (2013, December 11). W3C Semantic Web Activity Homepage. Retrieved January 27, 2016, from <https://www.w3.org/2001/sw/>

World Wide Web Consortium. (2015). W3C Standards: Semantic Web. Retrieved March 21, 2016, from <https://www.w3.org/standards/semanticweb/>

World Wide Web Consortium. (2016). W3C Data Activity: Building the Web of Data. Retrieved March 21, 2016, from <http://www.w3.org/2013/data/>

Zins, C. (2007a). Classification Schemes of Information Science: Twenty-Eight Scholars Map the Field. *Journal of the American Society for Information Science and Technology*, 58(4), 645–672. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007b). Conceptions of Information Science. *Journal of the American Society for Information Science and Technology*, 58(4), 335–350. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007c). Conceptual Approaches for Defining Data, Information, and Knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), 479–493. <https://doi.org/10.1002/asi.20508>

Zins, C. (2007d). Knowledge Map of Information Science. *Journal of the American Society for Information Science and Technology*, 58(4), 526–535. <https://doi.org/10.1002/asi.20505>