



Ceci n'est pas une pipe.

Magnin

Luís Miguel Oliveira Machado

A presença da Web Semântica no *output* dos cursos de
Mestrado/Doutoramento em Ciência da Informação
(Portugal e Brasil, 2005-2015)

Dissertação de Mestrado em Ciência da Informação, orientada pela Doutora Maria da Graça Simões e
coorientada pelo Doutor Renato Rocha Souza, apresentada ao Departamento de Filosofia, Comunicação e
Informação da Faculdade de Letras da Universidade de Coimbra

2016



UNIVERSIDADE DE COIMBRA

Imagem da capa: Rene Magritte. 1929. *La Trahison des images (Ceci n'est pas une pipe)*. 63,5×93,98cm. Óleo s/tela. Los Angeles County Museum of Art [fonte: <http://www.wikiart.org/en/rene-magritte/the-treachery-of-images-this-is-not-a-pipe-1948>]

Faculdade de Letras

**A presença da Web Semântica no output dos cursos de
Mestrado/Doutoramento em Ciência da Informação
(Portugal e Brasil, 2005-2015)**

Ficha Técnica:

Tipo de trabalho	Dissertação de Mestrado
Título	A PRESENÇA DA WEB SEMÂNTICA NO <i>OUTPUT</i> DOS CURSOS DE Mestrado/DOCTORAMENTO EM CIÊNCIA DA INFORMAÇÃO (PORTUGAL E BRASIL, 2005-2015)
Autor/a	Luís Miguel Oliveira Machado
Orientador/a	Maria da Graça Simões
Coorientador/a	Renato Rocha Souza
Júri	Doutora Cristina Freitas (Presidente) Doutor Maurício Barcellos Almeida (1º vogal) Doutora Maria da Graça Simões (2º vogal) Doutor Renato Rocha Souza (3º vogal)
Identificação do Curso	2º Ciclo em Ciência da Informação
Área científica	Ciência da Informação
Data	20-10-2016
Classificação	19 valores



A consciência surgiu para que possamos conhecer a vida.

António Damásio (2004)

Agradecimentos	VII
RESUMO	IX
ABSTRACT	X
Lista de Siglas e Abreviaturas	XI
INTRODUÇÃO	1
PARTE I: CONTEXTUALIZAÇÃO TEÓRICA	9
1. O CONCEITO DE WEB SEMÂNTICA	10
1.1. Metodologia considerada na análise conceitual	11
1.2. Web Semântica: análise conceitual	13
1.2.1. Dimensão de análise: características descritivas	13
1.2.2. Dimensão de análise: base estrutural	16
1.2.3. Dimensão de análise: funções operativas	18
1.3. Vários termos para um conceito?	21
1.3.1. <i>Semantic Web, Web of Data, Linked Data</i> : uma possível definição	21
1.3.2. Da Web <i>com</i> dados ligados à Web de Dados	23
2. O CONCEITO DE CIÊNCIA DA INFORMAÇÃO	29
2.1. Metodologia considerada na análise concetual	29
2.1.1. Uma abordagem em três momentos	29
2.1.2. Seleção de obras	30
2.2. Ciência da Informação: análise concetual	34
2.2.1. Dado, Informação e Conhecimento	34
2.2.2. Domínio e âmbito da Ciência da Informação	36
2.2.3. Mapa do conhecimento da Ciência da Informação	38
2.3. Relações (inter)disciplinares	42
2.3.1. Da <i>Pré-CI</i> à <i>CI inicial</i>	42
2.3.2. Da <i>CI inicial</i> à <i>contemporânea</i>	45
2.3.3. Na passagem de século	48
2.4. Cursos de Ciência da Informação	53
2.4.1. Em Portugal	54
2.4.2. No Brasil	57

PARTE II: ESTUDO DE CASO: A PRESENÇA DA WEB SEMÂNTICA NO <i>OUTPUT</i> DOS CURSOS DE MESTRADO E DOUTORAMENTO EM PORTUGAL E NO BRASIL EM CIÊNCIA DA INFORMAÇÃO (2005 – 2015)	63
3. OBJETIVOS DO ESTUDO DE CASO	64
4. METODOLOGIA DO ESTUDO DE CASO	65
4.1. Fontes de obtenção de dados	66
4.2. Critérios usados na recolha de dados	68
4.3. Universo e amostra	72
4.3.1. Universo	72
4.3.2. Amostra	73
4.4. Critérios usados na análise dos dados	75
4.4.1. A presença da WS enquanto tema dos trabalhos recuperados	76
4.4.2. A presença da WS nos trabalhos dos cursos de CI recuperados	76
4.4.3. O perfil académico dos orientadores e coorientadores	78
5. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS	78
5.1. A presença da Web Semântica enquanto tema dos trabalhos recuperados	79
5.2. A presença da Web Semântica nos trabalhos dos cursos de Ciência da Informação	81
5.2.1. Análise dos resumos dos trabalhos em Ciência da Informação	81
5.2.2. Análise dos objetivos dos trabalhos em Ciência da Informação	87
5.2.3. Confrontação da análise dos resumos com a dos objetivos	88
5.3. O perfil académico dos orientadores e coorientadores dos trabalhos considerados neste estudo	91
CONCLUSÕES	95
Referências Bibliográficas	106
Índice de Figuras	111
Índice de Tabelas	113
Lista de apêndices	114

Agradecimentos

Não interessa chegar à meta, qualquer meta, se a não podermos partilhar com alguém. É com esse sentimento que os meus agradecimentos vão, em primeiro lugar, aos meus *chus* (Beatriz e Joana). Um agradecimento mais que devido pois um trabalho desta natureza retira-nos tempo e atenção.

Um agradecimento, também especial, à Doutora Graça Simões pela sua presença e investimento ao longo deste trabalho, que me fez sentir verdadeiramente orientado e ao Doutor Renato Souza, também pela sua disponibilidade no acompanhamento da orientação. A ambos pelas proficuas conversas, essenciais para (re)estruturar e consolidar o que se sabe e, muito importante, tomar consciência do que não se sabe!

A todos os Professores do Mestrado de Ciência da Informação com quem tive o privilégio de contactar, em especial à Doutora Maria Manuel Borges e à Doutora Cristina Freitas pelo seu voto de confiança.

Ao Professor Alexandre Pinto do Departamento de Engenharia Informática, pelas aulas sobre Web Semântica.

Aos investigadores Luciana Gracioso, Maurício Almeida e Fernanda Ribeiro pela disponibilização de respetivos artigos.

A todos os que se disponibilizaram a fornecer informações necessárias à realização deste trabalho, como: Simone Torres, Blanca Rodríguez Bravo, Pedro Pêgo (Departamento de Economia Gestão, Engenharia Industrial e Turismo da Universidade de Aveiro), Raquel Truta (RCAAP) e equipa do OASISbr.

Last but not least, a todos os amigos cuja presença em momentos oportunos foram essenciais: J. Carecho, A. Galvão, M. Inês, Maria José M., Marisa M., Edgar R., entre outros...

O meu sincero agradecimento.

RESUMO

Introdução: A *Web* é o local privilegiado da procura de informação. As potencialidades de uma *web semântica* para recuperar informação eficazmente afigura-se essencial. No âmbito da Ciência da Informação (CI), esta ideia surge com Otlet ao desenvolver uma “web semântica analógica”, justificação para a ligação entre a Web Semântica (WS) e as fundações teóricas da CI. **Objetivos:** Apresenta-se como objetivo geral, verificar e refletir sobre a presença da WS no *output* dos cursos de mestrado/doutoramento em CI, Portugal e Brasil (2005/2015); objetivos específicos, identificar: i) cursos em funcionamento no ano (2014/2015); ii) número de dissertações de mestrado (DM) e teses de doutoramento (TD) em CI associado à WS (2005/2015); iii) pontos de confluência entre a CI e a WS nestes trabalhos; e iv) caracterizar o perfil académico dos orientadores dos mesmos. **Metodologia:** Partiu-se de uma metodologia qualitativa: revisão bibliográfica (contextualização teórica dos objetos de estudo: WS e CI) e estudo de caso, centrado na análise de conteúdo dos trabalhos recuperados (RCAAP e OASISbr). **Resultados:** Identificaram-se 4 cursos de mestrado e 3 de doutoramento em Portugal e, no Brasil, 16 de mestrado e 9 de doutoramento; recuperaram-se 0 DM/TD (Portugal) e 17 DM, 5 TD (Brasil); pontos de confluência identificados: contribuição mútua (62%), WS enquanto tema interdisciplinar (29%) e complementaridade de conhecimentos (9%); dos 20 orientadores: 2 (10%) apresentam formação só em CI, 12 (60%) em CI e outras áreas e 6 (30%) só em outras áreas. **Conclusões:** o estudo não reflete o carácter interdisciplinar da WS enquanto trabalho conjunto da CI/Informática. Reforça a ideia que a WS se configura como um campo de direito próprio da CI. A WS é um constructo que incluiu a *Web of Data* e a *Web of Linked Data*; a CI define-se como o estudo da mediação dos fenómenos: dado, informação, conhecimento e respetiva implementação no domínio cultural, definição que não é consensual, especialistas da área apresentam outras conceções baseadas nas relações disciplinares desta, particularmente com a Informática no que respeita à *organização do conhecimento*.

Palavras-chave: Web Semântica; Ciência da Informação; cursos de Mestrado/Doutoramento; repositórios; Portugal; Brasil.

ABSTRACT

Introduction: The Web is the privileged place for seeking information; thus, the idea of structuring a semantic web to enhance and leverage the retrieval of relevant information is essential. This work was developed out of the original idea presented by Paul Otlet of what could be considered an analogy of the semantic web, justifying the connection between the Semantic Web (SW) and the theoretical foundations of the Information Science (IS). **Objectives:** We aimed at examining and analysing the presence of the SW in the intellectual production derived from the IS master and PhD courses, in Portugal and Brazil, in the period from 2005 to 2015. Specifically, we aimed at identify i) courses that are active nowadays; ii) the number of master's dissertations (MD) and PhD thesis (PD) on IS associated with the SW in the period; iii) identify the points of confluence between the IS and the SW in these works; and (iv) characterize the academic profile of these works' advisors. **Methodology:** As a methodology, we performed a literature review a theoretical contextualization of the objects of research: SW and IS - and a case study focused on content analysis of the retrieved works in the repositories (RCAAP and OASISbr). **Results:** As results, we have identified 4 Master's and 3 PhD courses in Portugal and 16 masters and 9 PhD courses in Brazil. We have retrieved no MD/PD in Portugal, 17 MD and 5 PD in Brazil. We have identified confluence points as the mutual contribution between the areas (62%), WS as an interdisciplinary theme (29%) and complementary knowledge (9%). Of the 20 advisors, 2 (10%) have training only in IS, 12 (60%) in IS and other areas, and 6 (30%) only in other areas. **Conclusions:** We concluded that the study does not reflect the interdisciplinary nature of SW as a joint work of IS / Computer Science (CS). It reinforces the idea that the SW is a field, in its own right, independent of the IS. SW is a construct that includes the Web of Data and the Web of Linked Data; and the IS is defined as the study of the mediation of the phenomena: data, information, knowledge and its consequences in the cultural field, a definition that is not consensual, and coexist with other conceptions mostly based in disciplinary relations of IS, particularly with CS in the knowledge organization domain.

Keywords: Semantic Web; Information Science; Master's / PhD courses; repositories; Portugal; Brazil.

Lista de Siglas e Abreviaturas

A3ES	Agência de Avaliação e Acreditação do Ensino (Portugal)
apBAD	Associação Portuguesa de Bibliotecários Arquivistas e Documentalistas
AB/AC	Áreas Básica ou Área do Conhecimento
AEF	Áreas de Formação e Educação
BAD	Biblioteconomia, arquivo e documentação (área de educação e formação)
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Brasil)
CC	Ciência da Computação
CI	Ciência da Informação
CNAEF	Classificação Nacional das Áreas de Educação e Formação
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brasil)
DGEEC	Direção-Geral de Estatísticas da Educação e Ciência (Portugal)
DGES	Direção Geral do Ensino Superior (Portugal)
DM	Dissertações de Mestrado
FINEP	Financiadora de Estudos e Projetos (Brasil)
IBCT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IES	Instituição de Ensino Superior
IOMRC	Introdução, Objetivos, Metodologia, Resultados e Conclusões (elementos constituintes da estrutura interna dos resumos)
OASISbr	Portal Brasileiro de Acesso Aberto à Informação Científica
PPG	Programa de Pós-Graduação
Plesp	Processo de inserção esperado (de documentos nos Repositórios)
PResp	Processo de recuperação esperado (de documentos dos Repositórios)
RCAAP	Repositório Científico de Acesso Aberto de Portugal
RI	Recuperação de informação
rIES	Repositório de Instituição de Ensino Superior
TAC	Tabela de Áreas do Conhecimento
TD	Teses de Doutoramento
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura
WWW	World Wide Web
W3C	World Wide Web Consortium

ESPECIFICAÇÕES/PADRÕES E OUTRAS SIGLAS INFORMÁTICAS

DTD	Document Type Definitions
GRDDL	Gleaning Resource Descriptions from Dialects of Languages
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
R2RML	RDB to RDF Markup Language
RDB	Relational DataBase
RDF	Resource Description Framework
RDFa	Resource Description Framework in attributes
RDF-S	Resource Description Framework Schema
IRI	International Resource Identifier
RIF	Rule Interchange Format
SPARQL	SPARQL Protocol and RDF Query Language
SQL	Structural Query Language
OWL	Web Ontology Language
URI	Uniform Resource Identifier
URL	Uniform Resource Location
URN	Uniform Resource Name
XML	eXtensible Markup Language

INSTITUIÇÕES DE ENSINO SUPERIOR

FCRB	Fundação Casa de Rui Barbosa
FEC-FP	Fundação Ensino e Cultura "Fernando Pessoa"
UAlg	Universidade do Algarve
UA	Universidade de Aveiro
UnB	Universidade de Brasília
UC	Universidade de Coimbra
UDESC	Universidade do Estado de Santa Catarina
UNESP/MAR	Universidade Estadual Paulista Júlio de Mesquita
UNIRIO	Universidade Federal do Estado do Rio de Janeiro
UFBA	Universidade Federal da Bahia
UFF	Universidade Federal Fluminense

UFG	Universidade Federal de Goiás
UFMG	Universidade Federal de Minas Gerais
UFPB/JP	Universidade Federal da Paraíba/João Pessoa
UFPE	Universidade Federal de Pernambuco
UFRN	Universidade Federal do Rio Grande do Norte
UFRJ	Universidade Federal do Rio de Janeiro
UFSC	Universidade Federal de Santa Catarina
UL	Universidade de Lisboa
UNICAMP	Universidade Estadual de Campinas
USP	Universidade de São Paulo
UP	Universidade do Porto

INTRODUÇÃO

Contextualização e problemática

Nas últimas décadas tem-se assistido a um aumento exponencial de informação produzida no espaço da *World Wide Web* (WWW), espaço também denominado simplesmente de *Web*. Frequentemente o termo *internet* também é utilizado como sinónimo de *Web*, mas tal não será o mais correto, nas palavras do próprio criador da WWW, Tim Berners-Lee (2010, p. 83): “The Web is an application that runs on the Internet, which is an electronic network that transmits packets of information among millions of computers according to a few open protocols.”¹

Nesta transição informacional, do analógico para o digital, uma carência começou a evidenciar-se: a necessidade de uma recuperação mais eficaz dessa informação, tanto a nível da precisão como da revocação (Caro-castro, 2010, p. 35; Patel-Schneider & Horrocks, 2007; R. R. Souza & Alvarenga, 2004, p. 133). Neste contexto espera-se uma evolução das ferramentas disponíveis no ambiente digital para uma gestão mais eficaz dos recursos que aí se encontram. Uma “nova” *Web*, denominada de *Web Semântica*, apresenta-se como a solução para essa situação. Esta *Web Semântica* (WS) sendo, ainda, novidade para alguns, encontra-se já em fase “adolescente”, não só no que respeita à idade (cerca de 17 anos)² como também, com a devida distância, no que concerne ao aumento da capacidade de processamento de informação característico desta fase de desenvolvimento humano³.

Muito se escreveu sobre esta “criança prodígio”, muitas esperanças e promessas nela foram colocadas, nomeadamente, como sendo capaz de criar um ambiente onde *agentes*, um tipo específico de software “inteligente”, processando e partilhando “conhecimento” nessa *Web*, conseguem realizar tarefas complexas de acordo com as necessidades dos utilizadores (Berners-Lee, 2006; Berners-Lee & Fischetti, 2000, p. 237; Berners-Lee, Hendler, & Lassila, 2001). A grande 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 (IA). Uma associação que os criadores da *Web Semântica*, logo nas publicações iniciais, procuraram afastar:

The concept of machine-understandable documents does not imply some magical artificial intelligence allowing machines to comprehend human mumblings. It relies solely on a machine’s ability to solve well-defined problems by performing well-defined operations on well-defined data. (Berners-Lee & Hendler, 2001, p. 1023)

¹ A separação entre os dois conceitos, *Internet* e *Web*, principalmente no que diz respeito a questões de propriedade torna-se necessária para a manutenção do “princípio da universalidade” e dos padrões abertos que caracterizam a WWW e impulsionam o desenvolvimento de novos serviços (Berners-Lee, 2010).

² Considera-se o ano 1999 como marco do “nascimento” da *Web Semântica* de acordo com os documentos mais antigos considerados na Tabela 1, em particular devido às referências explícitas ao conceito da WS no livro de Tim Berners-Lee *Weaving the Web*.

³ De acordo com Sprinthall e Collins (1999, p.118): “os adolescentes funcionam como se tivessem uma capacidade mental superior às crianças”, apresentando alterações qualitativas significativa no processamento de informação. (*Psicologia do adolescente*, 2.a ed., Lisboa: Fundação Calouste Gulbenkian).

Todavia essa associação não é necessariamente descabida, devido ao conceito de ontologia utilizado por ambas, IA e WS, com a mesma finalidade: como uma estrutura de conceitos representada por um vocabulário lógico que, utilizando meios computacionais, possibilita inferências automáticas (Almeida, 2013, p. 1687).

Sendo estas inferências automáticas a meta, para se chegar a este ponto é necessário percorrer um caminho que passa pela representação e, ainda antes, pela conceptualização da estrutura de onde as referidas inferências serão derivadas. Neste ponto, a abordagem da WS, segundo Shirky, difere da IA no sentido em que:

No lugar de tentar fazer com que as máquinas pensem como pessoas, podemos descrever o mundo de um modo em que as máquinas serão capazes de processar, mudando da tentativa de construir cérebros em silicone para o processo de modelar a informação do "mundo real".⁴ (Ray, 2010, pt. 8:00)

A ideia de disponibilizar a informação sobre o mundo ou, por outras palavras, o conhecimento humano, potenciando o trabalho colaborativo, não é original. Embora esta ideia esteja na base do desenvolvimento da WWW (Berners-Lee, 2009, pt. 2:40), uma outra “web semântica” de madeira e cartão (Laaff, 2011) foi desenvolvida por Paul Otlet meio século antes da *Web* digital de Berners-Lee:

Otlet’s idea of a Universal Network of Documentation as an infrastructure for “the registration of ideas” in all sorts of documentary formats comes close to the definition of the Semantic Web by the World Wide Web Consortium (W3C) as a “Web of data.” (Heuvel & Rayward, 2011, pp. 9–10)

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) (A. M. da Silva & Ribeiro, 2008, p. 47). Nesse sentido, existe uma ligação entre o propósito da Web Semântica e as fundações teóricas da Ciência da Informação (R. R. Souza & Alvarenga, 2004, p. 140). Além desta ligação, há autores que, ao analisar as práticas e processos, assim como os princípios e conceitos, da WS concluíram que “todas essas práticas se enquadram nos paradigmas de uma Ciência da Informação revisitada e expandida”, (Robredo, 2010, p. 33).

Neste contexto, a referência a uma CI “revisitada e expandida” parece sugerir uma mudança em relação à CI decorrente da Documentação. Ao procurar uma clarificação relativa à mudança sugerida, surgem novos pontos de interrogação decorrentes de relações disciplinares complexas entre a CI e outras áreas, nomeadamente o trio de disciplinas designado por BAD (Biblioteconomia, Arquivística e Documentação). Se, com esta tríade BAD, a CI apresenta uma relação intrínseca a ponto das suas fronteiras se confundirem, as relações com outras áreas aparentemente mais afastadas não são, por vezes, menos “intrusivas”:

Frequentemente as pessoas pensam tratar-se de Informática ou algo ligado ao Jornalismo. Alguns já ouviram falar de sua relação com a Biblioteconomia e pensam tratar-se apenas de um “novo nome” para ela. Cursos de “gestão da informação” causam ainda mais dúvidas. Recentemente, no Brasil e em outros países, cursos de Arquivologia e

⁴ Tradução livre do autor.

Museologia foram criados em faculdades ou departamentos de Ciência da Informação, ampliando ainda mais a confusão. (Araújo, 2013, p. 2)

Esta “confusão” assume maior relevo quando a mesma aparenta não se confinar a observadores externos à área da CI, atendendo à manifesta dificuldade na definição epistemológica e demarcação do campo de conhecimento e de intervenção da mesma (Araújo, 2013; Barreto, 2008; Galvão, 1998; Gracioso & Saldanha, 2011; Pinheiro, 1997; Saracevic, 1996; J. L. C. Silva & Freire, 2012). Associado a esta problemática encontra-se a apropriação de objetos de estudo do âmbito da CI por outras ciências, das quais se destaca a Ciência da Computação (CC). Entre os objetos de estudo, pela sua relevância na organização do conhecimento e o acesso ao mesmo, destacam-se as taxonomias e as, já referidas, ontologias (R. R. Souza, Almeida, & Baracho, 2013, p. 163). Ainda neste sentido, e de acordo com Almeida (2013, p. 1532), estes objetos de estudo mostram um historial mais longo na sua associação à CI que à CC.

No caso particular das ontologias, a sua relação inicial com a WS (Berners-Lee et al., 2001) é, atualmente, reconsiderada. Alguns investigadores, como James Hendler (2014, pt. 57:55), consideram, até, ser necessário afastar o conceito de ontologia do desenvolvimento da WS, uma vez que o tipo de organização do conhecimento da primeira não se “encaixa” na segunda. Desse modo, o desenvolvimento da WS passa necessariamente por sistemas de organização do conhecimento adaptáveis ao ambiente específico da *Web*.

Nesse sentido, a Ciência da Informação está numa posição privilegiada para fornecer a experiência e os conhecimentos necessários uma vez que a *organização do conhecimento* é um campo que, de modo algum, é novidade para ela (Sandhaus, 2011, pt. 25:05; R. R. Souza, Tudhope, & Almeida, 2012, p. 180).

Se há vinte anos atrás já se vaticinava que o “centro de gravidade das práticas informacionais desloca-se inexoravelmente de um polo constituído pelo papel para um polo eletrónico” (Le Coadic, 1996, p. 110), atualmente a *Web* consolida-se como o local onde a informação é procurada. Nesse contexto, as potencialidades de uma *web semântica*, no que concerne ao auxílio que a mesma prestaria na recuperação da informação em linha, afigura-se como algo essencial. No entanto, o conceito de Web Semântica encontra-se “enublado” por interpretações “abusivas” associadas ao termo *semântica* (Setzer, 2015) e a associação a outros termos tais como: *Web of data*, *Linked Data* ou ainda *Web of Linked Data*. Ao contrário de autores como Feigenbaum (2011, pt. 11:00) considera-se existir diferenças, tanto intrínsecas como contextuais, entre estes três termos que justificam uma demarcação conceitual dos mesmos e, como tal, merece a devida clarificação.

Outra clarificação que se mostra necessária é a ligação entre a CI e a WS, não tanto a nível conceitual pois a esse nível a relação afigura-se óbvia, tendo em conta o descrito a cima. Esta relação poderá, de uma forma simplificada e metafórica, ser descrita na seguinte ideia: *a procura de uma descrição do mundo todo e sua disponibilização para todo o mundo*. Neste contexto, coloca-se como questão verificar e refletir sobre essa ligação (CI-WS) a um nível essencialmente pragmático.

O cenário descrito fornece um enquadramento complexo quando o pretendido é refletir sobre a ligação das duas áreas, a CI e a WS, a um nível operacional, concreto. Para o efeito, o contexto teórico-prático que se consubstancia nas Dissertações de Mestrado (DM) e Teses

de Doutorado (TD) originárias de cursos da primeira área (a CI) cujo tema é a segunda (a WS), mostra ser um campo de estudo profícuo.

Estas produções científicas, as DM e as TD, são, como é sabido, indicadores importantes na aferição do desenvolvimento de uma área. Neste sentido, a reflexão efetuada fornece indícios relevantes para a clarificação da relação CI-WS no eixo pretendido, o prático, como também em termos teóricos. Tal entendimento decorre da combinação da análise intrínseca e da análise extrínseca aos referidos trabalhos.

Por fim, dada a variedade de designações da área da CI que se manifesta em diferentes regiões, em especial no que concerne à Europa e à América, afigura-se pertinente incidir o estudo sobre cursos de Portugal e do Brasil, dado as suas afinidades culturais.

Objetivos da investigação

Com este estudo pretende-se, como objetivo geral:

- Verificar e refletir sobre a presença da Web Semântica no *output* dos cursos de mestrado e doutoramento em Portugal e no Brasil em Ciência da Informação, entre o período de 2005 e 2015.

Como objetivos específicos:

1. Identificar os cursos de mestrado e doutoramento em Portugal e no Brasil em Ciência da Informação certificados, respetivamente, pela A3ES e pela CAPES⁵, em funcionamento no ano letivo 2014-2015;
2. Verificar o número de dissertações de mestrado e teses de doutoramento em Ciência da Informação associados à Web Semântica, produzidas em Portugal e no Brasil no período de 2005 e 2015, recuperadas de acordo com a metodologia desenhada para este estudo, nos repositórios RCAAP e OASISbr⁶;
3. Identificar os pontos de confluência da Ciência da Informação com a Web Semântica, presentes nos trabalhos recuperados;
4. Caracterizar o perfil académico dos orientadores e coorientadores dos trabalhos recuperados identificados no estudo como sendo de Ciência da Informação.

Os objetivos específicos articulam-se entre si para a obtenção do resultado geral pretendido e, embora tenham essa amplitude final, partem de um ponto-chave do objetivo geral. Assim, enquanto os objetivos específicos 2 a 4 partem das ações chave *verificar* e *refletir*, diretamente relacionadas com os resultados esperados, o primeiro dos objetivos específicos tem como função fornecer uma referência para a inclusão ou exclusão dos trabalhos recuperados de acordo com o curso em cujo contexto o mesmo foi gerado. Esta relação é explicitada no esquema apresentado na Figura 1.

⁵ Para os cursos de mestrado e doutoramento poderem funcionar é necessário a aprovação, em Portugal, da A3ES (Agência de Avaliação e Acreditação do Ensino Superior) e, no Brasil, da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), para mais detalhes ver ponto 2.4.

⁶ O RCAAP (Repositório Científico de Acesso Aberto de Portugal) e o OASISbr (Portal Brasileiro de Acesso Aberto à Informação Científica), configuram-se como pontos de acesso à produção científica e académica nacional, multidisciplinar e de acesso aberto, para mais detalhes ver ponto 4.1.

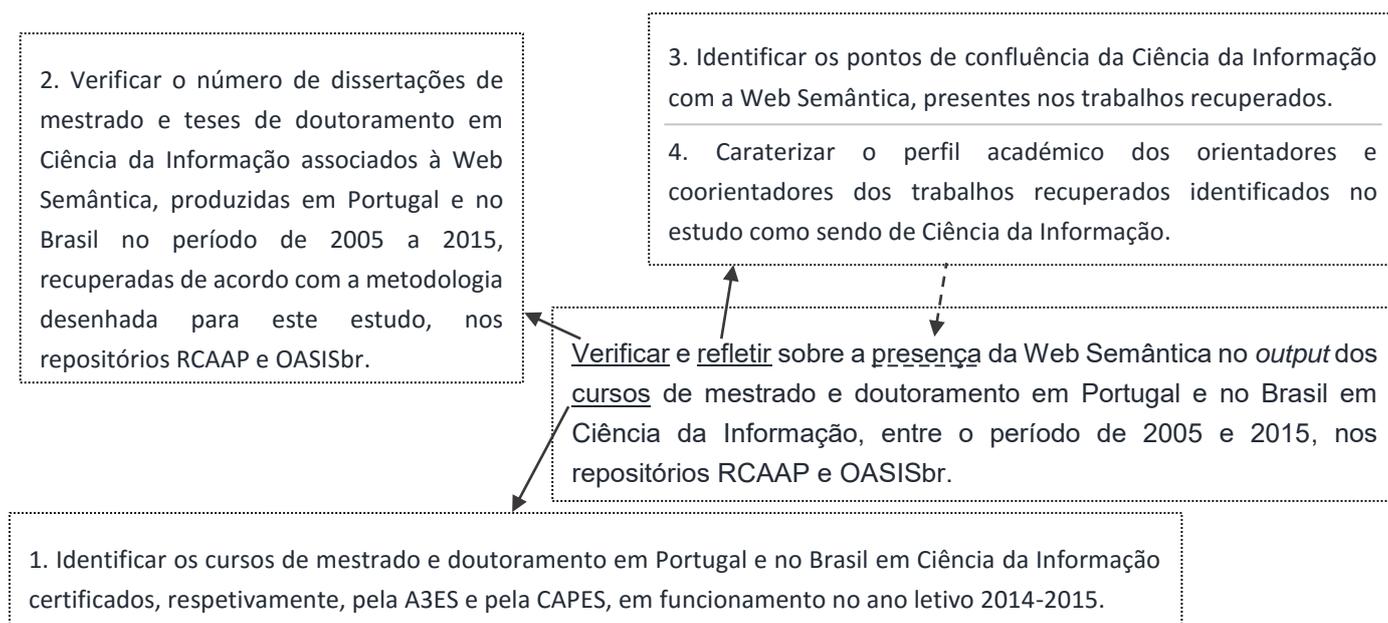


Figura 1. Relação dos objetivos específicos com objetivo geral a partir do ponto-chave inicial do mesmo.

Abordagem metodológica

Para dar cumprimento aos objetivos propostos desenhou-se uma metodologia de natureza qualitativa dividida em duas partes. Na primeira recorreu-se a uma revisão bibliográfica alicerçada num estudo comparativo da bibliografia consultada, na segunda a um estudo de caso.

Os conteúdos apresentados na primeira parte têm o intuito de contextualizar teoricamente os objetos de estudo de modo a sustentar o estudo de caso. Assim, no que concerne ao conceito de **Web Semântica**, foi considerado o seu desenvolvimento conceptual ao longo do tempo; quanto ao conceito de **Ciência da Informação** a abordagem foi no sentido da obtenção de uma referência relativamente aos cursos a incluir no estudo de caso. A abordagem concetual dos conceitos estudados foi precedida por uma breve introdução metodológica. Esta opção prendeu-se com o facto de a sua construção concetual se caracterizar por um nível significativo de complexidade, a qual poderia causar alguma entropia no seu entendimento.

A segunda parte do trabalho constitui-se no estudo de caso, centrado nas dissertações de mestrado e teses de doutoramento dos cursos de Ciência da Informação em Portugal e no Brasil, entre o período de 2005 a 2015, que abordam o tema da Web Semântica. A abordagem metodológica foca-se, essencialmente em dois pontos: i) recolha dos dados, ii) apresentação e análise dos resultados.

Uma vez que o complexo contexto abordado assim o exigiu, o desenho metodológico foi concebido usando etapas interdependentes como se apresenta na *Figura 2*.

Para o desenho desta abordagem metodológica tiveram um contributo essencial as seguintes obras: *Manual de investigação em ciências sociais* (de Quivy e Campenhoudt), *Análise de conteúdo* (de Bardin) e *Estudo de Caso: Planejamento e Métodos* (de Yin).

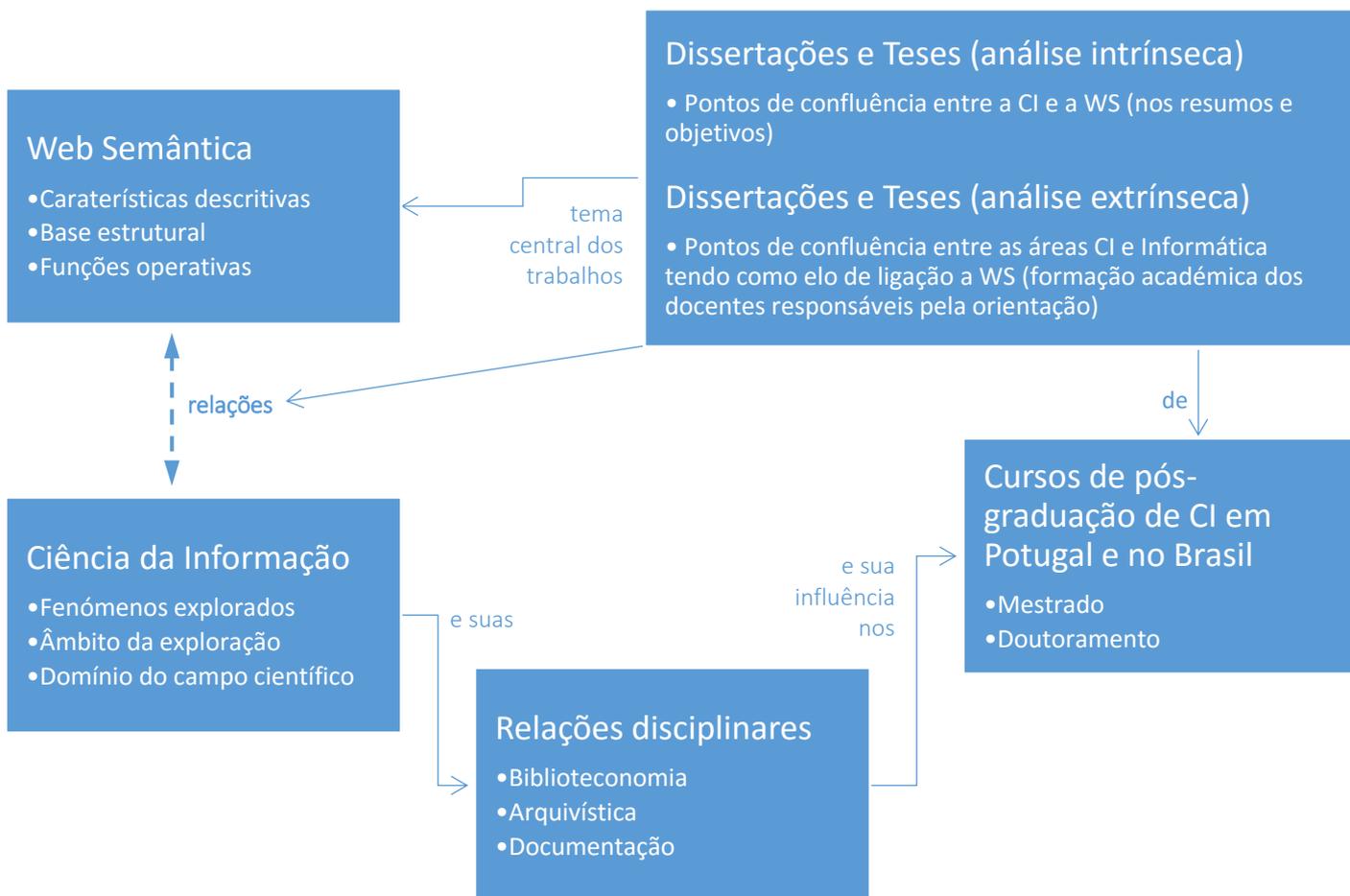


Figura 2. Esquema do relacionamento das etapas da análise concetual efetuada, com descrição dos respetivos focos da mesma.

Importa, ainda, referir que a abordagem aos termos específicos, presentes nos documentos analisados, seguiu uma de duas opções de acordo com a intenção pretendida. Por um lado, optou-se por manter os termos na língua nativa dos respetivos documentos de origem quando o respetivo conceito a eles associado foi objeto de análise. Por outro lado, quando a abordagem é interpretativa, efetuou-se a tradução dos termos em questão para português colocando subsequentemente os termos na língua original entre parêntesis. Em situações onde o termo sujeito à tradução está enquadrado em transcrições do texto original que o contextualizam, prescinde-se da repetição do mesmo entre parêntesis.

Estrutura

Além da presente introdução, na qual se expõe, de um modo geral, as ideias chave do estudo o trabalho é constituído, como foi referido, por duas partes. A primeira, a contextualização teórica, subdivide-se em dois pontos relativos aos conceitos Web Semântica e Ciência da Informação. Por sua vez, cada um destes pontos é subdividido em duas etapas: i) Metodologia considerada na análise concetual; ii) Análise concetual dos respetivos conceitos. Esta segunda etapa é, por sua vez, subdividida de acordo com os objetivos pretendidos para cada um dos conceitos. No caso da WS, a clarificação dos conceitos a ela associados, e, para a CI, além da análise conceitual propriamente dita, são também

apresentadas as suas relações disciplinares e os cursos de pós-graduação (mestrado e doutoramento).

A segunda fase, composta pelo estudo de caso, contextualizado entre o período de 2005 a 2015, sobre a presença da Web Semântica no *output* dos cursos de mestrado e doutoramento em Ciência da Informação, em Portugal e no Brasil, apresenta, igualmente duas partes. A primeira é constituída por uma breve introdução ao estudo de caso, na qual se contemplam os objetivos e a metodologia de forma pormenorizada, dado os vários passos efetuados no sentido de precisar a mesma e clarificar os limites do estudo. Na segunda parte apresenta-se os dados recolhidos e a análise dos resultados.

O estudo termina com a conclusão, na qual se registam as ideias chave e as limitações encontradas no desenvolvimento deste estudo. Apresentam-se, ainda nesse ponto, os contributos para a área da CI e novos rasgos para futuros trabalhos.

Bibliografia

Este último ponto da introdução tem a ver com a bibliografia consultada, dado o interesse e o peso que este elemento teve no presente estudo. As obras consultadas podem ser agrupadas em dois conjuntos: as relativas ao conceito da WS, que abordam a sua origem, percurso e desenvolvimento, e as referentes ao conceito de CI, com foco na teorização da sua gênese e desenvolvimento enquanto área científica e relações disciplinares com outros campos de estudo. Adicionalmente, refere-se um terceiro grupo onde se incluem obras que abordam a ligação entre estes dois conceitos, a WS e a CI. Um grupo não tão expressivo quanto o esperado principalmente devido à carência de obras de autores que não pertencem à área da CI.

Os documentos utilizados, no que diz respeito à WS, situam-se maioritariamente no presente século, sendo as exceções à regra referentes à década de 90 do século vinte. Uma vez que se trata de uma abordagem essencialmente atual, este conjunto de documentos concentra-se no período entre 2010 e 2016. Relativamente aos seus autores, estes são maioritariamente de origem anglo-saxónicos, dado o facto de ter sido nos países desta expressão linguística que se verificou a origem e o desenvolvimento mais pronunciado da WS.

No que concerne ao conceito de CI, a abordagem retrospectiva implicou uma abrangência temporal mais alargada, com cerca de um quarto das obras deste conjunto datadas entre 1934 e 1998 (embora a década de 90 englobe metade deste subconjunto). As restantes obras distribuem-se de forma bastante regular entre os anos 2002 e 2013. Também no que respeita aos seus autores, a origem geográfica dos mesmos apresenta uma distribuição mais alargada que a do primeiro conjunto. Encontra-se autores de vários países europeus (como Portugal, Espanha, França, Inglaterra e, até, a ex-URSS) e ainda do continente americano tanto do Norte como do Sul (Brasil).

Conclui-se este ponto referindo que, das obras consultadas, apenas aquelas que são citadas no corpo do texto se encontram registadas na respetiva secção do presente trabalho, estando as mesmas organizadas de acordo com a norma APA (6ª edição). Outras obras houve que tiveram um papel relevante na elaboração deste trabalho, contudo, não se encontra registadas nas referências bibliográfica dado o facto de não terem sido citadas. Entre estas últimas encontram-se os artigos apresentados no Apêndice 1.

PARTE I: CONTEXTUALIZAÇÃO TEÓRICA

Introdução à contextualização teórica

Neste capítulo irão ser abordados dois conceitos que se constituem o objeto de estudo. O conceito de web semântica e o conceito de ciência da informação. Dado o facto de as duas noções consideradas apresentarem características dinâmicas, entre as quais se destacam a sua linha evolutiva e a complexidade concetual, na medida em que ambas agregam na sua definição diversos aspetos do conhecimento, entendeu-se introduzir um ponto metodológico relativo à construção concetual de cada uma delas, para um melhor entendimento do seu significado. Assim, à definição de cada conceito precederá o ponto metodológico respetivo. Pretende-se, além da contextualização metodológica acima referida, explicitar claramente os processos utilizados nas análises pois considera-se que a exposição das suas delimitações, o mais detalhadamente possível, vai ao encontro do “critério de ‘refutabilidade’ da hipótese” exposto por Quivy e Campenhoudt (1992, p. 147).

Na contextualização sobre a Web Semântica não será abordado a questão das “versões” da *World Wide Web*, as chamadas “Web 2.0” e “Web 3.0”, uma vez que se entende não ser possível abordar essa questão sem um estudo mais aprofundado da *Web*, tanto à escala micro, enquanto “infraestrutura composta por linguagens e protocolos artificiais”, como à escala macro, enquanto “máquina social” integrante de um sistema mais amplo de interação humana (Hendler, Shadbolt, Hall, Berners-Lee, & Weitzner, 2008, pp. 63, 66). Um estudo dessa natureza extravasa o âmbito do presente trabalho.

1. O conceito de Web Semântica

Apesar das primeiras referências explícitas ao conceito Web Semântica (WS) terem surgido no final da centúria de 1900⁷, i.e., cerca de dez anos após a data da invenção da *World Wide Web* (WWW) em 1989⁸, as “sementes” da ideia do que viria a ser conhecido como a *Web Semântica* podem já ser encontradas na proposta inicial da WWW⁹ e em artigos anteriores a 1998¹⁰ (Bizer, Heath, & Berners-Lee, 2009, Capítulo 6). É, contudo, após 1998 que o termo Web Semântica começa a ser referido com uma certa recorrência em documentos de carácter científico.

⁷ Referências explícitas ao conceito web semântica podem ser encontradas nas notas pessoais de Tim Berners-Lee, nomeadamente em: *A roadmap to the Semantic Web*, datada de setembro de 1998, disponível em: <https://www.w3.org/DesignIssues/Semantic.html> (acedido a 24-3-2016).

⁸ Em 1989 Tim Berners-Lee trabalhava no CERN (Organização Europeia para a Pesquisa Nuclear) onde apresentou um sistema de gestão de informação considerado a matriz do que viria a ser a WWW (Bizer et al., 2009, Capítulo 6).

⁹ Na nota final da proposta submetida por Berners-Lee ao CERN pode ler-se: “A second phase would almost certainly involve some programming in order to set up a real system at CERN on many machines. An important part of this, discussed below, is the integration of a hypertext system with existing data, so as to provide a universal system, and to achieve critical usefulness at an early stage.”, (<https://www.w3.org/History/1989/proposal.html>, acedido a 22-03-2016).

¹⁰ No artigo de 1994 intitulado *The World-Wide Web*, como trabalho a desenvolver previa-se: “Evolution of objects from being principally human-readable documents to contain more machine-oriented semantic information, allowing more sophisticated processing” (Berners-Lee, Cailliau, Luotonen, Nielsen, & Secret, 1994). Também em 1994, a visão para a WS de Tim Berners-Lee foi expressa pelo próprio na primeira *World Wide Web Conference* (Shadbolt et al., 2006, p. 96).

1.1. Metodologia considerada na análise conceitual

Atendendo ao contexto da origem do conceito web semântica (Berners-Lee, 1998; Berners-Lee & Fischetti, 2000, p. 177) objetivou-se uma análise do mesmo a partir de um conjunto de referências tendo como denominador comum a participação do seu criador, Sir Timothy Berners-Lee (cf. Tabela 1).

Os critérios para a seleção das referências alicerçam-se nas principais regras de Bardin (2011, pp. 126–128) para a constituição de um *corpus* de análise, nomeadamente a representatividade, homogeneidade e pertinência. Assim, tomou-se por base a lista de “publicações selecionadas” que consta na biografia oficial de Tim Berners-Lee¹¹. A essas dez publicações com datas entre 1999-2009¹², acrescentaram-se mais sete referências datadas entre 2009-2015, seis descrições presentes em outras tantas páginas *web* do *World Wide Web Consortium* (W3C)¹³ e uma entrevista ao seu diretor Tim Berners-Lee, referenciada em uma dessas páginas.

A inclusão das referências mais atuais (na Tabela 1 com o prefixo *SW*)¹⁴ foi considerada pertinente atendendo à volatilidade do campo em estudo e à relação entre Tim Berners-Lee e o W3C. Simultaneamente, foi-se ao encontro de uma quarta regra expressa por Bardin (2011, p. 126), a regra da exaustividade.

Durante a pré-análise dos documentos foi-se construindo uma hipótese segundo a qual o conceito da Web Semântica teria sofrido uma evolução, ao longo do tempo, no sentido de tornar a sua operacionalização mais concreta e imediata. Hipótese reforçada pela chamada de atenção de Marcondes (2012, p. 26) para uma “mudança significativa da proposta original” passando de uma “web de significados” para uma “web de dados” a propósito da “visão bem mais realista” da WS apresentada pelo W3C em 2010. A análise do conceito Web Semântica foi, assim, no sentido da verificação da hipótese formulada. Formulação necessária para a explicitação da direção e respetivas dimensões de análise (Bardin, 2011, p. 130; Quivy & Campenhoudt, 1992, p. 135).

¹¹ Biografia disponível no URL: <https://www.w3.org/People/Berners-Lee/Overview.html>, acessado a 26-3-2016)

¹² A lista *Selected Publications* (disponível no URL: <https://www.w3.org/People/Berners-Lee/Longer.html>, acessado a 26-3-2016) é composta por 14 itens (na realidade apresenta 15 mas uma referência é repetida). Destes 14, 4 não foram incluídos devido a: 2 serem publicações anteriores a 1998 (1992 e 1994) e não apresentam referências diretas à WS; e outros 2 (de 2008 e 2010) não se encontrarem em acesso aberto na sua forma final ou em *preprint*.

¹³ O W3C, fundado por Tim Berners-Lee em 1-10-1994, é um consórcio internacional que desenvolve e estabelece padrões abertos para a criação e interpretação de conteúdos para a WWW. (Informação recolhida em: <https://www.w3.org/Consortium/>, acessado a 24-3-2016).

¹⁴ As fontes referenciadas foram consultadas em 26-3-2016 nos seguintes URLs, (entre parêntesis está a respetiva referência da Tabela 1: <https://www.w3.org/2001/sw/#publs> (refs. *SW.01*); <https://www.w3.org/2001/sw/SW-FAQ> (ref. *WS.02*); https://www.w3.org/wiki/Semantic_Web (ref. *SW.03*); <https://www.w3.org/2001/sw/> (ref. *SW.04*); <https://www.w3.org/2013/data/> (ref. *WS.05*); https://www.w3.org/2001/sw/wiki/Main_Page (ref. *WS.06*); <https://www.w3.org/standards/semanticweb/> (ref. *WS.07*).

Tabela 1

Referências documentais usadas na análise concetual da “Web Semântica”

Ref.	Data	Autoria	Documento	Contexto
SP.01	1999 jun.7	Tim Berners-Lee, Dan Connolly, Ralph R. Swick	<i>Web Architecture: Describing and Exchanging Data</i>	W3C Note
SP.02	1999 Set.22	Tim Berners-Lee, Mark Fischetti	<i>Weaving the Web</i>	Livro
SP.03	2001 abr.26	Tim Berners-Lee, James A. Hendler	<i>Publishing on the Semantic Web</i>	Nature (artigo)
SP.04	2001 mai.	Tim Berners-Lee, James A. Hendler, Ora Lassila	<i>The Semantic Web</i>	Scientific American (artigo)
SP.05	2002 out.	James A. Hendler, Tim Berners-Lee, Eric Miller	<i>Integrating Applications on the Semantic Web</i>	The Journal of The Institute of Electrical Engineers of Japan (artigo)
SP.06	2005 set.13	Tim Berners-Lee, Wendy Hall, James A. Hendler, Nigel Shadbolt, Daniel Weitzner	<i>The Emerging Science of the Web</i>	Web Science Workshop (relatório)
SP.07	2006 jun.	Hall Shadbolt, Tim Berners-Lee	<i>The Semantic Web Revisited</i>	IEEE Intelligent Systems Journal (artigo)
SP.08	2006 ago.11	Tim Berners-Lee, Wendy Hall, James A. Hendler, Nigel Shadbolt, Danny Weitzner	<i>Creating a Science of the Web</i>	Science (artigo)
SP.09	2006 set.	Tim-Berners Lee, Wendy Hall, James A. Hendler, Kieron O'Hara, Nigel Shadbolt, Daniel J. Weitzner	<i>A Framework for Web Science</i>	Foundations and Trends in Web Science (artigo)
SP.10	2009 mar.	Bizer, Heath, Berners-Lee	<i>Linked data - The story so far</i>	International Journal on Semantic Web and Information Systems (artigo)
SW.01	2009 out.22	Tim Berners-Lee	<i>Web 2.0 Summit 09 discussion: Conversation with Tim Berners-Lee</i>	W3C Semantic Web Activity: Publications / Articles / Interviews ^a
SW.02	2009 nov.12	W3C - World Wide Web Consortium	<i>W3C Semantic Web</i>	W3C Semantic Web: Frequently Asked Question
SW.03	2013 jun.27	W3C - World Wide Web Consortium	<i>SemanticWeb</i>	W3C Wiki
SW.04	2013 Dez.11	W3C - World Wide Web Consortium	<i>W3C Semantic Web Activity</i>	W3C SW Activity (introdução e "What is the Semantic Web?")
SW.05	2013 ^b	W3C - World Wide Web Consortium	<i>W3C Data Activity: Building the Web of Data</i>	W3C Data Activity ("Context & Vision")
SW.06	2014 jan.28	W3C - World Wide Web Consortium	<i>Semantic Web Wiki</i>	SW Wiki ("Manin Page")
SW.07	2015	W3C - World Wide Web Consortium	<i>Semantic Web</i>	W3C Standards (introdução e "Linked Data")

Nota. SP: referências incluídas na secção "Selected Publications" da Biografia de Tim Berners-Lee; SW: referências extraídas das páginas do *website* do consórcio W3C dedicadas à Web Semântica. ^a A secção "Publications / Articles / Interviews", (coluna *contexto* linha SW.01), para além da entrevista listada contém mais três itens: o artigo listado com a ref. SP.04 e outras duas referências sem a participação de Tim Berners-Lee; ^b dois meses após a consulta, em 24-3-2016, para a análise esta página atualizou a data de 2013 para 2016, embora continue com 2013 no seu url: <https://www.w3.org/2013/data/>.

As três dimensões de análise concebidas para a definição do conceito web semântica foram: i. as *caraterísticas descritivas*; ii. a *base estrutural* e iii. as *funções operativas*, correspondendo, respetivamente, às seguintes perguntas de partida: i. *O que se pode entender por web semântica?*; ii. *O que é necessário para que a web semântica seja uma realidade?*; e iii. *O que se pretende atingir com a web semântica?*. Com esta categorização pretendeu-se verificar se os documentos a analisar patenteavam uma mudança na forma como o conceito foi exposto, tanto nos elementos descritos como estruturais, e ainda nas funcionalidades que se esperava obter da WS.

Numa segunda fase, procedeu-se a uma categorização por “acervo” (Bardin, 2011, p. 149), i.e., a constituição dos grupos e respetivos títulos concetuais foram estabelecidos posteriormente à análise das unidades de registo¹⁵ consideradas em cada dimensão, tendo-se seguido, como linha orientadora para essa análise, a “orientação ao padrão” (Yin, 2001, p. 136-143). Para a confrontação dessas unidades com a linha temporal adotou-se um procedimento baseado na “análise de séries temporais” (Yin, 2001, pp. 143-145).

1.2. Web Semântica: análise conceitual

A partir do *corpus* de análise selecionado (cf. Tabela 1), que abrange os anos de 1999 a 2015, e das três dimensões propostas para o seu estudo: 1- *Caraterísticas descritivas*, 2- *Base estrutural* e 3- *Funções operativas*, verifica-se que o conceito, associado à utilização do termo Web Semântica, foi sofrendo uma mutação no sentido de tornar a sua operacionalização mais concreta e imediata, como é demonstrado na análise que se segue.

1.2.1. Dimensão de análise: características descritivas

Dentro da dimensão *Características descritivas*, as unidades de registo foram organizadas em três grupos: *Web of Data* (1a), *Linked Data* (1b) e *descrições que incluem “semântica”* (1c). Esta categorização das unidades de registo decorre da constatação da sua relação com os conceitos que emprestam o respetivo termo aos títulos atribuídos aos grupos 1a e 1b e, no caso do grupo 1c, da relação entre os conceitos: semântica, sentido, significado, lógica, compreensão e conhecimento (Almeida et al., 2011). Dentro dos grupos foi ainda efetuada uma distinção entre as unidades de registo tendo em conta os termos utilizados para a representação dos conceitos (cf. Tabela 2).

Esta primeira etapa de análise aponta três possíveis definições para o que se poderá entender por Web Semântica. As duas primeiras, grupo 1a e grupo 1b, identificam, respetivamente, a WS como sendo uma *Web* de dados (*the Web of Data*) e, de uma forma mais específica, uma *Web* de dados ligados (*the Web of Linked Data*). A que corresponde ao grupo 1c destaca-se daquelas pela dificuldade na atribuição de uma “etiqueta” identificativa, mas poder-se-ia atribuir a designação de *Web de informação* tendo por referência a distinção de Setzer (2015) relativamente aos conceitos de *dado* e de *informação*: “Uma distinção fundamental entre dado e informação é que o primeiro é puramente sintático e a segunda contém

¹⁵ Utilizou-se como limitação semântica o “tema” para a definição das unidades de significação, aqui chamadas de “unidades de registo” segundo a terminologia que consta na obra consultada de Bardin (2011, p. 134).

necessariamente semântica (implícita nas palavras ‘significativo’ e ‘significação’ usada em sua caracterização).”¹⁶

Tabela 2

Análise do conceito Web Semântica: Grupos e respectivas unidades de registo consideradas na dimensão de análise 1- Características descritivas

Grupos	Unidades de registo e respectivas referências da Tabela 1
(1a) <i>Web of Data</i>	<ul style="list-style-type: none"> i. <i>A Web of Data</i> (SP.10, SW.02, SW.03, SW.04, SW.07); ii. <i>One extension of the Web moving from text documents to data resources</i> (SP.08); <i>Is intended to function in the context of the relational model of data</i> (SP.09); iii. <i>Part of the Web of Data</i> (SW.05).
(1b) <i>Linked Data</i>	<ul style="list-style-type: none"> i. <i>The Web of linked data</i> (SW.06, SW.07); ii. <i>A new data model to support the linking of data from many different models</i> (SP.01); <i>The web of connections between different forms of data</i> (SP.02); <i>A world of trusted information shared along collaborating groups of users</i> (SP.03); <i>An open web of inter-referring resources</i> (SP.08); <i>A type of extension of the Web to extend the Web to cover linked data</i> (SP.09); <i>The world of linked data</i> (SW.010). iii. <i>Linked Data provides the means</i> (SP.10).
(1c) descrições que incluem “semântica”	<ul style="list-style-type: none"> i. <i>The Web of understanding</i> (SP.01); <i>A universal web of knowledge</i> (SP.03); ii. <i>An extension of the current Web in which information is given well-defined meaning</i> (SP.04, SP.05); <i>A web of logic</i> (SP.06); <i>A Web of actionable information derived from data through a semantic theory for interpreting the symbols</i> (SP.07); iii. <i>A web of data with meaning</i> (SP.02).

Nota. O documento com a referência *SW.05* (inserido no ponto iii do grupo 1a) marca o fim das atividades autónomas em torno da WS, desenvolvidas pelos grupos de interesse ligados ao W3C, fundindo-se essas atividades na *W3C Data Activity*, de maior âmbito: “The Data Activity merges and builds upon the eGovernment and Semantic Web Activities.” (W3C, 2013a). No documento com a referência *SP.10* (inserido no ponto iii do grupo 1b) o termo “Linked Data” não surge como uma descrição para a Web Semântica, contudo considerou-se relevante a sua inclusão dado o termo ser usado posteriormente como definição para a WS.

Relativamente à hipótese formulada sobre a evolução do conceito Web Semântica, de que este se direciona para uma operacionalização mais concreta e imediata, a análise da forma como as diferentes “descrições” se articulam ao longo do tempo aponta para a confirmação da mesma (cf. Figura 3).

Como indícios de uma mudança para uma definição mais concreta e restrita podem apontar-se quatro:

- a. As descrições do grupo 1c, que apresentam um maior grau de complexidade, atendendo à utilização de termos como *understanding*, *knowledge*, *meaning* ou *logic*, encontrarem-se em documentos situados na primeira metade do período analisado – entre 1999 e 2006;
- b. A concentração das referências explícitas aos conceitos *web of data* e *linked data* nos documentos mais recentes, a partir de 2009 para o primeiro conceito e 2014 para o segundo;

¹⁶ Importa ressaltar que o autor citado considera ser “impossível introduzir e processar semântica em um computador, porque a máquina mesma é puramente sintática” (Setzer, 2015).

- c. A integração das *Semantic Web Activities* na *W3C Data Activity* cujo objetivo é a construção da Web de Dados¹⁷ (W3C, 2013a);
- d. Nos dois documentos mais recentes listados na Tabela 1 (ref. *SW.06* e *SW.07*), surge a especificação do termo Web Semântica restringindo o seu âmbito: “The term ‘Semantic Web’ refers to W3C’s vision of the Web of linked data.” (W3C, 2014, 2015b).

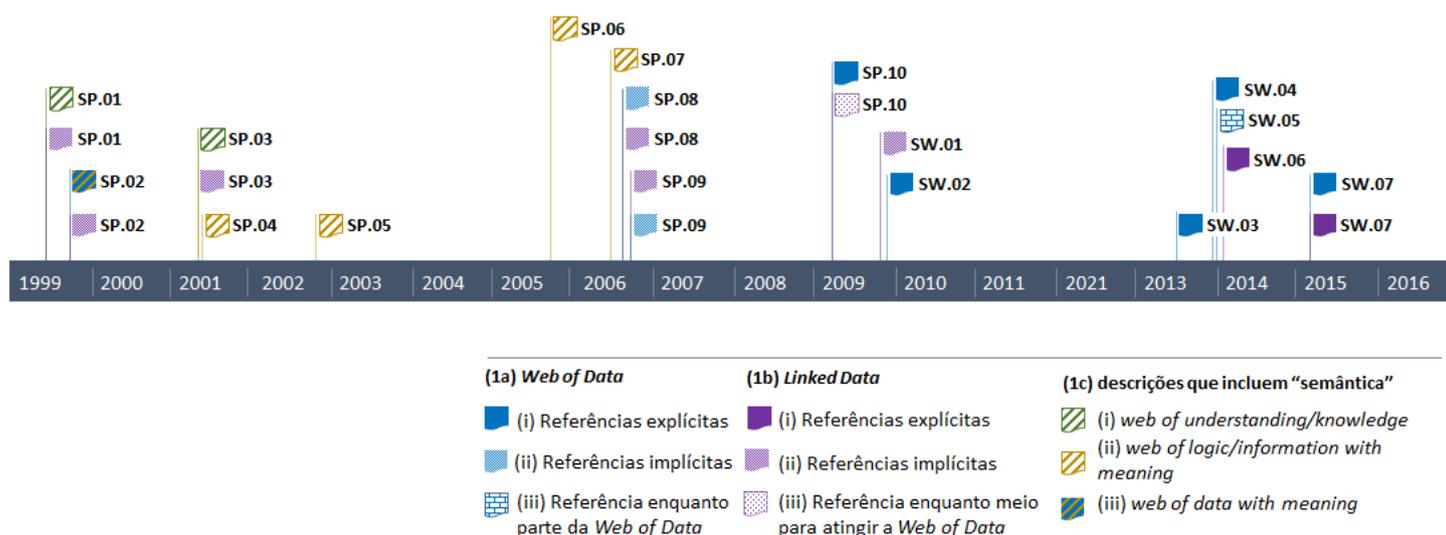


Figura 3. Distribuição temporal das unidades de registo consideradas na dimensão de análise *Caraterísticas descritivas* de acordo com a data do documento de proveniência.

Torna-se necessário ressaltar, a respeito destas considerações, que uma descrição da WS menciona explicitamente o conceito *web of data* antes de 2009, mais especificamente dez anos antes, no livro *Weaving the Web* (ref. *SP.02* da Tabela 1). No entanto, essa menção foi incluída no grupo *descrições que incluem “semântica”* (1c), observando a descrição contida no glossário do referido documento: “Semantic Web: The Web of data with meaning in the sense that a computer program can learn enough about what the data means to process it.”¹⁸ (Berners-Lee & Fischetti, 2000, p. 237).

É interessante verificar, neste percurso, a evolução concetual do conceito *Linked Data*. Em 2009 *Linked Data* é considerado *um meio* para atingir a Web Semântica: “while the Semantic Web, or Web of Data, is the goal or the end result of this process, Linked Data provides the means to reach that goal.” (Bizer et al., 2009), já em 2014 e 2015, atentando à visão do W3C (tal como exposto no ponto d., este aparenta ser *a própria* Web Semântica. Esta interpretação concorre para alguma perplexidade uma vez constar no mesmo documento de 2015 (ref. *SW.07*) essa descrição acompanhada de uma outra que diz claramente: “The Semantic Web is a Web of data.” (W3C, 2015b).

¹⁷ Apesar da página *web* onde se encontra a referida descrição estar, à data da consulta para análise, referenciada com o ano 2013, foi posteriormente atualizada para 2016 (tal como referido na Nota 3 da Tabela 1), uma atualização que pode ser entendida como um reforço indiciário da hipótese expressa nesta análise.

¹⁸ Além da descrição contida no glossário essa associação explícita ocorre uma vez mais no documento: “This creates what I call a Semantic Web—a web of data that can be processed directly or indirectly by machines” (Berners-Lee & Fischetti, 2000, p. 177).

Embora esse não seja caso único de dupla descrição no mesmo documento (como se poderá verificar na Tabela 2 e Figura 3), é-o no sentido de parecer indicar serem os termos *semantic web*, *web of data* e *linked data* referentes ao mesmo conceito. Ideia contrária àquela que é apresentada nas referências anteriores no que diz respeito ao primeiro e último conceitos, pois o âmbito da WS é exposto como mais abrangente que o de *Linked Data*. Uma outra possibilidade é os termos *Linked Data* e *Web of Linked Data* representarem conceitos com abrangências diferentes, sendo o último equiparado ao conceito *Web of Data* que aparece como uma definição mais concreta para a WS substituindo o termo *semântica* por *dados*. Embora esta possibilidade apresente uma possível explicação para o desfasamento descrito, concorre para o aumento da confusão concetual entre os termos.

Apesar dessa falta de clarificação é possível verificar, sintetizando o que foi apresentado neste domínio de análise, uma passagem da abstrata semântica, associada ao “entendimento” da informação (*semantic web*), para o processamento sintático, utilizando dados concretos ligados entre si (*web of linked data*), próprio dos meios computacionais.

1.2.2. Dimensão de análise: base estrutural

Na segunda dimensão de análise as unidades de registo aí consideradas foram descrições de procedimentos, de conceitos e/ou abstrações e de especificações e/ou padrões que, de forma direta ou contextualmente, são apresentados como necessários, ou mesmo essenciais, para o desenvolvimento da WS. As descrições de procedimentos foram inseridas num grupo homónimo, *Procedimentos* (2a). Analisando as unidades de registo, aí incluídas, identificou-se um padrão nas descrições sendo possível uma distribuição das mesmas entre: *Procedimentos relacionados com a estruturação de dados para processamento automático* (i); *Procedimentos focados na criação de interoperabilidade entre sistemas e ligação de dados* (ii); e *Descrições que apresentavam as duas tipologias estabelecidas* (iii).

Tabela 3

Análise do conceito Web Semântica: Grupo 2a (Procedimentos) e respetivas unidades de registo consideradas na dimensão de análise 2- Base estrutural

Unidades de registo e respetivas referências da Tabela 1
Grupo 2a: Procedimentos
<ul style="list-style-type: none"> i. <i>describing explicit relationships between things and containing semantic information intended for automated processing (SP.01); The first step is putting data on the Web in a form that machines can naturally understand, or converting it to that form (SP.02); More information needs to be in a form that the machine can “understand” rather than simply display (SP.03);</i> ii. <i>requires increasing the amount of data exposed in RDF (SP.07); Linking is key to the SW. In particular, although the publishing of data and the use of RDF is essential, in many cases the practice has been the conversion of data into RDF (SP.09); with Linked Data as a foundation (SP.10); put the data out there and using standards URIs to things and make links (SW.01); requires a focus not just on the interoperability of data but of communities (SW.05);</i> iii. <i>It is based on the idea of having data on the Web defined and linked (SP.05); This means creation of a common framework that allows data to be shared and reused across application, enterprise, and community boundaries, to be processed automatically by tools as well as manually, including revealing possible new relationships among pieces of data (SW.02); It is about common formats for integration and combination of data drawn from diverse sources (...) It is also about language for recording how the data relates to real world objects (SW.04).</i>

Nota. A ênfase a negrito corresponde ao segmento com maior peso na categorização das respetivas unidades de registo.

A descrições de conceitos e/ou abstrações e de especificações e/ou padrões foram incluídas num segundo grupo intitulado de *Arquitetura* (2b). Nesse grupo as unidades de registo foram naturalmente associadas a um dos dois subgrupos que o constituem, conceitos/abstrações (i) ou especificações/padrões (ii) (cf. Tabela 4/Tabela 3).

Tabela 4

Análise do conceito Web Semântica: Grupo 2b (Arquitetura) e respetivas unidades de registo consideradas na dimensão de análise 2- Base estrutural

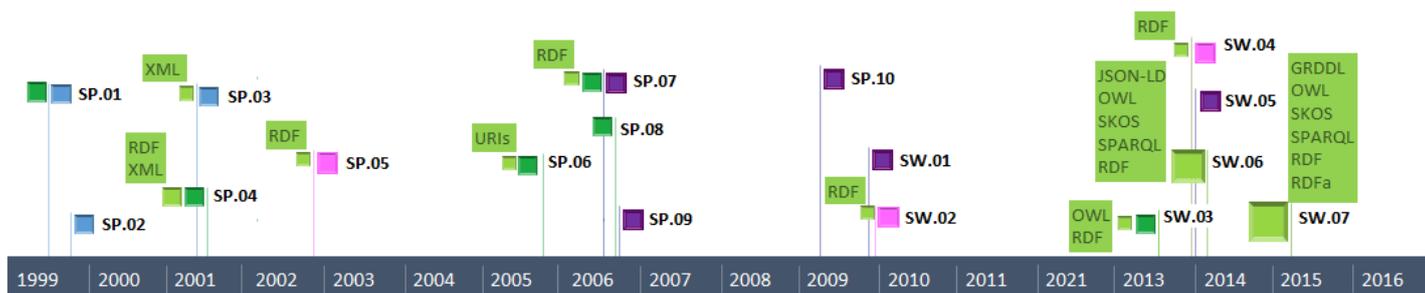
Unidades de registo e respetivas referências da Tabela 1
Grupo 2b: Arquitetura
<p>i. <i>The Semantic Web must be based on a facility that can expand as human understanding expands (SP.01); the third basic component of the Semantic Web [is] collections of information called ontologies (SP.04); Logic and ontologies will suffice to extract much of the value from the data held in structured relational databases (SP.06); information derived from data through a semantic theory for interpreting the symbols (SP.07); the "Semantic Web" of relational data and logical assertions, computer logic is in its element (SP.08); The Semantic Web is characterized by the use of Graphs and Networked Data (SW.03);</i></p> <p>ii. <i>Layers of the semantic web are built as new languages and tools anchored in XML (SP.03); Two important technologies (...) are already in place: XML and RDF (SP.04); a fundamental component of the Semantic Web is the RDF; URIs is still the critical architecture (SP.06); The Semantic Web can't exist without carefully developed and agreed standards [like RDF] (SP.07); RDF, which is one of the fundamental building blocks of the Semantic Web, plays the role of a common model, as a kind of a "glue" to integrate the data (SW.02); what is called "Big-S Semantic Web"- data formats such as RDF and OWL that are being defined by the W3C (SW.03); It is based on the RDF (SW.04); Linked data are empowered by technologies such as RDF, SPARQL, JSON-LD, OWL, and SKOS (SW.06); Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS (...) RDF provides the foundation for publishing and linking your data. Various technologies allow you to embed data in documents (RDFa, GRDDL) or expose what you have in SQL databases, or make it available as RDF files (SW.07).</i></p>

Nota. A ênfase a negrito corresponde aos conceitos/abstrações e às especificações/padrões identificados.

Observando os elementos assinalados na linha temporal, de acordo com as datas dos respetivos documentos, verifica-se uma concentração dos procedimentos focados na criação de interoperabilidade entre sistemas e ligação de dados (grupo 2a ii.) a partir de 2006, deixando de se registar os exclusivamente relacionados com a estruturação de dados para processamento automático (grupo 2a i., cf. Figura 4). Constatação compatível com um direcionamento do conceito da WS para a *web of linked data*, como já foi referido.

Também no grupo 2b verifica-se nas referências mais recentes a presença do conceito de *web of linked data*. Tanto ao nível dos conceitos/abstrações, com na utilização de um termo alternativo para o conceito *Linked Data – Networked Data*¹⁹ (ref. SW.03), como nas especificações associadas a esse conceito: “Linked data are empowered by technologies such as”, patente nas referências SW.06 e SW.07 (cf. Tabela 3). Poder-se-á relacionar o facto dessas duas referências serem as mais recentes, com a maior quantidade de especificações/padrões presente nessas descrições, 5 na ref. SW.06 e 6 na ref. SW.07. Tal facto estará relacionado com o tempo necessário para o desenvolvimento e maturação das especificações em causa. Todavia, a relação direta estabelecida, nesses dois documentos, entre as especificações e o termo *linked data* poderá ser interpretada como mais uma forma de operacionalizar a WS reduzindo a sua abrangência.

¹⁹ Na referência de origem o termo está ligado a uma página *web* com o nome *Networked Data* que explica: “One of the important ideas behind the concept of the SemanticWeb is ‘Networked Data’, aka LinkedData”, <https://www.w3.org/wiki/NetworkedData> (acedido a 7-5-2016).



(2a) Procedimentos

- (i) estruturação de dados para processamento automático
- (ii) criação de interoperabilidade entre sistemas e ligação de dados
- (iii) as duas tipologias de procedimentos

(2b) Arquitetura

- (i) conceitos/abstrações
- (ii) especificações/padrões

Figura 4. Distribuição temporal das unidades de registo consideradas na dimensão de análise *Base estrutural* de acordo com a data do documento de proveniência. As dimensões dos signos utilizados para representar as *Especificações/padrões* variam de acordo com a quantidade de elementos mencionados nas respetivas unidades de registo.

1.2.3. Dimensão de análise: funções operativas

Por último, na dimensão *Funções operativas*, a partir das descrições das funções e/ou objetivos da WS constituintes das unidades de registo consideradas, foi possível verificar a presença de uma finalidade operacional recorrente: “to use the Web like a single global database” (Bizer et al., 2009). Conceito que, assumindo diferentes expressões, é possível identificar em 7 das 17 referências, (cf. Tabela 5).

Tabela 5

Análise do conceito Web Semântica: Unidades de registo, dentro da dimensão de análise Funções operativas, que apontam como objetivo geral da WS a utilização da Web como uma base de dados global e não apenas um repositório de documentos

Ano	Unidades registo e respetivas referências da Tabela 1
1999	<i>to support the linking of data from many different models</i> (SP.01)
2002	<i>[for] linking of databases</i> (SP.05)
2005	<i>to exploit the large amount of data in structured relational databases</i> (SP.06)
2006	<i>to allow independent consistent data systems to be connected locally without requiring global consistency</i> (SP.08)
2009	<i>to use the Web like a single global database</i> (SP.10)
2013	<i>[to] allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing</i> (SW.04)
2013	<i>to lead the use of the Web as an exchange medium for data as well as documents</i> (SW.05)

Nota. O ano apresentado corresponde à data do documento de origem.

Pode, ainda, encontrar-se uma referência explícita à ideia da “base de dados global” no documento de 2015 (ref. SW.07), a propósito da necessidade da linguagem de consulta

SPARQL (*Protocol and RDF Query Language*)²⁰, para a Web Semântica uma vez esta ser entendida como uma *base de dados global* (W3C, 2015b).

Nas duas referências documentais mais recentes (referências *SW.06* e *SW.07*) a finalidade da WS é expressa de uma forma um pouco mais vaga: “to enable computers to do more useful work and to develop systems that can support trusted interactions over the network” (W3C, 2014, 2015b). Uma possível interpretação para esta opção será a mesma apresentar uma posição mais cautelosa, dado a complexidade da ideia da “base de dados global”. A menção a “trabalho mais útil efetuado pelos computadores” e a “interações em rede confiáveis”, sendo bastante genérica, não se compromete com objetivos específicos potencialmente mais dúbios em termos da sua realização a curto ou médio prazo, ou mesmo de todo, atendendo à opinião de autores como Lluís Codina (2003, p. 151), Clay Shirky ou Alon Halevy (Ray, 2010, pt. 7:00). É possível encontrar esse tipo de objetivos em algumas das funções, esperadas para a WS, expressas nos documentos até 2009, dentro do *corpus* analisado (cf. Tabela 6).

Tabela 6

Análise do conceito Web Semântica: Unidades de registo, dentro da dimensão de análise Funções operativas, que apontam funções esperadas da WS relacionadas com o processamento automático e/ou destinado aos “utilizadores finais” do sistema computacional

Ano	Unidades registo e respetivas referências da Tabela 1
1999	<i>to capture information that links independent representations of overlapping areas of knowledge</i> (SP.01)
1999	<i>[to] learn enough about what the data means to process it (...) to do something [with the data] it wasn't able to do directly"</i> (SP.02)
2001	<i>[for] new and better search engines, and users will thus be able to issue significantly more precise queries (...) [for] automated methods for helping users to understand the content produced by those in other scientific disciplines</i> (SP.03)
2001	<i>[to] bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users</i> (SP.04)
2006	<i>to facilitate the use of data as well as their discovery, to go beyond Google in this respect</i> (SP.09)
2009	<i>[to do] data oriented queries</i> (SW.01)
2009	<i>in resource discovery and classification to provide better, domain specific search engine capabilities (...); by intelligent software agents to facilitate knowledge sharing and exchange; in content rating</i> (SW.02)

Nota. Entende-se por “utilizador final” o utilizador que apenas utiliza o sistema completamente desenvolvido e finalizado, usando-o para o fim para o qual foi concebido²¹. O ano apresentado corresponde à data do documento de origem.

²⁰ SPARQL é um protocolo de comunicação e uma linguagem de consulta para dados em RDF (Resource Description Framework), análogo à linguagem SQL (Structured Query Language) para base de dados relacionais, (informação disponível em: <https://www.w3.org/TR/ld-glossary/#sparql> e <https://www.w3.org/TR/rdf-sparql-protocol/#intro>, URLs acedidos a 15-5-2016). Em algumas fontes surge a designação *Simple Protocol and RDF Query Language*, uma vez que essa era a proposta inicial para o nome desta especificação, posteriormente Dave Beckett, um dos responsáveis pelo seu desenvolvimento, considerando a mesma não ser propriamente “simples”, propôs a substituição do termo *Simple* pelo próprio acrónimo “SPARQL”, (informação recolhida em <https://lists.w3.org/Archives/Public/public-rdf-dawg/2004OctDec/0453.html>, acedido a 15-5-2016).

²¹ Definição adotada a partir das expostas em: http://www.webopedia.com/TERM/E/end_user.html e <http://whatis.techtarget.com/definition/end-user>, (URL acedidos a 12-5.2016).

Confrontando os três domínios, este último aparenta contrariar o direcionamento para uma operacionalização mais concreta e restrita apontado pelos outros dois domínios analisados. Todavia, vendo sob a perspectiva exposta acima – a descrição de funções genéricas como forma de evitar metas potencialmente mais longínquas – as funções operativas mais atuais, analisadas no terceiro domínio, não contradizem totalmente a hipótese formulada uma vez que esta aponta para uma “operacionalização, do conceito de WS, mais concreta e *imediate*”.

Relativamente à *operacionalização concreta* a questão relaciona-se com os destinatários das funcionalidades oferecidas pela WS. Se não for restringido aos utilizadores finais, consoante definido na Nota 1 da Tabela 6, pode-se considerar funções da WS os seguintes procedimentos concretos, incluídos no documento com a referência SW.07 (W3C, 2015b):

- i. Publicar e ligar dados utilizando o formato RDF²²;
- ii. Usar “tecnologias” como RDFa²³ ou GRDDL²⁴ para inserir dados em documentos, expor o conteúdo de uma base de dados SQL²⁵, ou ainda torná-los disponíveis no formato RDF.

A inclusão deste tipo de funcionalidades ocorre, também, noutros documentos mais antigos, como no documento com a referência *SP.05* onde, como funções operacionalizadas por “tecnologias fornecidas pela Web Semântica” (Hendler, Berners-Lee, & Miller, 2002) já em uso nessa altura, é apontado:

- i. A ligação entre bases de dados;
- ii. A partilha de conteúdo entre aplicações usando diferentes XML DTDs ou outros *Schemas*²⁶;

²² O RDF (Resource Description Framework) pode ser entendido como um conjunto de regras internacionais que estabelecem a forma como um recurso deverá ser descrito em modo legível por máquinas. Um recurso pode ser qualquer coisa, incluindo documentos, pessoas, objetos físicos e conceitos abstratos. O RDF permite fazer afirmações sobre recursos usando triplas Sujeito-Predicado-Objeto (S-P-O), onde S e O são recursos *Web* e P a natureza da sua relação, (informação recolhida em: <http://whatis.techtarget.com/definition/framework> e <https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/>, URLs acedidos a 15-5-2016).

²³ RDFa (Resource Description Framework in attributes) oferece um conjunto de marcação de atributos para os dados contidos em páginas web legíveis apenas por humanos, na forma de dados estruturados em RDF. (informação recolhida em <https://www.w3.org/TR/rdfa-primer/>, URL acedido a 15-5-2016).

²⁴ GRDDL (Gleaning Resource Descriptions from Dialects of Languages) é um mecanismo de recolha de dados expressos em RDF incluídos em documentos XML (ver Nota 26) de diferentes linguagens de domínios específicos, chamados de "dialetos XML", combinando-os para uma apresentação conjunta. (informação recolhida em <https://www.w3.org/TR/grddl/> e [https://www.w3.org/2001/sw/grddl-wg/tut7/gtut2.html#\(13\)](https://www.w3.org/2001/sw/grddl-wg/tut7/gtut2.html#(13)), URLs acedidos a 15-5-2016).

²⁵ No contexto o termo *base de dados SQL* é utilizado para designar uma Base de Dados Relacionais (RDB: *Relational Database*). Na relação entre os dados contidos numa RDB e o formato RDF a linguagem de marcação R2RML (RDB to RDF Mapping Language) tem um papel fulcral permitindo mapear esses dados para o modelo RDF. A estrutura do gráfico RDF resultante desse processo reflete diretamente a estrutura da base de dados original. (informação disponível em <https://www.w3.org/TR/r2rml/>, URL acedido a 10-6-2016).

²⁶ XML (*eXtensible Markup Language*) é uma linguagem de marcação largamente utilizada para partilha de informação estruturada: entre humanos, entre máquinas e entre ambos, uma vez que foi concebida para ser legível pelos dois tipos de utilizadores, (informação disponível em: <https://www.w3.org/standards/xml/core>, acedido a 9-5-2016). XML DTDs (Document Type Definitions) é um dos possíveis XML *Schemas* utilizados para expressar restrições em documentos XML, nomeadamente no que diz respeito à estruturação dos seus elementos e atributos, (informação disponível em: <https://www.w3.org/standards/xml/schema>, acedido a 9-5-2016).

iii. A descoberta e combinação de *web services*²⁷.

Apesar da distância que os separa, 13 anos, a diferença entre estes dois “conjuntos” de funções está nas especificações expressas, uma vez que a finalidade das mesmas é igual – interligar várias bases de dados na *Web*, tal como é expresso na página específica do *website* do W3C para *Linked Data*:

This collection of interrelated datasets on the Web can also be referred to as Linked Data. To achieve and create Linked Data, technologies should be available for a common format (RDF), to make either conversion or on-the-fly access to existing databases (relational, XML, HTML, etc). It is also important to be able to setup query endpoints to access that data more conveniently. W3C provides a palette of technologies (RDF, GRDDL, POWDER, RDFa, the upcoming R2RML, RIF, SPARQL) to get access to the data. (W3C, 2015a).

1.3. Vários termos para um conceito?

A análise efetuada confirma a hipótese formulada inicialmente: *o conceito da Web Semântica sofreu uma evolução, ao longo do tempo, no sentido de tornar a sua operacionalização mais concreta e imediata*, no entanto, o próprio conceito Web Semântica ficou pouco claro. Uma clarificação dificultada pelo relacionamento conceitual confuso com os termos *Web of Data* e *Linked Data*, uma vez ter-se verificado, nos três domínios analisados, uma alteração da utilização, terminológica e conceitual, de *semantic web* para *web of linked data*, ou simplesmente para *linked data* e, paralelamente a estes termos, a utilização de *web of data* para o mesmo conceito²⁸.

1.3.1. *Semantic Web, Web of Data, Linked Data*: uma possível definição

Apesar da fonte mais atual analisada (documento *SW.07* da Tabela 1) apontar para uma relação de sinonímia entre os três conceitos, outras referências obscurecem essa relação ou, até, a contradizem. Como exemplo paradigmático dessa situação aponta-se a página *web* atual sobre *Linked Data* do W3C²⁹, onde as relações entre os três conceitos se mostram contraditórias como é ilustrado na Tabela 7.

O facto do exemplo usado na Tabela 7 ser uma página *web* hierarquicamente dependente (ver Nota 29) da página onde é expresso que o termo *Semantic Web* refere-se à visão do

²⁷ Um *Web service* é um *software* de sistema concebido para funcionar como suporte na interoperabilidade entre máquinas em rede, (informação disponível em: <https://www.w3.org/TR/ws-gloss/#webservice>, acessado a 5-5-2016).

²⁸ A propósito da falta de clarificação e possíveis interpretações das relações de sinonímia entre estes conceitos veja-se a troca de argumentos causado pela publicação em linha de Tom Heath, investigador na área de *Linked Data*, sobre o significado desses termos no URL: <http://tomheath.com/blog/2009/03/linked-data-web-of-data-semantic-web-wtf/> (acessado a 2-6-2016). A posição de alguns autores é considerar todos estes termos, e outros, simplesmente como sinónimos (Feigenbaum, 2011, pt. 10:10).

²⁹ A página em questão (url: <https://www.w3.org/standards/semanticweb/data>) encontra-se no *website* do W3C com a seguinte hierarquia: *W3C (homepage do consórcio) – Standards – Semantic Web – Data* (página com o título “Linked Data”). No momento da sua consulta, em 16-5-2016, as duas páginas iniciais (*W3C* e *Standards*) apresentam a data de 2016 e as outras duas (*Semantic Web* e *Data*) estão datadas com o ano 2015.

W3C da *Web of linked data* (W3C, 2015b), torna ainda mais complexo as relações entre os três conceitos, contradizendo a inferência *d.* da referida Tabela. Por outro lado, a “taxonomia” utilizada no *website* do W3C, para a organização das páginas em questão, confirma essa mesma inferência ao colocar a página sobre *linked data* hierarquicamente subordinada à página *Semantic Web*.

Tabela 7

Exemplo das relações contraditórias de sinonímia e de pertença entre os conceitos Semantic Web, Web of Data e Linked Data encontradas numa mesma fonte

Afirmações retiradas do texto ^a	Inferências
<i>a. The Semantic Web is a Web of Data</i>	a. <i>Semantic Web</i> e <i>Web of Data</i> são termos diferentes para o mesmo conceito.
<i>b. to make the Web of Data a reality, it is important to have the huge amount of data on the Web available in a standard format, reachable and manageable by Semantic Web tools</i>	b. Se a <i>Semantic Web</i> possui ferramentas para atingir a <i>Web of Data</i> não fará sentido serem termos diferentes para o mesmo conceito.
<i>c. relationships among data should be made available to create a Web of Data. This collection of interrelated datasets on the Web can also be referred to as Linked Data</i>	c. <i>Web of Data</i> e <i>Linked Data</i> podem ser descritos como uma coleção de bases de dados interligadas na Web, pelo que haverá uma relação de sinonímia entre os dois conceitos.
<i>d. Linked Data lies at the heart of what Semantic Web is all about: large scale integration of, and reasoning on, data on the Web</i>	d. <i>Linked Data</i> está no cerne da <i>Semantic Web</i> , pelo que será a parte mais importante, mas não o todo.

Nota. Fonte utilizada – página *web* do W3C sobre *Linked Data*, com data de 2015 no momento da sua consulta em 16-5-2016 no url: <https://www.w3.org/standards/semanticweb/data>. ^a Excertos retirados das secções “What is Linked Data?” e “What is Linked Data Used For?”.

Perante esta complexidade inter-relacional dos três conceitos considera-se necessário clarificar o que, no presente trabalho, se entende por Web Semântica e sua relação com os outros dois termos, *Web of Data* e *Linked Data*. Assim, por Web Semântica entende-se um constructo abrangente que inclui:

- i. a grande finalidade de tornar a *Web* numa base de dados global, incluindo as aplicações/interfaces necessárias para seu uso;
- ii. as especificações (padrões, linguagens, vocabulários, protocolos) desenvolvidas e/ou associadas à sua operacionalização;
- iii. os princípios e respetivos procedimentos associados à prática de ligação de dados em formato aberto na *web*, enquanto base de sustentação de todo o conceito.

Por *Web of Data* (Web de Dados), por sua vez, entende-se a *Word Wide Web* resultante do desenvolvimento da Web Semântica, i.e., uma *Web* onde é possível realizar consultas similares às efetuadas em bases de dados, independente do tipo, formato e fonte desses dados, esperando-se, como resultado, dados semanticamente interrelacionados, ou seja, *informação* e não apenas uma listagem de documentos, muitas vezes sem ligação semântica entre si, como acontece na atual “Web de documentos” (W3C, 2013b, 2015b). A Web de Dados está, dessa forma, diretamente relacionada com o *ponto i.* enunciado a cima para a Web Semântica, dado o facto de constatar-se que o termo *Web of Data* é utilizado com mais frequência quando se fala do objetivo da WS (Berners-Lee, 1998; Bizer et al., 2009, Capítulo 1; Hendler et al., 2002; W3C, 2013a). Por outro lado, quanto surgem menções a

“ferramentas”, “tecnologias” ou algo similar ao exposto no *ponto ii.* é o termo *Semantic Web* o utilizado e não *Web of Data* (Berners-Lee et al., 2006, p. 108; Hendler et al., 2008, p. 68; W3C, 2014, 2015b).

Por fim, o conceito de *Linked Data* é entendido como um conjunto de procedimentos e regras para a publicação de dados na *Web* de forma a possibilitar a ligação dos mesmos a outros dados de diferentes fontes (Bizer et al., 2009, Capítulo 2). Muito embora o termo *Linked Data* também possa ser associado à ligação de dados fora do ambiente *Web*, utilizando “tecnologias da web semântica”, a grande mais valia desse conjunto de procedimentos, regras e especificações associadas é a sua adaptação à *Web*: “The primary value of Linked Data, of RDF and related technologies, is that these technologies have the Web at their ‘core’, providing a unique means of integrating data at Web scale.” (W3C, 2013a).

O conceito de *Linked Data* está, desse modo, associado ao explicitado no *ponto iii.* para a definição do que se entende por Web Semântica. Nesse contexto, atendendo aos “princípios da WWW” (ver Nota 1), a questão do formato aberto faz todo o sentido, daí a introdução do conceito *Linked Open Data*³⁰.

A ligação entre os termos *Web of Data* e *Linked Data*, e os conceitos apresentados, respetivamente, nos pontos i. e iii. descritos a cima, assim como a associação entre o termo *Semantic Web* e o referido no ponto ii. pode, ainda, ser encontrada cumulativamente em trechos de documentos como os que se exemplificam:

Semantic web technologies [ii.] and Linked Data principles [iii.] are paving the way for the Web of Data, a global data space [i.] that relies on a stack of technologies like URIs, HTTP, and RDF to empower information retrieval. (Parreiras, 2012, p. 189).

In the context of the emergent Web of Data [i.], a large number of organizations, institutes and companies (e.g., DBpedia, Geonames, PubMed ACM, IEEE, NASA, BBC) adopt the Linked Data practices [iii.] and publish their data utilizing Semantic Web technologies [ii.]. (Bikakis, Tsinaraki, Gioldasis, Stavrakantonakis, & Christodoulakis, 2013, p. 2).

1.3.2. Da Web com dados ligados à Web de Dados

A clarificação pretendida sobre estes conceitos, não ficaria completa sem abordar o que se entende por *Web of Linked Data*. No Ponto 1.2.1 foi referido a possibilidade dos termos *Linked Data* e *Web of Linked Data* representarem conceitos com abrangências diferentes, tendo o primeiro uma abrangência menor que o último que seria mais um quase-sinónimo³¹ do conceito *Web of Data*, especificando o tipo de dados “dessa *Web*”. Atendendo ao exposto nesta análise relativamente ao âmbito dos três conceitos (*semantic web*, *web of data* e *linked data*), faz-se um entendimento mais restrito do conceito *Web of Linked Data* relativamente ao associado ao termo *Web of Data*.

³⁰ A página *web* onde se encontram as regras, estabelecidas por Berners-Lee em 2006, para *Linked Data* foi alvo de uma atualização em 2010, com a introdução da “abordagem cinco estrelas” para a ligação de dados em formato aberto, no capítulo “Is your Linked Open Data 5 Star?”. Página disponível no URL: <https://www.w3.org/DesignIssues/LinkedData.html> (consultada em 27-5-2016).

³¹ Entende-se por *quase-sinónimo* “cada um dos termos de uma dada língua que designam uma mesma noção, mas que se situam em níveis de língua ou em níveis de conceptualização diferentes, ou então que se empregam em situações de comunicação diferentes”, definição expressa no *website* do Centro de Estudos de Linguística Geral e Aplicada da Universidade de Coimbra, (acedido a 2-6-2016. URL: <http://www.portaldalinguaportuguesa.org/?action=terminology&act=view&id=3223>).

Corroborando com essa visão mais restrita, pode observar-se num tutorial do W3C de 2014 que, por *Web of Linked Data* dentro da *Semantic Web Standard Stack*, se considera apenas as especificações SPARQL e RDF. Especificações essas relacionadas, respetivamente, com os componentes: consulta (*query*) e representação (*representation*) (cf. Figura 5). Já a *Web of Data*, tal como apresentada no ponto anterior, necessita do componente raciocínio (*reasoning*), i.e., necessita de um conjunto de regras³², para o processamento automático de dados, sua integração consistente e descoberta de novas relações entre eles, num processo denominado de *inferência*³³.

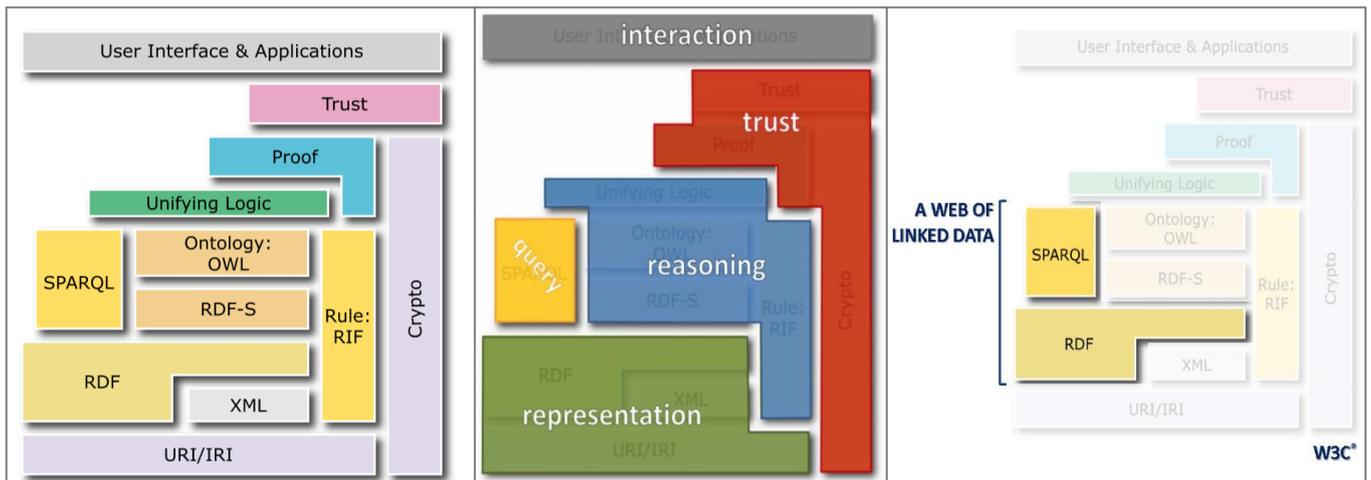


Figura 5. Três diapositivos do tutorial sobre *Linked Data* apresentado na 23ª conferência internacional da *World Wide Web* em abril de 2014. Fabien Gandon, *Semantic Web and Linked Data or how to link data and schemas on the web: A W3C tutorial at WWW2014*, http://www.slideshare.net/fabien_gandon/semantic-web-and-linked-data, (acedido a 4-5-2016).

Perspetiva semelhante à apresentada na Figura 5 pode ser encontrada na projeção isométrica da “pilha de tecnologias da Web Semântica”, onde apenas uma pequena seleção das mesmas é considerada como pertencentes à pilha *Linked Data* (cf. Figura 6). Essa representação é uma adaptação operada por Ioannis Parapontis, em 2012, ao original de 2009 realizado por Benjamin Nowack. Uma alteração que reforça o que acima foi referido sobre as componentes incluídas na *Web of Linked Data*.

A representação original não apresenta a descrição das “tecnologias” que fazem parte da pilha *Linked Data* e exclui a “camada de consulta” (*Query*). Já a alteração de 2012, como se pode verificar na Figura 6, inclui essa camada e não associando quaisquer especificações às camadas *Formats* e *Models*, apresenta como “tecnologias da pilha *Linked Data*” as

³² Dado a variedade de “paradigmas populares” (expressão usada na página do W3C consultada) para o uso de regras em representação de conhecimento e modelagem de negócios foi desenvolvido o padrão RIF (*Rule Interchange Format*). Este padrão faz o intercâmbio de regras entre diferentes sistemas, em particular entre mecanismos concebidos para a *Web*. A abordagem adotada foi conceber uma família de línguas, chamados dialetos, rigorosamente especificados em termos de sintaxe e “semântica”. Informação recolhida em <https://www.w3.org/TR/rif-overview/> (URL acedido a 10-6-2016).

³³ Segundo o W3C: “Inference on the Semantic Web is one of the tools of choice to improve the quality of data integration on the Web, by discovering new relationships, automatically analysing the content of the data, or managing knowledge on the Web in general. Inference based techniques are also important in discovering possible inconsistencies in the (integrated) data.”. Informação disponível no URL: <https://www.w3.org/standards/semanticweb/inference>, (acedido a 4-5-2016).

especificações: SPARQL, RDF e URI/IRI³⁴, ou seja, incluindo apenas as camadas *consulta* e *representação*.

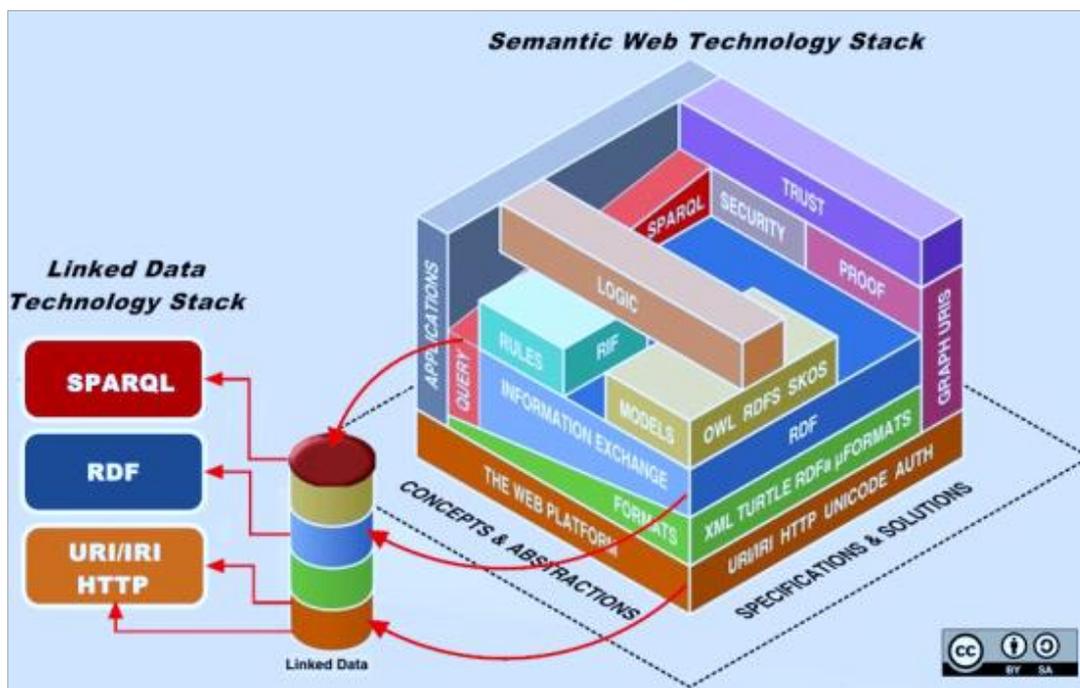


Figura 6. Projeção isométrica da “pilha tecnológica” da Web Semântica, com destaque das especificações associadas a *Linked Data*. Adaptação de Ioannis Parapontis,, <https://ioannis.parapontis.com/2012/12/07/semantic-web-technology-stacks/>, do original de Benjamin Nowack, <http://bnode.org/blog/2009/07/08/the-semantic-web-not-a-piece-of-cake>, (URLs acedidos em 8-6-2016).

Assim, por *Web of Linked Data* entende-se o que James Hendler (um dos investigadores que trabalha na área desde o início), chama de “the Linked Data view of the Semantic Web” que utiliza, essencialmente, as especificações associadas às camadas *representação* e *consulta*, uma vez que essa abordagem apenas necessita de “um pouco de acordo ontológico” (Hendler, 2014, pt. 25:50). Razão pela qual quem está envolvido no desenvolvimento de aplicações relacionadas com *Linked Data* utilizará maioritariamente RDF e SPARQL e

³⁴ Segundo o W3C, os URI/IRI funcionam como a “cola” que mantêm a *Web* unida. O URI (*Uniform Resource Identifier*) é uma cadeia de caracteres usada para identificar um recurso na *Web*, que podem ser do tipo *nome*, *local* ou ambos. Um URI do tipo *nome* é um *Uniform Resource Name* (URN), um número ISBN (*International Standard Book Number*), por exemplo, funciona como um URN. O URI do tipo *local* é designado por *Uniform Resource Locator* (URL), que identifica o recurso utilizando a representação do mecanismo de acesso ao mesmo como, por exemplo, a sua “localização” na rede. O IRI (*International Resource Identifier*) diferencia-se do URI apenas pelo conjunto de caracteres suportados, enquanto este último só aceita caracteres latinos ASCII (*American Standard Code for Information Interchange*), o IRI aceita o *Universal Character Set* (Unicode/ISO 10646), que pode ir até 2 biliões de símbolos, alargamento que permite a sua utilização independente da língua a representar, mas traz, também, problemas de segurança adicionais, nomeadamente a utilização de caracteres semelhantes para fins ilegítimos (<http://www.päypal.com> não é o mesmo que <http://www.paypal.com>). Informação recolhida nos seguintes URLs: <https://www.w3.org/2004/11/uri-iri-pressrelease>, <https://www.w3.org/TR/2001/NOTE-uri-clarification-20010921/> e <https://blog.4psa.com/url-urn-uri-iri-why-so-many/>, acedidos a 8-6-2016.

significativamente menos, RDF-S³⁵ ou OWL³⁶ (Cambridge Semantics, 2016; Hendler, 2011b, pt. 46:05).

No entanto, atendendo ao conceito abrangente que neste trabalho é adotado como sendo a Web Semântica, a *Web of Linked Data*, ou *Web de Dados Ligados*, não é aqui entendida como uma visão particular da mesma, antes a parte da WS que melhor se adequa ao ambiente essencialmente destruturado da *Web*. Uma parte em crescendo desde 2007, como é possível ver na evolução do digrama conhecido como a “nuvem de dados abertos ligados” (*Linked Open Data Cloud*) que se apresenta na Figura 7.

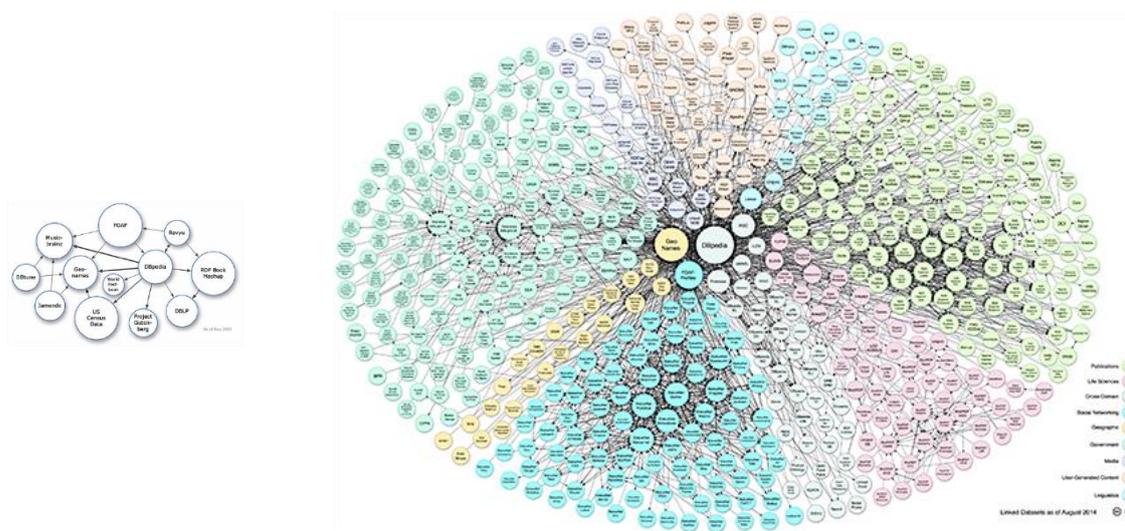


Figura 7. Evolução da *Linked Open Data Cloud*, de maio de 2007 (à esquerda) a agosto de 2014 (à direita). Richard Cyganiak e Anja Jentzsch, 2014, <http://lod-cloud.net/> (acedido a 17-7-2016).

Apesar de, atualmente, não apresentar o mesmo crescimento gigantesco dos primeiros anos (cerca de 50 vezes mais triplas RDF em 4 anos)³⁷, a razão para tal deve-se ao facto do gráfico apenas mostrar os dados ligados em *formato aberto*, não incluindo os maiores utilizadores comerciais deste conjunto específico de “tecnologias da web semântica” (Hendler,

³⁵ RDF-S (*Resource Description Framework Schema*) é um vocabulário para modelação de dados representados em RDF. Pode ser visto como uma extensão do modelo básico RDF, possibilitando a definição de hierarquias tanto de classes como de propriedades, assim como a definição de domínios e contradomínios para essas propriedades, permitindo um primeiro conjunto de restrições sobre as triplas definidas, além de inferências que deduzem triplas não declaradas de forma explícita. Informação recolhida em: <https://www.w3.org/TR/rdf-schema/> e <http://ceweb.br/guias/web-semantica//capitulo-4/> (URLs acedidos a 9-6-2016).

³⁶ OWL (*Web Ontology Language* e não *Ontology Web Language* como aparece em algumas fontes) é uma linguagem que estende o vocabulário RDF-S, tal como este estende o modelo básico RDF, oferecendo um vocabulário muito mais amplo assim como alarga os tipos de restrições possíveis de aplicar ao conjunto de triplas definidas. Além disso, são oferecidos diversos construtores que permitem, entre outros, a construção de classes complexas, o encadeamento de propriedades e a relação entre diferentes ontologias, aqui entendidas como artefactos computacionais, i.e., vocabulários expressos numa linguagem de representação do conhecimento que podem ser processados automaticamente e gerar inferências. (Almeida, 2013, p. 1688) Informação recolhida em: <http://ceweb.br/guias/web-semantica//capitulo-4/>, <https://www.w3.org/TR/owl-syntax/> e <http://www.cambridgesemantics.com/semantic-university/rdfs-vs-owl> (URLs acedidos a 9-6-2016).

³⁷ De 500 milhões de triplas RDF, contabilizados em 2007, passou a cerca de 25 bilhões em 2010, já em 2014 apresenta “apenas”, aproximadamente, 30 bilhões. Dados recolhidos das apresentações “Introduction to Semantic Web Technologies & Linked Data” de Oktie Hassanzadeh (2011) e “Semantic Web: The Inside Story” de James Hendler (2014).

2011a, pt. 34:15). Utilizadores como os “gigantes”: *Facebook, Google, Microsoft, Yahoo!, eBay e Oracle*³⁸ (Hendler, 2014, pt. 28:30).

Apesar dessa adoção, para o utilizador final a WS passa completamente, ou quase, despercebida. Sendo a explicação para tal, o facto do sucesso da Web Semântica ocorrer nos bastidores da *Web* “visível” ao comum utilizador (Hendler, 2014, pt. 1:05; Sletten, 2014). Já em 2011 Ivan Herman, líder da *Digital Publishing Activity* do W3C, afirmou: “we forget about the client side for web applications world”³⁹. Dir-se-ia que foi esquecido o “eixo P” (ver Figura 8) mencionado na página *web* do W3C intitulada “The Semantic Web Made Easy”:

The "perception" axis is to signify the abstraction/adaptation of the technologies in the tower, towards the people. To enable their perception by the people. In fact, this axis could have been well named as the "People axis", but perception is maybe a more detailed term. In any case, as both perception and people start with P, this can simply be named as the "P" axis. (Marchiori, 2003)

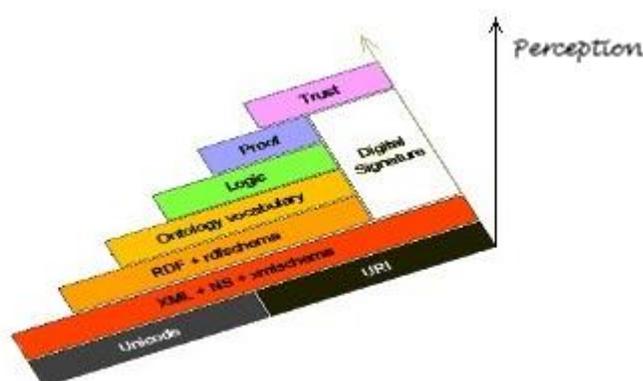


Figura 8. Perspetiva da “torre da web semântica” vista pelo “eixo P” (de Percepção e Pessoas). Massimo Marchiori, <https://www.w3.org/RDF/Metalog/docs/sw-easy>, (URL acedido a 12-6-2016)

Se se considerar que: “Linked Data tends to emphasize the *Web* part of *Semantic Web* and deemphasize the *Semantic* parts” (Cambridge Semantics, 2016), e que o objetivo global da *Web* é auxiliar a Humanidade (Berners-Lee, 2009, pt. 24:45), poder-se-á afirmar que uma peça importante está a faltar nesta equação. Uma peça complexa atendendo ao panorama atual e ao ano (2009) em que Berners-Lee afirmou: “my particular interest, for the past few years, has been the user interface to the semantic web, looking at how there’s a lot of stuff out there and how you can make best use of it all”⁴⁰.

Como relatado acima, a *Web de Dados Ligados* usa pouca expressividade semântica, além do RDF-S apenas utiliza “uma pequeníssima parte de OWL” (Hendler, 2011a, pt. 26:45) e

³⁸ Como exemplo da utilização de “especificações *Linked Data*” pode-se apontar o protocolo OGP (*Open Graph Protocol*) do *Facebook* (<http://ogp.me/>) e o vocabulário partilhado *Schema.org* (<http://schema.org>) patrocinado pelas empresas *Google, Microsoft, Yahoo!* e *Yandex*. (URLs acedidos a 12-6-2016).

³⁹ Afirmção utilizada por Ivan Herman como resumo da sua posição no evento *Semantic Web Death Match 2011: Industry vs. Academia vs. Standards*, http://videolectures.net/iswc2011_panel/, (URL acedido a 12-6-2016).

⁴⁰ Afirmção retirada da transcrição de uma entrevista efetuada na *British Computer Society* em novembro de 2009. Transcrição disponível no URL: <http://www.bcs.org/content/conWebDoc/33720>, (acedido a 12-6-2016).

sistemas leves de organização do conhecimento, também chamados de “ontologias leves”⁴¹ (*Lightweight Knowledge*) (Hendler, 2014, pt. 32:15). Estes sistemas são assim chamados uma vez que, de acordo com Isotani e Bittencourt: “A principal ênfase das ontologias leves é definir a taxonomia que representa a relação hierárquica entre conceitos.” (2015, Capítulo 3.2.3).

O termo “expressividade semântica” quando usado em associação a artefactos computacionais, tais como os vocabulários mencionados acima, baseia-se na *lógica formal* uma vez que o potencial semântico e de inferência computacional se baseia em: “identificar conteúdos ou ‘inferi-los’ a partir de conteúdos já reconhecidos, independentemente (e apesar) desses conteúdos.” (Marcondes, 2012, p. 63). Consistindo, o incremento de “semântica”, no aumento de restrições que limitem as possíveis interpretações dos conteúdos representados ao que se pretende efetivamente representar (Wassermann, 2014, pt. 53:00).

É essa, a *semântica formal*, a semântica da WS (Almeida et al., 2011, p. 202) que, de acordo com Steizer (2015) “é de fato, apenas um tratamento sintático expresso por meio de uma teoria axiomática ou de associações matemáticas de seus elementos com operações realizadas por um computador”.

Assim sendo, seria de esperar a utilização de sistemas mais poderosos em expressividade semântica, como as “ontologias pesadas”⁴², para ir além da *Web de Dados Ligados*. No entanto, esses sistemas não se adequam ao ambiente aberto da *Web*, pois este põe em risco a consistência dos seus modelos formais (Hendler, 2011b, pt. 53:00, 2014, pt. 21:10).

Há, portanto, a necessidade de “fazer a ponte” entre a *Web de Dados Ligados* já existente, ou talvez seja mais correto falar da existência de uma “*Web com dados ligados*”, e os sistemas de representação/organização do conhecimento de maior expressividade semântica para que se obtenha uma **efetiva Web de Dados** (Hendler, 2011b, pt. 1:01:50, 2014, pt. 47:35).

⁴¹ Sistemas como, por exemplo: SKOS - *Simple Knowledge Organization System* (<https://www.w3.org/2004/02/skos/>); Dublin Core (<http://dublincore.org>); FOAF - *Friend Of A Friend* (<http://xmlns.com/foaf/spec/>); SIOC - *Socially Interconnected Online Communities* (<http://sioc-project.org/>) e *Good Relations* (<http://www.heppnetz.de/projects/goodrelations/>), URLs acedidos a 13-6-2016.

⁴² Uma descrição dos vários tipos de ontologia pode ser consultada em <http://ceweb.br/livros/dados-abertos-conectados/capitulo-3/#sh3.2.3>, (URL acedido a 13-6-2016).

2. O Conceito de Ciência da Informação

A abordagem deste conceito divide-se em duas partes, uma primeira onde se explana a metodologia específica considerada na análise do conceito de Ciência da Informação (ponto 2.1) e uma segunda relativa à análise propriamente dita deste conceito (pontos 2.2, 2.3 e 2.4). São eles: Ciência da Informação: análise concetual, Relações (inter)disciplinares e Cursos de Ciência da Informação; respetivamente.

2.1. Metodologia considerada na análise concetual

Para o conceito Ciência da Informação (CI) considerou-se adequado uma abordagem metodológica que assentasse em três momentos, por forma a demonstrar a sua evolução continua. Tal opção tem a ver com a complexidade que envolve o referido conceito, nomeadamente às características que lhe são inerentes e que, de resto, são pouco consensuais. Entre elas destaca-se o carácter interdisciplinar da área, (Fernandes & Cedón, 2009, p. 116) o conceito de informação enquanto objeto de estudo (Zins, 2007c), e ainda a questão dos diferentes paradigmas, cujo conceito causa, *per se*, controvérsia suficiente (Rojas & Domínguez, 2009, p. 209). Perante tais questões, cuja análise transcende os limites do presente trabalho, a abordagem efetuada apresenta um cariz mais descritivo que analítico. A limitação imposta decorre, assim, da consciencialização da complexidade do assunto a abordar, mas também do que se pretende com a abordagem – a obtenção de uma referência relativamente aos cursos a incluir no estudo de caso, i.e., os cursos de Mestrado e Doutoramento de Ciência da Informação.

2.1.1. Uma abordagem em três momentos

Da pré-análise deste conceito sobressaiu uma hipotética divisão em três períodos temporais que facilitaria a abordagem concetual à CI. Não sendo possível definir com precisão a fronteira entre esses três períodos (nem tão pouco aconselhável, dado a continuidade e interseção dos acontecimentos), utilizou-se como demarcação das fases um período de transição, *grosso modo* coincidente com a passagem entre as décadas do século XX, de 50 – 60, para a passagem do primeiro para o segundo momento e, deste para o terceiro, entre as décadas de 70 – 80, não se fixando limites para o início da primeira fase nem para o final da terceira.

De acordo com a hipótese formulada a passagem do primeiro período para o segundo é marcado por eventos⁴³ como: os “projetos Cranfield” (1957 e 1963), experiências relacionadas com medidas de precisão e revocação classificadas como das mais relevantes em CI (Capurro & Hjørland, 2007, p. 179); a conferência internacional de informação científica (Washington, 1958), considerada por Shera e Cleveland (1977, p. 257) o acontecimento que marcou a passagem da Documentação em CI; as conferências do Georgia Tech (1961 e 1962), tidas como um marco, seja por se considerar que aí se efetuou a primeira formulação conceitual para a CI (L. V. R. Pinheiro, 2009, p. 110), seja porque foi a partir

⁴³ A listagem dos eventos não é de forma alguma exaustiva, apenas se apresenta alguns considerados pelos autores consultados como marcantes de uma passagem de estágio no desenvolvimento da área. Entre outros não mencionados destaca-se a Conferência Internacional sobre Princípios de Catalogação, realizada em Paris em 1961, pela importância que representou para a padronização internacional, fundamental para a interoperabilidade entre sistemas.

dessa formulação que Taylor, em 1963, e Borko, em 1968, redigiram as suas “clássicas” definições da CI (Garcia, 2002; A. M. da Silva & Ribeiro, 2008, p. 53).

Por sua vez, a delimitação da passagem da segunda fase para a terceira tem por base a opinião de alguns investigadores que apontam o final da década de 70 e princípio da de 80, como o início de estudos de revisão/consolidação da área, nomeadamente o artigo publicado em 1977 de Shera e Cleveland, "History and foundations of Information Science", considerado o primeiro artigo de revisão da área (Pinheiro, 1997, p. 8), e ainda a publicação, em 1981, da primeira edição do manual "Introduction générale aux sciences et techniques de information et de la documentation" de Guinchat e Menou. Este manual pode ser entendido como a consolidação do esforço da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) na criação de políticas de informação científica e tecnológica para os países em desenvolvimento (Araújo, 2013, p. 8). É também na passagem da década de 70 para a de 80 que, segundo Araújo (2013, p. 16), ocorrem os dois momentos de “relativo consenso” na comunidade científica ligada à CI no que diz respeito à definição do objeto de estudo do campo. O primeiro, no que concerne à necessidade de se incorporar o conceito de “conhecimento” na definição de “informação”, consenso obtido no evento “The Copenhagen Conference Theory and Application of Information Research”, em 1977 e, o segundo momento, nos anos seguintes, no que respeita à definição tríade de informação: “a *informação* é a medida da alteração que os *dados* provocam numa estrutura de *conhecimento* [ênfase adicionado]” (Araújo, 2013, p. 16).

Reforçando a escolha deste período como o de transição e fazendo a ponte para um dos pontos-chave do presente trabalho – os cursos de ensino superior, é no início da década de 80 que se assiste a uma mudança nos cursos relacionados com a área, tanto em Portugal como no Brasil. No primeiro (1982) com a extinção do curso de Bibliotecário-Arquivista da Universidade de Coimbra e subsequente criação do Curso de Especialização em Ciências Documentais, na mesma instituição nesse ano e nas Universidades de Lisboa e do Porto nos anos seguintes (A. M. da Silva & Ribeiro, 2008, pp. 147–148). No segundo, no Brasil, o Curso de Documentação Científica passou para Curso de Especialização em Documentação e Informação (1984). No que diz respeito ao Mestrado em Ciência da Informação, que se iniciou nesse país em 1970, na segunda metade da década de 80 dá-se início à fase de consolidação do mesmo (L. V. R. Pinheiro, 1997, pp. 87-88).

Cada um dos três momentos considerados na abordagem do conceito CI é identificado com um título, a saber: “Pré-CI”, “CI inicial” e “CI contemporânea”, respetivamente para o primeiro, segundo e terceiro momentos. Note-se que os três momentos estabelecidos e suas designações deverão ser encarados como um expediente operatório para facilitar a abordagem do conceito e não como uma divisão/rutura efetiva dos acontecimentos.

2.1.2. Seleção de obras

Ciente que uma abordagem conceptual, mesmo pouco aprofundada como a pretendida, imbrica em questões epistemológicas foi, nesse sentido, necessário encontrar critérios objetivos e uniformes para a seleção dos inúmeros trabalhos (empiricamente constatados durante a pré-análise) que se debruçam sobre a génese da CI⁴⁴ e suas relações disciplinares.

⁴⁴ Já em 1977 Shera e Cleveland apontavam para a quantidade de abordagens sobre as origens da CI: “It is amazing how many people discussed the foundations of information science, some qualified to do so and some not so qualified.” (1977, p. 260).

A CI é considerada por alguns autores como uma das áreas científicas mais introspectivas (R. R. Souza, Almeida, & Baracho, 2013, p. 160). A abundância de obras encontradas é propícia à “bulimia livresca” alertada por Quivy e Campenhoudt pelo que os procedimentos que a seguir se descrevem visam atender às recomendações desses autores sobre a “seleção cuidadosa de um pequeno número de leituras” (Quivy & Campenhoudt, 1992, p. 51).

Desde modo na seleção das obras, optou-se por procurar uma base de referência comum que reunisse as características: representatividade, homogeneidade, pertinência e exaustividade, de acordo com as regras de Bardin (2011, pp. 126–128) já aqui enunciadas. Para o efeito utilizou-se as referências bibliográficas dos 17 artigos reunidos sob o título “O Estatuto Epistemológico da Ciência da Informação” que constam na obra datada de 2009, *A Ciência da Informação criadora de conhecimento – volume I*, coordenada por Maria Manuel Borges e Elias Sanz Casado (cf. Apêndice 1).

O processo de seleção das obras presentes nas bibliografias referidas passou por várias etapas que, sucintamente, se passam a descrever:

1ª – mapeamento dos autores presentes nas referências bibliográficas dos 17 artigos – 256 autores;

2ª – seleção dos autores mais referidos utilizando a média de referências por autor (9 ou mais entradas)⁴⁵, independente do local do artigo em que consta e do número de artigos onde foram encontradas – 43 autores;

3ª – seleção de autores utilizando a média de artigos onde os autores são referidos (4 ou mais artigos)⁴⁶ – 23 autores;

4ª – das 60 obras dos 23 autores selecionados, optou-se por aquelas que são referidas em pelo menos 2 bibliografias dos 17 artigos de base – 25 obras⁴⁷;

5ª – nesta etapa, de natureza mais qualitativa, procurou descartar-se as obras cuja temática excedia claramente o âmbito pretendido para a presente análise (ver ponto 2.1) – 17 obras⁴⁸.

Relativamente às obras excluídas da lista, merece uma observação particular a obra “A Teoria Matemática da Comunicação” (1949) de Shannon e Weaver, dado a sua importância para a fundamentação teórica da CI numa fase inicial (até meados da década de 70) a que alguns autores chamam de “paradigma físico” (Capurro, 2003) ou “modelo físico” (Araújo, 2013, pp. 9, 22). Todavia, para o presente trabalho interessa em particular verificar o entendimento dos autores selecionados sobre a obra em questão mais do que propriamente a sua análise dado o facto de esta apresentar um cariz matemático bastante restrito. A esse

⁴⁵ Como é sabido a média é bastante influenciada pelos valores extremos o que, neste caso em particular, foi particularmente útil pois a intenção foi encontrar os autores “mais referidos” e, de facto, observando os valores da mediana e da moda, 2 e 1, respetivamente, poder-se-á afirmar que o valor da média, 9 (arredondado às unidades), marca a diferença.

⁴⁶ Neste caso pretendia-se um valor uniforme, pelo que o valor da média, 3,9, afigura-se adequado dado o valor 4 da mediana e a distribuição bimodal, 3 e 4, que a série apresenta.

⁴⁷ Apesar dos autores selecionados serem referidos em 4 ou mais artigos, em alguns casos os mesmos não eram mencionados nas bibliografias respetivas. As 60 obras iniciais apresentam uma presença, nas bibliografias dos 17 artigos, que varia de 1 a 4, pelo que se optou por selecionar, nesta fase, as obras com mais de uma presença.

⁴⁸ As obras retiradas da lista foram: “A sociedade em rede” de M. Castells; “A estrutura das revoluções científicas” de T. Kuhn; “Epistemologia da Interdisciplinaridade” e “Interdisciplinaridade: conceito, problema e perspectiva” de O. Pombo; “Ciência com consciência”, “Introdução ao pensamento complexo” e “Sobre a interdisciplinaridade” de E. Morin; e “A Teoria Matemática da Comunicação” de C. Shannon & W. Weaver.

propósito Shera e Cleveland (1977, p. 261), citando Fairthorne, referem existir uma extrapolação puramente retórica do âmbito bem mais estrito da teoria de Shannon⁴⁹.

Ao confrontar a data de publicação das 17 obras com os três momentos estabelecidos verificou-se uma concentração no terceiro período, mais especificamente após 1990, e a inexistência de obras na lista cuja data se enquadra no primeiro momento, i.e, anteriores a 1964. Tal constatação é compatível com a hipótese formulada, reforçando também os títulos atribuídos aos três momentos temporais.

Atendendo aos anos das obras e ao facto de existirem na lista quatro autores com mais de uma obra (Capurro, Pinheiro, Saracevic e Wersig), optou-se por seleccionar apenas uma obra de cada um desses 4 autores e introduzir 4 de outros autores que, embora não tenham sido considerados nas etapas anteriores, representassem alguma característica pertinente para a lista, como o país de origem ou a data da obra. Procedeu-se, dessa forma, a duas etapas extra, a da seleção das obras, com a mesma autoria, a permanecer na lista (6ª etapa) e a de inclusão das obras introduzidas (7ª etapa).

6ª – o critério para a seleção da obra, de cada autor repetido, a permanecer na lista relacionou-se com o período onde a mesma se inseria, optando-se pela obra correspondente ao período que apresentasse um menor número de trabalhos, desde que a diferença do número de presenças das obras em causa, nas diferentes bibliografias, não fosse superior a um⁵⁰. No caso das obras, do mesmo autor, que se situavam num mesmo momento temporal seleccionou-se a obra com o maior número de presenças nas bibliografias. Após a seleção a lista ficou reduzida a 11 obras⁵¹;

7ª – para a inclusão verificou-se, dos 43 autores referidos na 2ª etapa, ordenados pelo número de referências contabilizadas, quais eram aqueles que, não estando já incluídos, apresentavam as características referidas – país de origem ainda não incluído ou data das obras anterior a 1964. Da aplicação deste critério sobressaíram 4 autores: Otlet, Briet, Mikhailov e López Yepes⁵². Os dois primeiros reunindo as duas características consideradas, os outros dois autores apresentavam a característica relacionada com o país⁵³. Com esta inclusão a lista passou a 15 obras (cf. Tabela 8).

⁴⁹ Claude Shannon é considerado o autor da teoria uma vez que publicou uma primeira versão em 1948 no artigo intitulado "A Mathematical Theory of Communication", i.e., um ano antes da obra publicada em conjunto com Warren Weaver, cujo título apresenta uma subtil diferença, o artigo inicial passou de indefinido a definido: "The Mathematical Theory of Communication".

⁵⁰ Sendo a variação do número de presenças nas bibliografias de referência entre 2 e 4, este critério apenas não se aplicaria numa situação: uma das obras com 2 e a outra com 4.

⁵¹ A razão pela qual restaram 11 obras e não 13, ou seja, 17 iniciais menos 4 (uma de cada autor repetido), deve-se ao facto de um dos quatro autores, Pinheiro, apresentar um total de 4 trabalhos na lista, resultando a soma em 6 obras a retirar. As obras retiradas foram: de Pinheiro – "Traçados e limites da Ciência da Informação" (1995), "Campo interdisciplinar da Ciência da Informação: fronteiras remotas e recentes" (1998), "Ciência da Informação: desdobramentos disciplinares, interdisciplinaridade e transdisciplinaridade" (2006); de Capurro – "The Concept of Information" (2003); de Saracevic – "Information Science" (1999); e de Wersig – "Information science: the study of postmodern knowledge usage" (1993).

⁵² Os autores não incluídos na lista das 25 obras resultantes da 4ª etapa e que, apresentando um número igual ou superior ao dos 4 seleccionados, não o foram são: Japiassu, Popper, Bachelard, Bunge e Le Goff.

⁵³ As obras de Mikhailov e de López Yepes analisadas não correspondem às referidas nas bibliografias de base uma vez não ter sido possível ter acesso às mesmas em tempo útil. Na escolha das obras de substituição procurou-se trabalhos que, atendendo à informação disponível, abordam o assunto numa perspectiva disciplinar.

Tabela 8

Obras selecionadas para servirem de base à abordagem ao conceito de Ciência da Informação

Ref.	Autoria	TR	Ano	Obra	QB
Obras selecionadas que constam nas bibliografias utilizadas como base de referência:					
1A	Otlet	34	1934	<i>Traité de documentation</i>	1
1B	Briet	17	1951	<i>Qu'est-ce que la documentation?</i>	2
2A	Borko	18	1968	<i>Information science: What is it?</i>	3
2B	Wersig	28	1975	<i>The phenomena of interest to Information science</i>	2
2C	Belkin	15	1976	<i>Information science and the phenomenon of information</i>	3
2D	Shera	17	1977	<i>History and foundations of Information science</i>	3
2E	Brookes	16	1980	<i>The Foundations of Information science</i>	2
3A	Saracevic	51	1991	<i>Ciência da informação: origem, evolução e relações</i>	4
3B	Le Coadic	11	1994	<i>A Ciência da Informação</i>	4
3C	Pinheiro	48	1997	<i>A Ciência da Informação entre sombra e luz</i>	4
3D	Silva	31	2002	<i>Das "Ciências" Documentais à Ciência da Informação</i>	3
3E	Araújo	11	2003	<i>A ciência da informação como ciência social</i>	2
3F	Capurro	42	2003	<i>Epistemologia e Ciência da Informação</i>	3
Obras selecionadas de substituição:					
2s	Mikhailov	17	1971	<i>An Introductory Course on Informatics/Documentation</i>	-
3s	López Yepes	19	1995	<i>La Documentación como disciplina</i>	-
Obras substituídas:					
x1	Mikhailov	-	1973	<i>Fundamentals of Informatics</i>	1
x2	López Yepes	-	1978	<i>Teoría de la Documentación</i>	1

Nota. TR: total de referências contabilizadas nos 17 artigos independente do local onde a referência foi encontrada; QB: quantidade de bibliografias onde a obra se encontra. As obras realizadas em coautoria foram as seguintes (Ref – coautor(es)): 2C – Neveling; 2D – Cleveland; 2E – Robertson; 3D – Ribeiro; 2s – Giljarevskij; x1 – Chernyi, e Giljarevskij. Os títulos encontram-se na língua da edição consultada o que, nos casos das obras com as referências: 3A, 3B e 3F, corresponde a uma tradução e não ao original. Foram omitidos os títulos secundários de forma a encurtar a tabela, os títulos completos poderão ser consultados nas referências bibliográficas do presente trabalho. O ano apresentado corresponde à primeira edição das respectivas obras.

Estas 15 obras forneceram a base para a abordagem ao longo da linha evolutiva, mas não constituíram as únicas fontes de apoio, tendo-se recorrido a outras sempre que se considerou oportuno. A distribuição das 15 obras, relativamente aos três momentos e respetivos eventos descritos no ponto 2.1.1, pode ser visualizada na Figura 9.

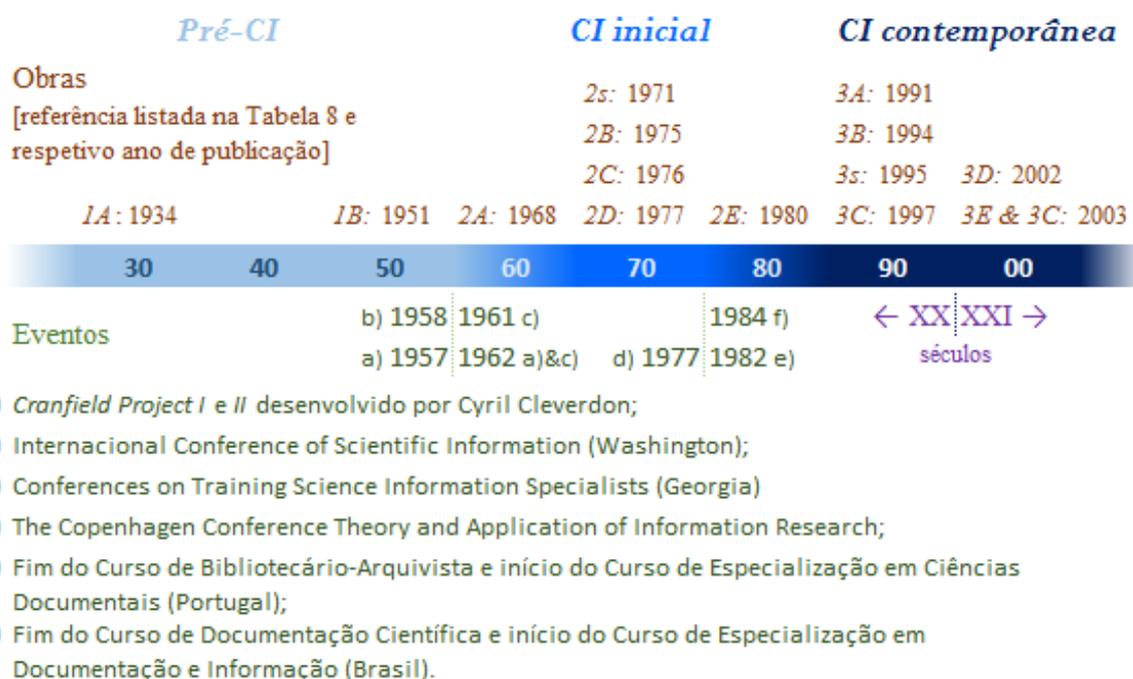


Figura 9. Distribuição das 15 obras selecionadas relativamente aos três momentos e respetivos eventos descritos.

2.2. Ciência da Informação: análise concetual

Estabelecida a linha temporal e obras de referência para análise, considerou-se útil partir de uma definição do conceito de CI que pudesse ser considerada representativa e, assim, centrar a análise das obras na contextualização disciplinar da área.

Para o efeito recorreu-se ao estudo “Knowledge Map of Information Science”, efetuado por Chain Zins entre 2003 e 2005, baseado num painel composto por 57 investigadores de renome, relacionados com a CI, de 16 países. Os resultados do referido estudo foram publicados em quatro artigos⁵⁴, em 2007, onde é apresentada uma visão *mainstream* da área no início do século XXI (Zins, 2007a, 2007b, 2007c, 2007d).

Como referencial para a abordagem ao mapa do conhecimento da CI, elaborado por Zins, interessa contextualizar o mesmo com os modelos conceituais quer da própria Ciência da Informação, quer dos seus componentes elementares: Dado (D), Informação (I) e Conhecimento (C) (Zins, 2007c, p. 479).

2.2.1. Dado, Informação e Conhecimento

De acordo com os resultados apresentados por Zins a perspetiva mais comum no campo da CI no que se refere aos conceitos dado, informação e conhecimento (D-I-C), pode ser

⁵⁴ Os artigos são: “Conceptions of Information Science”; “Conceptual Approaches for Defining Data, Information, and Knowledge”; “Classification Schemes of Information Science: Twenty-Eight Scholars Map the Field”; e “Knowledge Map of Information Science”. Todos os quatro artigos foram publicados em 2007 no *Journal of the American Society for Information Science and Technology*, vol.58(4).

caraterizada como uma *abordagem não metafísica, centrada no ser humano, de base cognitiva tipo proposicional*:

At this point, we can characterize the most common conceptual approach for defining data, information, and knowledge in the context of information science. Undoubtedly, the most common conceptual approach that represents the mainstream of the field is characterized as the non-metaphysical, human-centered, cognitive-based, propositional approach. (Zins, 2007c, p. 488).

Especificando ainda mais essa abordagem, de acordo com as definições propostas pelo painel de investigadores, Zins considerou uma outra divisão: os domínios subjetivo e universal⁵⁵:

The division between D-I-K in the subjective domain (SD), namely, D-I-K as inner phenomena bound in the mind of the individual knower versus the universal domain (UD), namely, D-I-K as external phenomena to the mind of the individual knower, (Zins, 2007c, p. 488).

O posicionamento, por parte dos investigadores, dos fenómenos⁵⁶ D-I-C como internos (domínio subjetivo) ou externos (domínio universal) ao indivíduo, ou nas duas situações simultaneamente, levou Zins à construção de 5 modelos (cf. Tabela 9).

Tabela 9

Distribuição dos fenómenos Dado (D), Informação (I) e Conhecimento (C) nos domínios universal e subjetivo de acordo com os modelos identificados por Chain Zins

Modelos	Domínio Universal	Domínio Subjetivo
1	D-I	C
2	D	I-C
3	D-I-C	I-C
4	D-I	D-I-C
5	D-I-C	D-I-C

Nota. Adaptação do quadro: “FIG.1. Four models for defining data (D)–information (I)–knowledge (K).” (Zins, 2007c, p. 489).

Zins apresenta o modelo 1 como o mais comum entre os investigadores que participaram no estudo embora posicione a sua própria definição no *modelo 5*. Os passos descritos pelo autor na apresentação do modelo mais comum são apresentados na Figura 10.

⁵⁵ Zins optou por essa designação, universal, em detrimento de *objetivo* por considerar esta última mais atreita a interpretações diferentes da pretendida (2007c, p. 486).

⁵⁶ Utiliza-se o termo *fenómeno* para caracterizar os conceitos de dado, informação, conhecimento e, também, mensagem, tal como o faz Zins (2007b, p. 340). Embora o autor não explicito o conceito de fenómeno faz-se o mesmo entendimento que o empregue por Silva e Ribeiro (2008): “Se olharmos para um objeto, o objeto tal como é experimentado por mim é um fenómeno.” (p. 32, em nota de rodapé).

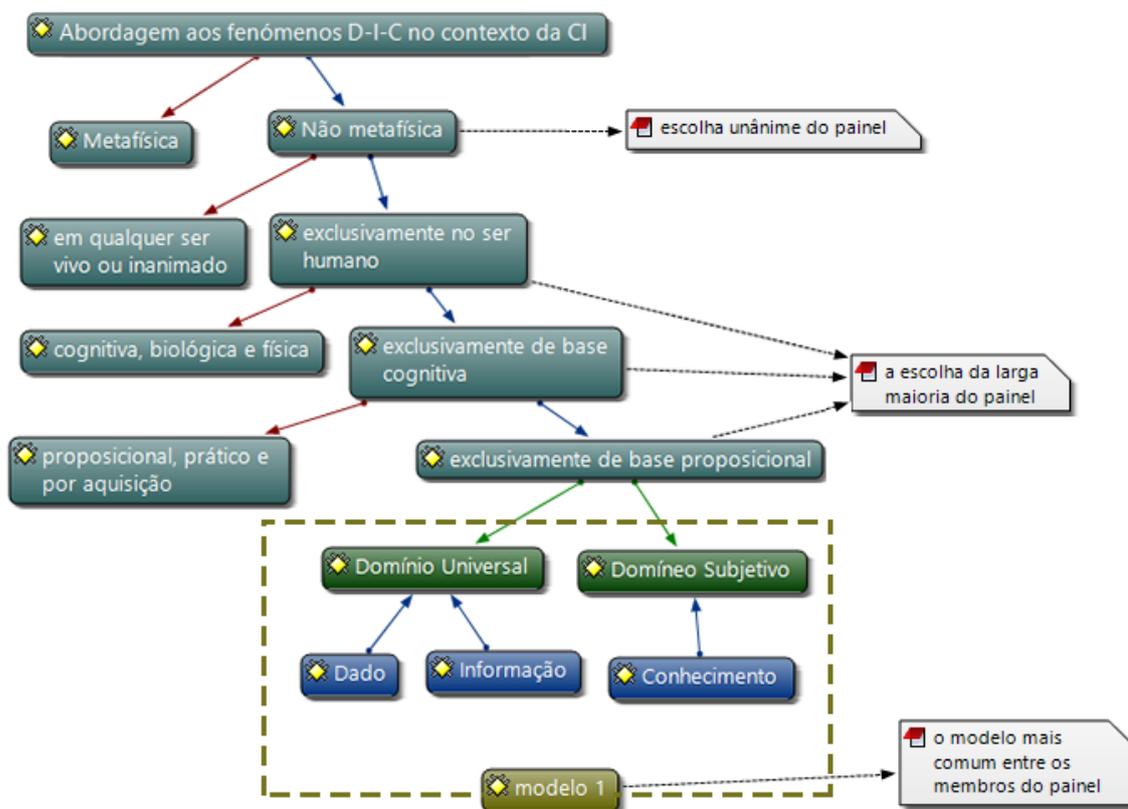


Figura 10. Esquema da abordagem dos fenômenos D-I-C no contexto da CI elaborado a partir da análise de Chain Zins (2007c, pp. 487–489)

A existência de diferentes modelos, mesmo após o filtro da abordagem mais comum, mostra uma diversidade de perspectivas quanto às definições dos conceitos de D-I-C, complexificando a determinação do foco da CI no que respeita aos “fenômenos explorados” (Zins, 2007b, p. 340). Complexidade que é aumentada por dois fatores: a entrada de um quarto fenômeno – Mensagem (M); e o uso “abusivo” ou “indevido” de terminologia nas definições dos respectivos conceitos, razão pela qual Zins optou por uma definição abrangente e não específica no que respeita a esta questão:

First, agreement has to be reached on the explored phenomena: data versus information versus knowledge versus message (D-I-K-M). Nevertheless, analysis of the panel’s definitions of D-I-K-M made it clear that the wording can be deceptive. Panel members often misused the terminology. Therefore, I adopt Begthol’s definition as an ad hoc position that IS explores D-I-K-M phenomena, without differentiating, however defined and in whatever relation to each other. (Zins, 2007b, p. 340).

2.2.2. Domínio e âmbito da Ciência da Informação

Para além dos fenômenos explorados, as outras duas questões, consideradas por Zins como essenciais para a definição do conceito de CI, são: o domínio do campo científico e o âmbito da exploração:

Based on the panel discussions, conceptions of information science differ mainly on three key issues: phenomena, domain, and scope: What are the explored phenomena? What is the domain of the field? What is the scope of the exploration? (...) Resolving the three

issues is crucial. It underlies six generic conceptions, or models, of Information Science. (Zins, 2007b, p. 340).

Relativamente ao âmbito, Zins identificou dois tipos: um mais restrito, o âmbito do processo de mediação (*mediating scope*), que se foca nos aspetos intermediários dos fenómenos D-I-C-M e não nos fenómenos em si, e outro mais abrangente, o âmbito inclusivo (*inclusive scope*), que abrange todos os aspetos dos fenómenos D-I-C-M. Cada um destes dois âmbitos incluem três modelos com diferentes domínios de ação, perfazendo seis modelos conceituais para a CI.

O autor relaciona estes dois âmbitos com duas CIs bastante díspares. No *âmbito do processo de mediação*, ao focar-se nos aspetos de intermediação como, por exemplo, naqueles aspetos envolvidos no ato de facilitar a ligação entre os fenómenos D-I-C-M e os utilizadores, a CI distingue-se de outros campos científicos. Já no *âmbito inclusivo*, a CI torna-se um nome genérico, um “guarda-chuva”, que abrange todos os campos⁵⁷ que estudam os aspetos dos fenómenos D-I-C-M:

To summarize, according to the three mediating conceptions information science is one field next to other fields, which explore the various perspectives of the D-I-K-M phenomena, whereas according to the three inclusive conceptions information science is a name for a generic field that comprises all the other fields that explore the various perspectives of the D-I-K-M phenomena. (Zins, 2007b, p. 341).

Os seis modelos identificados por Chain Zins podem ser hierarquizados em categorias de abrangência⁵⁸, do mais restrito ao mais abrangente: 1 – Tecnológico Digital (*Hi-Tech*); 2 – Tecnológico Global (*Technology*)⁵⁹; 3 – Cultural/Social (*Culture/Society*); 4 – Mundo Humano (*Human World*); 5 - Mundo dos Seres Vivos (*Living World*); 6 - Mundo Natural (*Living & Physical World*), (cf. Figura 11).

Destes seis modelos, quatro demonstraram ser mais significativos em termos da sua presença nas definições apresentadas pelo painel de investigadores, os modelos: Tecnológico Digital, Cultural/Social; Mundo Humano; e Mundo dos Seres Vivos. A presença dos outros dois modelos, o Tecnológico Global e o Mundo Natural, é bastante residual, sendo o modelo 3, o modelo Cultural/Social, o que representa a visão *mainstream* da área:

In the analysis of the panel members’ responses, the hi-tech model, the culture model, the human world model, and the living world model emerged as more significant. The vast majority of the panel responses—as well as myself—represent the culture model.

⁵⁷ Segundo Zins o “guarda-chuva” (o autor utiliza esse mesmo termo) da CI de âmbito inclusivo cobre *todos* os campos científicos relacionados com o estudo dos fenómenos D-I-C-M: “in the inclusive models, cognitive sciences, epistemology, philosophy of science, sociology of knowledge, education, linguistics, semiotics, logic and the like are subfields of information science. Accordingly, all these fields are **information sciences** [ênfase adicionado].” (Zins, 2007b, p. 341).

⁵⁸ Embora Zins apenas explicita a hierarquia entre os três modelos de domínio mais restrito: “hi-tech is a sub-category of technology, and technology is a sub-category of culture.” (2007b, p. 349), da descrição dos seis modelos infere-se o esquema apresentado.

⁵⁹ A opção pelas designações “tecnológico global” e “tecnológico digital” deriva da explicitação apresentada por Zins: “*Technology* is the various tools created by humans. *Hi-tech* is computer/electronics based technology” (2007b, p. 349).

Although the study is qualitative, it seems that the culture model represents the mainstream of contemporary information science. (Zins, 2007b, p. 341)

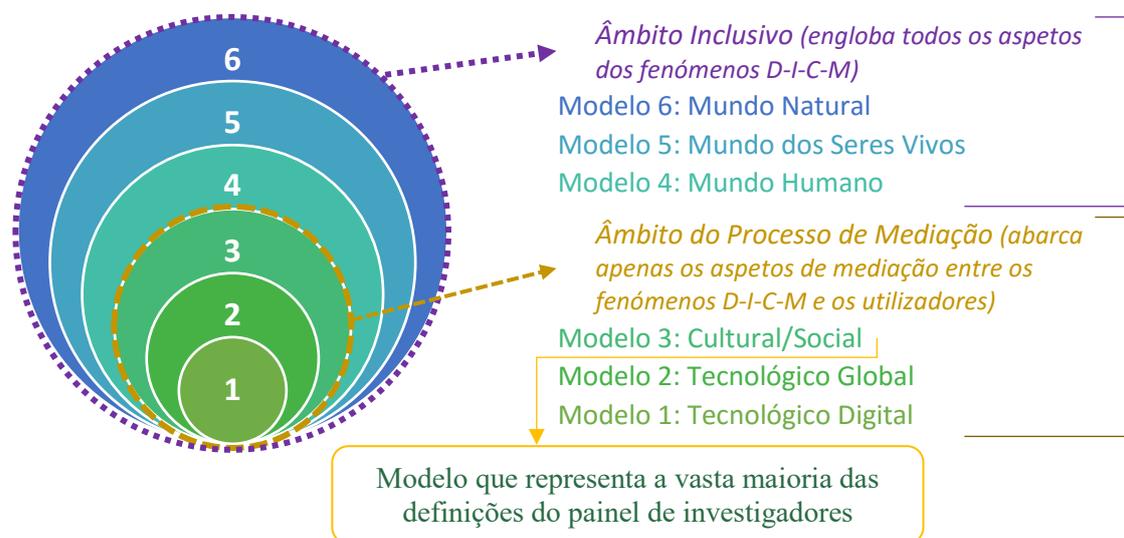


Figura 11. Esquema dos modelos conceituais para a definição de CI, identificados por Chain Zins, elaborado a partir da descrição dos mesmos (Zins, 2007b, pp. 340–350).

De acordo com o modelo que representa a visão *mainstream* da CI, o modelo Cultural/Social, a Ciência da Informação pode ser definida como o estudo dos aspetos de mediação dos fenómenos: dado, informação, conhecimento e mensagem, e sua implementação no domínio cultural (Zins, 2007b, p. 340). Cultural, neste contexto, é entendido como as várias formas da sociedade encarar a realidade (Zins, 2007b, p. 349).

2.2.3. Mapa do conhecimento da Ciência da Informação

O mapa do conhecimento da área da CI elaborado por Chain Zins, que a seguir se apresenta, reflete de forma “sistemática e abrangente” (Zins, 2007d, p. 526) o modelo Cultural/Social referido acima, pretendendo ser um instrumento para a compreensão da estrutura do domínio do conhecimento da CI e das relações conceituais entre os seus constituintes principais (Zins, 2007d, p. 531).

De acordo com o referido mapa, a CI comporta dois domínios de conhecimento: o metaconhecimento (*meta-knowledge*) e o conhecimento específico da área (*subject-based knowledge*). Esses domínios encontram-se divididos em 10 categorias principais, sendo a primeira, Fundações (*Foundations*), atribuída ao primeiro domínio, e ao segundo as restantes nove: Recursos/Fontes (*Resources*); Trabalhadores do Conhecimento (*Knowledge Workers*); Conteúdos (*Contents*); Aplicações (*Applications*); Operações & Processos (*Operations & Processes*); Tecnologias (*Technologies*); Contextos (*Environments*); Organizações (*Organizations*); e Utilizadores (*Users*) (Zins, 2007d, p. 528-529).

O primeiro domínio, Metaconhecimento, inclui as linhas de orientação, os pressupostos funcionais e princípios epistemológicos, metodológicos, conceituais, teóricos e históricos relevantes para o conhecimento da própria CI. O segundo domínio, Conhecimento Específico, diz respeito ao estudo dos aspetos e condições de mediação inerentes ao conhecimento humano no domínio universal.

As categorias principais (1º nível) contêm subcategorias organizadas em dois níveis. A categoria Fundações apresenta quatro categorias de 2º nível: Teoria (*Theory*); Investigação (*Research*); Educação (*Education*); e História (*History*). As outras nove categorias de 1º nível têm duas subcategorias, Questões (*Issues*) e Tipologias (*Types*), que se repetem em cada uma delas. Quanto ao terceiro nível, o mesmo é apresentado pelo autor como estando em desenvolvimento: “In many cases the third level is not fully developed, and is left for further development in future studies by the IS academic and professional community” (Zins, 2007d, p. 528).

O mapa exhibe, também, campos de intervenção da CI exemplificativos, associando-os a algumas categorias: Teoria da CI (*Theory of IS*) – Teoria (1.1); Metodologia da Investigação (*Research Methodology*) – Investigação (1.2); Ensino da CI (*LIS Education*) – Educação (1.3) e Profissionais do Conhecimento (3.); História da CI (*History of CI*) – História (1.4); Qualidade da informação e Sistemas de Informação (*Information Quality & Information Systems*) – Recursos/Fontes (2.); Ética da Informação (*Information Ethics*) – Profissionais do Conhecimento (3.) e Contextos (8.); Sociedade da Informação (*Social Informatics*) Contextos (8.); Estudos de usuários e Comportamento Informacional (*User Studies & Information Types Behavior*) – Utilizadores (10.), tal como se pode observar na Figura 12.

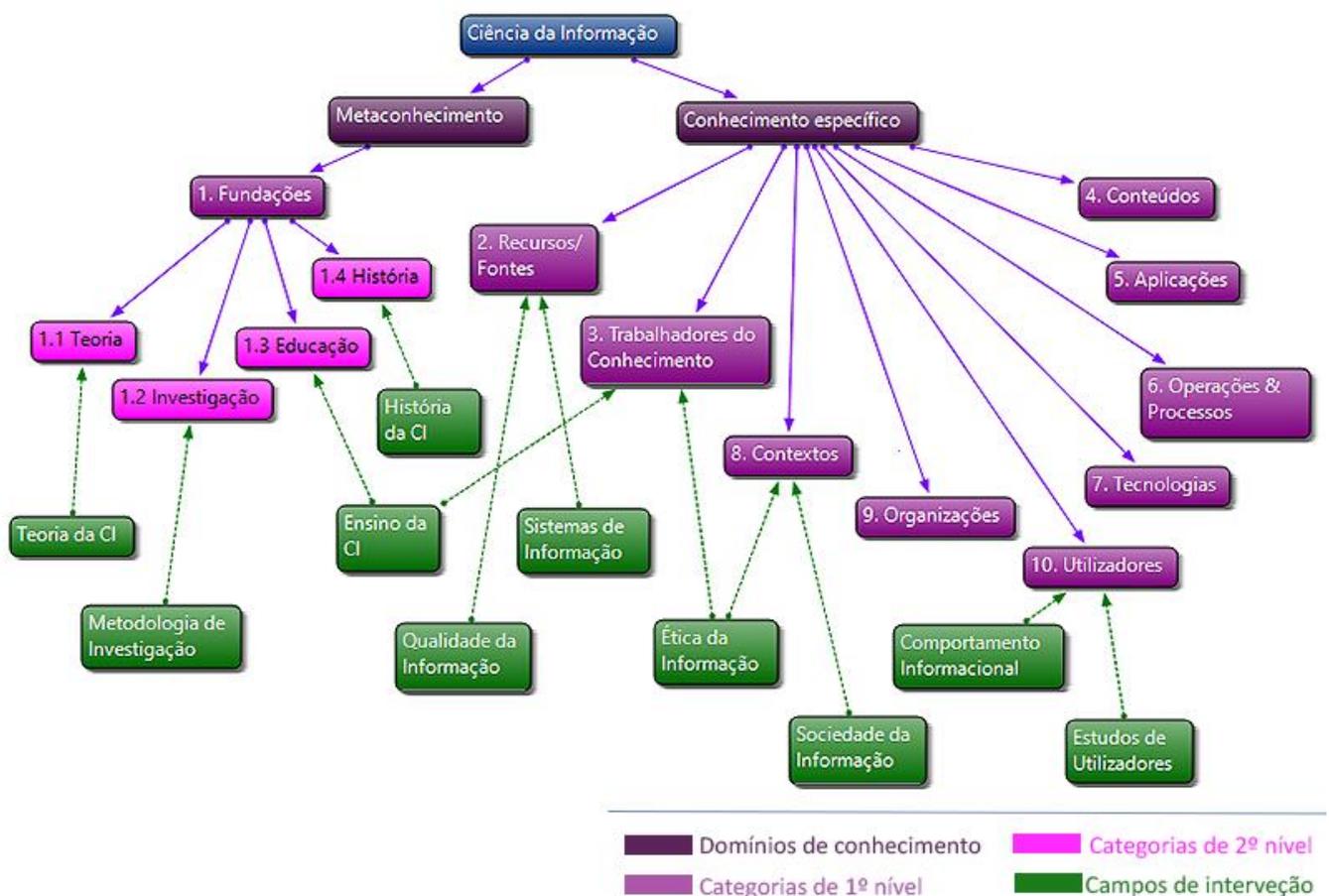


Figura 12. Mapeamento das relações entre os componentes apresentados no modelo de Chain Zins para o mapa de conhecimento da CI, elaborado a partir do quadro: “FIG. 2. Knowledge map of information science” (Zins, 2007d, p. 529)

A ordem das categorias principais não é aleatória, sendo evidente o posicionamento da categoria Fundações no início, já as restantes seguem a ordem de um plano de ação delineado

a partir das questões: “Quem? O quê? Porquê? Como? Onde? Quando?”, para a compreensão do processo de mediação entre as fontes, e outros recursos informacionais, e os destinatários finais, os utilizadores:

The nine elements are based on the following rationale. Information science explores the various conditions relevant for connecting resources (section 2) with users (section 10). (...) Sections 3 through 9 represent seven building blocks of the mediating process. To simplify the explanation of their order let us group them into two parallel sets of characteristics. The first set follows the “who, what, why, how, where, and when” order. The second set follows the equivalent “6 Ms” order, which is “mediator, matter, motive, method, means, and milieu.” (Zins, 2007d, p. 528)

A relação entre as 6 questões e os 6 “Ms” definidos por Zins⁶⁰ não é linear, nem tão pouco o é na relação com as categorias principais do domínio Conhecimento Específico, como é possível verificar na Figura 13.

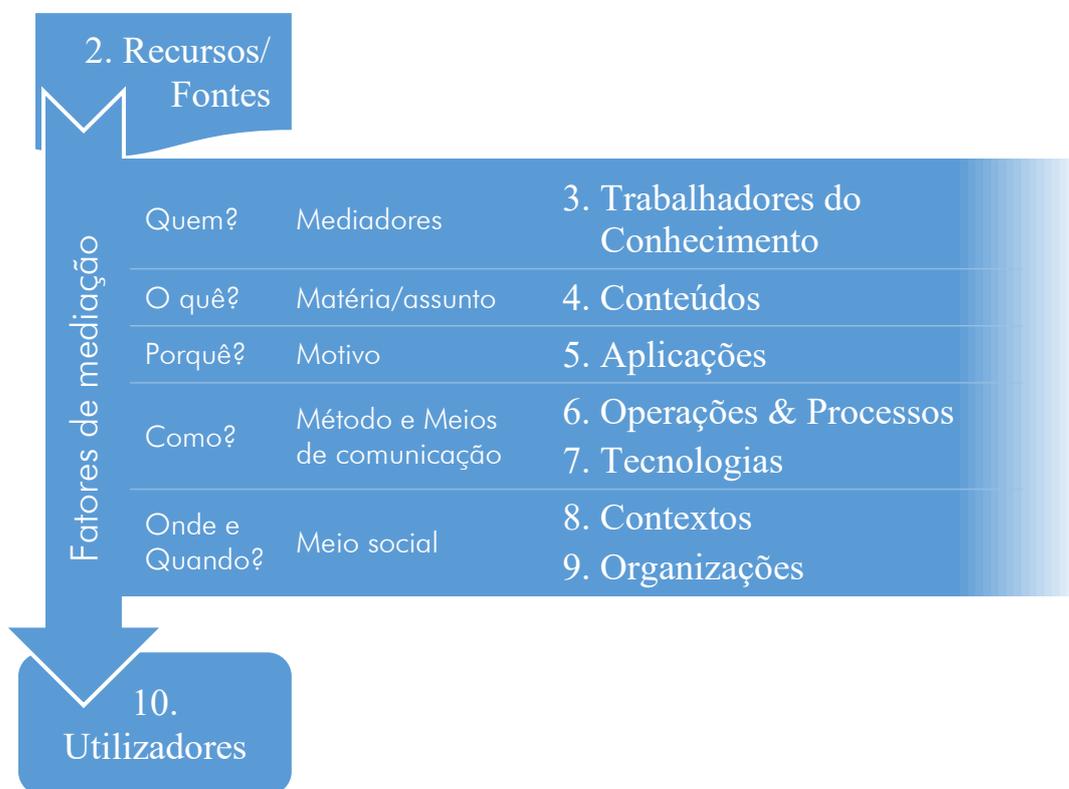


Figura 13. Esquema de relacionamento das características associadas às categorias principais do domínio Conhecimento Específico do Mapa do Conhecimento da CI, elaborado a partir do quadro: “FIG. 2. Knowledge map of information science” (Zins, 2007d, p. 529)

⁶⁰ A interpretação que se faz dos 6 “Ms” merece alguns reparos, nomeadamente no que diz respeito a: *matter* – matéria/assunto (o termo *assunto* tem como função associar o termo *matéria* ao significado *conjunto de assuntos*); *means* – meios de comunicação (no quadro “Knowledge map of information science” este “M” tem a seguinte apresentação “means (media)” (Zins, 2007d, p. 529)); *milieu* – meio social (tendo em conta as categorias associadas: Contextos (*Environments*) e Organizações (*Organizations*), a especificação do meio, o social, afigura-se adequado e mais explícito, sendo que, na relação com as questões “onde?” e “quando?”, apresenta uma abordagem em duas dimensões: a espacial e a temporal, necessária a uma correta contextualização desse meio.

A estrutura apresentada neste mapa é, segundo Zins (2007d, p. 528), capaz de representar as categorias e subcategorias de 26 esquemas de classificação da CI, de um total de 28⁶¹, elaborados pelos participantes no estudo. Esquemas onde, na maioria, uma das categorias principais inclui o conceito *conhecimento*, levando Zins a considerar reforçada a sua sugestão da mudança do foco do campo, assim como do próprio nome, de *informação* para *conhecimento* (Zins, 2007a, p. 668).

Assim, de acordo com o estudo de Chain Zins, a Ciência da Informação, ou a “Ciência do Conhecimento” como proposto pelo investigador (Zins 2007b, p. 335), poderá ser definida como: o estudo dos aspetos de mediação dos fenómenos: dado, informação e conhecimento enquanto artefactos humanos, e sua implementação no domínio cultural/social.

a. A questão da interdisciplinaridade

Relativamente à questão da interdisciplinaridade, que sobressaiu na pré-análise como uma característica marcante da CI, esta é apenas referida explicitamente em 4 das 50 definições apresentadas pelo painel de investigadores⁶². Não deixando de ser um facto relevante para uma reflexão sobre esta questão, é necessário ter em conta que alguns investigadores, como Le Coadic (1996, p. 22), apesar de não mencionarem a característica da interdisciplinaridade nas suas definições para o estudo de Zins são defensores da mesma.

Zins também não menciona o termo interdisciplinar, a sua abordagem vai no sentido de uma fundamentação teórica da CI com base em outros campos de conhecimento que inclui na categoria “Fundações”, dentro da subcategoria “Teoria”⁶³ (ver Figura 12). A considerar-se interdisciplinaridade seria dentro da alínea b) da enumeração efetuada por Lucinéia Bicalho:

A CI pode ser considerada interdisciplinar se forem considerados: a) a complexidade de seu objeto e dos problemas que deve tratar; b) a utilização diversificada de teorias e metodologias que têm origem em outras áreas para complementação dos estudos que realiza; c) a contribuição de várias áreas na sua constituição; e, d) a diversidade de origem de seus pesquisadores. (Bicalho, 2009, p. 248)

Importa referir que os resultados da investigação desta autora “não sustentam a característica interdisciplinar conferida à área” (Bicalho, 2009, p. 248), indo ao encontro do grupo de autores que considera que a CI não é interdisciplinar de todo ou, pelo menos, não da forma como esta é proposta (Fernandes & Cedón, 2009, p. 118). Por outro lado, Edivanio Souza e Eduardo Dias afirmam existir no Brasil “uma aproximação consensual por parte dos autores a respeito da natureza interdisciplinar da Ciência da Informação.” (2009, p. 140). Entre esses autores encontra-se um dos quatro participantes do estudo de Zins que explicitaram essa característica no seu conceito de CI, Lena Pinheiro, que advoga ser a CI uma “interciência” por natureza (2009, p. 110).

⁶¹ A razão para a exclusão dos dois modelos está relacionada, segundo Zins (2007a, p. 668), com o facto de um deles refletir o modelo 6, o Mundo Natural, e o outro ser “demasiado” generalista.

⁶² Os quatro investigadores são: Charles H. Davis, Donald Hawkins, Scott Seaman e Lena Vania Pinheiro. Os três primeiros definem a CI como um “campo interdisciplinar” já Pinheiro refere ser uma “abordagem interdisciplinar” (Zins, 2007b, pp. 337–339).

⁶³ Os 17 campos do conhecimento incluídos são: antropologia; artes; comunicação; ciência da computação; economia; educação; engenharia; história; direito; linguística; epistemologia; ética; ciência política; psicologia; metodologia da investigação; e sociologia (Zins, 2007d, pp. 337–339).

Tal como foi referido no ponto 2.1, embora esta característica não seja objeto de análise conceitual, dado a sua complexidade exceder os limites impostos ao presente trabalho, far-se-á uma breve retrospectiva histórica focada essencialmente no relacionamento entre a CI e as disciplinas associadas descritas nas obras enumeradas na Tabela 8.

2.3. Relações (inter)disciplinares

A revisão da literatura mostrou existir uma relação ancestral entre a Ciência da Informação e a tríade de disciplinas Biblioteconomia, Arquivística e Documentação, com frequência designadas pelo acrónimo BAD. Uma relação algo controversa resultando em diferentes conceções da própria CI e consequentemente nos cursos superiores a ela associada (Ribeiro, 2006; Stumpf, 2009).

2.3.1. Da *Pré-CI* à *CI inicial*

Segundo Silva e Ribeiro (2008, p. 47) a Ciência da Informação deriva “naturalmente” da *Documentação* existindo, de acordo com os mesmos autores, uma “ideia bastante consensual” de uma linha de continuidade entre ambas⁶⁴. Por outro lado, segundo Jesse Shera, citado por López Yepes, (1995, p. 116) essa mesma Documentação tem as suas raízes na Biblioteconomia começando, inicialmente, por ser essencialmente a mesma coisa para, na passagem do século XIX para XX, se assistir a uma demarcação da Documentação em relação à Biblioteconomia: “an entirely new body of techniques for organization, subject analysis, bibliographic description, and annotation that differed markedly from conventional library practice; and this new technology came to be called documentation.” (Shera & Cleveland, 1977, p. 250)

Na opinião de Shera, a Documentação não sugere uma nova ciência (*apud* López Yepes, 1995, p. 117). Neste sentido, López Yepes, (1995, p. 117), entende que Shera subordina totalmente a Documentação à Biblioteconomia. No entanto, observando as fases do desenvolvimento da Documentação descritas por Paul Otlet, verifica-se que este autor considera o contrário, i.e., que a Documentação integra a Biblioteconomia:

Et maintenant voici qu'une nouvelle phase est commencée. Ce n'est plus ni celle de la Bibliothéconomie, ni celle de la Bibliographie, c'est celle de l'ensemble du Livre et du Document, la Documentation. L'une et l'autre en sont des parties, mais des parties rattachées à un corps plus vaste, dont l'existence les élargit, les élève, les transforme. (Otlet, 1934, p. 17)

Apesar desta exposição, Otlet (1934, p. 6-7) não inclui explicitamente o termo “biblioteconomia” na apresentação das sete componentes da “sua” Documentação, ao contrário da Bibliografia que aí é incluída (item C). Os sete componentes são: A. *Les Documents particuliers*; B. *La Bibliothèque*; C. *La Bibliographie*; D. *Archives documentaires*; E. *Les Archives administratives*; F. *Les Archives anciennes*; G. *Les*

⁶⁴ Nas definições de CI, apresentadas pelos participantes do estudo de Chain Zins, os dois investigadores que abordam a história da área referem essa origem: Michael Buckland – “Information science is the field formerly known as *Documentation*, and now commonly referred to as *Information Science*.” e Birger Hjørland – “Historically, IS developed out of special librarianship and documentation.” (Zins, 2007b, pp. 336, 338).

documents autres que bibliographiques et graphiques; H. Les Collections Muséographiques; e I. L'Encyclopédie.

Embora, como foi referido, a Biblioteconomia não conste da lista, a descrição do item B, *La Bibliothèque*, aponta para uma relação direta com o conceito⁶⁵. Assim como também é possível relacionar, atendendo não só à designação das componentes como das respetivas descrições, as disciplinas Arquivística (itens F e E) e Museologia (item H).

A relação entre as quatro áreas, Documentação, Biblioteconomia, Arquivística e Museologia, referida acima, é explicitada por Suzanne Briet, cerca de 20 anos depois do “Tratado de Documentação” de Otlet: “Archiviste, bibliothécaire, conservateur de collection, notre documentaliste est tout à la fois”. (Briet, 1951, p. 20) E a autora faz a ponte para os “bibliotecários especialistas” que trabalham nos *information/intelligence officers* nos Estados Unidos ou na Grã-Bretanha, considerando-os “primos diretos” dos documentalistas franceses (Briet, 1951, p. 19).

A relação integradora de Otlet e Briet implica uma subordinação das outras disciplinas à Documentação. Como seria de esperar essa visão está longe de ser consensual. Referindo-se especificamente à relação entre a Biblioteconomia e a Documentação, López Yepes afirma ser esse “conflito” uma das fontes de ambigüidade terminológica relativamente à designação da área: “La situación se convirtió en conflictiva provocando ambigüedades terminológicas y conceptuales todavía vigentes con mayor o menor virulencia.” (López Yepes, 1995, p. 102)

Uma situação que, segundo Barreto, terá sido despoletada, em última análise, com a publicação do artigo “As we may think” de Vannevar Bush, em 1945:

[Vannevar Bush] advertiu, em seus escritos, que a base teórica utilizada na construção dos sistemas de classificação da informação além de ultrapassada estava errada. (...) Na Inglaterra e no resto do mundo os acontecimentos, desde a publicação do “*As we may think*” até a Conferência de 1948 da Royal Society, provocaram uma cisão com a Biblioteconomia que durou perto de 40 anos. (Barreto, 2008, p. 8)

Assim, as ambigüidades terminológicas, no que diz respeito às designações da área, foram aumentando de acordo com a importância dada ao conceito de *informação* em detrimento de *documento*: “cuando en el curso de la historia reciente de nuestra disciplina el contenido documental se ha visto desplazado por el contenido informativo, han irrumpido nuevas denominaciones, como Information Science, Informations-und Dokumentationswissenschaft e Informatika,” (López Yepes, 1995, p. 103). Nessa “continuidade” da Documentação para a CI Shera e Cleveland distinguem os Estados Unidos da América (EUA), da Europa e não só:

In Europe, in India (through the influence of S.R. Ranganathan), and to a large extent in Latin America, the understanding of what documentation is has remained relatively constant and fundamentally French. In the United States, however, documentation took a rather different turn, due largely to the development of fine grain photographic emulsions and the miniature camera using 35 millimeter motion picture film with an acetate, non explosive, base. (Shera & Cleveland, 1977, p. 252)

⁶⁵ O item B, “La Bibliothèque” apresenta a seguinte descrição: “C'est la collection des documents eux-mêmes maintenus chacun dans leur intégrité individuelle (Livres et publications diverses de toutes espèces). La collection est disposée en des réceptacles adéquats et rendus facilement accessibles (rayons, livres, magazines); elle est classé et cataloguée.” (Otlet, 1934, p. 6).

As técnicas de microfilmagem e posteriores desenvolvimentos tecnológicos no âmbito da recuperação de informação⁶⁶ (Shera & Cleveland, 1977, p. 256), no entender daqueles autores, marcarão a diferença da CI norte-americana relativamente à europeia. Contudo, mesmo nos EUA, outras disciplinas são consideradas como estando na gênese da CI. Borko e Doyle, em 1964, além da Documentação e da Recuperação de Informação, mencionam explicitamente a Comunicação e as Ciências Comportamentais (Shera & Cleveland, 1977, p. 266). A exclusão das outras duas disciplinas BAD, Biblioteconomia e Arquivística, nesta enumeração, parece refletir duas posições distintas as quais se passa a apresentar.

A Arquivística, no entender de Silva e Ribeiro (2008, p. 135), manteve-se, na maioria dos casos, à margem do desenvolvimento da CI⁶⁷, sendo entendida por muitos como uma disciplina auxiliar da História. Esta perspetiva pode ser observada na enumeração das quatro disciplinas consideradas por Le Coadic como as “primeiras disciplinas” relacionadas com a CI – biblioteconomia, museologia, documentação e jornalismo:

O que caracteriza as quatro disciplinas que foram atuantes, até o presente, no campo da informação - a biblioteconomia, a museologia, a documentação e o jornalismo é que todas atribuíram um interesse particularmente grande aos suportes da informação e não à própria informação. [e, em nota de rodapé, associada ao termo biblioteconomia:] A arquivística, disciplina auxiliar da história preocupa-se com a conservação dos documentos que resultam da atividade de uma instituição ou de uma pessoa física ou jurídica. Os arquivos não passam de documentos conservados, enquanto as bibliotecas são constituídas de documentos por elas reunidos. (Le Coadic, 1996, p. 14)

Nos EUA, país de origem de Borko e Doyle, tal perspetiva terá sido a razão para situar o ensino da Arquivística nos departamentos de História até meados do século XX, altura em que se iniciou uma progressiva integração nas escolas de Biblioteconomia e CI embora mais por questões economicistas do que epistemológicas (A. M. da Silva & Ribeiro, 2008, pp. 136–137). Conforma essa questão um estudo da Organização para a Cooperação e Desenvolvimento Económico (OCDE), realizado em 1970, sobre a harmonização dos cursos de CI norte-americanos no qual, de acordo com Silva e Ribeiro, não existe qualquer menção a temas arquivísticos nas matérias aí enumeradas, concluindo os autores que: “os arquivistas não eram considerados ‘especialistas em informação’ e que a sua formação *continuava* [ênfase adicionado], ainda nesta altura, arredada da área da CI.” (A. M. da Silva & Ribeiro, 2008, p. 138).

Relativamente à não inclusão da Biblioteconomia na enumeração de Borko e Doyle, esta poderá advir de uma “perspetiva outliana” de subordinação daquela disciplina à Documentação ou, em sentido contrário, desta última ser vista como um prolongamento específico da Biblioteconomia.

⁶⁶ O termo *information retrieval* terá sido cunhado por esta altura (1950) por Calvin Moores para designar “um novo tipo de pesquisa feita em fundos de informação, efetuada a partir da especificidade dos assuntos”, marcando a passagem de uma “indexação conceptual” para uma “indexação por palavra” (Simões, 2008, p. 31).

⁶⁷ É interessante verificar as relações disciplinares que o *clássico* manual “Fundamentos da disciplina Arquivística” apresenta: “A arquivística situa-se no cruzamento de novos contextos culturais, dos novos modos de gestão tal como as novas tecnologias. Ela está na confluência de várias disciplinas: informática, **ciências da informação** [ênfase adicionado], história, linguística, arqueologia, etnologia, etc. Ao serviço de algumas delas, ela tem por obrigação servir-se das outras, a fim de assegurar a evolução e o seu desenvolvimento.” (Rousseau & Couture, 1998, p. 54)

2.3.2. Da *CI inicial à contemporânea*

No artigo de 1968, “Information Science: What Is It?”, de Harold Borko, é possível interpretar uma relação do tipo de subordinação da Documentação, neste caso à CI. Uma relação diferente da que o mesmo artigo parece fazer para a CI e a Biblioteconomia. Nesse artigo a menção às duas disciplinas é feita no seguinte contexto:

[Information Science] is an interdisciplinary science derived from and related to such fields as mathematics, logic, linguistics, psychology, computer technology, operations research, the graphic arts, communications, library science, management, and other similar fields. (...) Librarianship and documentation are applied aspects of information science. (Borko, 1968, p. 3)

A Biblioteconomia surge como uma disciplina de origem e relacionada à CI⁶⁸ enquanto que a Documentação é, a par dos deveres do bibliotecário⁶⁹, considerada uma aplicação prática da CI. Aspeto reforçado pela definição apresentada no mesmo artigo para a Documentação: “Documentation is one of many applied components of information science.” (Borko, 1968, p. 5).

López Yepes tipificou as definições para o conceito de Documentação de acordo com a relação entre esta e a Biblioteconomia considerando existir três tipos de relação: de “subordinação”, de “sobreposição” ou de “justaposição” (López Yepes, 1995, p. 104). Na primeira a Biblioteconomia faz parte da Documentação, na segunda acontece o inverso e, na terceira, existe um “paralelismo” entre as duas disciplinas⁷⁰. É nesta última posição onde o investigador espanhol aparentemente se posiciona atendendo à corrente, da qual é líder (A. M. da Silva & Ribeiro, 2008, p. 52), defensora de uma “Ciência da Documentação” inserida nas Ciências da Comunicação Social ou Ciências (no plural) da Informação (López Yepes, 1995, p. 100). Esta corrente é algo contrária à “opinião consensual” de uma evolução do conceito e do âmbito da Documentação rumo à Ciência da Informação (Silva & Ribeiro, 2008, p. 51).

Se essa “opinião consensual” é possível em relação à Documentação já no que se refere à CI na sua relação com a Biblioteconomia tal não parece ser o caso. Tal situação prende-se, desde logo, do modo como a primeira aparentemente se “impôs” à segunda: “Institucionalmente, a Ciência da Informação se desenvolveu *ocupando* [ênfase adicionado] os espaços já constituídos pela Biblioteconomia (departamentos de universidades, periódicos, associações), embora dialogando apenas com uma parte dela,” (Araújo, 2013, p. 26). Se bem que essa “ocupação” se entende ter sido mais pela integração que pela substituição. Na década de 60 do século passado os “conteúdos de Ciência da Informação” começaram a ser inseridos nos cursos de Biblioteconomia (Araújo, 2013, p. 6) e, na década

⁶⁸ Comparando com a definição que saiu das conferências do Georgia Institute of Technology (1961-62) na qual a de Borko foi baseada (Shera & Cleveland, 1977, p. 265), é interessante verificar um pormenor na primeira definição, esta apenas apresenta as disciplinas como sendo *a origem* da CI e não como *a ela relacionadas*.

⁶⁹ Dado o contexto onde surge o termo *Librarianship* e seu uso específico no lugar do termo *library science*, optou-se pela tradução “deveres do bibliotecário” de acordo com a definição “the position or duties of a librarian”, apresentada em <http://www.dictionary.com/browse/librarianship> (consultado em 2-8-2016).

⁷⁰ López Yepes apresenta a sua tipologia da seguinte forma: Las definiciones discutidas pueden agruparse en dos tipos: 1) Definiciones en relación con la Biblioteconomía. 2) Definiciones sin relación con la Biblioteconomía. Las primeras pueden ser: a) Definiciones globales o de superposición; b) Definiciones paralelas o de yuxtaposición, y c) Definiciones subordinadas o de infraposición. (López Yepes, 1995, p. 104)

seguinte, as designações de cursos, escolas ou departamentos de instituições de ensino superior começaram a refletir essa mudança, adotando a denominação de *Libray and Information Science* (Araújo, 2013, p. 6; A. M. da Silva & Ribeiro, 2008, p. 134). Nome frequentemente usado nos EUA⁷¹ embora, segundo Araújo (2013, p. 6), seja principalmente na Inglaterra e nos países nórdicos que a área seja assim designada.

A renomeação de escolas e departamentos de Biblioteconomia ocorreu no Brasil apenas durante as décadas de 80 e 90 tendo um impacto diferente no nome dos cursos de acordo com a sua tipologia: “Os cursos de graduação em Biblioteconomia⁷² mantiveram, na grande maioria dos casos, sua denominação. Mas os cursos de pós-graduação tiveram também o nome alterado para Ciência da Informação.” (Araújo, 2013, p. 6-7) Já em Portugal essa alteração terá ocorrido mais tardiamente, já em pleno século XXI, tendo em conta o que Silva e Ribeiro em 2002 referem, a propósito do Curso de Especialização em Ciências Documentais (CECD):

As críticas formuladas nos diversos encontros em que o modelo de formação veiculado pelos CEDC foi discutido não têm contribuído para uma mudança de perspectiva, já que ele se tem mantido sem alterações até ao momento atual e, nos últimos anos, até inspirou a criação de múltiplos cursos, em variadas universidades do setor privado, com estrutura semelhante e *curriculum* idêntico (A. M. da Silva & Ribeiro, 2008, p. 149).

É interessante verificar que, ao contrário do Brasil, em Portugal a tendência da afirmação da designação *Ciência da Informação* de forma autónoma também se verifica ao nível da licenciatura⁷³. Essa afirmação é, no entender de Pinheiro, a opção correta uma vez que a união em termos de designação para uma só área implica negar a “independência científica da Ciência da Informação, com seu próprio estatuto científico” até porque, refere a autora, “a documentação surge da cisão com a Biblioteconomia, portanto, nasce da divergência” (L. V. R. Pinheiro, 1997, p. 246). Opinião semelhante é mostrada por Saracevic:

Embora a CI e a biblioteconomia sejam grandes aliadas, a ponto de muitos assumirem o termo *biblioteconomia e ciência da informação* para descrever um mesmo campo de estudos, na realidade, as diferenças apontadas são de tamanha importância qualitativa que desautorizam tal união (Saracevic, 1996, p. 49).

Essa posição de independência, ainda de acordo com Pinheiro, não nega as fortes relações interdisciplinares da CI. A par da relação que a CI assume com a Biblioteconomia a

⁷¹ A título de exemplo pode-se referir que a junção dos termos *Libray e Information* surge em metade dos nomes das escolas, colégios ou departamentos dos 20 cursos da área (dos EUA) melhores classificados em 2013, de acordo com a lista apresentada pela *U.S. News & World Report's*. Os restantes 10 (escolas, colégios ou departamentos) apenas apresentam o termo *Information* nas respetivas designações. (lista disponível em <http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-library-information-science-programs>, acessado a 3-8-2016).

⁷² Na listagem do Ministério de Educação do Brasil (consultada em <http://emec.mec.gov.br/>, a 3-8-2016), dos 49 cursos de Biblioteconomia e/ou CI: 41 mantêm a designação *Biblioteconomia* sem qualquer acrescento e apenas 1 ostenta o termo *Ciência da Informação* de forma autónoma. Dos restantes 7, 2 apresentam as duas designações (sendo que um deles utiliza o termo no plural *Ciências da Informação*); 3 juntam ao termo *Biblioteconomia* o termo *Documentação* e outros 2 o termo *Gestão de Unidades de Informação*.

⁷³ Veja-se a recente mudança da designação *Ciência da Informação Arquivística e Biblioteconómica* para *Ciência da Informação*, da licenciatura da Universidade de Coimbra, ocorrida no ano letivo 2015-2016, a designação anterior era única, em termos de nomenclatura similar, para os cursos desse grau académico de acordo com o sistema em linha da A3ES (acessado em <http://www.a3es.pt/pt/acreditacao-e-auditoria/resultados-dos-processos-de-acreditacao/acreditacao-de-ciclos-de-estudos> a 3-8-2016).

investigadora coloca a relação desta com a Ciência da Computação. Dado essa circunstância, em muitos casos, segundo Pinheiro (1997, p. 248), a CI foi “confundida” por aquelas duas disciplinas. Neste sentido, entende-se ser possível relacionar a identificação da CI, com a Biblioteconomia e com a Ciência da Computação, com as duas orientações ou “raízes” da CI: a primeira tem a ver com o comportamento informacional e a memória intelectual da civilização, e uma segunda, ligada às técnicas de recuperação de informação e aos sistemas de informação em ambiente digital (Capurro, 2003, Capítulo II; Pinheiro, 1997, p. 245; Saracevic, 2009, p. 15).

A orientação relacionada com as técnicas de recuperação de informação, muito vincada nos anos 60, leva Araújo (2003, p. 22) a afirmar que a CI não “nasce” como ciência social contribuindo para a mesma ser “confundida” com a Ciência da Computação. Esta “colagem” da CI à Computação/Informática foi reforçada pelo facto de Alexander Mikhailov ter adotado o termo *Informatika* para designar a CI na então União Soviética. Apesar das semelhanças terminológicas com o termo francês *Informatique*⁷⁴, o próprio Mikhailov apresenta o objetivo da disciplina como sendo o estudo da “informação científica” denominando os seus especialistas de “cientistas da informação”:

Informatics is a scientific discipline which studies the structure and characteristics (but not the specific contents) of scientific information, as well as the rules of scientific information work, its theory, history, methodology and organization. (...) We believe that both these terms – “documentation” and “documentalistics” - are inadequate, for they lay the stress on documents whereas the subject studied by the discipline in question is the structure and characteristics of scientific information. (...) Specialists in this new discipline - informatics - likewise must have a name. Our suggestions is “information scientists”. (Mikhailov & Giljarevskij, 1971, pp. 14-15)

A referência de Mikhailov à Documentação revela uma mesma ligação de origem, tal qual a CI ocidental e, como relações disciplinares, o autor refere: “mathematical information theory, cybernetics, semiotics, linguistics, psychology, library science, bibliography, book science, science of science, and several technical disciplines⁷⁵.” (Mikhailov & Giljarevskij, 1971, p. 17)

Pelo exposto, mais uma vez a Biblioteconomia é referida, ao contrário da Arquivística. Contudo, Wersig, um outro investigador natural da União Soviética (embora radicado na Alemanha), menciona a Arquivística, inserida no grupo das “suas” *Ciências da Informação* (Wersig & Neveling, 1975, Capítulo 3.3), no qual inclui: “information science (informatics), library science, museology, archivistics, education”. Estranha-se, contudo, a inclusão da disciplina Educação nesta lista pois, usando as palavras de Pinheiro (1997, p. 117), “existe uma diferença essencial entre os campos do conhecimento que se alimentam de informação,

⁷⁴ Curiosamente a entrada “Informatique” da enciclopédia em linha *Larousse* apresenta o título “La science de l'information” para a descrição da origem do termo: “Le terme *informatique*, qui désigne une discipline née avec l'ordinateur, est un néologisme français, introduit en 1962 par Philippe Dreyfus, condensant les mots *information* et *automatique*. Les Anglo-Saxons parlent de *computer science* et de *data processing*.”, informação recolhida em <http://www.larousse.fr/encyclopedie/divers/informatique/61302> a 4-8-2016.

⁷⁵ As *technical disciplines* ou *technical sciences*, como os autores também lhes chama, referidas são: *Microphotographs; Photography and electrical engineering; optics and electronics; electrography; e electronic digital computers*, (Mikhailov & Giljarevskij, 1971, p. 17).

praticamente todos, e aqueles cujo objeto de estudo é a informação, qualquer que seja a sua natureza.”

As *Ciências da Informação* de Wersig não incluem a Documentação, como acontece na área homónima referida por López Yepes pois, ao contrário deste, Wersig considera esta disciplina uma “área de trabalho prático” designando-a também de “recuperação de informação”. Acresce ainda, o referido autor, que foram as necessidades deste “trabalho prático” (Documentação) o motor do desenvolvimento da CI e não outros campos de estudo:

“Information science” did not develop out of another field of study (like psychology) or from the intersection of two fields (like bio-chemistry), but out of the needs of an area of practical work, called “documentation” or “information retrieval”. (Wersig & Neveling, 1975, Capítulo 1)

Esta delimitação da disciplina de Documentação a um trabalho prático intui ir ao encontro da afirmação de Shera e Cleveland (1977, p. 250): “In time, then, *documentation* came to be, through the efforts of FID, almost synonymous with classification, particularly the UDC⁷⁶.” Apesar da evidente ligação da documentação com a CDU pelo denominador comum – Paul Otlet (Simões, 2011, p. 178), uma tal restrição afigura-se demasiado limitadora da ação daquela disciplina. Revelando, os autores, uma posição convicta de subordinação da Documentação à Biblioteconomia evidente na seguinte passagem: “Finally, one waggish American librarian, whose name is now unhappily forgotten, has defined *documentation* as ‘Librarianship performed by amateurs’.” (Shera & Cleveland, 1977, p. 252)

Num cenário como o que foi apresentado não é de estranhar um panorama no ensino tal como o descrito por Brookes:

When visiting schools of information science in North America I have often been introduced to the faculty members in the following terms: “Here is Dr. A, he teaches *linguistics* for information science. And here is Prof. B who gives courses in *computer science* for the information scientists. Dr. C here is a statistician who has a course on *statistics* for information science.” And so it goes on until I am compelled to ask: “And who teaches *information science*?” (Brookes, 1980, p. 128)

Tal como se observa nesta citação (supra) à pergunta: “Quem ensina CI?”, a resposta comum naquela época era, segundo Brookes: “information science is a peculiar mix of linguistics, communication, computer science, statistics, research methods, together with some techniques from library science such as indexing and classification.” (Brookes, 1980, p. 128). Esta situação leva Brookes a afirmar que a CI, enquanto disciplina sem campo próprio, não tem futuro; neste sentido este investigador argumenta que a CI tem o seu próprio “território, problemas e visão dos assuntos humanos” necessitando, contudo, de ainda desenvolver os seus princípios e técnicas específicas (Brookes, 1980, p. 128).

2.3.3. Na passagem de século

Após cerca de quatro décadas, considerando o período de finais dos anos 50 a 1997, data da tese de Pinheiro, a investigadora considera que a CI “ainda pode estar no seu período de

⁷⁶ Afirmação feita a propósito da atuação da Federação Internacional de Documentação (FID) no seu esforço “quase exclusivo”, após a Segunda Guerra Mundial, de promoção da Classificação Decimal Universal (CDU ou UDC na língua inglesa) (Shera & Cleveland, 1977, p. 250).

emergência ou, no máximo, de evolução uniforme” (L. V. R. Pinheiro, 1997, p. 93). Apontando, nessa evolução, “novas articulações disciplinares, com a Comunicação, por exemplo, numa aproximação cada vez mais forte.” (L. V. R. Pinheiro, 1997, p. 248). A Comunicação junto com a Ciência da Computação formam com a CI, na opinião da investigadora, “um triângulo disciplinar altamente dependente da nova ordem tecno-cultural” (L. V. R. Pinheiro, 1997, p. 249).

Outro autor com opinião similar é Saracevic:

First, information science is interdisciplinary in nature. However, with various advances, relations with various disciplines are changing over time. The interdisciplinary evolution is far from over. Second, information science is inexorably connected to information technology. A technological imperative is compelling and encouraging the evolution of information science, as is the evolution of a number of other fields, and moreover, of the information society as a whole. (Saracevic, 2009, p. 15)

Como campos com os quais a CI desenvolve relações interdisciplinares mais “pronunciadas e significantes” esse autor aponta: “biblioteconomia, ciência da computação, ciência cognitiva (incluindo inteligência artificial - IA) e comunicação” (Saracevic, 1996, p. 48).

Com uma perspectiva semelhante no que se refere à característica da interdisciplinaridade, Le Coadic, nas treze disciplinas que enumera como sendo as que mais colaboram com a CI⁷⁷, apenas converge com Saracevic na Informática/Ciência da Computação (Le Coadic, 1996, p. 22). A ausência da Biblioteconomia naquela lista de treze itens pode estar no facto do investigador a considerar, juntamente com a Documentação, como as “sementes” da CI (cf. Figura 14).

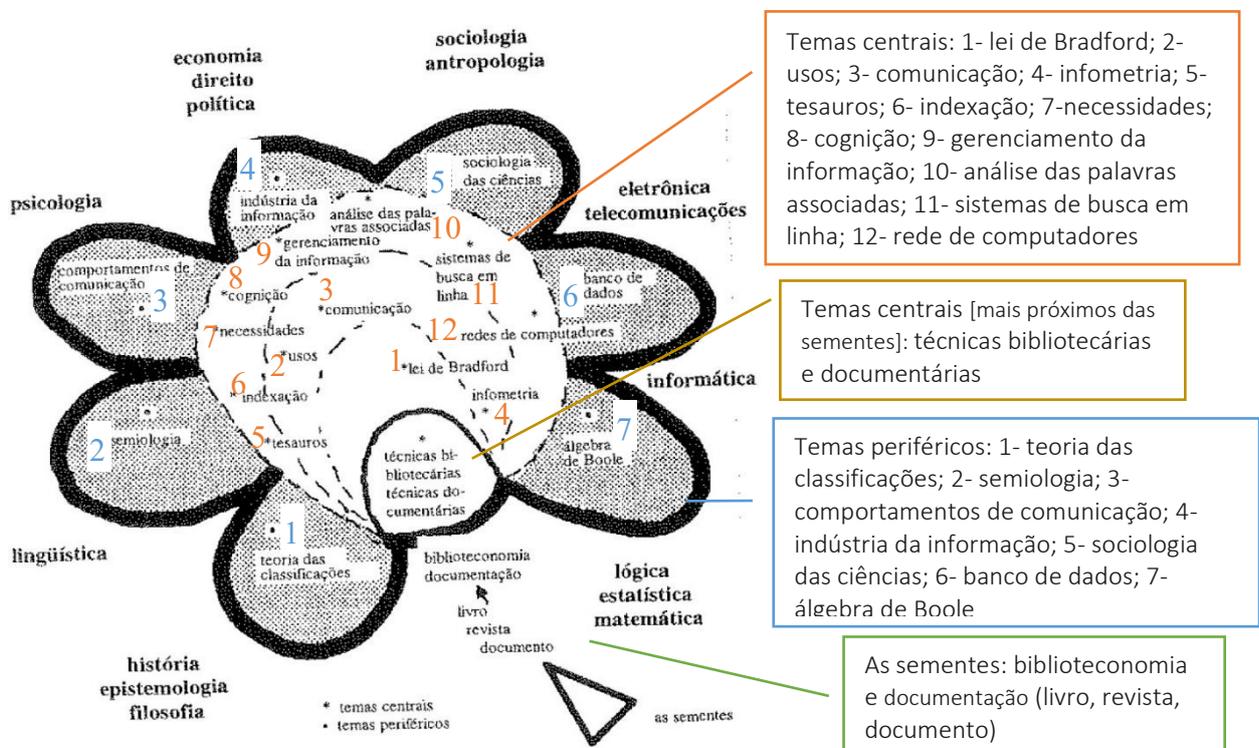


Figura 14. O mapa da Ciência da Informação segundo Yves-François Le Coadic (1996, p. 24).

⁷⁷ As disciplinas a que Le Coadic se refere são, pela ordem como surgem na fonte consultada: psicologia, linguística, sociologia, informática, matemática, lógica, estatística, eletrônica, economia, direito, filosofia, política e telecomunicações (Le Coadic, 1996, p. 22).

Contrariamente a Saracevic, que refere a Ciência Cognitiva e a Comunicação como disciplinas com ligações mais “pronunciadas e significantes” com a CI, Le Coadic não inclui essas duas disciplinas nos dezasseis campos do seu mapa da CI, contudo, coloca a “cognição” e a “comunicação” como “temas centrais” nesse mapa (itens 8 e 3, respetivamente, da Figura 14). Neste contexto, este investigador dá importância a cinco campos em particular: Psicologia, Sociologia, Economia, Informática e Telecomunicações, considerando que os estudos realizados por investigadores destas áreas “contribuíram em muito” para a tornar a CI uma “ciência social rigorosa” (Le Coadic, 1996, p. 22). Para Le Coadic estes relacionamentos podem ser entendidos do seguinte modo: “Seu conteúdo [da CI], marcado pelo selo da interdisciplinaridade, é uma sábia dosagem de ciências matemáticas e físicas, bem como ciências sociais e humanas.” (Le Coadic, 1996, p. 109).

Para finalizar esta questão apresenta-se a posição da investigadora brasileira Lena Pinheiro e dos investigadores portugueses Malheiro da Silva e Fernanda Ribeiro. A da primeira pode ser entendida na seguinte passagem introdutória:

Durante vinte anos de estudos de Ciência da Informação, nossa percepção é de que a Ciência da Informação tem seu próprio estatuto científico, como ciência social que é, portanto, interdisciplinar por natureza, e apresenta interfaces com a Biblioteconomia, Ciência da Computação, Ciência Cognitiva, Sociologia da Ciência e Comunicação, entre outras áreas, e suas raízes, em princípio, vêm da bifurcação da Documentação/Bibliografia e da Recuperação da Informação. (Pinheiro, 1997, p. 1)

A ausência da Arquivística nesta descrição poderá ter relação com a “visão americana” da CI, já para os investigadores portugueses: “a inevitável e natural simbiose da Biblioteconomia com a Arquivística e com um tipo especial de Informática, isto é, a aplicada de forma sistemática à gestão das/nas Organizações” formam o “núcleo duro” da CI, núcleo “homogéneo e portador de unidade e identidade” (A. M. da Silva & Ribeiro, 2008, p. 79), como se pode verificar na Figura 15.

A perspetiva “unitária e sistémica” dos referidos autores portugueses é fundada nos seguintes pressupostos:

Importa, desde já, sublinhar a nossa rejeição epistemológica da *library science* e da *archival science*, por estas constituírem aplicações ou especificações teórico-práticas ínsitas a um “território” uno e identificável como tal – a Ciência da Informação. Rejeitamos, também, o uso/abuso do conceito de interdisciplinaridade para caracterizar este campo específico, significando na maioria das vezes que se trata de uma “miscelânea” de perspectivas e de métodos de proveniência diversa e de articulação inexistente, logo caótica. (A. M. da Silva & Ribeiro, 2008, p. 79)

Pelo descrito e apresentado no diagrama da Figura 15, existe uma relação diferente entre os componentes do “núcleo” e entre estes e os restantes campos. Não sendo objeto de estudo do presente trabalho a diferente natureza destas relações, sublinha-se apenas o aspeto que se afigura importante nomeadamente pelo carácter mais “íntimo” que os três componentes nucleares partilham. Lançando pistas para o enquadramento de uma área ou disciplina, consoante a abordagem, ainda não explicitamente referida – a Gestão da Informação.

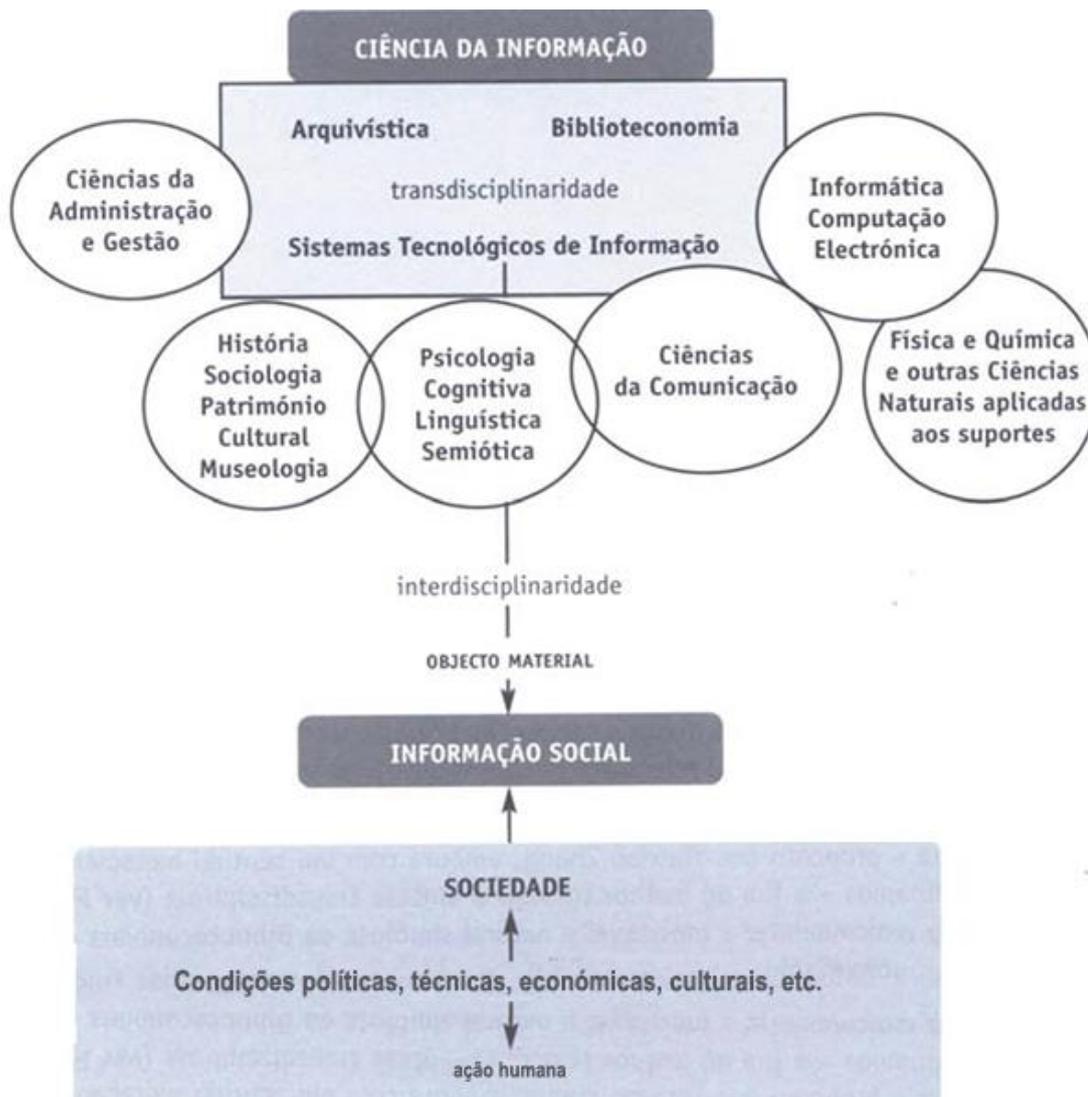


Figura 15. Diagrama do campo da Ciência da Informação segundo Armando Malheiro da Silva e Fernanda Ribeiro, (2008, p. 80).

Tendo como ponto de partida, para o enquadramento da Gestão da Informação em alguma área específica de educação/conhecimento, a descrição de Silva e Ribeiro: “um tipo especial de Informática, isto é, a aplicada de forma sistemática à gestão das/nas Organizações”, é possível que a mesma seja enquadrada em mais dois “setores” para além do da informação, na computação ou na administração/gestão. Tal situação “justificará” o enquadramento dos cursos com essa designação – Gestão da Informação, ou similar, em grupos diferentes de acordo com a ênfase do mesmo, dependendo, em última análise, dos critérios de quem classifica, tendo em conta que “qualquer classificação implica uma irreduzível arbitrariedade” (Pombo, 2002, p. 8).

A título de síntese visual das relações disciplinares referidas ao longo do ponto 2.3, elaborou-se a Tabela 10 e a Figura 16. Nelas são apresentadas as disciplinas que foram encontradas nas obras seleccionadas, tendo como referência o ano da respetiva obra. Uma vez que o foco, em termos disciplinares, foi a tríade BAD optou-se por um critério de frequência no qual foram consideradas as disciplinas referidas por três ou mais autores. Refere-se que, das três disciplinas BAD, a Arquivística foi a menos referida, apresentando apenas três menções.

Tabela 10

Disciplinas cuja relação com a Ciência da Informação é referida por três ou mais autores incluídos na Tabela 8 nas respectivas obras aí descritas

Ref.	Autoria	Ano	Disciplinas referidas
2A	Borko	1968	Library Science; Documentation (c); Computer Technology; Psychology; Communications; Linguistics; Mathematics.
2s	Mikhailov	1971	Library Science; Documentation (a); Electronic Digital Computers (t); Psychology; Linguistics; Mathematical Information Theory.
2B	Wersig	1975	Library Science; Documentation (c); Archivistics ; Technology (Information Technology); Psychology (Psychology of Information); Museology; Sociology (Sociology of Information).
2E	Brookes	1980	Library Science (t); Computer Science; Communication; Linguistics; Statistics.
3A	Saracevic	1991	Biblioteconomia ; Ciência da Computação; Ciência Cognitiva (incluindo Inteligência Artificial); Comunicação.
3B	Le Coadic	1994	Biblioteconomia (o); Documentação (o); Arquivística (o/t) Informática; Psicologia; Jornalismo (o); Linguística; Museonomia (o); Sociologia; Matemática e Estatística.
3C	Pinheiro	1997	Biblioteconomia; Documentação/Bibliografia (a); Ciência da Computação; Ciência Cognitiva; Comunicação; Sociologia da Ciência.
3D	Silva	2002	Biblioteconomia; Documentação (a); Arquivística ; Informática, Computação Eletrónica e Sistemas Tecnológicos de Informação; Psicologia Cognitiva; Ciências da Comunicação; Linguística; Museologia; Sociologia.

Nota. O nome das disciplinas é o que se encontra nas obras pelo que se encontra na língua aí utilizada. (a): disciplina referida enquanto antecedente da CI. (o): disciplina referida como estando associada à origem da CI. (c) disciplina referida como componente prática da CI. (t) disciplina referida como técnica(s) específica(s).

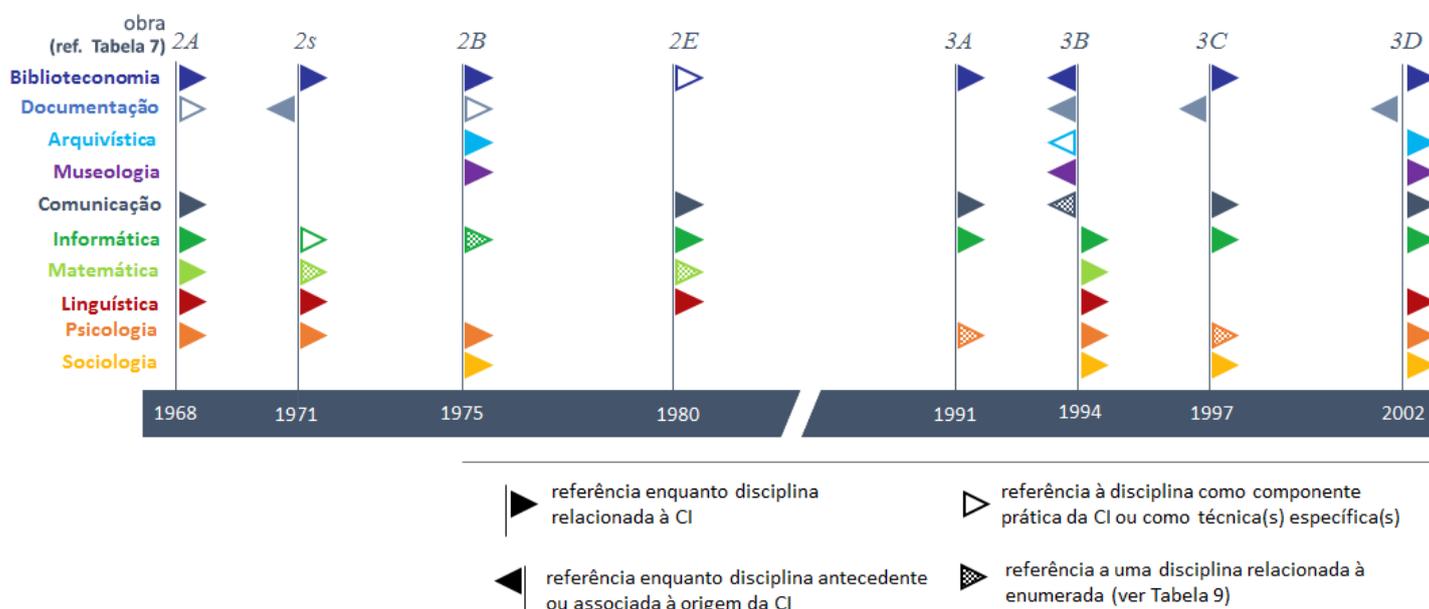


Figura 16. Distribuição temporal das referências às disciplinas mencionadas por três ou mais autores incluídos na Tabela 8 nas respectivas obras aí descritas.

Na Figura 16 foram associadas algumas disciplinas com designações diferentes como, por exemplo: Ciência da Computação à Informática; Ciência Cognitiva à Psicologia; Jornalismo à Comunicação; e Estatística à Matemática. Tal opção teve a ver com duas razões diferentes. Por um lado, durante o período entre 1968 a 2002, terão ocorrido “mutações” quanto às

designações das disciplinas e aos seus conteúdos, acrescentando, ainda, o facto de existirem variações nominativas regionais. Por outro lado, a opção pragmática que presidiu à construção desta síntese – a de ser uma apresentação mais coesa em termos disciplinares. Entrando neste ponto os referidos critérios do classificador, todavia, é esse confronto dicotómico que caracteriza o processo de classificar (Simões, 2011, p. 32) e, com ele, como refere a mesma autora (Simões, 2011, p. 31), não apenas se ordena os objetos como se os “domina”.

Analisando a distribuição apresentada na Figura 16 das disciplinas ao longo da linha cronológica, é possível verificar que as três disciplinas BAD apresentam presenças diferentes.

A primeira, a Biblioteconomia, é acompanhada pela Informática mostrando uma presença consistente nas menções às suas relações disciplinares com a CI (surge em todas as 8 obras). A segunda disciplina BAD, a Arquivística, apresenta, neste diagrama, uma relação direta com a Museologia – foram mencionadas nas mesmas três obras. No que diz respeito à Documentação, a terceira da tríade BAD, apresenta uma presença similar em número à Psicologia e à Comunicação, mas bastante diferente destas na forma como é mencionada. A Documentação é referida de duas formas: como disciplina que esteve na origem da CI (5 vezes) e, em número menor, como uma componente prática da CI (em 1968 e 1975), o que é consistente com a visão que a aponta como “natural” antecessora da CI.

Relativamente às outras disciplinas, a Psicologia e a Comunicação marcam também uma presença bastante consistente (a primeira com 7 menções e a segunda com 6). As restantes três, Linguística (com 5 menções), Matemática e Sociologia (ambas com 4 menções) apresentam, cada uma, particularidades diferentes na distribuição das suas menções. A primeira surge ao longo de toda a linha de tempo, a segunda mais concentrada nos primeiros anos (de 1980 para trás) e a terceira concentrada nos anos mais recentes (de 1994 em diante).

2.4. Cursos de Ciência da Informação

O panorama descrito nos pontos anteriores reforçou a necessidade de estabelecer critérios objetivos para a seleção dos cursos a considerar para o estudo de caso, como sendo de CI. Com esse objetivo procurou-se nas estruturas oficiais associadas ao respetivo sistema educativo de Portugal e do Brasil, uma classificação e/ou organização dos cursos que assistisse no processo seletivo.

A diversidade encontrada nas designações dos cursos, e até das áreas onde estes se encontram integrados, pode relacionar-se com o percurso dinâmico da área, como atesta o ponto 2.3. Apesar das classificações encontradas refletirem esse percurso, cujas características se encontram estreitamente associadas à contextualização geográfica, no caso a Europa e a América do Norte; este não explica, por si só, a dualidade de critérios encontrada no enquadramento dos cursos. Casos exemplificativos desta dualidade serão descritos nos pontos 2.4.1 e 2.4.2 de modo a explicitar essa situação.

Tendo em conta a contextualização histórica relativa às relações disciplinares da Ciência da Informação, a seleção dos cursos a considerar de CI foi fundamentada nas classificações encontradas nas entidades: Direção-Geral de Estatísticas da Educação e Ciência (DGEEC) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), respetivamente

para cursos portugueses e brasileiros, e os ajustamentos considerados necessários, que nos casos observados serão explicados em (2.4.1 e 2.4.2).

Para além do enquadramento institucional das classificações de cada país, apresentam-se as listagens de programas de mestrado e doutoramento considerados de CI, que se encontram em funcionamento no ano letivo 2015-2016 e cujos programas foram aprovados pelas entidades oficiais do respetivo país (Tabela 11 e Tabela 13).

Em Portugal os cursos necessitam da acreditação da Agência de Avaliação e Acreditação do Ensino Superior (A3ES) para funcionarem⁷⁸. No Brasil o processo passa pela obtenção de uma avaliação positiva (uma nota igual ou superior a 3) efetuada pela fundação CAPES para que os diplomas dos respetivos cursos tenham validade nacional⁷⁹.

Apesar das abordagens serem diferentes os efeitos da avaliação destas entidades são os mesmos – a desativação dos cursos. Em Portugal, a não obtenção, ou a cessação, da acreditação por parte da A3ES proíbe o funcionamento dos ciclos de estudos em questão⁸⁰. No Brasil, a obtenção de uma nota inferior a 3 implica o cancelamento da autorização de funcionamento assim como o respetivo reconhecimento da CAPES⁸¹.

2.4.1. Em Portugal

Em território nacional a entidade responsável pela avaliação e acreditação dos cursos do Ensino Superior, a A3ES, não é um organismo público, trata-se de uma fundação de direito privado, reconhecida como de utilidade pública. Como se verificou que no sistema em linha da A3ES⁸² os cursos não são organizados por área de conhecimento, ou de educação/formação, procurou-se em outras entidades oficiais, ou oficialmente reconhecidas, uma classificação que atendesse ao pretendido para o presente trabalho. Três organizações diferentes, em outras tantas entidades, foram encontradas, nomeadamente: na Associação Portuguesa de Bibliotecários Arquivistas e Documentalistas (apBAD), na Direção Geral do Ensino Superior (DGES) e na DGEEC⁸³.

⁷⁸ Artigo 27.º do Regulamento N.º 392/2013, da A3ES, publicado em Diário da República, 2.ª Série, de 16 de outubro, que aprova o regime dos procedimentos de avaliação e de acreditação das instituições de ensino superior e dos seus ciclos de estudos.

⁷⁹ Artigos 1.º e 10.º da Portaria N.º 91, de 29 de julho de 2015, publicada no Diário Oficial da União - Seção 1, N.º 145 de 31 de julho, que fixa normas e procedimentos para submissão, avaliação, divulgação e envio dos resultados da avaliação ao Conselho Nacional de Educação, e início de funcionamento dos programas novos de pós-graduação, em níveis de mestrado e doutorado.

⁸⁰ O ponto 3 do artigo 38.º do Regulamento N.º 392/2013, da A3ES, (publicado em Diário da República, 2.ª Série, de 16 de outubro), determina: “a cessação dos efeitos da acreditação implica, de forma imediata, a proibição de continuação do funcionamento dos ciclos de estudos em questão e a perda do reconhecimento dos graus académicos a que aqueles conduziam”.

⁸¹ Apesar de, tanto a Portaria N.º 91 de 29 de julho de 2015 como a Portaria N.º 1418 de 23 de dezembro de 1998 do Ministério da Educação, apenas fazerem referência aos efeitos da obtenção de notas superiores a 2, no *website* da CAPES esses efeitos são explicitados. Informação recolhida de <http://capes.gov.br/acessoainformacao/perguntas-frequentes/avaliacao-da-pos-graduacao/7421-sobre-avaliacao-de-cursos>, acessado a 12-8-2016.

⁸² Disponível no URL: <http://www.a3es.pt/pt/acreditacao-e-auditoria/resultados-dos-processos-de-acreditacao/acreditacao-de-ciclos-de-estudos> (acedido a 2-3-2016).

⁸³ A recolha de informação foi efetuada durante o mês de novembro de 2015, tendo-se consultado os *websites* das respetivas entidades: a apBAD (os dígitos “ap” é um acrescento do autor para diferenciar a presente abreviatura da homónima área de educação e formação), é uma entidade sem fins lucrativos com o estatuto de Pessoa Coletiva de Utilidade Pública, URL: http://www.apbad.pt/Formacao/formacao_cdisp.htm; as DGES e DGEEC, são direções do Ministério da Ciência, Tecnologia e Ensino Superior, URL:

Após a análise das referidas listagens (apBAD, DGES e DGEEC) foi considerado ser a da DGEEC aquela que apresenta uma organização consistente com o pretendido⁸⁴, até pelo documento que suporta a mesma: a Classificação Nacional das Áreas de Educação e Formação (CNAEF) de 2005. A CNAEF resultou da subclassificação das *Áreas de estudo* da *Classificação Internacional Tipo da Educação*, da UNESCO, elaborada a pedido e sob supervisão do Gabinete de Estatística das Comunidades Europeias e do Centro Europeu para o Desenvolvimento da Formação Profissional (Ministério das Actividades Económicas e do Trabalho (MAET), 2005, p. 2281).

A CNAEF está organizada em três níveis: 1º – *Grandes grupos*; 2º – *Áreas de estudo*; 3º – *Áreas de educação e formação*. Sendo o conteúdo programático dos cursos o critério para sua inclusão numa determinada área de 3º nível (MAET, 2005, p. 2283).

Nesta hierarquia, os cursos relativos à CI têm o seguinte enquadramento: 1º nível – *Ciências sociais, comércio e direito*; 2º nível – *Informação e jornalismo*; 3º nível – *Biblioteconomia, arquivo e documentação (BAD)*. É, assim, dentro da Área de Educação e Formação designada pelo acrónimo BAD que surge a referência ao “programa formativo” das “Ciências da informação” como é exposto na Portaria nº 256/2005 de 16 de Março, que aprova a atualização da CNAEF:

Os programas de formação em biblioteconomia, arquivo e documentação dizem respeito aos métodos que permitem seleccionar, obter, organizar e conservar a informação e facilitar a sua utilização. Esta área inclui os programas cujo conteúdo principal incida sobre as seguintes formações: Acervo museológico; Arquivo; Biblioteconomia; Ciências da informação; Documentação (MAET, 2005, p. 2291).

Contudo, no sistema da DGEEC, por um lado, não há referência aos cursos ativos e àqueles que foram descontinuados, por outro, infere-se que há cursos repetidos, dado o facto de os

<http://www.dges.mctes.pt/DGES/pt/OfertaFormativa/CursosConferentesDeGrau/CiclosAutorizados/> e URL: <http://www.dgeec.mec.pt/np4/171/>.

⁸⁴ A lista encontrada na apBAD apresenta vários dados desatualizados, tendo em conta a data de atualização exibida: 14-7-2015. Pode-se referir, a título de exemplo, a inclusão de cursos descontinuados em 2013, como os cursos de Ciências Documentais das universidades Lusófona e Autónoma de Lisboa, ou a não listagem do mestrado de Coimbra em Ciência da Informação, que deu início no ano letivo 2013-2014. Atendendo aos cursos incluídos na listagem, à tipologia dos mesmos (doutoramento, mestrado, pós-graduação, licenciatura, bacharelato e cursos médios) e ao título atribuído ao conjunto - “Formação Disponível em Portugal”, dir-se-ia que a intenção foi a de divulgar toda a formação existente em terras Lusas que tivesse alguma ligação, mesmo reduzida, ao contexto profissional da apBAD, independente da área formativa. Assim se infere pela inclusão de cursos como: o mestrado em História e Património - ramo Arquivos Históricos (da Universidade do Porto), o mestrado em Ciências da Educação - Área de especialização em Educação e Bibliotecas (da Universidade Lusófona do Porto, curso não acreditado pela A3ES em 10-07-2012 pelo que também foi descontinuado) ou ainda, o mestrado em Informação Empresarial (da Escola Superior de Estudos Industriais e de Gestão do Politécnico do Porto). Estas considerações levaram, assim, à não utilização desta possível classificação de cursos.

Relativamente à DGES, verificou-se que esta atribuiu uma codificação aos cursos que não se coadunava com o pretendido para o presente trabalho. Por um lado, aparenta depender da designação específica dos cursos, por exemplo: cursos com o nome *Ciências da Informação e da Documentação* é atribuído o código 6039 já o curso denominado *Ciências da Documentação e Informação* tem um código diferente, 6721, por outro lado, em contradição com essa especificidade, os cursos de designação *Ciência da Informação* e o curso *Educação e Bibliotecas* recebem o mesmo código, 6025. As listas da DGES apresentavam, aquando da sua consulta, a seguinte data de atualização: 23-7-2015.

quatro programas de doutoramento listados, por exemplo, três serem da Universidade de Évora e registarem a mesma designação *Ciências da Informação e da Documentação*⁸⁵.

O quarto programa listado pertence a uma instituição privada (ref. *Pt03.D* da Tabela 11). Interessa ainda considerar na listagem desta Direção que faltam dois programas de doutoramento na área BAD, o da Universidade de Coimbra e o das Universidades do Porto e da de Aveiro (programa de doutoramento conjunto), considerados pelas Instituições, nas quais se inserem, como sendo da área da CI. Infere-se que esta circunstância tenha a ver com a seguinte situação. O curso de Doutoramento de Coimbra (ref. *Pt07.D*) entrou apenas em funcionamento no presente ano letivo (2015/2016), facto que pode explicar a sua ausência na base de dados da DGEEC cujo *website* apresenta como atualizada em 31-7-2015, à data da recolha dos dados (de 1-11-2015 a 3-12-2015). Já no caso do curso da parceria das Universidades do Porto e Aveiro (ref. *Pt08.D*) a razão é outra, pois o mesmo consta da base de dados da DGEEC estando, contudo, classificado como pertencendo à área *Ciências informáticas*. Tal classificação terá ocorrido devido a ter sido considerado as “Ciências e Tecnologias da Comunicação”⁸⁶ como área científica predominante do programa. Todavia, essa designação não consta na CNAEF (nem como área nem como conteúdo programático). Contudo, analisando os objetivos e as competências a desenvolver pelo referido programa, é possível verificar a predominância das referências à Ciência da Informação⁸⁷, pelo que fará sentido a inclusão do referido programa nos cursos de CI.

Por último, refere-se que o curso Gestão da Informação e Bibliotecas Escolares (ref. *Pt06.M*) apesar de se encontrar listado na área BAD, as informações disponíveis na página do mesmo apontam para uma maior ligação à educação. Indicam, inclusive, o programa como de acesso a “Cursos de 3º ciclo na área das Ciências da Educação”⁸⁸. Assim, mantendo a lógica da inclusão do curso referido no parágrafo anterior, considera-se que este curso da Universidade Aberta não deveria ser considerado como sendo de CI.

⁸⁵ Essa instituição teve o referido programa em funcionamento até 1-9-2015 data da não acreditação do mesmo pela A3ES.

⁸⁶ No sistema em linha da Universidade do Porto (UP) pode encontrar-se a seguinte informação relativa às áreas científicas predominantes do 3º ciclo em *Informação e Comunicação em Plataformas Digitais*: “Fields of Science (FOS): **Outras humanidades**; Oficial: Ciências e Tecnologias da Comunicação”. Informação disponível no URL: https://sigarra.up.pt/flup/pt/cur_geral.cur_areas_cient_view?pv_curso_id=167&pv_ano_lectivo=2015&pv_tipo_cur_sigla=&pv_origem=CUR (acedido a 2-3-2016). Já na página de apresentação do mesmo curso, no *website* da Universidade de Aveiro (UA), são apresentadas 4 áreas científicas: “Ciências da Comunicação; **Ciências da Informação**; Ciências e Tecnologias da Comunicação; Ciências Sociais”, as mesmas são apresentadas pela ordem aqui usada, mas não relevando qualquer uma delas como a predominante. Informação disponível no URL: <http://www.ua.pt/course/181/?p=1> (acedido a 2-3-2016).

⁸⁷ Da análise dos objetivos e competências do programa doutoral Informação e Comunicação em Plataformas Digitais, disponíveis em linha nos sistemas das duas universidades (UP e UA), verifica-se que é a *Ciência da Informação* (no singular e não no plural como aparece na apresentação do curso no sistema da UA) a com mais referências (3), seguida das *Ciências da Comunicação* (com 2) e, finalmente, as *Ciências e Tecnologias da Comunicação*, (referida uma vez só). Completando as 4 áreas científicas referidas na nota anterior (Nota 86), falta apontar a ausência a qualquer referência às *Ciências Sociais* nos objetivos e competências daquele programa. Informação recolhida nos *websites* da UP e da UA, disponível, respetivamente, no URL: https://sigarra.up.pt/flup/pt/cur_geral.cur_areas_cient_view?pv_curso_id=167&pv_ano_lectivo=2015&pv_tipo_cur_sigla=&pv_origem=CUR e <http://www.ua.pt/course/181/?p=2> (accedidos a 2-3-2016).

⁸⁸ Para além da referida ligação à área das Ciências da Educação, pode ler-se nas condições de acesso ao curso: “O mestrado destina-se preferencialmente a candidatos que estejam qualificados profissionalmente para a docência ou que tenham prática de ensino”, <http://www2.uab.pt/guiainformativo/detailcursos.php?curso=37> (acedido a 2-3-2016).

Confrontando a base de dados da DGEEC com a listagem dos cursos do 2º e 3º ciclos do ensino superior, disponível no *site* da DGES, e estes com os dados disponibilizados pela A3ES e as informações recolhidas nos *sites* das instituições de ensino superior, elaborou-se a Tabela 11 onde se pretendeu sistematizar e organizar a informação recolhida nas diversas fontes.

Tabela 11

Cursos do 2º e 3º Ciclos do Ensino Superior (mestrado e doutoramento) acreditados pela A3ES e em funcionamento em Portugal no ano letivo 2015-2016, listados na área de educação e formação “Biblioteconomia, Arquivo e Documentação (BAD)” da CNAEF e/ou considerados no presente trabalho como sendo de Ciência da Informação

Ref.	Instituição / Departamento	Programa	Tipo	M/D
Cursos listados dentro da área “BAD” da CNAEF considerados, no presente trabalho, como sendo de CI:				
<i>Pt01.M</i>	Universidade de Coimbra (UC) / Faculdade de Letras	Ciência da Informação	Púb.	M
<i>Pt02.M</i>	Universidade do Porto (UP) / Faculdade de Letras & Faculdade de Engenharia	Ciência da Informação	Priv.	M
<i>Pt03.D</i>	Fundação Ensino e Cultura "Fernando Pessoa" (FEC-FP) / Faculdade de Ciências Humanas e Sociais	Ciência da Informação	Púb.	D
<i>Pt04.M</i>	Universidade de Lisboa (UL) / Faculdade de Letras	Ciências da Documentação e Informação	Púb.	M
<i>Pt05.M</i>	Universidade do Algarve (UAAlg) / Faculdade de Ciências Humanas e Sociais	Ciências Documentais	Púb.	M
Cursos listados dentro da área “BAD” da CNAEF não considerados, no presente trabalho, como sendo de CI:				
<i>Pt06.M</i>	Universidade Aberta (UAb) / Departamento de Educação e Ensino à Distância	Gestão da Informação e Bibliotecas Escolares	Púb.	M
Outros cursos não listados dentro da área “BAD” considerados, para o presente trabalho, como sendo de CI:				
<i>Pt07.D</i>	Universidade de Coimbra (UC) / Faculdade de Letras	Ciência da Informação	Púb.	D
<i>Pt08.D</i>	Universidade do Porto (UP) / Faculdade de Letras & Universidade de Aveiro (UA) / Departamento de Comunicação e Arte ^a	Informação e Comunicação em Plataformas Digitais	Púb.	D

Nota. Elaborado a partir da listagem obtida no *website* da DGEEC em 1-12-2015 (obtida dinamicamente no URL: <http://www.dgeec.mec.pt/np4/171/>) e de informações colhidas nos *websites* da DGES, da A3ES e das respetivas instituições de ensino superior (recolha efetuada entre os dias 1 de novembro e 3 de dezembro de 2015). Púb.: ensino público; Priv.: ensino privado; M: programas de 2º ciclo – mestrado; D: programas de 3º ciclo – doutoramento (D). ^a O curso é lecionado em parceria pelas duas universidades (Universidade do Porto e Universidade de Aveiro).

2.4.2. No Brasil

No Brasil, tal como foi referido, é a CAPES, uma fundação do Ministério da Educação, a responsável pela avaliação dos cursos de pós-graduação *stricto sensu* (programas de mestrado e doutorado). Ao contrário da sua congénere portuguesa, a A3ES, esta fundação apresenta uma classificação dos referidos cursos no seu sistema em linha, denominado de “Plataforma Sucupira”⁸⁹. Nessa plataforma, para além de outras opções, interessa referir a possibilidade de pesquisar os programas por duas áreas de organização temática, por *Área Básica* e por *Área de Avaliação*.

⁸⁹ Acedido em 2-3-2016 no URL:

<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/listaPrograma.jsf>

As Áreas de Avaliação fazem parte de uma classificação dos cursos utilizada pela CAPES, elaborada pela própria fundação, para auxiliar o processo de avaliação de sua competência. As Áreas Básicas, também denominadas por *Áreas do Conhecimento*, por sua vez, representam o segundo nível da *Tabela de Áreas do Conhecimento* (TAC) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). A CAPES utiliza estas duas classificações que apresentam algumas variações organizativas embora as Áreas de Avaliação sejam um constructo posterior baseado na TAC⁹⁰.

A organização das Áreas de Avaliação obedece a três níveis: 1º nível – *Colégios*; 2º nível – *Grandes Áreas*; 3º nível – *Áreas de Avaliação*. Os cursos relacionados com a CI estão enquadrados, nesta classificação, da seguinte forma: 1º nível – *Colégio de Humanidades*; 2º nível – *Ciências Sociais Aplicadas*; 3º nível – *Ciências Sociais Aplicadas I*⁹¹. Apesar de não conter níveis mais restritos, a CAPES aplica esta classificação à estabelecida na TAC abarcando cada Área de Avaliação várias Áreas Básicas desta.

A quantidade de grupos de segundo nível, áreas Básicas ou do Conhecimento, da TAC que a área Ciências Sociais Aplicadas I abarca varia de acordo com a versão da mesma que é observada, uma vez que existe um desfasamento entre a TAC que é disponibilizada nas fontes oficiais consultadas (CAPES, CNPq e o portal do Governo Federal do Brasil)⁹² e a versão que aparenta estar a ser utilizada na renovada *Plataforma Sucupira* da CAPES⁹³.

Apresenta-se na Tabela 12 as alterações registadas nas versões da TAC, ao longo dos seus cerca de 40 anos de existência, no que concerne à área CI, de forma a contextualizar as discrepâncias encontradas aquando da recolha de dados (de 1-11-2015 a 3-12-2015) relativos aos programas de mestrado e doutorado classificados como pertencendo à CI.

As alterações organizativas da TAC, no que diz respeito à CI, deram-se tanto nas designações dos vários níveis como no próprio nível onde foi inserida, observando-se estas, ainda, na relação hierárquica estabelecida entre a CI, a Biblioteconomia e a Arquivística. Como poderá verificar-se na Tabela 12 essa relação foi-se alterando ao longo das versões sem nunca se repetir. Uma análise mais detalhada dessas alterações sai fora do âmbito do presente

⁹⁰ Mais informação relativa ao processo de avaliação e sua relação com a TAC pode ser consultada em linha no portal do Governo Federal do Brasil, no URL: <http://dados.gov.br/dataset/tabela-de-areas-de-conhecimento-do-ensino-superior> (acedido a 5-3-2016).

⁹¹ Na página do *website* da CAPES dedicada às áreas de avaliação (URL: <http://www.capes.gov.br/avaliacao/sobre-as-areas-de-avaliacao>, acedido a 5-3-2016) a designação do 3º nível é igual ao 2º nível, sem o “I”, justificando-se por esta ser a única área com essa designação, ao contrário de outras como Ciências Biológicas, com 3, ou Engenharias, com 4. No entanto, a designação *Ciências Sociais Aplicadas I* é a que surge, no mesmo *website* da CAPES, na página específica da área (URL: <http://www.capes.gov.br/component/content/article/44-avaliacao/4662-ciencias-sociais-aplicadas-i>, acedido a 5-3-2016) assim como na *Plataforma Sucupira* da mesma fundação.

⁹² A TAC foi consultada nas seguintes fontes: CAPES - <http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>, com data de publicação: 11-7-2012; CNPq - <http://www.cnpq.br/documents/10157/186158/TabeladeAreasdoConhecimento.pdf> (ou em formato legível por máquina: http://cnpq.br/web/guest/dados_abertos), sem referência à data de publicação; portal do Governo Federal do Brasil (dados.gov.br) - <http://dados.gov.br/dataset/tabela-de-areas-de-conhecimento-do-ensino-superior>, com data de publicação: 3-2011. (URLs acedidos a 6-3-2016).

⁹³ No início de dezembro de 2015 a plataforma digital da CAPES, onde se pode obter informação referente aos programas de pós-graduação *stricto sensu*, foi atualizada, não só em termos estéticos e funcionais como a nível estrutural, nas *Áreas Básicas* onde os referidos programas são listados denotando, na versão atualizada (versão 1.8.42), uma organização classificativa não compatível com a versão da TAC disponibilizada nas fontes referidas na Nota 92.

trabalho⁹⁴ contudo, importa atentar à opinião de Sales e Murguia (2015, pp. 32–33) segundo os quais a TAC “é um dispositivo de governo que prioriza a gestão em prol do conhecimento” cujo “critério prevalecente não é nem filosófico nem conceitual mas, sim, administrativo”⁹⁵.

Tabela 12

A Ciência da Informação nas versões da “Tabela de Áreas do Conhecimento” do CNPq.

Versão	1º Nível - Grande Área	2º Nível - Área Básica ou Área do Conhecimento	3º Nível - Subárea(s)
1976 ^a	[nível não incluído]	Comunicação	Ciências da Informação
1982	Ciências Humanas, Sociais e Artes	Ciência da Informação, Biblioteconomia e Arquivologia	Teoria da Informação Tratamento da Informação
1984	Ciências Sociais Aplicadas	Ciência da Informação	Teoria da Informação Biblioteconomia Arquivologia
2005 ^b	Ciências Socialmente Aplicáveis	Ciência da Informação	Fundamentos da Ciência da Informação Gestão da Informação Tecnologias da Informação Informação Especializada

Nota. Adaptação do “Quadro 1 Ciência da Informação nas versões da Tabela de Áreas do Conhecimento” (R. F. de Souza & Stumpf, 2009, p. 45) aumentado com a proposta de 2005 (CNPq, CAPES, & FINEP, 2005, pp. 14–15). Existe ainda um 4º Nível denominado de “Especialidade” cuja finalidade é realizar a caracterização temática da atividade de pesquisa e ensino, sendo que uma mesma especialidade não é exclusiva de uma subárea independentemente da área ou grandes áreas a que está subordinada⁹⁶. ^a Nesta primeira versão a Biblioteconomia e a Documentação surgem como especialidades da subárea Ciências da Informação. ^b Nesta versão a Biblioteconomia e a Arquivologia (designação em uso no Brasil) passam novamente para áreas do conhecimento mas, ao contrário da versão de 1982 em que estão agrupadas com a CI, agora cada uma é tratada como área autônoma. Além disso, as especialidades aumentaram de 865, da versão anterior, para 1400, deixando de estar associadas às subáreas para serem listadas alfabeticamente (CNPq et al., 2005, p. 5).

Das quatro versões da TAC, apresentadas na Tabela 12, é a de 1984 aquela que se encontra disponibilizada nos *websites* das entidades CNPq e CAPES e no portal de dados abertos, *dados.gov.br*, do Governo Federal do Brasil⁹⁷. Apesar de esta versão não ser a mais atual,

⁹⁴ Duas perspectivas sobre essas alterações podem ser encontradas nos artigos: Sales, R. de, & Murguia, E. I. (2015). Determinações políticas na produção científica da Ciência da Informação do Brasil: Impacto da Tabela de Áreas de Conhecimento (TAC) do CNPq. *Scire*, 21(1), 27–34; e Souza, R. F. de, & Stumpf, I. R. C. (2009). Ciência da Informação como área do conhecimento: Abordagem no contexto da pesquisa e da Pós-Graduação no Brasil. *Perspectivas em Ciência da Informação*, 14(spe), 41–58.

⁹⁵ Também na introdução à TAC, disponível no *website* da CAPES, pode ler-se: “A classificação das Áreas do Conhecimento **tem finalidade eminentemente prática** [ênfase adicionado], objetivando proporcionar às Instituições de ensino, pesquisa e inovação uma maneira ágil e funcional de sistematizar e prestar informações concernentes a projetos de pesquisa e recursos humanos aos órgãos gestores da área de ciência e tecnologia.”, URL: <http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao> (acedido a 5-3-2016).

⁹⁶ Informação colhida no *website* da CAPES no URL: <http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao> (acedido a 5-3-2016).

⁹⁷ A versão da TAC que a CNPq disponibiliza é a versão original de 1984 enquanto que as versões disponibilizadas pela CAPES e no portal *dados.gov.br* contêm algumas ligeiras alterações à mesma. Alterações justificadas pela “necessidade operacional de para [sic] atender a necessidade de organização do processo de avaliação e fomento realizado pela Capes”, que passou pela criação da “Grande Área Multidisciplinar e, dentro dela as áreas Interdisciplinar, Ensino de Ciências e Matemática, Materiais e Biotecnologia. Além disso, criou dentro da área Interdisciplinar, as subáreas Meio-Ambiente e Agrárias; Engenharia/Tecnologia/Gestão; Saúde

infere-se ser a oficial, uma vez ser a que consta nas entidades competentes. A versão de 2005 surge como uma proposta preliminar da Comissão Especial de Estudos das entidades: CNPq, CAPES e FINEP (Financiadora de Estudos e Projetos) (CNPq et al., 2005). No entanto, comparando os programas de mestrado e doutorado listados na Área Básica *Ciência da Informação*, na plataforma da CAPES, em novembro de 2015 com os listados em dezembro do mesmo ano verifica-se, na primeira lista, um ajuste à TAC de 1984, já a lista de dezembro aponta para uma aproximação às alterações propostas na versão de 2005⁹⁸. A Tabela 13 apresenta uma confrontação dessas listas para uma visualização mais clara da relação efetuada entre as listas e as respetivas versões da TAC.

Pela análise dessa Tabela infere-se que a exclusão dos cursos Gestão da Informação e do Conhecimento (ref. *Br13.Mp*) e Memória e Acervos (ref. *Br14.Mp*) da lista consultada em novembro não estará relacionada com as versões da TAC dado a novidade dos cursos relativamente à data de atualização da plataforma digital da CAPES⁹⁹.

No que diz respeito à exclusão dos dois cursos da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), da lista de dezembro, dentro da Área Básica *Ciência da Informação*, tal afigura-se como resultado da alteração da classificação usada, a TAC de 1984, para a versão da mesma de 2005. Os referidos cursos da UNIRIO, Biblioteconomia (ref. *Br15.Mp*) e Gestão de Documentos e Arquivos (ref. *Br16.Mp*), surgem na lista de novembro dentro da Área Básica *Ciência da Informação* em concordância com a TAC de 1984 que apresenta, como *Subáreas* dessa Área – Biblioteconomia e Arquivologia, onde os referidos cursos estavam, respetivamente, enquadrados. Por sua vez, na listagem extraída da *Plataforma Sucupira* em dezembro os mesmos cursos não surgem dentro dessa mesma Área Básica (*Ciência da Informação*). Surgem, antes, de forma autónoma dentro das suas respetivas Subáreas, que passam a Áreas Básicas tal como é proposto na alteração à TAC datada de 2005.

Para o presente trabalho todos os 16 programas que constam da Tabela 13 serão considerados como sendo de CI por três ordens de razões:

- i. A classificação anterior (TAC de 1984) ainda aparenta estar em vigor¹⁰⁰;
- ii. A Biblioteconomia e a Arquivística fazem parte da área BAD de Portugal obtendo-se uma maior uniformização de critérios com a inclusão das Subáreas brasileiras

e Biológicas; e Sociais e Humanidades”, (informação disponível no URL: <http://dados.gov.br/dataset/tabela-de-areas-de-conhecimento-do-ensino-superior>, acessado a 6-3-2016).

⁹⁸ Também no SEER do IBCT (Sistema Eletrónico de Edição de Revistas do Instituto Brasileiro de Informação em Ciência e Tecnologia) se infere ser a versão preliminar de 2005, da TAC, a utilizada devido às designações das Áreas do Conhecimento aí apresentadas, tais como “Ciências Socialmente Aplicáveis” ou “Engenharias e Computação”, renomeações propostas na TAC de 2005 em lugar de, respetivamente, “Ciências Sociais Aplicadas” e “Engenharias”. A aplicação das designações da TAC de 2005 no SEER pode ser observada no URL: http://seer.ibict.br/index.php?option=com_mtree&task=listcats&cat_id=121&Itemid=109, (accedido a 6-3-2016).

⁹⁹ O sistema em linha da CAPES apresentava em 1-11-2015 como data de atualização 20-3-2015 e os cursos “Gestão da Informação e do Conhecimento” (ref. *Br13.Mp*) e “Memória e Acervos” (ref. *Br14.Mp*) só entraram em funcionamento no ano letivo 2015-2016 (em 8-9-2015 e em 10-7-2015, respetivamente) tendo também obtido a recomendação da CAPES posteriormente à data de atualização (o curso ref. *Br13.Mp* em 19/06/2015 e o curso ref. *Br14.Mp* em 26/03/2015).

¹⁰⁰ Efetuou-se um contato com a CAPES, em dezembro de 2015, via correio eletrónico no sentido de esclarecer a situação relativa à adoção da TAC proposta em 2005, tendo-se obtido como resposta a remissão para as “áreas de conhecimento que estão na página da Capes ou CNPq” (CAPES, comunicação pessoal, dezembro 15, 2015).

correspondentes (Biblioteconomia e Arquivologia);

- iii. Apesar de, no sistema em linha atual da CAPES, a *Ciência da Informação*, a *Biblioteconomia* e a *Arquivologia* serem tratadas como Áreas Básicas autônomas, a interpenetração das áreas é visível nos programas que incluem¹⁰¹.

Tabela 13

Cursos de pós-graduação (mestrado e doutorado) em funcionamento no Brasil no ano letivo 2015-2016, listados dentro da “Área do Conhecimento” Ciência da Informação no website da CAPES (confrontação entre as listagens de 1-11-2015 e de 1-12-2015)

Ref.	Instituição	Programa	Mest.	Dout.
Cursos listados em ambos os momentos na “Área” Ciência da Informação:				
Br01.MD	Universidade de São Paulo (USP)	Ciência da Informação	Aca.	Aca.
Br02.MD	Universidade Estadual Paulista Júlio de Mesquita Filho/Marília (UNESP/MAR)	Ciência da Informação	Aca.	Aca.
Br03.Ma	Universidade Estadual de Londrina (UEL)	Ciência da Informação	Aca.	-
Br04.MD	Universidade Federal da Bahia (UFBA)	Ciência da Informação	Aca.	Aca.
Br05.MD	Universidade Federal da Paraíba/João Pessoa (UFPB/JP)	Ciência da Informação	Aca.	Aca.
Br06.Ma	Universidade Federal de Pernambuco (UFPE)	Ciência da Informação	Aca.	-
Br07.MD	Universidade Federal de Santa Catarina (UFSC)	Ciência da Informação	Aca.	Aca.
Br08.MD	Universidade Federal Fluminense (UFF)	Ciência da Informação	Aca.	Aca.
Br09.MD	Universidade Federal do Rio de Janeiro (UFRJ)	Ciência da Informação	Aca.	Aca.
Br10.MD	Universidade de Brasília (UnB)	Ciência da Informação ^a	Aca.	Aca.
Br11.MD	Universidade Federal de Minas Gerais (UFMG)	Ciência da Informação ^a	Aca.	Aca.
Br12.Mp	Universidade do Estado de Santa Catarina (UDESC)	Gestão da Informação	Prof.	-
Cursos que passaram a estar listados, em 1-12-2015, dentro da “Área” Ciência da Informação:				
Br13.Mp	Universidade Federal do Rio Grande do Norte (UFRN)	Gestão da Informação e do Conhecimento	Prof.	-
Br14.Mp	Fundação Casa de Rui Barbosa (FCRB)	Memória e Acervos	Prof.	-
Cursos listados dentro da “Área” Ciência da Informação em 1-11-2015, ausentes na lista de 1-12-2015:				
Br15.Mp	Universidade Federal do Estado do Rio de Janeiro (UNIRIO)	Biblioteconomia	Prof.	-
Br16.Mp		Gestão de Documentos e Arquivos	Prof.	-

Nota. Elaborado a partir das listagens obtidas no *website* da CAPES nas datas 1-11-2015 e 1-12-2015 (a de novembro já não se encontra *online* e a mais recente é obtida dinamicamente no URL: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/listaPrograma.jsf>). Os mestrados estão qualificados como acadêmicos (Aca.) ou profissionais (Prof.). ^a Embora apareça na lista da CAPES com a nomenclatura de “Ciências da Informação” nos *websites* das respectivas instituições os nomes dos programas encontram-se no singular.

¹⁰¹ Na listagem da Área Básica *Biblioteconomia* está incluído um curso denominado “Ciências da Informação”, curso desativado da UFPB/JP que tem um programa de nome “Ciência da Informação” na lista da Área Básica *Ciência da Informação*. Por sua vez esta última tem na sua lista o recente curso “Memória e Acervos” que também poderia ser posicionado na área *Arquivologia* (área esta que apenas conta com um representante o curso “Gestão de Documentos e Arquivos”). Para finalizar aponta-se o caso do curso “Gestão da Informação” da UDESC que se posiciona como sendo da área de “Ciência da Informação/Biblioteconomia”, (URL da informação relativa ao curso da UDESC: <http://www.faed.udesc.br/?id=674> (acedido a 6-3-2015)).

PARTE II: ESTUDO DE CASO: A PRESENÇA DA WEB SEMÂNTICA NO
OUTPUT DOS CURSOS DE MESTRADO E DOUTORAMENTO EM PORTUGAL E
NO BRASIL EM CIÊNCIA DA INFORMAÇÃO (2005 – 2015)

Introdução ao estudo de caso

O presente estudo reflete um conjunto de quatro hipóteses a explorar. A sua introdução prende-se com a importância das mesmas para a construção do modelo de análise adotado. Segundo Quivy e Campenhoudt:

A organização de uma investigação em torno de hipóteses de trabalho constitui a melhor forma de a conduzir com ordem e rigor, sem por isso sacrificar o espírito de descoberta e de curiosidade que caracteriza qualquer esforço intelectual digno deste nome. (1992, p. 119)

As hipóteses formuladas foram as seguintes:

- i. Os trabalhos finais de mestrado e doutoramento inseridos em cursos de CI, cujo tema central é a Web Semântica, terão aumento em número de forma gradual desde 2005 até à data, embora se preveja uma diminuição dos mesmos nos anos mais recentes;
- ii. Os trabalhos referidos estarão concentrados em cursos lecionados no Brasil, prevendo-se uma relação de quatro para um relativamente aos cursos lecionados em Portugal;
- iii. Observando os trabalhos referidos será possível constatar uma diferença, de acordo com a data da sua realização (mais antiga – mais recente) e a tipologia dos mesmos (dissertação – tese), na abordagem da relação entre a CI e a WS (mais geral – mais específica).
- iv. Os orientadores e/ou coorientadores dos referidos trabalhos terão formação na área da Informática/Ciência da Computação;

A partir destas hipóteses determinaram-se os objetivos e respetiva metodologia para a sua consecução.

3. Objetivos do estudo de caso

Com este estudo pretendeu-se verificar e refletir sobre a presença da Web Semântica nas dissertações de mestrado (DM) e teses de doutoramento (TD) em Portugal e no Brasil em Ciência da Informação, entre o período de 2005 e 2015. Para o seu desenvolvimento, enquadrados nas hipóteses formuladas, definiu-se os seguintes objetivos específicos¹⁰²:

- a) Verificar o número de dissertações de mestrado e teses de doutoramento em Ciência da Informação associados à Web Semântica, produzidas em Portugal e no Brasil no período de 2005 e 2015, recuperadas de acordo com a metodologia desenhada para este estudo, nos repositórios RCAAP e OASISbr;
- b) Identificar os pontos de confluência da Ciência da Informação com a Web Semântica, presentes nos trabalhos recuperados;
- c) Caracterizar o perfil académico dos orientadores e coorientadores dos trabalhos recuperados identificados no estudo como sendo de Ciência da Informação.

¹⁰² Aqui apresenta-se apenas os três objetivos específicos, que dizem respeito a esta segunda parte em particular, dos quatro estabelecidos para todo o trabalho

O objetivo da alínea a) relaciona-se com as duas primeiras hipóteses (i. e ii.) e os outros dois objetivos, alíneas b) e c), relacionam-se com as outras duas hipóteses, respetivamente (iii. e iv.).

4. Metodologia do estudo de caso

O estabelecimento de limites de incidência do estudo, sendo essencial a qualquer investigação, é crucial para os *estudos de caso* tendo em conta os “preconceitos tradicionais”, utilizando a expressão de Robert Yin (2001, p. 28), relativos a este tipo de estratégia.

Nesse sentido, limitaram-se os trabalhos a analisar, aos incluídos nos portais digitais: Repositório Científico de Acesso Aberto de Portugal (RCAAP) e Portal Brasileiro de Acesso Aberto à Informação Científica (OASISbr). Ambos se configuram como ponto de acesso à produção científica e académica nacional, multidisciplinar e de acesso aberto¹⁰³. No ponto 4.1 far-se-á uma caracterização mais detalhada dos referidos portais.

Relativamente à tipologia dos trabalhos a analisar, esta limitou-se ao estudo das Dissertações de Mestrado (DM) e das Teses de Doutoramento (TD), dado tratar-se de documentos que, por um lado, requerem um considerável grau de aprofundamento e, por outro, como estão vinculados a um dado curso de ensino superior são passíveis de identificar como trabalhos produzidos, ou não, na área da Ciência da Informação.

Por fim, foram estabelecidos como limites temporais o período considerado entre 2005 e 2015. A escolha desse período prende-se com o facto de, no seu desenrolar, se ter assistido à evolução conceptual da Web Semântica, assim como da sua aplicabilidade em vários domínios do conhecimento, entre os quais na CI. Refere-se ainda que não se optou pelo ano 2001, data da publicação do artigo “The Semantic Web” na *Scientific American*, devido ao facto do próprio Tim Berners-Lee afirmar em 2006 que ainda muito continuava por realizar (Shadbolt, Hall, & Berners-Lee, 2006, p. 96). Como reforço da seleção deste marco cronológico (2005) está o facto dos repositórios eleitos para a recolha de dados terem iniciado funções após 2005 (o OASISbr em 2006 e o RCAAP em 2008).

Na seleção dos trabalhos que integram a amostra, i.e., considerados como sendo de CI, foram utilizadas as tabelas incluídas no ponto 2.4 (Tabela 11, para os cursos portugueses e Tabela 13 para os brasileiros) no caso dos cursos em funcionamento. Para os cursos desativados comparou-se a identificação que consta nos respetivos registos dos repositórios de origem com a informação disponível nos sistemas em linha da DGEEC e da CAPES, respetivamente para os cursos de Portugal ou do Brasil. Nos casos dúbios procedeu-se à verificação da identificação nos próprios trabalhos (abriram-se os respetivos registos)¹⁰⁴.

Em relação aos pontos de confluência entre a CI e a WS, analisaram-se os resumos dos trabalhos que constituíram a amostra no sentido de verificar esses mesmos pontos no discurso dos investigadores. Nos objetivos dos referidos trabalhos foi ainda observada a

¹⁰³ Síntese da informação disponibilizada nos *websites* dos respetivos repositórios RCAAP: <http://www.rcaap.pt/about.jsp>; e OASISbr: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao/portal-brasileiro-de-acesso-aberto-a-informacao-cientifica-oasisbr> (accedidos a 7-5-2016).

¹⁰⁴ A opção deste procedimento ser apenas aplicado em casos excecionais e não de forma sistemática a todos os registos deve-se à morosidade do mesmo sendo que, na maioria dos registos, tal processo é notoriamente desnecessário dado a informação procurada estar disponível sem proceder ao mesmo.

presença do termo Web Semântica e o contexto no qual ela ocorre relativamente aos verbos de ação utilizados para expressar os objetivos dos trabalhos em causa. Pretendendo-se, deste modo, determinar a relação de proximidade entre a WS, enquanto objeto de estudo da amostra e a CI, representada pelos autores dos trabalhos analisados neste estudo de caso¹⁰⁵.

No que concerne ao perfil académico dos orientadores e coorientadores dos trabalhos que constituem a amostra, este incidu na formação académica dos referidos docentes. Foram consideradas como fontes para a recolha desses dados as plataformas digitais: *DeGóis*, propriedade da Fundação para a Ciência e a Tecnologia¹⁰⁶, para docentes portugueses e a *Lattes*, da responsabilidade da CNPq, para docentes brasileiros¹⁰⁷.

Após esta abordagem geral à metodologia, passa a apresentar-se de um modo mais pormenorizado as fontes de obtenção de dados, os critérios usados na recolha de dados, o universo e a amostra e os critérios usados na análise dos dados.

4.1. Fontes de obtenção de dados

A seleção dos portais digitais RCAAP e OASISbr como fontes de obtenção de dados para o presente estudo de caso afigurou-se a opção óbvia dado reunirem quatro características consideradas essenciais: i. são promovidos por entidades oficiais¹⁰⁸; ii. são pontos de acesso agregadores da produção científica dos respetivos países; iii. são promotores do movimento de Acesso Aberto¹⁰⁹; iv. são, ainda, interoperáveis, i.e., a pesquisa efetuada em um dos portais permite recuperar também registos do outro, garantindo uma “malha” mais apertada na pesquisa, devido aos diferentes sistemas com que operam¹¹⁰.

Para uma melhor contextualização destes repositórios apresenta-se uma breve informação extraída dos respetivos *websites*. Esta informação descreve com clareza os serviços que os mesmos prestam:

O RCAAP é um portal agregador (meta-repositório) dos documentos depositados nos vários repositórios institucionais existentes em Portugal e no Brasil. No RCAAP são

¹⁰⁵ Os trabalhos analisados com este procedimento constituem um conjunto de 22 documentos, (ponto 4.3).

¹⁰⁶ A Fundação para a Ciência e a Tecnologia, tutelada pelo Ministério da Ciência, Tecnologia e Ensino Superior, é a agência pública nacional de apoio à investigação em ciência, tecnologia e inovação, em todas as áreas do conhecimento. Informação recolhida em <https://www.fct.pt/fct> (acedido a 7-8-2016).

¹⁰⁷ Informações recolhidas nos respetivos *websites*: <http://www.degois.pt/index.jsp?id=1> e <http://lattes.cnpq.br/>, acedidos a 7-8-2016.

¹⁰⁸ O RCAAP é promovido por três entidades: a UMIC – Agência para a Sociedade do Conhecimento, IP; a FCCN – Fundação para a Computação Científica Nacional; e a Universidade do Minho (<http://projeto.rcaap.pt/index.php/lang-pt/sobre-o-rcaap/promotores>, acedido a 7-8-2016). O OASISbr é gerido pelo IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia enquanto unidade de pesquisa do Ministério da Ciência, Tecnologia e Inovação do governo federal brasileiro (http://www.mcti.gov.br/noticia/-/asset_publisher/epbV0pr6eIS0/content/ibict-reformula-portal-de-publicacoes-cientificas;jsessionid=8C54470E38DD43EA76C7EC790A8DBFBB, acedido a 7-8-2016).

¹⁰⁹ Sobre a relação entre os portais e o Acesso Aberto veja-se, no caso do RCAAP, a descrição que se encontra em <http://projeto.rcaap.pt/index.php/lang-pt/sobre-o-rcaap/enquadramento> (acedido a 7-8-2016) e, para o OASISbr, a nota “Sobre” e o texto “Critérios de Qualidade” do portal *LA Referencia*, do qual o Portal Brasileiro depende, em <http://lareferencia.info/vufind/> (acedido a 7-8-2016).

¹¹⁰ Apesar da interoperabilidade entre os dois repositórios significar, em teoria, que os registos que se encontram num dever-se-ão encontrar usando o outro, na prática, os testes prévios efetuados apresentaram resultados díspares para pesquisas com os mesmos parâmetros. (Ver, a título exemplificativo, a gravação efetuada a 1-11-2015 de um desses testes, disponível no URL: <http://www.screencast.com/t/yKnFFzq8quuh>).

reunidos e indexados não só a descrição (metadados) mas também o texto integral. (RCAAP, 2016)¹¹¹.

O Portal brasileiro de publicações científicas em acesso aberto - OASISbr é um mecanismo de busca multidisciplinar que permite o acesso gratuito à produção científica de autores vinculados a universidades e institutos de pesquisa brasileiros. Por meio do oasisbr é possível também realizar buscas em fontes de informação portuguesas. (OASISbr, 2016)¹¹².

Em termos de enquadramento legislativo, a situação não é igual para os dois repositórios. No que respeita ao OASISbr, essa diferença poderá estar relacionada com a organização federal do Brasil, circunstância que concorre para que o processo legislativo a nível nacional seja mais complexo.

Em Portugal, o RCAAP está enquadrado pelo Decreto-Lei nº 115/2013, que determina a obrigatoriedade do depósito de todas as DM e TD produzidas no país, num repositório da sua rede.

No Brasil as iniciativas legislativas com teor similar à portuguesa, ainda não passaram do estatuto de projeto. A primeira, Projeto de Lei 1120/2007, obteve os vários pareceres favoráveis necessários, mas foi arquivada por não ter sido aprovada até final de uma legislatura. A segunda, Projeto de Lei do Senado 387/2011, elaborada no ano em que a primeira foi arquivada, ainda se encontra em tramitação (M. P. Costa & Leite, 2016, p. 41).

Apesar de não existir, no Brasil, uma imposição normativa para o depósito em acesso aberto da produção científica, existem políticas institucionais em vigor nas universidades brasileiras nesse sentido, muito por ação do IBICT (S. M. S. Costa, Kuramoto, & Leite, 2013, p. 139). Neste contexto a integração dos trabalhos nos repositórios institucionais, onde o OASISbr vai buscar os registos, está apenas dependente das respetivas instituições e das suas regras particulares.

Em Portugal a situação é diferente, na medida em que as normas, já referidas, estipulam para a inserção das DM e TD em repositórios da rede do RCAAP, um prazo máximo de sessenta dias após a conceção dos respetivos graus. Importa referir que essa obrigatoriedade aplica-se apenas aos trabalhos defendidos após agosto de 2013. No entanto, a introdução dos registos no RCAAP de esses mesmos trabalhos, defendidos desde essa data até setembro de 2015, foram contemplados com um prazo alargado até ao corrente ano pela Portaria nº 285/2015 (DM até julho e TD até dezembro).

Toda a situação descrita, tanto em Portugal como no Brasil, concorre para que a quantidade de registos seja altamente dinâmica independentemente do ano ao qual esses registos dizem respeito.

Relativamente às funcionalidades dos repositórios considerados, existem também diferenças entre os dois; nomeadamente no que diz respeito aos campos de “pesquisa avançada” disponíveis e, ainda, à respetiva quantidade de campos que é possível usar em simultâneo. Assim, o RCCAP oferece 5 campos enquanto o OASISbr oferece 6. Quanto à possibilidade

¹¹¹ Texto retirado do URL: <http://projeto.rcaap.pt/index.php/lang-pt/como-pesquisar-documentos/introducao-4>, acessado a 7-8-2016.

¹¹² Texto retirado do URL: <http://oasisbr.ibict.br/vufind>, acessado a 7-8-2016.

de seu uso simultâneo, o primeiro oferece 3 e o segundo um número virtualmente ilimitado¹¹³.

Os 5 campos presentes no RCAAP, onde é possível realizar a procura direcionada, são: i. Título; ii. Autor; iii. Assunto; iv. Descrição e v. Texto integral, tendo estes uma correspondência com os elementos constituintes dos documentos, sendo no caso das DM e TD a seguinte: i. título; ii. autoria; iii. palavras-chave; iv. resumo; e v. todo o documento agregado.

O OASISbr, tal como foi referido, apresenta 6 campos, contudo, não se trata da simples adição de um sexto campo pois apenas os 3 primeiros, Título, Autor; e Assunto, apresentam uma correspondência isomórfica entre os campos dos dois repositórios. Além desses 3, o portal brasileiro oferece, como campos de pesquisa específica: Ano da publicação; Recurso; e Todos os campos. Sendo que, relativamente à ligação com os elementos das DM e TD, o primeiro corresponde à data da defesa, o segundo ao repositório de origem, (onde o documento se encontra) e o terceiro faz uma procura simultânea em todos os 5 elementos anteriores (título; autoria; palavras-chave; data da defesa; e repositório de origem).

Comparando as funcionalidades de “pesquisa avançada” nos dois repositórios é evidente que o OASISbr oferece um maior número. Este repositório oferece, ainda, a possibilidade de se efetuar uma pesquisa em todos os campos simultaneamente, ao contrário do RCAAP. O campo Texto integral, do portal português, não oferece semelhante função. Os testes efetuados mostraram que os registos recuperados dessa forma não correspondem ao esperado, não só em termos de quantidade como também no que se refere aos próprios documentos recuperados¹¹⁴. Todavia, apesar dessa evidência, dois aspetos da interface do RCAAP merecem destaque: o primeiro, tem a ver com a possibilidade de efetuar pesquisas apenas nos resumos das DM/TD (campo Descrição), funcionalidade que não existe no OASISbr; o segundo diz respeito à relação booleana que se estabelece entre campos. Ao contrário do portal brasileiro onde esta é pouco intuitiva¹¹⁵ no português ela é evidente.

Estes dados, e outros adicionais como os filtros pré e pós pesquisa, são apresentados no Apêndice 2 para uma comparação mais facilitada.

4.2. Critérios usados na recolha de dados

A recolha de dados, incluindo a relativa ao perfil académico referido no ponto 4, foi efetuada entre os dias 2 e 12 de fevereiro de 2016. A opção por esta data prende-se com o propósito de recuperar os trabalhos produzidos entre 2005 e 2015¹¹⁶.

Para o registo dos elementos elaboraram-se duas folhas de recolha de dados (FRD), a saber:

¹¹³ A interface do OASISbr não só permite acrescentar campos à descrição como oferece a possibilidade de criar inúmeros grupos de pesquisa, cada um com a mesma opção de acréscimo indiscriminado de campos. O apuramento do limite de grupos/campos de pesquisa não se mostrou relevante para o presente trabalho.

¹¹⁴ Os testes efetuados não foram conclusivos quanto ao padrão de recuperação, no entanto, como o campo Texto integral não foi utilizado para o presente estudo de caso a questão não foi aprofundada.

¹¹⁵ No OASISbr a seleção dos operadores booleanos entre os campos de pesquisa faz-se através da seleção das seguintes opções: “Todos os termos” para *And*; “Qualquer termo” para *Or*; e “Nenhum termo” para *Not*.

¹¹⁶ Apesar da data escolhida para a recolha, a natureza dinâmica da inserção de registos nos portais usados no presente estudo (RCAAP e OASISbr) condiciona qualquer tentativa de controle relativamente ao ano de realização dos trabalhos e respetiva data de sua inclusão nos repositórios (ver ponto 4.1).

- a. FRD geral – onde foram registadas as informações consideradas pertinentes de todos os trabalhos (251. Ver Apêndice 3a);
- b. FRD específica – onde foram registados os dados complementares provenientes de trabalhos de CI, (22, aqueles que constituem a amostra. Ver Apêndice 3b).

Para além do termo de pesquisa, *web semântica*, procurou avaliar-se a pertinência da utilização de algum outro termo semanticamente associado. Para tal, efetuou-se uma recolha dos termos associados à Web Semântica em quatro glossários em linha: W3C, *Open Semantic Framework* (OSF), *Base22.com* e *ws.org.br*¹¹⁷.

Como critério de relevância para a seleção dos 128 termos coletados nos quatro glossários referidos foi estabelecido que estes deveriam encontrar-se registados em mais de dois dos mesmos. Aplicado esse critério a lista ficou reduzida a 14 termos (ver Apêndice 4).

Com base nessa recolha procuraram-se esses 14 termos nos programas, ementas e conteúdos das disciplinas, assim como nas linhas de pesquisa dos cursos considerados no presente estudo como sendo de CI¹¹⁸ (cf. Tabela 11 e Tabela 13). De forma a manter-se a fidelidade às fontes dos termos não foi aplicada uma uniformização na diversidade de nomenclaturas encontradas. Dos 14 termos iniciais, 5 não obtiveram qualquer correspondência na comparação com as fontes de CI¹¹⁹, reduzindo a lista a 9 itens (cf. Tabela 14).

Seguindo o mesmo critério de relevância e tendo em conta que foram 11 os cursos nos quais se encontraram correspondência dos termos procurados, apenas 2 termos apresentaram um número de ocorrências superior a 5: *Ontologies/Ontologias* com 10 e *Semantic web/Web semântica* com 7 (coluna [B], a negrito, da Tabela 14).

Perante os resultados obtidos, para além do termo *web semântica*, apenas o termo *ontologia* se afigura relevante. No entanto, no contexto referido, (programas, ementas, conteúdos das disciplinas, etc) este termo, *ontologia*, não se encontra totalmente relacionado com a WS; embora, *ontologia*, enquanto conceito, encontra-se estreitamente relacionado com a CI¹²⁰. Nesta perspetiva, a sua utilização como termo de pesquisa acarretaria um número elevado de potenciais relações espúrias.

¹¹⁷ A primeira fonte, W3C, será o mais próximo de fonte “oficial”, a segunda, OSF, é a fonte do glossário sugerido pela *Thomson Reuters* (<http://www.opencalais.com/glossary-of-semantic-technology-terms/>), a terceira, *Base22*, e a quarta, *ws.org*, são os restantes glossários específicos (relativos à Web Semântica) obtidos numa pesquisa efetuada em 6-11-2015 no *Google* usando, respetivamente, a expressão «glossary + "semantic web"» e «glossário + "web semântica"», analisando os resultados até à quinta página. Os termos foram recolhidos dos seguintes *websites*, (entre parênteses retos está a data da atualização apresentada nas respetivas páginas): W3C - https://www.w3.org/2001/sw/wiki/Semantic_Web_terminology [10-11-2011] e https://www.w3.org/2001/sw/wiki/Main_Page [28-1-2014]; *Open Semantic Framework* - http://wiki.opensemanticframework.org/index.php/Semantic_Technologies_Glossary [6-1-2014]; *Base22.com* - <https://wiki.base22.com/display/btg/Glossary+of+Semantic+Web+Terms> [20-6-2013]; *ws.org.br* - <http://ws.org.br/index.php/Gloss%C3%A1rio> [2014].

¹¹⁸ A busca foi efetuada entre 9 e 13 de novembro de 2015 na informação disponível nos *websites* oficiais das respetivas Instituições de Ensino Superior dos cursos incluídos na Tabela 11 e na Tabela 13.

¹¹⁹ Os 5 termos excluídos por não registarem ocorrências nas fontes consultadas referentes aos cursos de CI foram: *Class*; *GRDDL* (*Gleaning Resource Descriptions from Dialects of Languages*); *Property*; *RDFa* (*Resource Description Framework in attributes*) e *SPARQL* (*SPARQL Query Language for RDF*).

¹²⁰ Refira-se que *ontologia* é objeto de estudo em CI (entendendo esta como área abrangente não como disciplina científica formalizada) desde o século XIX (Almeida, 2013, p. 1532).

Tabela 14

Listagem de termos com três ou mais ocorrências nas fontes de WS e com uma ou mais ocorrências nos cursos de CI

Termos e suas variações encontradas nas fontes consultadas	Ocorrências			Somadas	
	fontes WS [A]	Cursos CI		[A] +	[A] +
		[B]	[C]	[B]	[C]
<i>Inference/ Inferred hierarchy/ Inferência/ Sistemas de classificação inferenciais</i>	3	2	3	5	6
<i>Linked data/ Open data/ Linked open data (LOD)/ Dados abertos interligados</i>	3	5	8	8	11
<i>Ontologies/ Ontologias</i>	3	10	33	13	36
<i>OWL (Web Ontology Language)</i>	3	2	4	5	7
<i>RDF (Resource Description Framework)</i>	3	1	3	4	6
<i>Reasoner/ Semantic reasoner/ Reasoning engine/ Rules engine/ Raciocinadores/ Automatização do raciocínio</i>	3	2	2	5	5
<i>Relation/ Relationships among those things/ Entidades e relações</i>	3	1	1	4	4
<i>Semantic web/ Web semântica</i>	4	7	17	11	21
<i>SKOS (Simple Knowledge Organization System)</i>	3	1	2	4	5

Nota. [A]: quantidade de ocorrências encontradas nos glossários/listas terminológicas; [B]: quantidade de ocorrências encontradas por curso, i.e., cada termo foi contabilizado uma única vez por curso, independentemente das vezes que ocorreram no mesmo curso; [C] quantidade de ocorrências encontradas nos programas, conteúdos ou ementas das disciplinas e linhas de pesquisa, todas as ocorrências foram contabilizadas.

Como indicadores da relação entre os cursos de CI e as ontologias fora do contexto da WS, podem apontar-se dois: o primeiro indicador relaciona-se com os quatro cursos onde não foram encontradas qualquer menção ao termo Web Semântica apresentando menções ao termo Ontologias (destacando-se um desses cursos com 6 menções a este último termo¹²¹); o segundo, com a diferença expressiva do número total de ocorrências entre os dois termos em confronto, 33 para o primeiro e 17 para o segundo, (cf. Tabela 15).

A inclusão do termo ontologia implicaria, ainda, uma clarificação conceitual do termo atendendo à sua utilização, por vezes dúbia, em contextos e áreas alargados (Almeida, 2013, p. 1532; R. R. Souza et al., 2012, p. 187). Uma tal clarificação complexa¹²² que, a somar ao inevitável alargamento do universo do estudo, afastaria o foco do mesmo da presença da web semântica para um conceito mais abrangente, importante sem dúvida, mas que se situa fora do âmbito e dos objetivos do presente trabalho.

¹²¹ Adicionalmente ainda se pode indicar ter-se encontrado, em 8 dos 9 cursos do Brasil listados na Tabela 15, disciplinas que se debruçam sobre organização e representação do conhecimento onde as ontologias fazem parte da ementa, mas não há qualquer referência à WS.

¹²² A este propósito veja-se o artigo: Almeida, M. B. (2014). Uma abordagem integrada sobre ontologias: Ciência da Informação, Ciência da Computação e Filosofia. *Perspectivas em Ciência da Informação*, 19(3), 242–258. <http://doi.org/10.1590/1981-5344/1736>.

Tabela 15

Número total de ocorrências dos termos web semântica e ontologias nos cursos listados na Tabela 11 (Portugal) e na Tabela 13 (Brasil) onde foram encontradas menções aos termos da Tabela 14

Ref.	Instituição	Curso	Web Semântica		Ontologias	
			ocorrências	M/D	ocorrências.	M/D
Portugal						
Pt01.M	UC	Ciência da Informação	1	M	0	M
Pt02.M	UP	Ciência da Informação	2	M	4	M
Brasil						
Br01.MD	USP	Ciência da Informação	5	M, D	7	M, D
Br02.MD	UNESP/MAR	Ciência da Informação	3	M, D	4	M, D
Br04.MD	UFBA	Ciência da Informação	1	M	1	M, D
Br06.Ma	UFPE	Ciência da Informação	0	M	1	M
Br07.MD	UFSC	Ciência da Informação	3	M, D	5	M, D
Br08.MD	UFF	Ciência da Informação	2	M, D	2	M, D
Br09.MD	UFRJ	Ciência da Informação	0	M, D	2	M, D
Br11.MD	UFMG	Ciência da Informação	0	M, D	6	M, D
Br15.Mp	UNIRIO	Biblioteconomia	0	M	1	M
soma			17		33	

Nota. Todas as ocorrências encontradas nos programas, conteúdos ou ementas das disciplinas e linhas de pesquisa dos respetivos cursos, foram contabilizadas. M – cursos de mestrado; D – cursos de doutoramento. Os casos onde as ocorrências se verificaram em disciplinas comuns aos dois tipos estão identificados com a notação “M, D”.

Definido o termo de pesquisa, web semântica, procedeu-se a testes prévios nos dois repositórios, RCAAP e OASISbr, no sentido de aferir e ajustar os procedimentos a aplicar, nomeadamente no que respeita à forma de recuperar possíveis variações do termo eleito, assim como relativamente aos campos e filtros a utilizar nas respetivas plataformas digitais.

Salienta-se que os critérios foram estabelecidos de acordo com os objetivos específicos dado a articulação intrínseca destes com o processo de recolha de dados (Quivy & Campenhoudt, 1992, p. 184). Nesse sentido foram aplicados os seguintes procedimentos:

- Utilizou-se a expressão de pesquisa *web AND sem** por forma, a abranger variações do termo web semântica, e deste modo aumentar a revocação. O uso do operador AND leva a que não se observe a discriminação da ordem de apresentação dos dois termos simples, *web* e *sem*, e o auxiliar de pesquisa truncatura (*) permitir a recuperação de todas as palavras com a raiz *sem*¹²³. Esta estratégia de pesquisa concorre para a recuperação do termo na língua inglesa (*semantic web*) que, de outra forma, obrigaria a uma pesquisa a dois tempos;
- Aplicou-se a expressão selecionada, *web AND sem**, nos campos de “pesquisa avançada” *Título* e *Assunto*, comuns aos dois repositórios, utilizando como ligação entre

¹²³ A razão para a utilização da expressão de pesquisa *sem** e não *semantic** ou *semantic?* não se deve ao sinal ortográfico presente na palavra *semântica*, ambos repositórios lidam com o mesmo como se não estivesse presente, relaciona-se, antes, com uma questão técnica de alguns repositórios de origem, detetada nos testes prévios, que substituem o carácter *â* por dois pontos de interrogação (??), levando à exclusão dos respetivos trabalhos.

os dois campos o operador booleano OU. A escolha destes dois campos garante uma uniformização de critérios entre repositórios e direciona a pesquisa a trabalhos onde se assume que a WS tenha um papel central uma vez que os trabalhos recuperados deverão apresentar o respetivo termo em um, ou nos dois elementos relevantes de identificação/caraterização das DM e TD (título do documento e/ou palavras-chave);

- c. Efetuou-se a pesquisa de forma faseada, obedecendo à ordem cronológica (de 2005 a 2015) e separadamente de acordo com a tipologia de trabalhos (DM e TD), utilizando para tal os filtros prévios de refinamento de pesquisa *Data/Ano de publicação* e *Tipo de documento*;
- d. Registou-se na FRD geral (ver Apêndice 3a) os seguintes dados: título do documento (1); autoria (2); designação do curso (3); instituição (4); ano (5); tipologia (6); repositório onde foi recuperado (7); se é curso de CI (8); país (9); variação do termo recuperado (10); URL do documento (11); e outras observações (12). Importa salientar que o ano anotado nesta FRD foi o apresentado no registo do respetivo repositório;
- e. Na FRD específica, destinada aos trabalhos que constituem a amostra (ver Apêndice 3b), para além dos dados correspondentes aos campos 1 a 6 da FRD geral, recolheram-se os seguintes: departamento/escola; orientador e, nos casos aplicáveis, o coorientador. Nesta FRD foi registado o ano que consta no documento recuperado para confirmar ou retificar o mesmo;
- f. Ainda na FRD específica, finalizada a recolha dos trabalhos, registou-se a filiação e a formação académica dos orientadores e coorientadores dos trabalhos que constituem a amostra.

4.3. Universo e amostra

Dada a importância que este item assume na metodologia entendeu-se dividi-lo em dois pontos: universo e amostra.

4.3.1. Universo

Relativamente ao universo do presente estudo de caso, este constituiu-se pelos registos recuperados de acordo com os procedimentos descritos no ponto anterior, resultando em 519 registos (RCAAP com 245 e OASISbr com 274). Destes, 61 correspondem a registos duplicados reduzindo assim o total de registos para 458. Os 61 duplicados correspondem à soma de 34 registos do RCAAP com 27 do OASISbr¹²⁴. Destes 458 registos uma parte significativa encontra-se duplicada nos dois repositórios (207), como é suposto¹²⁵. Conferindo os registos em cada repositório obteve-se 251 (200 DM e 51 TD) não duplicados.

Dos 251 registos recuperados, 28 mostraram ser falsos positivos. Deste número, 24, apesar de não apresentarem o termo web semântica nos campos pesquisados (título e assunto), foram recuperados por serem abrangidos pela expressão de pesquisa usada (web AND

¹²⁴ Um esquema com a relação entre estes valores e os repositórios é apresentada no Apêndice 5.

¹²⁵ Devido à questão da interoperabilidade entre os dois repositórios, já aqui abordada.

sem*)¹²⁶. Os restantes 4, que foram apenas recuperados no RCAAP, apresentam datas anteriores ao período estabelecido (2005-2015), sendo 3 deles publicados em 2000 e o outro a 2003, pelo que não deveriam ter sido recuperados. Com a exclusão dos 28 falsos positivos restaram 223 que, após a aplicação do critério de seleção qualitativa (ser curso de CI), de acordo com o descrito na presente metodologia (ponto 4), resultou numa amostra de 22 trabalhos, (17 DM e 5 TD).

Face ao número que a constituiu, se analisada numa perspetiva quantitativa ele é reduzido, contudo, como a abordagem metodológica é de âmbito qualitativo este número não concorrerá para constrangimentos de maior, quer de resultados, quer na respetiva análise.

4.3.2. Amostra

Os 22 trabalhos que constituem a amostra dividem-se, quanto à sua tipologia, em 17 Dissertações de Mestrado (DM) e 5 Teses de Doutoramento (TD). Estes provêm, na sua totalidade, de cursos de 11 universidades brasileiras. Destas, destaca-se a Universidade Estadual Paulista (UNESP/MAR) com 7 trabalhos (6 DM e 1 TD), quantidade que representa uma percentagem expressiva da amostra, (32% arredondando às unidades), i.e., perto de um terço da amostra. Os restantes 15 encontram-se distribuídos de forma bastante equitativa pelas restantes instituições representadas (cf. Figura 17).

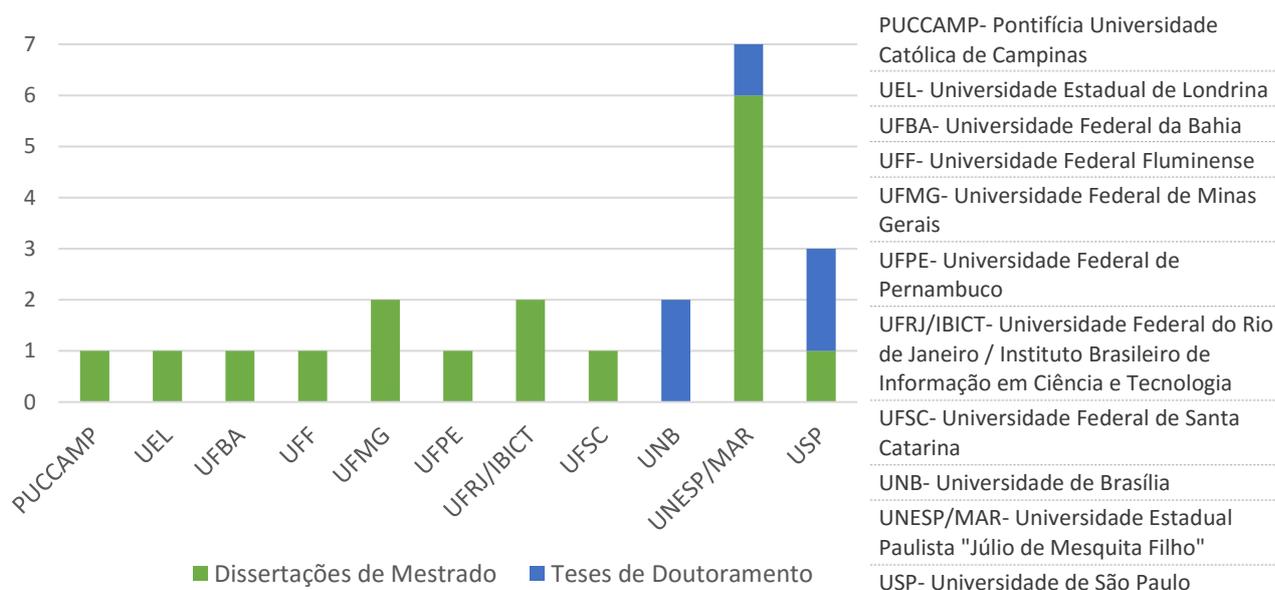


Figura 17. Distribuição da amostra pelas Instituições de Ensino Superior aí representadas.

Refira-se que os cursos das universidades de Campinas (PUCCAMP) e de Londrina (UEL) encontram-se desativados¹²⁷. O curso desta última é, da amostra, o único com uma designação diferente de Ciência da Informação, apresentando-se sob a nomenclatura de Gestão da Informação.

¹²⁶ Os termos pelos quais os registos foram recuperados e respetiva contabilização de ocorrências são: sem-fio (8, alguns no plural); semiótica (6 com algumas variações); (multi)semiose (4); semi-(automático/ automatizado/presencial/supervisionado, 4); 1- *web-semp* (1); *step semantic* (1).

¹²⁷ Informação recolhida em <https://sucupira.capes.gov.br/> acessado a 14-8-2016.

Os trabalhos que foram recuperados, produzidos em instituições de ensino superior portuguesas, são originários de cursos de outras áreas que não a CI, apesar de, nesses cursos, constar um trabalho pertencente à Universidade de Aveiro, cuja designação do curso é também: Gestão da Informação. Esse curso, desativado desde 2012, encontrava-se enquadrado na Área de Educação e Formação – Gestão e Administração, do Departamento de Economia, Gestão e Engenharia Industrial¹²⁸. Na Tabela 16 são listados os 22 trabalhos que constituem a amostra do estudo.

Tabela 16 (continua)

Identificação dos trabalhos que compõem a amostra do presente estudo de caso

Ref.	Autoria	Título do trabalho	Universidade
<i>Dissertações de Mestrado</i>			
dm.01.2005	Rachel Alves	Web semântica: uma análise focada no uso de metadados	UNESP/MAR
dm.02.2005	Marcelo Jorge	Ontologias no suporte a portais semânticos	UFMG
dm.03.2006	Rogério Ramalho	Web semântica: aspetos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação	UNESP/MAR
dm.04.2008	Cintia Pinheiro	A construção do conhecimento científico: a web semântica como objeto de estudo	UNESP/MAR
dm.05.2008	Sinara Molossi	Inserção da biblioteca digital de teses e dissertações no contexto da web semântica	UFSC
dm.06.2008	Fabiano Castro	Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da ciência da informação: uma abordagem do <i>marcont initiative</i> na era da web semântica	UNESP/MAR
dm.07.2009	Rodrigo Carvalho	Perspetivas na web semântica para a Ciência da Informação	PUCAMP
dm.08.2010	Maria Menezes	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	UFBA
dm.09.2011	Eliana Demarques	Estudo exploratório sobre compatibilização semântica de vocabulário utilizando ontologia: fundamentação teórica-metodológica visando ao aperfeiçoamento do MHTX	UFMG
dm.10.2011	Ana Araújo	Registros de marcas da junta comercial do estado de Pernambuco: um olhar semântico para a organização de um acervo histórico	UFPE
dm.11.2012	Déborah Ambinder	Artigos científicos digitais na web: novas experiências para apresentação, acesso e leitura	UFF
dm.12.2012	Ilza Andrade	As dimensões semântica e pragmática da web e dos mecanismos de busca no ciberespaço	UEL
dm.13.2012	Amanda Sousa	Organização em sistema caótico: uso das <i>tags</i> para classificação da informação pelos usuários da rede social <i>twitter</i>	UFRJ/IBICT
dm.14.2013	Francisco Gomes	Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária	UFRJ/IBICT
dm.15.2013	Renata Silva	As tecnologias da web semântica no domínio bibliográfico	UNESP/MAR
dm.16.2013	Camila Ribeiro	Representação da informação dinâmica em ambientes digitais	USP
dm.17.2014	Jaidier Ferreira	<i>Wikis</i> semânticos: da web para a web semântica	UNESP/MAR

Nota. A referência atribuída a cada documento identifica o ano que consta no mesmo nos quatro últimos dígitos.

¹²⁸ Apesar do curso não se encontrar listado em qualquer das entidades consultadas (DGEEC, A3ES ou DGE), ao contrário de outros cursos desativados, procedeu-se a contatos com o departamento da UA, do qual o curso dependia, que confirmaram a inclusão do mesmo na área Gestão e Administração (Pêgo, comunicação pessoal, abril, 14, 2016). Outros cursos com essa mesma designação, como os da Universidade Nova de Lisboa, apresentam o mesmo enquadramento.

Tabela 16 (continuação)

Identificação dos trabalhos que compõem a amostra do presente estudo de caso

Ref.	Autoria	Título do trabalho	Universidade
<i>Teses de Doutorado</i>			
td.01.2006	Edgard Oliveira	Autoria de documentos para a web semântica: um ambiente de produção de conhecimento baseado em ontologias	UNB
td.02.2010	José Segundo	Representação iterativa: um modelo para repositórios digitais	UNESP/MAR
td.03.2013	Fernanda Monteiro	Web semântica e repositórios digitais educacionais na área de saúde: uma modelagem com foco no objetivo de aprendizagem para refinar resultados de busca	UNB
td.04.2015	Julietti Andrade	Interoperabilidade e mapeamentos entre sistemas de organização do conhecimento na busca e recuperação de informações em saúde: estudo de caso em ortopedia e traumatologia	USP
td.05.2015	Denyson Mota	Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática: um estudo crítico	USP

Nota. A referência atribuída a cada documento identifica o ano que consta no mesmo nos quatro últimos dígitos.

Relativamente à distribuição da amostra ao longo do período considerado, esta apresenta uma distribuição irregular (cf. Figura 18).

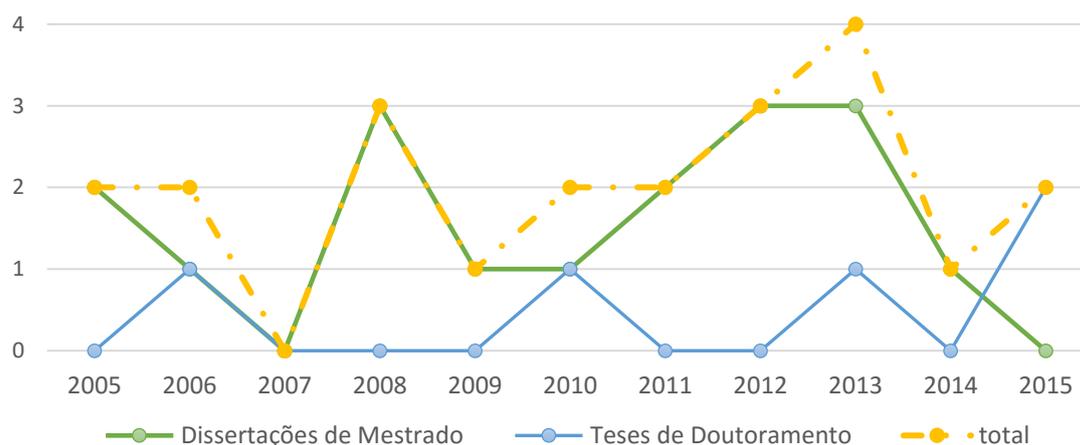


Figura 18. Distribuição da amostra pelo período considerado.

4.4. Critérios usados na análise dos dados

O conjunto de técnicas e procedimentos metodológicos aplicado ao presente estudo foi a análise de conteúdo. Segundo a definição de Bardin a análise de conteúdo é:

Um conjunto de técnicas de análise das comunicações visando obter por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) dessas mensagens. (Bardin, 2011, p. 40)

Afirma o mesmo autor que: “a técnica de análise de conteúdo adequada ao domínio e ao objetivo pretendidos tem de ser reinventada a cada momento,” (Bardin, 2011, p. 36). Partindo deste contexto, sentiu-se a necessidade de se adotarem duas abordagens diferentes, dado o facto de o cumprimento dos objetivos considerados assim o exigirem. Desse modo,

no que respeita à concretização dos objetivos a) e b)¹²⁹ o nível analítico desta técnica foi mais exaustivo, relativamente ao cumprimento do objetivo c)¹³⁰, cuja análise de conteúdo se pautou apenas pela contabilidade de frequência e a aferição de categorias.

No presente estudo de caso a análise de conteúdo incidiu sobre dois conjuntos de dados. O primeiro relativo aos 223 trabalhos que resultaram após os procedimentos descritos no ponto 4.3; o segundo conjunto referente à amostra (22 trabalhos). Salienta-se que a análise do primeiro teve a ver apenas com a presença da Web Semântica enquanto tema dos trabalhos recuperados; no segundo, a análise incidiu sobre a presença da Web Semântica nos resumos e nos objetivos dos trabalhos dos cursos de CI e ainda, no que respeita à segunda abordagem referida a cima, sobre a relação entre as áreas CI e Informática na formação académica dos orientadores. Deste modo, apresenta-se uma breve descrição da análise aplicada aos dois conjuntos de dados.

4.4.1. A presença da WS enquanto tema dos trabalhos recuperados

No que concerne aos dados referentes aos 223 trabalhos, a análise incidiu sobre os cursos que os mesmos representam efetuando-se, quando aplicável, uma comparação entre os subconjuntos formados pelos trabalhos de cada país (49 de Portugal e 174 do Brasil).

O facto de a metodologia utilizada na recolha de dados não filtrar os trabalhos por cursos específicos, permitiu efetuar uma comparação entre a amostra, enquanto conjunto representativo da presença da Web Semântica na área da Ciência da Informação, com a mesma presença nos restantes cursos de diferentes áreas. Para o efeito usou-se uma adaptação simplificada do método de *adequação a padrão* exposto por Yin (2001, pp. 136-140), criando-se um esquema de classificação contendo apenas três categorias para organizar os cursos: *Ap*- CI; *Bp*- Informática/Computação; e *Cp*- outros cursos. Estes grupos constituem os “padrões de base prognóstica” (Yin, 2001, p. 136) os quais foram comparados aos inferidos pela observação empírica.

Como indicadores de referência para a classificação dos cursos, representados pelos 223 trabalhos, utilizou-se as Áreas de Educação e Formação (AEF), no caso dos cursos portugueses e as Áreas Básicas ou do Conhecimento (ABC), para os brasileiros. O conjunto das AEF e ABC será designado, no presente trabalho por *áreas de educação/conhecimento*¹³¹. Essas classificações (AEF e ABC) são as que se encontram aplicadas nos sistemas em linha das respetivas agências nacionais: DGEEC (Portugal) e CAPES (Brasil).

4.4.2. A presença da WS nos trabalhos dos cursos de CI recuperados

Partindo da metodologia exposta por Bardin (2011, pp. 134–144) no que diz respeito às unidades de registo, unidades de contexto e regras de enumeração foi adotado o método que agora se apresenta para a análise dos resumos e objetivos dos 22 trabalhos que constituem a amostra.

¹²⁹ Refere-se aos pontos: 4.4.1 e 4.4.2, respetivamente.

¹³⁰ Refere-se ao ponto: 4.4.3.

¹³¹ Importa clarificar que as Áreas de Educação e Formação representam as áreas de 3º nível da CNAEF (Portugal) e as Áreas Básicas ou do Conhecimento correspondem às de 2º nível da TAC (Brasil), conforme descrito nos pontos 2.4.1 e 2.4.2, assim, o conjunto referido engloba todas essas áreas.

a. Resumos dos trabalhos

No que respeita aos resumos, a análise partiu de uma categorização em duas classes: uma referente aos pontos de confluência entre a CI e a WS, a outra aos elementos constituintes da estrutura interna dos resumos. A diferente natureza das duas classes condicionou a seleção do processo para a constituição das suas subclasses. Segundo Bardin, esse processo pode ser efetuado à priori, “quando a organização do material decorrer diretamente dos funcionamentos teóricos hipotéticos”; ou, de forma inversa, quando “o sistema de categorias não é fornecido, antes resulta da classificação analógica e progressiva dos elementos.” (Bardin, 2011, p. 149). Estes processos são designados pelo autor como: “procedimento por *caixas*”, relativamente ao primeiro, e “procedimento por *acervo*”, no que concerne ao segundo (Bardin, 2011, p. 149).

Nesse contexto, para a classe dos pontos de confluência entre a CI e a WS, dado as subclasses dependerem da análise do discurso dos autores, optou-se pelo *procedimento por acervo*; para a outra classe mencionada, a estrutura interna do resumo, utilizou-se o *procedimento por caixas*, uma vez as subclasses se encontrarem previamente definidas pelos respetivos elementos: introdução, objetivos, metodologia e conclusões (IOMRC).

Como unidades de registo elegeram-se a limitação semântica *tema*, de acordo com a definição apresentada por Bardin. Para este autor:

O tema é a unidade de significação que se liberta naturalmente de um texto analisado segundo certos critérios relativos à teoria que serve de guia à leitura. O texto pode ser recortado em ideias constituintes, em enunciados e em proposições portadores de significações isoláveis. (2011, p. 135).

A essas unidades de registo foi aplicada uma “análise de séries temporais”, uma vez que esta permite determinar os eventos causais ao longo do tempo (Yin, 2001, p. 147).

b. Objetivos dos trabalhos

Para a análise dos objetivos dos trabalhos, utilizou-se o termo Web Semântica como unidade de registo à qual foi aplicada uma análise quantitativa da seguinte forma: i. em termos da sua *presença*¹³² nos objetivos gerais; ii. na sua *frequência* nos objetivos específicos. Relativamente à *intensidade*, no que diz respeito a estes dois critérios em particular, foi dado um maior peso à presença do termo considerado quando se encontrava registado nos objetivos gerais, na medida em que é o objetivo geral que, usualmente, traduz o tema do trabalho.

Às mesmas unidades de registo foi aplicada uma análise da sua relação com as unidades de contexto consideradas, que correspondem às proposições onde as unidades de registo se inserem. Para tal, identificou-se o verbo da ação, o objeto dessa ação e, nos casos aplicáveis, o complemento contextual. Nesta fase da análise o objeto da ação foi o elemento com maior peso, sendo que, obedecendo à mesma lógica da primeira fase da análise, aos objetos das ações dos objetivos gerais foi dado um peso superior relativamente aos dos objetivos específicos.

A combinação das duas análises determinou a *direção* expressa no que se designou por “grau de proximidade” de acordo com os critérios explicitados na Figura 19.

¹³² Os termos em itálico referem-se a conceitos expressos na obra citada de Bardin (2011, pp. 134–144).

Como é possível constatar na mesma figura, optou-se por uma *escala de intensidade negativa*, correspondendo o valor zero à máxima proximidade entre os objetivos do trabalho e a WS enquanto tema do mesmo.

Critérios usados na análise do termo Web Semântica quanto à			Grau de proximidade atribuído	
função desempenhada nas unidades de contexto consideradas.	presença/frequência nos objetivos.		quantitativo	qualitativo
	Geral	Específicos		
Configura-se como o próprio objeto da ação.	Apresenta.	Em 50% ou mais.	0	maior ↓ menor
Enquanto objeto da ação apenas se encontra nos objetivos específicos.			-1	
Surge como auxiliar do objeto da ação ou como elemento contextual.		Em menos de 50%.	-2	
	Surge no geral ou em um dos específicos e não nos dois em simultâneo.	-3		
Relação conceitual não explicitada.	Não apresenta.	Igual a 0%.	-4	

Figura 19. Relação dos critérios usados na análise do termo Web Semântica nos objetivos dos 22 trabalhos que constituem a amostra.

Os objetivos foram extraídos do corpo dos trabalhos, mais especificamente do ponto relativo aos mesmos ou, na ausência deste, da introdução.

4.4.3. O perfil académico dos orientadores e coorientadores

No que respeita ao perfil académico dos orientadores e coorientadores dos trabalhos que constituem a amostra, foram estabelecidas três categorias para a aferição da formação dos mesmos, a saber: i. pós-doutoramento e doutoramento (pD/D); ii. mestrado e especializações após a licenciatura (M/Esp); iii. licenciatura e bacharelato (L/B). Como indicador adicional foi usado o ano de conclusão das formações que se encontrava registado nos respetivos currículos (que constam na plataforma *Lattes*). Este indicador foi apenas usado em situações nas quais a designação da formação suscitava dúvidas, dada a heterogeneidade de percursos formativos apresentados pelos docentes em questão.

Após a classificação referida foi efetuada a contabilização da frequência dos docentes que possuíam formação na área da CI e em outras áreas. Para a classificação das áreas de formação foi aplicada a descrita no ponto 4.4.1.

5. Apresentação e análise dos resultados

Neste ponto, tal como o título refere, serão apresentados e analisados os resultados deste estudo. Sempre que seja pertinente far-se-á uma articulação entre os resultados propriamente ditos e os conhecimentos explanados na Parte I, relativa à contextualização teórica.

5.1. A presença da Web Semântica enquanto tema dos trabalhos recuperados

Após a comparação entre os grupos utilizados como padrões de base prognóstica¹³³ (*Ap*- CI, *Bp*- Informática/Computação, *Cp*- outros cursos) e os dados recolhidos, considerou-se necessária a introdução de um quarto grupo que se situou na fronteira entre os grupos *Bp* e *Cp*. Desse modo, a divisão dos cursos não foi efetuada por três, mas sim por quatro grupos¹³⁴, cuja distribuição se apresenta na tabela que se segue (cf. Tabela 17).

Tabela 17

Distribuição dos 223 trabalhos, de acordo com o curso que representam, pelas Áreas de Educação e Formação (Portugal) ou pelas Áreas Básicas ou do Conhecimento (Brasil) onde os respetivos cursos estão classificados

Áreas de Educação e Formação (em uso na DGEEC - Portugal)	Σ .Pt	Áreas Básicas ou do Conhecimento (em uso na CAPES - Brasil)	Σ .Br	Σ .Tt
<i>Grupo A</i>	0		22	22
BAD (Biblioteconomia, Arquivo e Documentação)		Ciência da Informação		
<i>Grupo B</i>	30		113	143
Ciências Informáticas		Ciência da Computação		
<i>Grupo C</i>	15		32	47
Eletrónica e Automação		Automação Eletrônica de Processos Elétricos e Industriais	1	
		Engenharia Elétrica	23	
		Engenharia/Tecnologia/Gestão	8	
<i>Grupo D</i>	4		7	11
Ciências da Educação	1	Administração	1	
Ciências Dentárias	1	Desenho Industrial	1	
Gestão e Administração	1	Engenharia Mecânica	1	
Tecnologia dos Processos Químicos	1	Linguística	1	
		Saúde e Biológicas	1	
		Sociais e Humanidades	1	
		Tratamento e Prevenção Psicológica	1	
soma	49		174	223

Nota. As áreas incluídas nos grupos referem-se às de 3º nível, no caso das de Portugal, e de 2º nível, para o Brasil, de acordo com as classificações em uso nas respetivas entidades nacionais (DGEEC e CAPES). Nos grupos com mais de uma área a distribuição dos trabalhos pelas mesmas é apresentada na respetiva linha. Σ .Pt: somas relativas às áreas de Portugal; Σ .Br: somas relativas às áreas do Brasil; Σ .Tt somas totais de cada grupo.

Refere-se que os grupos *A* e *B* apresentam uma só Área de Educação/Conhecimento para cada país e os outros dois (grupos *C* e *D*), no caso do Brasil apresentam mais de uma e, para Portugal, apenas o último grupo apresenta mais de uma Área.

¹³³ Conforme apresentado no ponto 4.4.1.

¹³⁴ A propósito da classificação de “saberes” é interessante observar o que Olga Pombo refere: “Também as classificações das ciências são classificações reais e, enquanto tal, não escapam às irregularidades e anomalias de que só as classificações ideais estariam isentas.” (2002, p. 11)

O *Grupo C*, que na abordagem inicial se encontrava diluído nos grupos *Bp* e *Cp*, revelou a coesão e a pertinência suficiente para se apresentar como autónomo tornando, assim, a classificação mais exaustiva, não comprometendo, todavia, a simplicidade pretendida da mesma.

A distribuição percentual dos quatro grupos concebidos, relativamente ao total de trabalhos de cada país, (49 de Portugal (100%) e 174 do Brasil (100%)), pode ser observada nos gráficos da Figura 20.

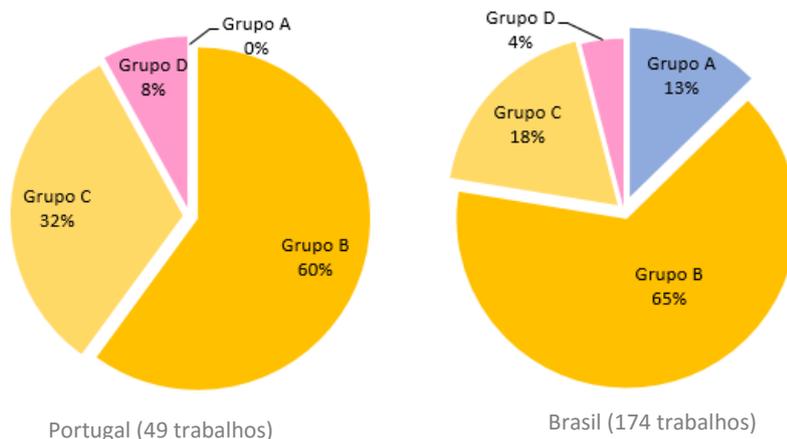


Figura 20. Distribuição dos 223 trabalhos recuperados (DM e TD), de acordo com o país de origem (49 de Portugal e 174 do Brasil), pelos grupos estabelecidos no presente trabalho para organização das áreas de educação/conhecimento em uso pelas entidades DGEEC e CAPES.

É possível verificar uma relação evidente, entre os dois países, na percentagem de trabalhos incluídos no grupo *B*, Informática/Computação, 61%¹³⁵ em Portugal e 65% no Brasil. Percentagens que, além de serem expressivas (superiores a 50%) são também previsíveis dado o tema dos trabalhos ser a Web Semântica, assunto naturalmente abordado nesta área. Verifica-se que a percentagem do grupo *C* (Eletrónica/Automação) em Portugal apresenta um valor equivalente à soma da percentagem deste grupo com a do grupo *A* (CI) no Brasil¹³⁶.

A distribuição percentual da totalidade dos trabalhos relativa aos dois países, distribuída pelos quatro grupos apresenta-se na Figura 21, composta por dois gráficos: o Gráfico *Trabalhos* e o Gráfico *Cursos*. Importa referir que, ao total de trabalhos (223) representados o respetivo gráfico corresponde a 41 cursos com diferentes designações (ver Apêndice 6) representados, também, no gráfico respetivo.

Verifica-se, nos grupos *B*, *C* e *D*, uma distribuição mais uniforme na quantidade de cursos relativamente à dos trabalhos, pesa embora o grupo *B* continuar a ser o que apresenta um maior número de cursos com diferentes designações. É importante salientar que o grupo considerado apresenta apenas uma área em cada país (Ciências Informáticas em Portugal e Ciência da Computação no Brasil¹³⁷). Interessa, ainda, observar que este grupo revela maior

¹³⁵ As percentagens apresentadas ao longo do presente ponto e seus subpontos serão arredondadas às unidades.

¹³⁶ Brasil: 13% (grupo A) + 18% (grupo C) do = 32% e Portugal 31% (grupo C).

¹³⁷ É interessante verificar que, na listagem da CAPES, a Área *Ciência da Computação* não inclui os cursos com a designação Engenharia da Computação das universidades de São Paulo e da estatal de Rio de Janeiro, o primeiro incluído na Área *Engenharia Elétrica* e o segundo na Área *Engenharia/Tecnologia/Gestão* (ver Apêndice 6).

dispersão de designações¹³⁸ que a da Área onde a CI é posicionada dentro das classificações analisadas (CNAEF e TAC, conforme exposto no ponto 2.4).

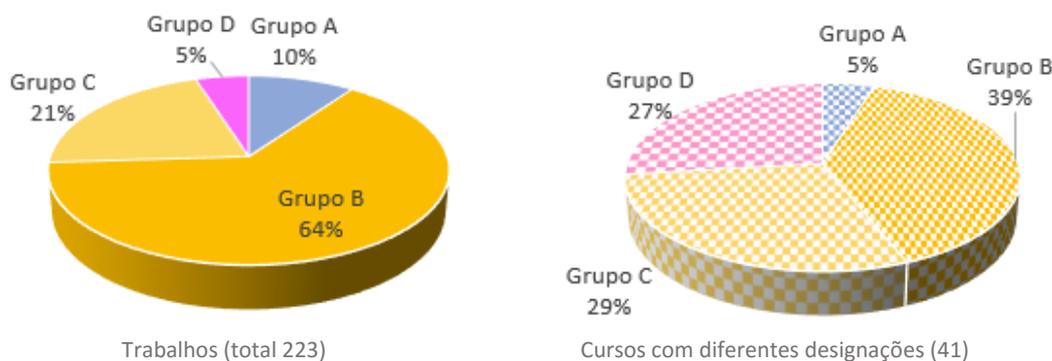


Figura 21. Distribuição dos 223 trabalhos recuperados pelos grupos estabelecidos no presente trabalho para organização das Áreas de Educação/Conhecimento e dos 41 cursos com designações diferentes (representados nos 223 trabalhos) pelos grupos estabelecidos no presente trabalho para organização das Áreas de Educação/Conhecimento.

5.2. A presença da Web Semântica nos trabalhos dos cursos de Ciência da Informação

A análise inicial dos 22 trabalhos que constituem a amostra incidu sobre os resumos dos mesmos (ver Apêndice 7) uma vez ser esse o elemento que permite ter uma visão global dos respetivos trabalhos. Essa pré-análise reforçou a opção metodológica tida no presente estudo no que se refere à seleção, num tempo posterior, dos *objetivos* dos trabalhos¹³⁹ (ver mesmo Apêndice 7). Dos cinco elementos da estrutura interna dos resumos, foi no elemento *objetivos* que se verificou a maior percentagem dos pontos de confluência entre a WS e a CI (42%). Este e outros resultados serão devidamente explanados nos respetivos subpontos que se seguem.

Os resultados da análise das duas variáveis (resumos e objetivos) serão apresentados, em primeiro lugar, de forma separada no respetivo subponto e, num segundo momento, de forma conjunta.

5.2.1. Análise dos resumos dos trabalhos em Ciência da Informação

Como primeiro resultado refere-se que, dos 22 (100%) resumos analisados, apenas em 13 (59%) destes foram detetadas referências claras a pontos de convergência entre a WS e a CI. Uma percentagem inferior à esperada, representando pouco mais de metade da amostra. É sobre esta subamostra, composta por 13 trabalhos, que a análise incidu. A distribuição das referências pela amostra e subamostra é apresentada na Figura 22.

¹³⁸ No caso de Portugal a designação no plural *Ciências Informáticas* pode funcionar como uma “justificação” para a dispersão no nome dos cursos, no entanto, no caso do Brasil a situação é idêntica à da CI pois a área é designada no singular *Ciência da Computação*.

¹³⁹ Uma vez que na pré-análise dos resumos se verificou falta de clareza quanto aos objetivos de alguns dos trabalhos e dado a importância que estes indicavam ter para os objetivos do presente trabalho, optou-se por analisar os objetivos dos trabalhos constituintes da amostra expressos fora dos resumos, i.e., explanados no corpo do trabalho.

Tal como referido na alínea a. do ponto 4.4.2, para a classe *pontos de confluência* entre a CI e a WS, utilizou-se o *procedimento por acervo* para determinar as respectivas subclasses. Atendendo à análise do discurso dos autores considerou-se pertinente agrupar as unidades de registo em quatro subclasses, a saber: *contributo da WS para a CI* (a); *contributo da CI para a WS* (b); *WS como tema na fronteira entre CC (Ciência da Computação) e CI* (c); *uso conjunto do conhecimento CI-WS* (d).

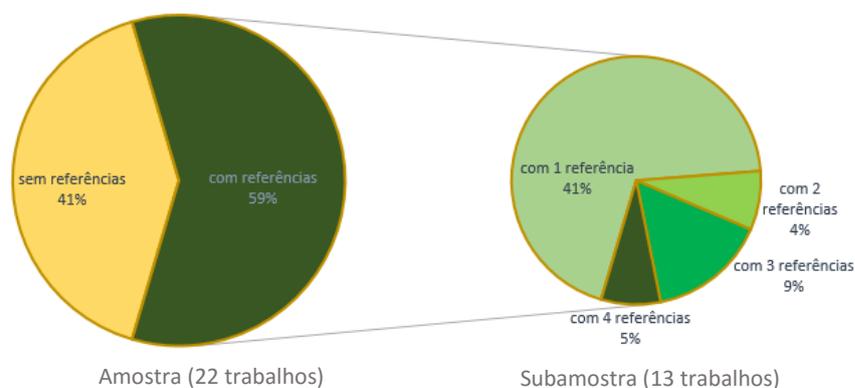


Figura 22. Distribuição das referências a pontos de convergência entre a WS e a CI pela amostra (22 trabalhos) e subamostra (13 trabalhos).

Apesar das quatro classes definidas, em rigor, os pontos de confluência entre a CI e a WS encontrados foram três:

1. A existência de uma contribuição mútua (subclasses (a) e (b));
2. A WS enquanto tema interdisciplinar (subclasse (c));
3. Os conhecimentos, da CI e da WS, são complementares (subclasse (d)).

No entanto, uma vez que se observou uma orientação para a contribuição entre estas duas áreas, considerou-se oportuna a especificação da mesma, resultando o ponto (1.) *contribuição mútua* nas duas subclasses referidas.

No mesmo ponto da metodologia (ponto 4.4.2.a) foi referido a aplicação de uma segunda classe, a *estrutura interna do resumo*, da qual se determinaram cinco subclasses: introdução (i.); objetivos (ii.); metodologia (iii.); resultados (iv.); conclusões (v.). A aplicação desta classe em conjunto com a *pontos de confluência* às unidades de registo¹⁴⁰ extraídas dos resumos produziu uma rede de relações que será apresentada de forma parcelar (pelas quatro subclasses *pontos de confluência*) dado a complexidade do esquema, no seu todo, ser contrária ao seu objetivo enquanto instrumento de análise.

Inicia-se a apresentação com a subclasse *contributo da WS para a CI* (a), onde é possível verificar a concentração das ligações (a vermelho) entre as unidades de registo¹⁴¹ e a subclasse *objetivos* (ii.) (cf. Figura 23). A quantidade de ligações, 6, mostra-se relevante dado essa subclasse apresentar 8 ligações ao todo (as outras duas ligações encontram-se em outras duas subclasses, relativas aos pontos de confluência). Essas ligações provêm de unidades de registo pertencentes a 7 trabalhos diferentes (ligações a cinza), o que poderá ser

¹⁴⁰ As unidades de registo consideradas são identificadas nos respetivos resumos a negrito no Apêndice 7.

¹⁴¹ Dado a extensão das unidades de registo, as mesmas são representadas por uma adaptação sucinta do seu conteúdo nas figuras correspondentes à análise dos pontos de confluência (da Figura 23 à Figura 26).

interpretado como uma intenção, materializada nos objetivos, dos seus autores em procurar o contributo da WS para a CI, com esta orientação em particular (da WS para a CI).

Analisando essas 6 “intenções” é possível detetar uma especificação dos contributos procurados:

1. O contributo da WS para as tecnologias de recuperação de informação (RI) (P6: dm.05.2008)¹⁴²;
2. O contributo das tecnologias usadas para a construção da WS para a representação de informação, sob a perspectiva dos fundamentos da catalogação (P7: dm.06.2008);
3. O contributo da WS para a CI (P8: dm.07.2009);
4. O contributo da WS para o processo de RI (P10: td.02.2010);
5. O contributo de “temas” relacionados com a WS (em particular metadados e ontologias) para a descrição semântica, sob a perspectiva da CI (P12: dm.10.2011);
6. O contributo dos conceitos e tecnologias utilizados pela WS para a modelagem de metadados de catálogos bibliográficos CI (P17: dm.15.2013).

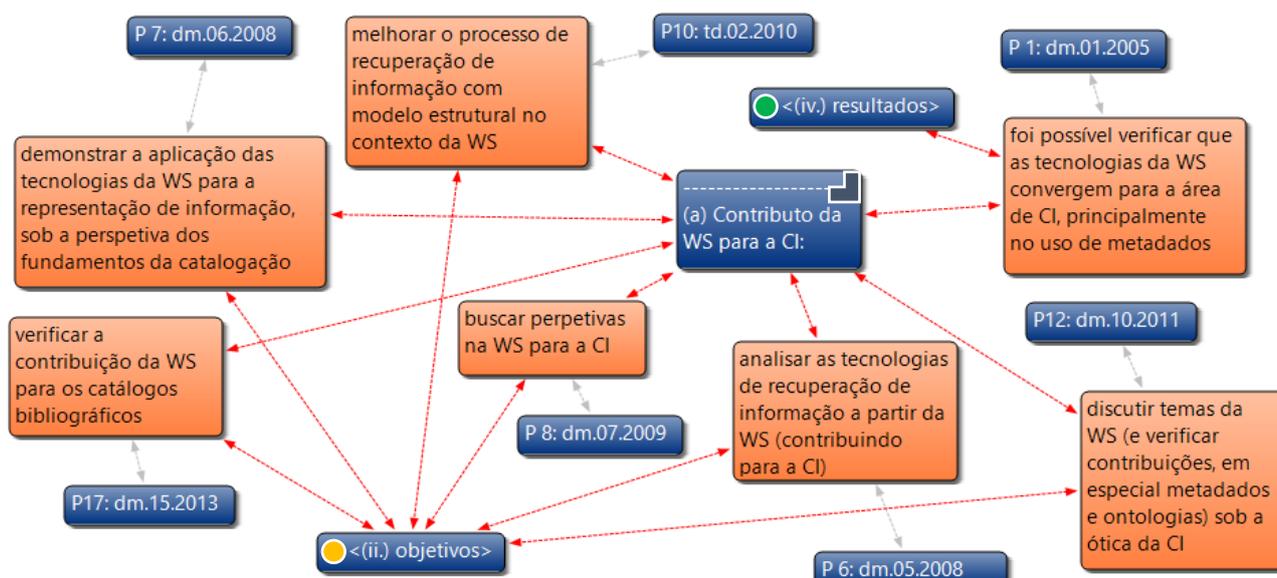


Figura 23. Diagrama de relacionamento entre a subclasse *contributo da WS para a CI* (a) e respectivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respectivos documentos de origem, identificados pela referência da Tabela 16.

À exceção do objetivo do trabalho *dm.07.2009*, todos os restantes expressam uma especificidade (que varia de trabalho para trabalho¹⁴³), a qual poderá explicar a ausência, nos respectivos resumos, de referências explícitas a pontos de convergência como resultados e/ou conclusões.

Ainda relativamente a este documento (*dm.07.2009*), cujo objetivo revela uma abordagem mais geral, o mesmo apresenta, de forma coerente, dois pontos de confluência no elemento

¹⁴² A referência numérica com o prefixo “P” corresponde ao código atribuído pelo *software* usado na análise dos documentos (*Atlas.ti* v.7.5.12).

¹⁴³ As especificidades referidas, para os contributos, são: os processos de recuperação (*dm.06.2008* e *td.02.2010*) e de representação (*dm.06.2008*) de informação; a descrição semântica de recursos informacionais (*dm.10.2011*); e a modelagem de metadados de catálogos bibliográficos (*dm.15.2013*).

conclusões em outras duas subclasses. Uma delas (subclasse (b)) refere-se ao mesmo ponto de convergência que a primeira subclasse – a contribuição mútua, a outra (subclasse (c)) refere-se a um outro ponto – a WS enquanto tema interdisciplinar. Esse documento apresenta, ainda, na metodologia do seu resumo, um quarto ponto de convergência incluído na subclasse (b), somando um total de 4 referências, sendo o único trabalho onde tal ocorre.

Também com a referência a um ponto de convergência mais geral (o direcionamento das “tecnologias da Web Semântica” para a CI), o resumo de um outro documento, *dm.01.2005*, apresenta uma situação inversa aos anteriores, dado que esse ponto surge como um resultado e não como um objetivo.

Na subclasse *contributo da CI para a WS* (b), todos os elementos encontram-se representados nas ligações às unidades de registo com a subclasse mencionada (cf. Figura 24). O primeiro elemento do resumo, *introdução*, tem nesta subclasse a sua única ligação às unidades de registo consideradas. O autor desse documento (*dm.17.2014*) não procura a relação entre as duas áreas, situação que se observa nos outros trabalhos, mas pelo contrário, considera esse facto uma premissa. Note-se que este trabalho faz parte dos 3 que foram publicados mais recentemente (este de 2014 e os outros dois de 2015).

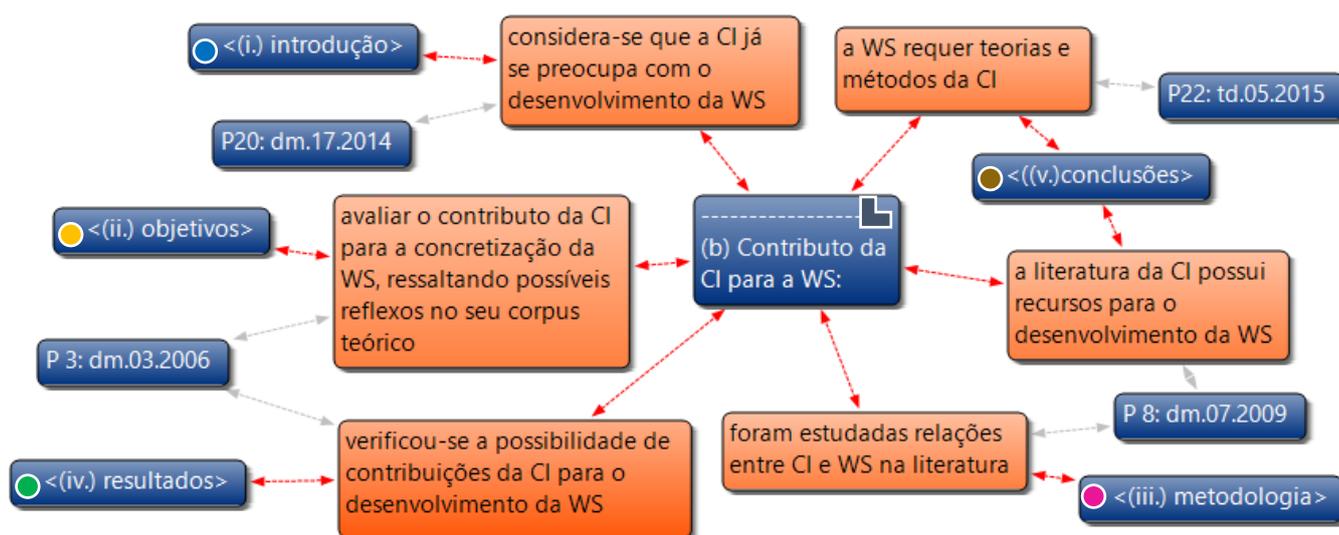


Figura 24. Diagrama de relacionamento entre a subclasse *contributo da CI para a WS* (b) e respetivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respetivos documentos de origem, identificados pela referência da Tabela 16.

O entendimento da contribuição da CI para a WS, referido no parágrafo anterior, é reforçado na conclusão do trabalho *td.05.2015*, onde se refere que o desenvolvimento “de forma adequada” da WS requer esse contributo. Estas duas afirmações, feitas em datas recentes (2014 e 2015) contrastam, de forma assertiva, com as efetuadas em datas anteriores, como é observado nos trabalhos de 2006 (*dm.03.2006*) e de 2009 (*dm.07.2009*). No trabalho datado de 2006 pretende-se avaliar o contributo da WS para a CI; no de 2009 efetua-se o estudo entre as duas áreas.

A subsecção (c) (esquema da Figura 25) envolve uma questão já abordada neste trabalho – a interdisciplinaridade (ver ponto 2.2.3.a). É interessante verificar a consistência da sua presença nos elementos do resumo, excetuando-se apenas a introdução.

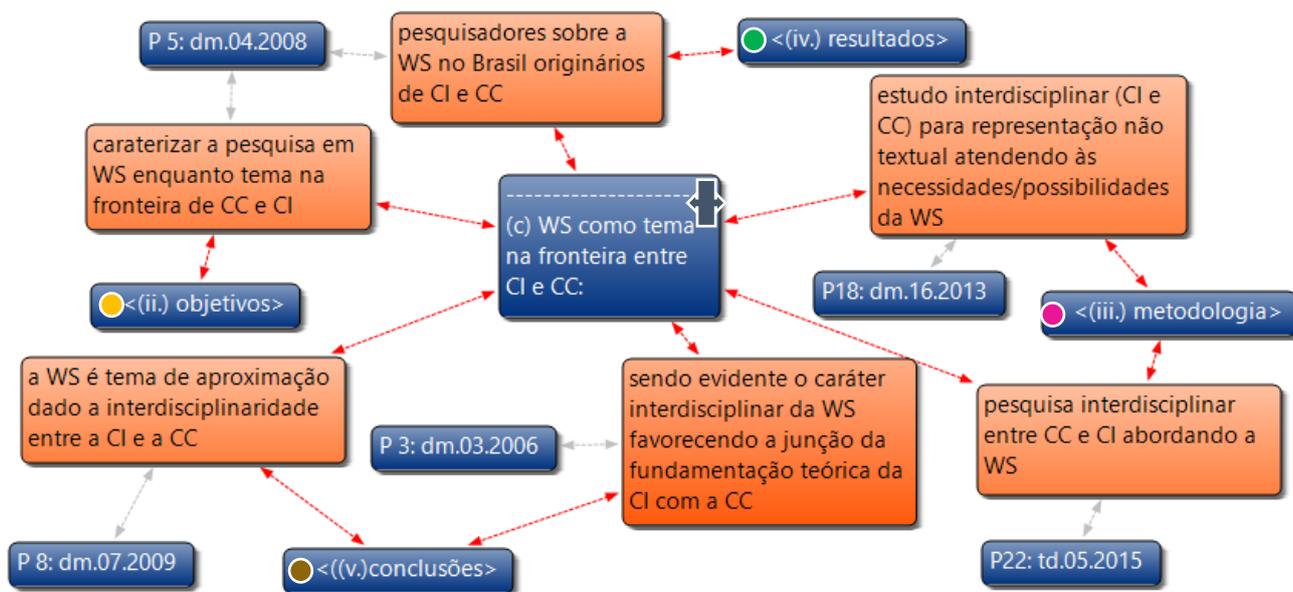


Figura 25. Diagrama de relacionamento entre a subclasse *WS como tema na fronteira entre CI e CC* (c), e respectivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respectivos documentos de origem (DM/ TD, identificados pela referência da Tabela 16).

Sendo o foco deste ponto de confluência entre a WS e a CI, a relação entre esta última e a Ciência da Computação (CC), dois dos quatro trabalhos representados nestas relações referem igualmente a Linguística (*dm.04.2008* e *dm.07.2009*). No trabalho *dm.07.2009* é ainda mencionado a Inteligência Artificial. A CC, a Linguística e a Inteligência Artificial, esta última associada à CC ou à Ciência Cognitiva, encontram-se presentes no histórico do relacionamento disciplinar da CI, apresentado no ponto 2.3 (ver Figura 16).

Por fim, a subsecção (d), *uso conjunto do conhecimento CI-WS*, apresenta duas unidades de registo de duas TD diferentes, (*td.04.2015* e *td.05.2015*), ambas do último ano do período seleccionado para o presente estudo (cf. Figura 26).

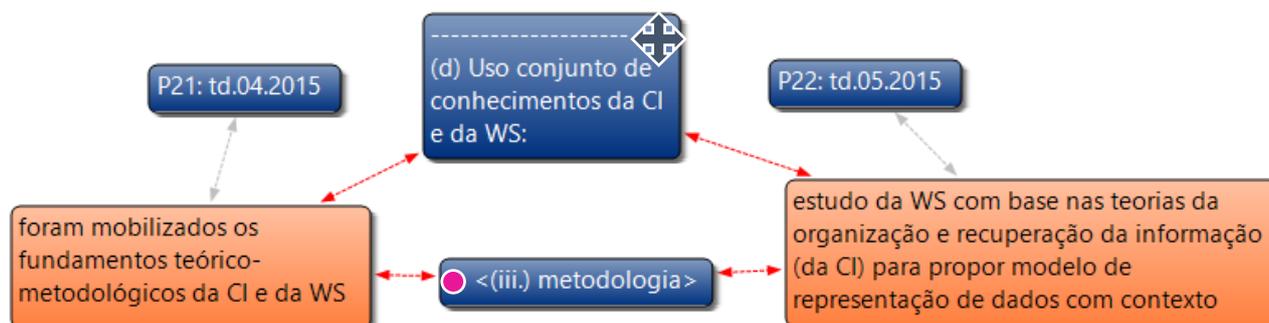


Figura 26. Diagrama de relacionamento entre a subclasse *uso conjunto de conhecimentos da CI e da WS* (d) e respectivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respectivos documentos de origem, identificados pela referência da Tabela 16.

O elemento do resumo onde estas unidades de registo se inscrevem (metodologia) e a complexidade da ação pretendida (a utilização conjunta dos conhecimentos de CI e WS) são coerentes com a tipologia dos documentos (TD), com o ano dos mesmos (2015) e com o próprio desenvolvimento de todo o contexto associado à WS, conforme foi apresentado ao longo do ponto 1.

A distribuição percentual das 21 (100%) unidades de registo extraídas dos resumos, pelas duas classes consideradas na análise efetuada, *pontos de confluência* e *estrutura interna do resumo*, pode ser observada na Figura 27.

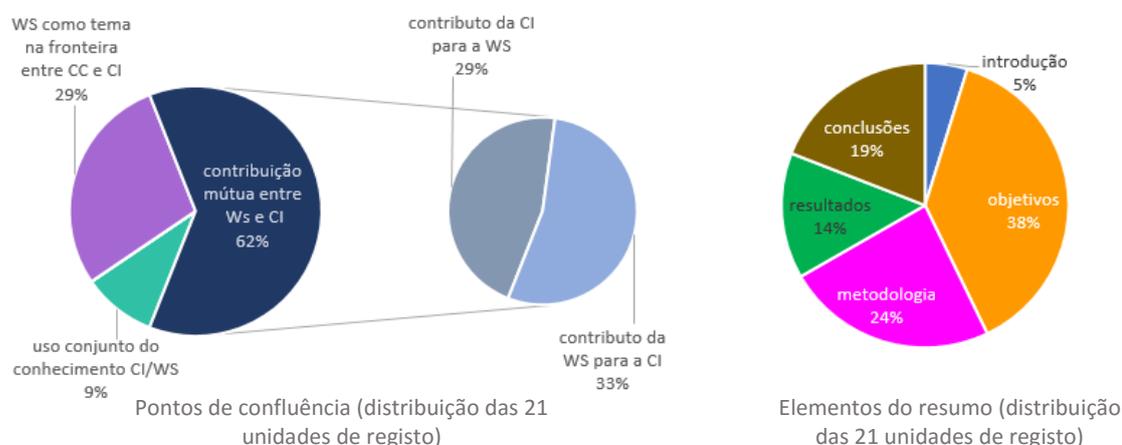


Figura 27. Distribuição percentual das 21 unidades de registo pelas duas classes de análise: *pontos de confluência* e *elementos do resumo*.

No gráfico *Pontos de confluência* da Figura 27 é evidente o “peso” do ponto *contribuição mútua* (62%), no entanto, repartindo esta percentagem pelas duas subclasses, *contributo da WS para a CI* e *contributo da CI para a WS*, consideradas neste mesmo ponto, a distribuição percentual configura-se mais equitativa¹⁴⁴.

No gráfico *Elementos do resumo* da mesma figura, destaca-se a percentagem desproporcional das unidades de registo incluídas no elemento *objetivos* (38%) relativamente às percentagens dos outros elementos. Verifica-se, neste contexto, um desajuste entre as “intenções” (*objetivos*) explicitadas e o alcance declarado das mesmas, representado nas percentagens menores obtidas nos elementos *resultados* (14%) e *conclusões* (19%). Apenas somando as percentagens obtidas por estes dois elementos (33%) se obtém um valor próximo do obtido pelo elemento *objetivos* (38%).

A diferença descrita no parágrafo anterior torna-se ainda mais significativa ao observar-se a Figura 28, onde se apresenta uma visão diacrónica da relação entre as duas classes em análise (*pontos de confluência* e *estrutura interna do resumo*). Nessa imagem é possível verificar que, em relação aos trabalhos analisados, o elemento *objetivos* não apresenta qualquer repetição nas 8 ocorrências aí representadas, enquanto que, no elemento *conclusões*, repete-se um trabalho (*dm.07.2009*). Esta situação concorre para uma maior distância entre os elementos do resumo em confronto, dado a diferença entre a quantidade de trabalhos distintos onde cada um se encontra (*objetivos* surge em 8 e *conclusões* em 3)¹⁴⁵.

¹⁴⁴ Em vez de 63%, 29% e 9% passa para 33%, 29%, 29% e 9%.

¹⁴⁵ Mesmo somando a quantidade de trabalhos diferentes onde os elementos *conclusões* e *resultados* se encontram o valor obtido, 5, é inferior aos 8 do elemento *objetivos*.

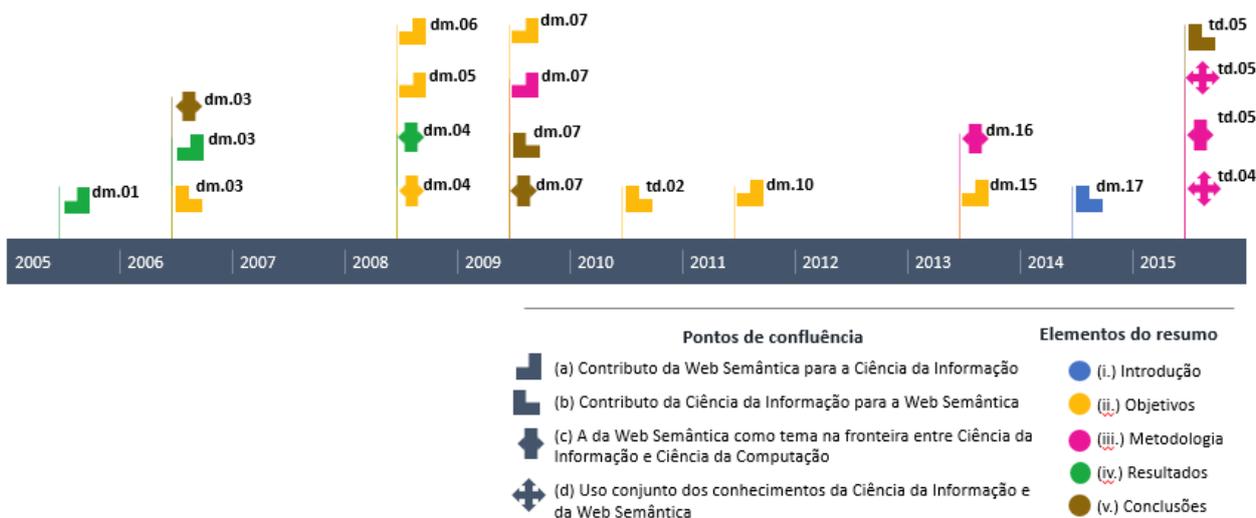


Figura 28. Distribuição das unidades de registo pela linha de tempo, de acordo com o ano dos respetivos documentos. Visão integrada das duas classes: *pontos de confluência* e *elementos do resumo*.

5.2.2. Análise dos objetivos dos trabalhos em Ciência da Informação

No que se refere à análise dos objetivos dos trabalhos constituintes da amostra após a aplicação os critérios de acordo com o exposto na Figura 19 determinou-se o grau de proximidade de cada trabalho relativamente à WS enquanto tema do mesmo (ver Apêndice 8). O grau que corresponde à maior proximidade (zero) foi atribuído a 23% dos trabalhos. No extremo oposto, o grau correspondente ao maior afastamento (menos quatro) obteve a percentagem de 32% dos trabalhos (ver Figura 29).

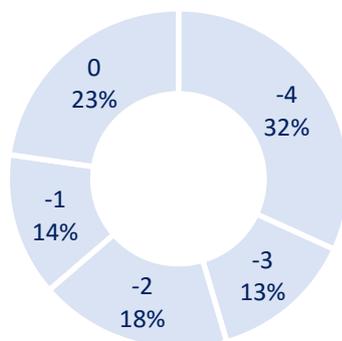


Figura 29. Distribuição da percentagem dos 22 (100%) trabalhos da amostra pelos graus de proximidade relativa à WS enquanto tema dos mesmos.

Verifica-se uma concentração de trabalhos nos extremos do espetro contrariando o esperado relativamente ao extremo negativo. De acordo com os critérios estabelecidos, esse extremo corresponde a uma ausência total, de forma explícita, do termo Web Semântica nos objetivos. Assim, tendo em conta a metodologia usada para a recuperação dos trabalhos, nomeadamente os campos usados na pesquisa (título e assunto, ver ponto 4.1), esse extremo deveria apresentar um valor residual e não o contrário – a maior percentagem.

A distribuição dos 22 trabalhos pelos graus de proximidade, ao longo do período considerado (2005 a 2015), permite verificar que os documentos classificados nesse extremo (maior afastamento) concentram-se na segunda metade do referido período, (de 2010 a 2015, cf. Figura 30).

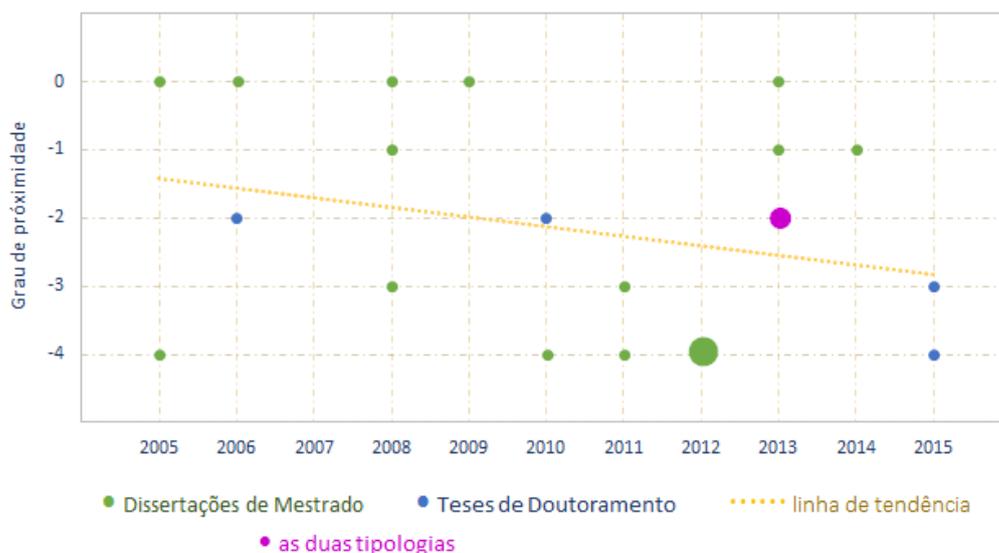


Figura 30. Distribuição dos 22 trabalhos pelos graus de proximidade ao longo do período considerado.

A tendência para o afastamento dos trabalhos em relação à WS enquanto tema dos mesmos é visível na respetiva linha representada no gráfico da Figura 30. Para essa tendência contribuiu em muito o “peso” das três DM de 2012 (ponto de maior dimensão), às quais foi atribuído o grau -4, assim como as duas TD de 2015, uma situada no nível -4 e a outra no -3. No mesmo gráfico é visível, também, a classificação -2 ou inferior a este valor que todas as TD obtiveram, assumindo este nível o valor médio da escala estabelecida (0 a -4).

5.2.3. Confrontação da análise dos resumos com a dos objetivos

As duas análises efetuadas (resumos e objetivos) abordam aspetos que supostamente estariam associados. Eles são: a explicitação de pontos de confluência entre a WS e a CI, nos resumos dos trabalhos e o grau de proximidade, entre as duas áreas, detetado nos objetivos dos mesmos trabalhos. Entre eles seria previsível que se observasse uma proporcionalidade direta entre o número de pontos de confluência e o grau de proximidade.

No intuito de verificar essa hipótese foi realizada a comparação entre os valores das duas análises (frequência de pontos de confluência e grau de proximidade) expressa no gráfico que se apresenta na Figura 31.

A linha *tendência esperada*, representada na Figura 31, evidencia a hipótese descrita. De acordo com esta seria previsível que os trabalhos sem referências a pontos de confluência entre a WS e a CI (nos resumos dos trabalhos) apresentariam (nos seus objetivos) um grau de proximidade, entre essas áreas, de valor -3. Relativamente ao extremo oposto, os resumos com três referências a pontos de confluência demonstrariam o grau máximo de proximidade (zero).

Apesar do valor 4 do eixo y se encontrar representado na escala medida, contudo, ele encontra-se ausente da escala esperada, facto que é reflexo da estimativa efetuada aquando da pré-análise dos resumos. No que respeita à ausência do valor -4 do eixo x (da escala esperada), observa-se a mesma situação descrita para o valor 4 do eixo y. Salienta-se que, neste último caso, a ausência do valor -4, relaciona-se com a situação referida no ponto 5.2.2 relativa aos critérios usados na recolha de dados, cuja aplicação deveria filtrar os trabalhos com aquele grau de afastamento do tema WS.

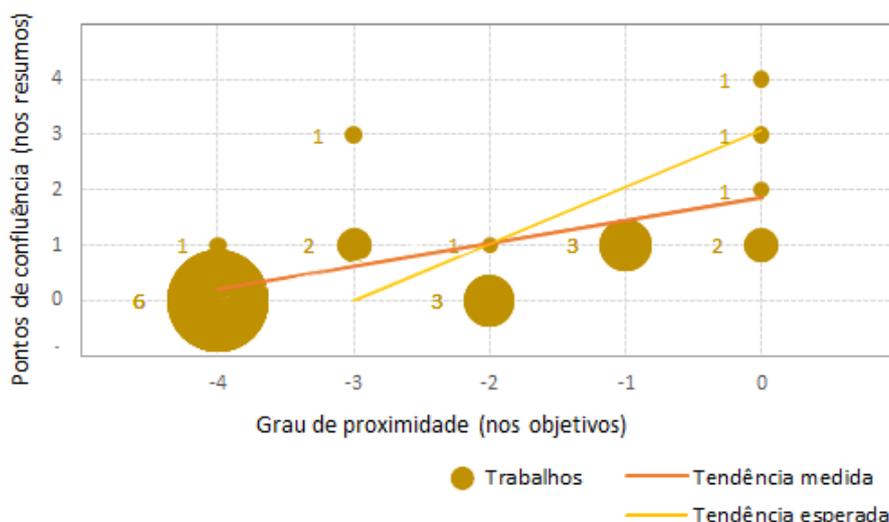


Figura 31. Comparação entre os valores: frequência de pontos de confluência nos resumos – grau de proximidade nos objetivos, determinados pelas respectivas análises aos 22 trabalhos.

Importa referir que a *tendência esperada* expressa a progressão linear do modelo de racionalidade, de acordo com a metodologia explanada por Quivy e Campenhoudt (1992, p. 264). Nesse contexto, a comparação entre a *tendência esperada* e a *tendência medida* revela a adequação e as limitações do modelo construído. Observando-se o gráfico da Figura 31 verifica-se um desajuste entre a *tendência esperada* e a medida que se analisa de seguida.

Observando a *tendência medida*, a relação entre as duas variáveis (eixos x e y) configura-se como moderada, (com um coeficiente de correlação de valor $0,576$)¹⁴⁶, indicando este a existência da associação prevista. Só dois valores se afastam de forma pronunciada da linha de *tendência medida*, representando estes os trabalhos *dm.07.2009* ($x=0$; $y=4$) e *td.05.2015* ($x=-3$; $y=3$). O valor apresentado pelo primeiro trabalho, apesar de ser uma exceção nos resultados (é o único caso que manifesta a conjugação dos valores máximos), não contradiz a hipótese, pelo contrário reforça-a¹⁴⁷. Todavia, o segundo documento revela resultados contraditórios, apresentando um nível de aproximação “fraco” (nível -3) e um número elevado de referências a pontos de confluência (3).

Neste último caso, os valores apresentados poderão indiciar um desajuste na hipótese formulada ou no modelo construído, tal como é apontado por Quivy e Campenhoudt:

Se os pontos não se distribuírem em torno desta diagonal, isso significa que a nossa hipótese não é confirmada, ou porque o nosso modelo é demasiado simples e deveria ser enriquecido com critérios de racionalidade e hipóteses suplementares, (1992, p. 265).

A necessidade de ajustes ao modelo é, também, indiciada pela presença de 2 documentos com um grau de proximidade máximo e um só ponto de confluência ($x=0$; $y=1$): *dm.01.2005* e *dm.15.2013*.

Por outro lado, não contrariando a hipótese, mas indiciando um desajuste entre a intenção que orientou os critérios usados na recolha dos documentos e a recuperação dos mesmos,

¹⁴⁶ Atendendo ao valor apurado se encontrar ligeiramente mais próximo de 1 que zero a menção *moderado* afigura-se como adequada.

¹⁴⁷ Relativamente à hipótese formulada, a conjugação dos valores 0 (x) e 4 (y) comprova a relação entre as variáveis diferindo somente no aumento de 1 valor relativamente à estimativa para os pontos de confluência.

apontam-se os 6 documentos (quantidade que se afigura excessiva) com os valores mínimos ($x=-4$; $y=0$): *dm.02.2005*, *dm.08.2010*, *dm.09.2011*, *dm.11.2012*, *dm.12.2012* e *dm.13.2012*.

Uma análise aos fatores que determinaram a recuperação dos referidos documentos aponta para a necessidade de exclusão de 2 dos 6 (os trabalhos: *dm.02.2005* e *dm.09.2011*); devido ao facto, por diferentes razões, estes não serem representativos do tipo de trabalho pretendido para o presente estudo. Em relação aos 4 restantes, se os mesmos fossem sujeitos a uma reavaliação, após os ajustes necessários ao modelo de análise, poderiam, eventualmente, apresentar uma nova distribuição ao longo da linha de tendência.

Quanto aos 2 casos referidos, o primeiro trabalho, *dm.02.2005*, foi considerado para a amostra por reunir as condições estabelecidas: ser proveniente de um curso de CI e o seu registo no OASISbr apresentar o termo Web Semântica nos campos título e/ou no assunto. No entanto, apenas nesta fase, analisando as possíveis implicações para os valores apresentados, verificou-se que os termos *web* e *semântica*, apesar de aparecerem nos resultados da pesquisa efetuada no OASISbr¹⁴⁸, não se encontram conjugados em um termo composto (*web semântica*) nos campos do respetivo registo. Face a esta situação, infere-se que a ausência desse termo composto no título e nas palavras-chave desse trabalho, justifica a abordagem periférica ao tema da WS.

O segundo trabalho, *dm.09.2011*, referido como não representativo, apresenta uma situação diferente. O termo composto web semântica surge no campo assunto do OASISbr, assim como do respetivo repositório de origem, contudo o mesmo não se encontra presente nas palavras-chave do documento em questão. Esta situação contraria o processo esperado para a recuperação de trabalhos com o termo web semântica nas palavras-chave definidas pelos respetivos autores (ver Figura 32).

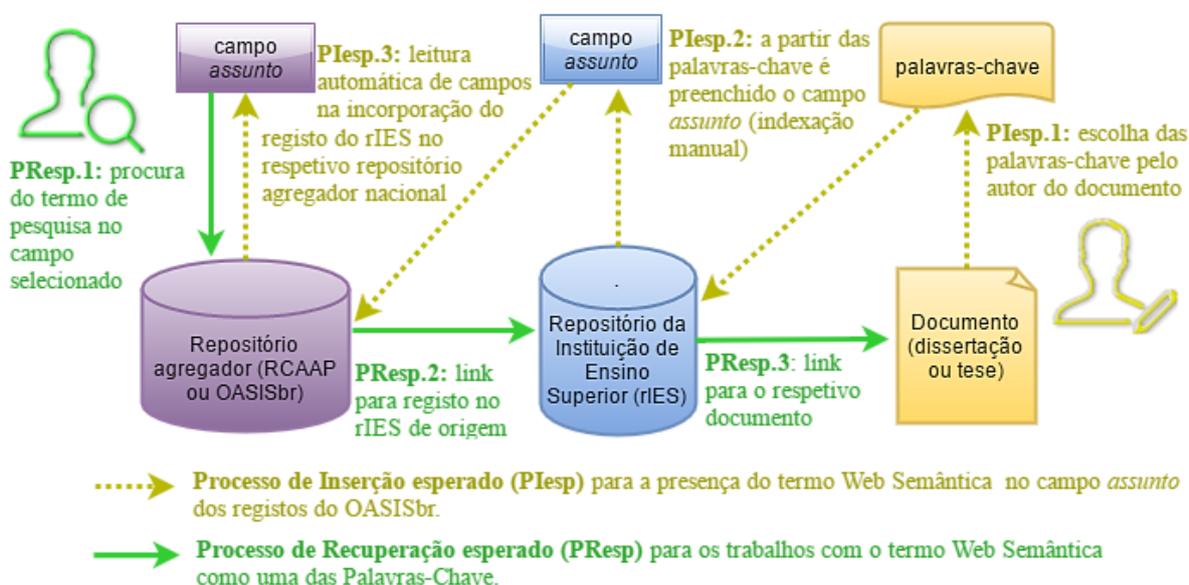


Figura 32. Representação dos processos esperados para: a inserção do termo Web Semântica no campo *assunto* dos registos das DM e TD, nos repositórios RCAAP e OASISbr, e para a posterior recuperação dos respetivos trabalhos.

¹⁴⁸ O registo deste documento, *dm.02.2005*, assim como o de *dm.09.2011*, foram dois dos 40 não recuperados no RCAAP.

O caso apresentado, da não correspondência entre as palavras-chave do documento e os termos presentes no campo assunto, não foi único, encontrou-se situações onde o termo web semântica aparecia como palavra-chave do documento, mas não estava registado no campo assunto dos repositórios OASISbr e/ou no de origem¹⁴⁹.

5.3. O perfil académico dos orientadores e coorientadores dos trabalhos considerados neste estudo

No que diz respeito ao perfil académico dos orientadores e coorientadores, foi necessário, no caso de um dos dois coorientadores, recolher dados mais pormenorizados em outra fonte¹⁵⁰ que não a plataforma *Lattes*, dado o facto da informação apresentada nesta ser insuficiente¹⁵¹.

Dos 22 trabalhos apenas dois foram realizados com coorientação, um de mestrado e outro de doutoramento. Refere-se, ainda, que, deste conjunto de 22 trabalhos, 4 são orientados por uma mesma pessoa e outros 2 por uma outra. Assim sendo, o número de orientadores é de 18 que, acrescido dos 2 coorientadores, resulta num total de 20 docentes, (11 do sexo feminino e 9 de sexo masculino, ver Figura 33).

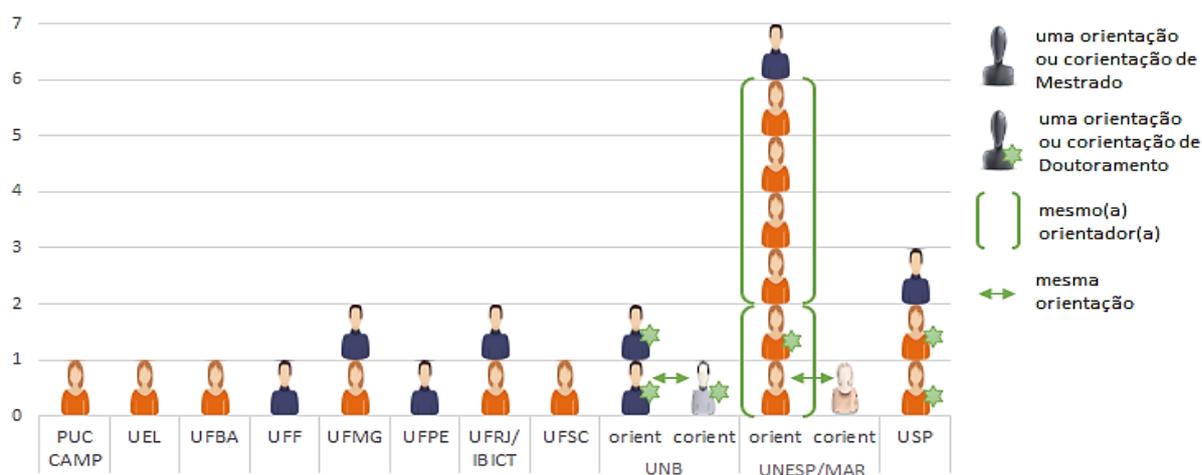


Figura 33. Distribuição dos orientadores e coorientadores pelas universidades associadas às dissertações de mestrado e teses de doutoramento que constituem a amostra.

Para a categorização das formações dos 20 docentes foi utilizada a mesma classificação por áreas apresentada no subponto 5.1: grupo A- CI¹⁵²; grupo B- Informática/Computação; grupo C- Eletrónica/Automação; grupo D- outros cursos. Agrupando os valores encontrados, em

¹⁴⁹ Como exemplos de documentos com o termo web semântica nas palavras-chave e o mesmo termo ausente nos campos *assunto* dos repositórios OASISbr (i) e/ou no de origem (ii) pode-se apontar: ausente em (i) – *td.01.2006* e *td.03.2013*; ausente em (ii) – *dm.11.2012*; ausente nos dois – *dm.05.2008* e *dm.14.2013*. No caso deste último documento (*dm.14.2013*) em cujo título não consta o termo web semântica, o mesmo foi recuperado pela presença, no campo assunto, do termo em inglês (*semantic web*).

¹⁵⁰ Os dados relativos ao coorientador em causa, Frank van Harmelen, foi recolhido do seu perfil no *LinkedIn*.

¹⁵¹ Salienta-se que apenas esse coorientador se encontrava filiado a uma outra instituição de ensino superior (IES) que não a associada à respetiva DM ou TD sendo, neste caso, de uma IES não brasileira.

¹⁵² Nas formações incluídas neste grupo, além das designadas pelo termo CI, foram encontradas as seguintes: Arquivologia (1 pD e Esp.); Biblioteconomia (3 B, 3 G e 1 M e L); Biblioteconomia e Documentação (3 G e 1 M e G); Documentação Científica (3 Esp.); Science in Library Service (1 M). A contabilização foi efetuada de acordo com os 3 estágios definidos (pD/D, M/Esp. e L/B) estando a Graduação (G) incluída no grupo L/B.

cada um dos grupos mencionados, foram identificados três estágios de formação: i. pós-doutoramento e doutoramento (pD/D); ii. mestrado e especializações (M/Esp); iii. licenciatura e bacharelato (L/B), como é mostrado na Figura 34. Contabilizando os valores absolutos das formações de cada grupo, mesmo se representadas pelo mesmo docente, obteve-se a contagem apresentada na mesma figura.

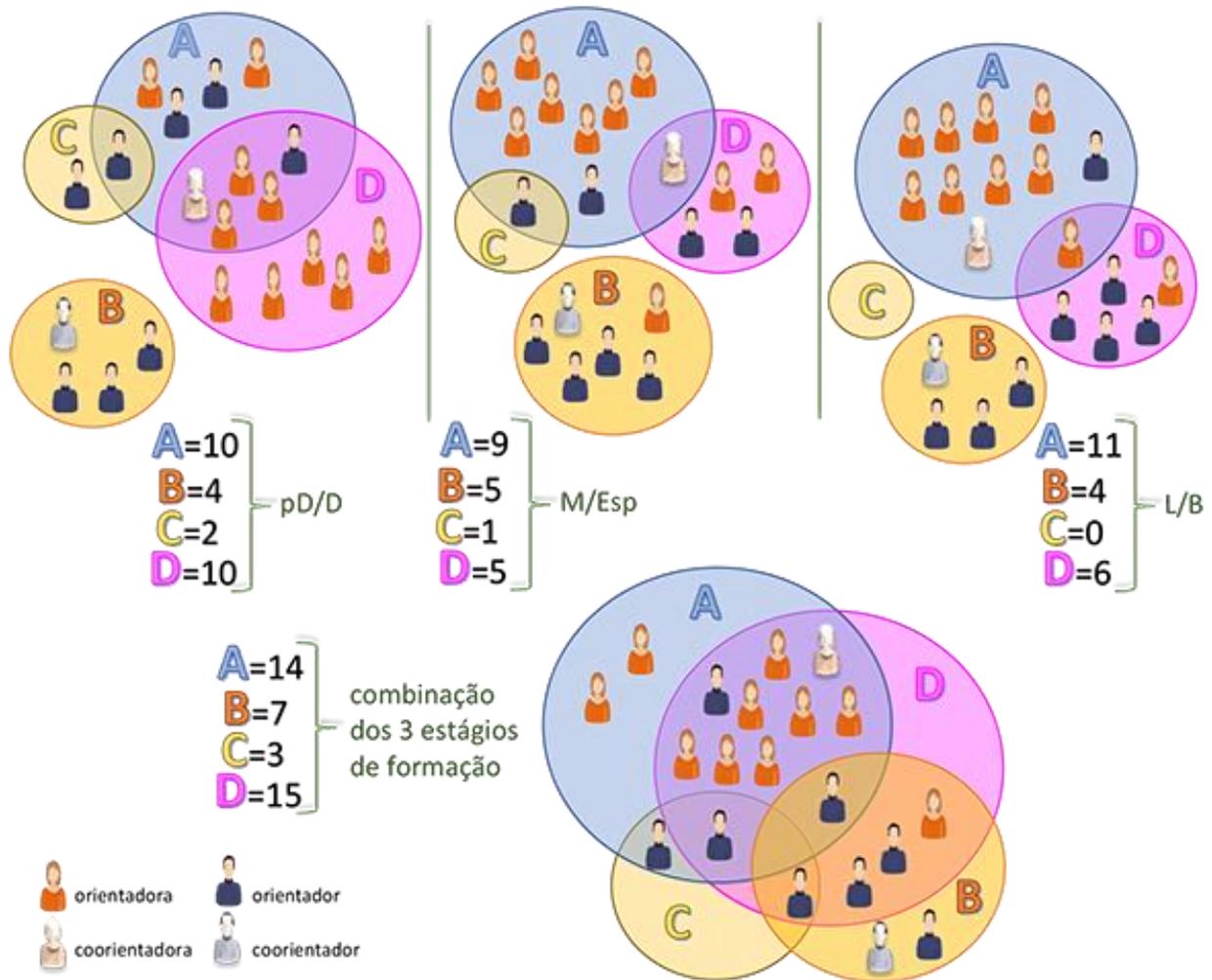


Figura 34. Distribuição da formação acadêmica dos 20 docentes, responsáveis pelas orientações das DM/TD da amostra, de acordo com os 4 grupos de áreas: A- CI; B- Informática/Computação; C- Eletrônica/Automação; D- outros cursos. Apresentação por: pós-doutoramento e doutoramento (pD/D); mestrado e especializações (M/Esp); licenciatura e bacharelato (L/B) e soma dos 3 estágios.

Da categorização e contabilização efetuada destacam-se três pontos:

- i. O número mais reduzido que o esperado de docentes com formação na área da Ciência da Computação/Informática (grupo B)¹⁵³;
- ii. Apenas um docente apresenta formação conjunta nos grupos A e B, ou seja, em Ciência da Informação e Ciência da Computação;

¹⁵³ Tendo em conta o tema dos trabalhos orientados que, como já referido anteriormente, está naturalmente relacionado com a Ciência da Computação.

- iii. Um aumento significativo nas formações de pós-Doutoramento e Doutorado (pD/D) do grupo *D*, relativamente à quantidade desse grupo nos outros estágios de formação.

Relativamente a este último ponto (iii.) e atendendo a que o grupo *D* representa formações de um espectro mais alargado de áreas, esse aumento revela-se consistente com a abrangência disciplinar dos profissionais relacionados com a área de CI (ver ponto 2.3). Esta abrangência disciplinar é reforçada pela constatação de um total de 15 docentes com formação em 12 áreas diferentes incluídas no grupo *D*, como apresenta a Figura 35 (etiquetas em tom rosa).

A figura referida representa o mapa das relações entre a formação académica dos orientadores e coorientadores com os trabalhos por eles orientados. Nesta figura é, ainda, possível verificar os pontos de convergência entre a CI e a CC tendo como elo de ligação a WS enquanto tema dos referidos trabalhos, (cf. Figura 35).

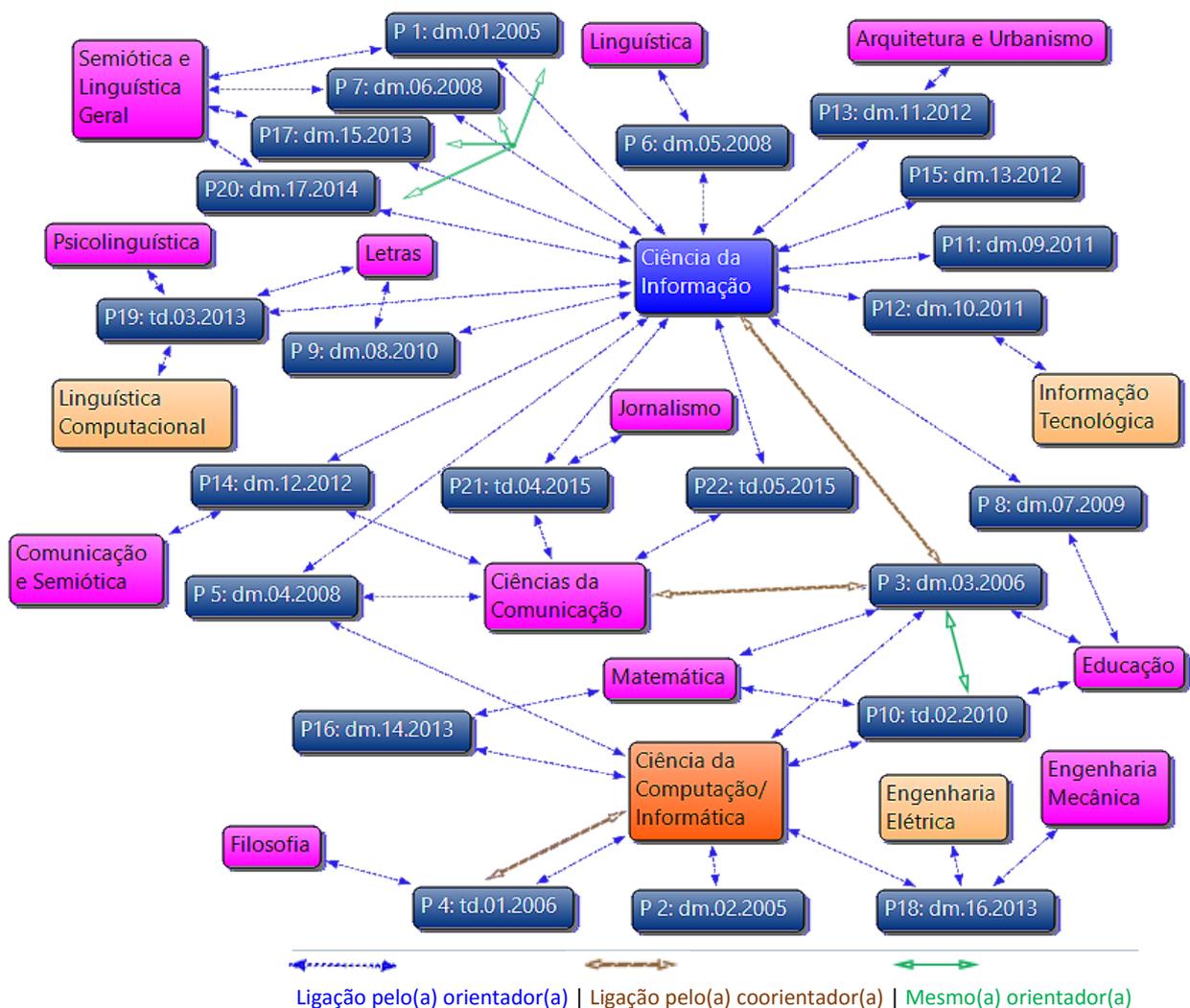


Figura 35. Mapa das relações entre a formação académica dos orientadores e coorientadores e os respetivos trabalhos que constituem a amostra orientados.

Ainda na mesma figura é possível observar que, das 12 áreas do grupo *D*, as Ciências da Comunicação, a Educação e a Matemática revelam ser a “ponte” entre a CI e a CC, ao

contrário das 3 áreas incluídas no grupo C (etiquetas em tom amarelo) que, teoricamente, têm uma maior relação com a CC.

Observa-se, também, que, além do docente que apresenta formação nas áreas da CI e da CC (orientador do trabalho *dm.04.2008*), existe uma outra ligação direta efetuada pela combinação das formações do orientador e coorientador (no trabalho *dm.03.2006*).

CONCLUSÕES

A primeira conclusão a extrair deste trabalho prende-se com o cumprimento dos objetivos do mesmo; quer no que se refere ao objetivo geral, quer no que se refere aos específicos. No que respeita ao geral: verificar e refletir sobre a presença da Web Semântica no *output* dos cursos de mestrado e doutoramento em Portugal e no Brasil em Ciência da Informação, entre o período de 2005 e 2015, salienta-se que este foi alcançado dentro dos limites impostos à mesma investigação de acordo com a metodologia.

Relativamente aos objetivos específicos, no que concerne ao *primeiro*: a identificação dos cursos de mestrado e doutoramento em Portugal e no Brasil em Ciência da Informação certificados, respetivamente, pela A3ES e pela CAPES, em funcionamento no ano letivo 2014-2015, o mesmo materializou-se na Tabela 11, para os cursos portugueses e na Tabela 13 para os brasileiros, com o devido enquadramento efetuado, respetivamente, nos pontos 2.4.1 e 2.4.2.

Quanto ao *segundo* objetivo: verificar o número de dissertações de mestrado (DM) e teses de doutoramento (TD) em Ciência da Informação associados à Web Semântica, produzidas em Portugal e no Brasil no período de 2005 e 2015, recuperadas nos repositórios RCAAP e OASISbr, este foi também cumprido uma vez que foram encontrados 22 trabalhos, a saber: 17 DM e 5 TD, todos provenientes do Brasil (ponto 4.3.2).

No que concerne ao *terceiro* objetivo: identificar os pontos de confluência da Ciência da Informação com a Web Semântica, presentes nos trabalhos recuperados, o seu comprimento permitiu determinar três pontos de confluência entre a CI e a WS, a saber: i. a presença de uma contribuição mútua entre as duas áreas (CI e WS); ii. a WS enquanto tema interdisciplinar (CI-CC), constituindo-se a própria WS como ponto de convergência; e iii. a complementaridade de conhecimentos entre a CI e a WS (ponto 5.2.1).

Finalmente, no que diz respeito ao *quarto* objetivo: caraterizar o perfil académico dos orientadores e coorientadores dos trabalhos recuperados identificados no estudo como sendo de Ciência da Informação, o mesmo foi cumprido através da categorização da formação dos referidos docentes, apresentada na Figura 34, assim como nas conclusões que serão apresentadas neste ponto (Conclusões) aquando dos resultados do estudo de caso.

No que respeita à contextualização teórica dos objetos de estudo foram extraídas conclusões relativas à Web Semântica e também à Ciência da Informação.

Conceito de Web Semântica

A análise efetuada ao conceito de Web Semântica permitiu, não só, confirmar a hipótese inicialmente formulada (ponto 1.3), como também, avançar com uma proposta de definição do mesmo, no sentido da sua clarificação e, conseqüentemente, a distinção de outros conceitos associados e, comumente, utilizados como sinónimos, tais como: *Web of Data*, *Linked Data* (ponto 1.3.1) e *Web of Linked Data* (ponto 1.3.2).

A hipótese formulada: *o conceito da Web Semântica sofreu uma evolução, ao longo do tempo, no sentido de tornar a sua operacionalização mais concreta e imediata*, foi atestada pela análise do conceito Web Semântica em três dimensões: *caraterísticas descritivas, base estrutural e funções operativas*.

Na primeira dimensão concluiu-se haver uma passagem do conceito *semântica*, de cariz mais abstrato associado ao “entendimento” da informação (*semantic web*), para o processamento sintático, utilizando dados concretos ligados entre si (*web of linked data*), próprio dos meios computacionais (ponto 1.2.1).

A análise da segunda dimensão, *base estrutural*, confirmou o direcionamento referido, da *semantic web* para a *web of linked data*, através da combinação de três fatores: i. os procedimentos relacionados exclusivamente com a estruturação de dados para processamento automático apenas se observaram entre 1999 e 2001; ii. a partir de 2006 registou-se uma concentração dos procedimentos focados na criação de interoperabilidade entre sistemas e ligação de dados; e iii. uma relação direta entre as especificações mais recentes e o termo *linked data* (ponto 1.2.2).

Finalmente, na terceira dimensão analisada relativa às *funções operativas*, concluiu-se que é necessário considerar a operacionalização da WS no aspeto formal (como se operacionaliza) e no temporal (quando se operacionaliza). Conclui-se, ainda neste ponto, que estes dois aspetos estão estreitamente relacionados uma vez que a questão está associada com os destinatários das funcionalidades oferecidas pela WS. Se as mesmas não forem restringidas aos utilizadores finais, pode considerar-se funções concretas da WS (aspeto formal), já implementadas no terreno (aspeto temporal), os seguintes procedimentos: i. publicar e ligar dados utilizando o formato RDF; ii. usar o RDFa para tornar o conteúdo de documentos legível pelas máquinas; e iii. usar o padrão GRDDL para expor o conteúdo de uma base de dados relacional (RDB) em formato RDF (ponto 1.2.3). Estas funções relacionam-se diretamente com o grande objetivo da WS que, de acordo com a análise realizada e de uma forma simplificada, pode ser descrito como a utilização da *Web* como uma base de dados global.

O objetivo da WS referido faz parte de um todo que, além dessa grande finalidade de tornar a *Web* numa base de dados global, incluindo aí as aplicações/interfaces necessárias para seu uso, engloba as especificações desenvolvidas e/ou associadas à sua operacionalização e os procedimentos necessários para a ligação de dados em formato aberto na *web*. Conclui-se assim, que é todo este constructo abrangente, o que define a Web Semântica (ponto 1.3.1).

Esta noção distingue-se pela sua abrangência dos restantes três conceitos, *Web of Data*, *Web of Linked* e *Linked Data*. A abrangência destes conceitos apresenta uma ordem descendente, a saber: *Web of Data*, é a finalidade da noção apresentada para a Web Semântica; *Web of Linked Data*, é uma parte específica dessa finalidade e *Linked Data*, é um aspeto específico dessa parte.

Iniciando por este último, o conceito de *Linked Data*, refere-se ao conjunto de procedimentos e regras para a publicação de dados de forma a possibilitar a ligação dos mesmos a outros dados de diferentes fontes. Estes procedimentos podem ser aplicados tanto dentro como fora do ambiente *Web*; assim como podem ser usados com dados em formato aberto (*Linked Open Data*) ou em formato proprietário (ponto 1.3.1).

Conclui-se, ainda, que é a combinação destes procedimentos, aplicados em ambiente *Web* a dados em formato aberto, com as especificações da WS associadas às camadas *representação* e *consulta* (de dados) e sistemas leves de organização do conhecimento (“ontologias leves”), que produz a chamada *Web of Linked Data* (ponto 1.3.2).

Por fim, outra ideia extraída sobre estas relações conceituais refere-se à diferença entre a *Web of Linked Data* e a *Web of Data*, dado as limitações da primeira, quer na abrangência de ligações, quer nas relações semânticas das mesmas, dado a *Web of Data* ser definida, neste trabalho, como a *Web*, no seu todo, resultante do desenvolvimento da Web Semântica, na qual (*Web of Data*) é possível realizar consultas similares às efetuadas em bases de dados (ponto 1.3.1). Para que esta “visão” se concretize é necessário a união de dois “mundos” ainda tenuemente conectados: a possibilidade, virtualmente ilimitada, de ligações entre dados – a parte *web*, com a potencialidade da inferência automatizada dos sistemas “inteligentes” – a componente *semântica*.

Conceito de Ciência da Informação

O recurso aos modelos apresentados por Chain Zins¹⁵⁴ para a definição do conceito de Ciência de Informação (ponto 2), ao mesmo tempo que representam uma visão atual e abrangente desta área, mostraram-se essenciais para sistematizar e estruturar um conceito com o elevado grau de complexidade com que se apresenta o de CI.

Com a aplicação destes modelos chegou-se a um conjunto de conclusões que se passam a apresentar. Assim, pode definir-se a CI como sendo o estudo dos aspetos de mediação dos fenómenos: dado, informação e conhecimento enquanto artefactos humanos, e a sua implementação no domínio cultural/social (ponto 2.2.3). Desta definição observam-se as três dimensões analisadas por Zins, designadas “questões chaves” por este autor. Através da sua análise é possível decompor o conceito de CI tornando-o mais inteligível, a saber: i. os fenómenos explorados – dado, informação e conhecimento, enquanto artefactos humanos; ii. o âmbito da exploração – estudo dos aspetos de mediação entre estes fenómenos e os utilizadores e iii. o domínio do campo científico – domínio cultural/social, entendido como a forma da sociedade encarar a realidade (ponto 2.2.2).

A partir da definição descrita no parágrafo anterior, que é apresentada como a visão *mainstream* dos profissionais de CI, reunindo um razoável consenso à sua volta, foi possível concluir que permanecem, atualmente, concepções bastante díspares do que se entende por CI pelos próprios especialistas da área. Concepções como, por exemplo: uma “disciplina matemática” (Luciana Duranti); uma parte, juntamente com a Biblioteconomia, da Ciência da Comunicação mais geral (Maria Biagetti); um outro termo para descrever o estudo da informação e seu uso na Informática (William Hersh) ou uma “autopromoção” da Biblioteconomia (H.M. Gladney) (Zins, 2007b, pp. 336–338).

Relativamente às designações desta área, conclui-se ainda, existir um grupo razoável de investigadores, entre os quais Zins, que considera a mudança da Ciência da Informação, tanto no foco da mesma como na própria designação, para “Ciência do Conhecimento”. Este entendimento resulta da importância que este conceito, conhecimento, apresenta, assumindo o mesmo uma posição dominante relativamente aos outros dois fenómenos explorados, dado e informação. Não é de todo indiferente a esta posição, o facto de a definição “extraída” do estudo deste autor mencionar os fenómenos explorados (dado, informação e conhecimento) enquanto artefactos humanos. Esta particularidade não coincide com o modelo mais comum (modelo 1, ver Figura 10) no que concerne ao posicionamento dos três fenómenos no

¹⁵⁴ Apesar dos riscos associados ao uso de modelos, devido à necessária simplificação conceitual inerente à sua construção, como lembra Le Coadic (1996, p. 73), o grau de complexidade do conceito de CI justifica a sua utilização.

domínio universal ou no subjetivo (ponto 2.2.1), uma vez que este modelo coloca o conhecimento no domínio subjetivo, posição contrária ao conceito de artefacto humano¹⁵⁵.

Paralelamente às questões levantadas pela definição da CI, foi possível concluir que a diversidade de perspectivas perante esta área não é alheia às suas relações disciplinares que, desde a sua gênese, se encontram em contínua evolução (ponto 2.3.3). Um desenvolvimento onde a relação entre a CI e as três disciplinas: Biblioteconomia, Arquivística e Documentação mostrou matizes bastante díspares entre elas, tal como se observa nas ideias que se seguem.

A Biblioteconomia mostrou uma presença constante a par da Informática/Ciência da Computação nas relações com a CI, embora quase antagónicas, dado o facto de a primeira ser associada às ciências sociais e a segunda às ciências “exatas”. Relativamente à “natureza científica” da CI, refere Pinheiro que: “de acordo com os pesquisadores da área, oscila entre ciência social, tecnologia, aplicabilidade ou prática,” (1997, p. 7). Esta “dualidade” da CI contribui para os diferentes conceitos apontados para a mesma, consoante a ênfase que se dá a uma ou a outra componente¹⁵⁶ (ponto 2.3.2). A CI, em alguns casos é mesmo identificada com uma dessas duas disciplinas, dependendo do seu contexto. Tal situação pode ser verificada nos exemplos mencionados das diferentes concepções da CI, dos participantes no estudo de Zins referido a cima.

Por sua vez, a Documentação manifestou uma presença diretamente relacionada com a gênese da CI. Salvo casos pontuais, é consensual a linha de continuidade da Documentação para a CI (pontos 2.3.1 e 2.3.3). Menos consensual mostrou-se a relação entre a Documentação e a Biblioteconomia adquirindo, esta, três perspectivas diferentes de acordo com a visão de três autores abordados: Otlet, Shera e López Yepes. Para o primeiro a Documentação engloba a Biblioteconomia (ponto 2.3.1), para o segundo, tem um entendimento exatamente oposto e, para o terceiro, há uma relação de paralelismo entre ambas (ponto 2.3.2).

Por fim, a Arquivística, apresentou uma relação mais afastada relativamente à CI. Embora se infira, pelo estudo efetuado, ser esse afastamento mais em termos de prática profissional do que epistemológico. Além disso, sendo das três disciplinas BAD a que menos foi referida nas obras consultadas, apresentou uma associação direta com a Museologia (ponto 2.3.3). Inferindo-se de todo este contexto, uma posição algo subalterna nomeadamente pela sua ligação à História enquanto disciplina auxiliar.

Estes relacionamentos, da CI com as três disciplinas BAD e destas três últimas, entre si, refletem-se nas diferentes designações dos cursos de ensino superior encontrados.

¹⁵⁵ Este posicionamento de Zins relaciona-o com a teoria dos três mundos de Popper, no entanto, o próprio investigador distancia-se dessa concepção: “I am not ‘Popperian.’ In fact, I am a phenomenologist. Generally, I follow Edmund Husserl’s phenomenology.” (Zins, 2007c, p. 493). Ao contrário de Brookes que, explicitamente, coloca o fenómeno estudado pela CI na interação entre o segundo e o terceiro mundo de Popper: “Furthermore, in adopting the interaction between Worlds 2 and 3 as our field of study we are laying claim to a territory which no other discipline has already claimed.” (Brookes, 1980, p. 128). O conceito de conhecimento enquanto fenómeno explorado pela CI, na ótica de Zins, distancia-se, igualmente, da visão de “conhecimento formalizado” de Belkin e Robertson, tendo em conta que estes consideram aquele como o limite do espectro de interesse da CI (Belkin & Robertson, 1976, p. 201).

¹⁵⁶ O conflito *social vs. tecnológico* afigura-se como um falso dualismo uma vez que conceitualmente o segundo oponente esvai-se sem o primeiro, dado a característica omnipresente nas definições de *tecnologia* da sua utilidade para as atividades humanas, ao ponto de se poder inferir que os *conhecimentos tecnológicos* só o são se forem *socialmente* considerados úteis para a realização dessas atividades humanas.

Nomenclaturas que, por sua vez, espelham o foco dos mesmos cursos embora, por vezes, esse reflexo se apresente algo difuso ou distorcido contribuindo para uma classificação dos mesmos em diferentes áreas do saber, apesar da sua similaridade, ou o inverso (ponto 2.4).

Outra ideia a reter do estudo, é o facto de em Portugal, o panorama encontrado se apresentar como estando em mudança de atitude¹⁵⁷ no que concerne aos cursos da área de CI. A posição inicial pode ser descrita como uma associação entre a visão de Otlet e a posição de López Yepes, i.e., uma visão da Documentação enquanto área que abarca as outras áreas relacionadas (ponto 2.3.1) “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 Ciência da Informação enquanto entidade que incorpora aqueles saberes (das disciplinas BAD) e outros, num todo maior que a soma das suas partes.

Um indicador do entendimento que se faz no parágrafo anterior é evidente na comparação entre a quantidade de cursos de mestrado e doutoramento com designações como: Ciências Documentais, Ciências da Documentação e da Informação ou Ciências da Informação e da Documentação, encontrando-se 2 ainda ativos (no ano letivo de 2014-15) e 7 foram descontinuados (ponto 2.4.1). Acresce que, destes 7 cursos, apenas 2 foram descontinuados por não terem obtido a acreditação necessária da A3ES, sendo que os restantes 5 o foram por iniciativa das próprias instituições de ensino superior no qual estavam integrados. Como indicadores suplementares pode apontar-se: a renomeação dos cursos de licenciatura da área para a designação Ciência da Informação, sem qualquer acrescento (ponto 2.3.2). Considera-se, ainda, como indicador a ausência de cursos (dos três graus de ensino superior) com designações relacionadas com a Arquivística e a Biblioteconomia¹⁵⁸.

No Brasil, o histórico da área CI, no contexto da TAC (Tabela de Áreas do Conhecimento) é significativamente elucidativo do complexo relacionamento entre as disciplinas BAD e a CI. Embora com uma distinção, no tipo de relacionamento, entre a Documentação e a CI, e esta com as outras duas, a Biblioteconomia e a Arquivística, que no Brasil assume o nome de Arquivologia (ponto 2.4.2).

Quanto à disciplina Documentação, no Brasil, da investigação efetuada relativamente aos cursos de ensino superior, conclui-se que existiu um percurso fluído desta para a CI. Como indicadores, apresenta-se a reduzida presença do termo documentação nas designações dos cursos de graduação¹⁵⁹ e a sua ausência total nos cursos de pós-graduação *stricto sensu* (ver Tabela 13).

¹⁵⁷ Embora o termo *atitude* não seja, porventura, o mais adequado é aqui usado como significando uma *demonstração de uma intenção* ou *modo de proceder*. Por seu lado, o termo *paradigma* traz consigo um “peso” que se entende ser demasiado intenso para a ideia que se pretende transmitir.

¹⁵⁸ Os 7 cursos, dos três graus de ensino superior, com o termo *biblioteca* na sua designação, apenas dois estão ativos, *Promoção da Leitura e Bibliotecas Escolares* incluído na área Língua e Literatura Materna e o *Gestão da Informação e Bibliotecas Escolares* da Universidade Aberta (UAb), referido na Tabela 11. Os 5 desativados, apesar de incluídos na área BAD, seguem o padrão do da UAb na sua ligação com as Ciências da Educação (ponto 2.4.1). Verifica-se, assim, em Portugal uma maior ligação da Biblioteconomia à educação que à CI. (A listagem dos cursos referidos foi consultada no sistema em linha da DGEEC a 26-8-2016 no URL: <http://www.dgeec.mec.pt/np4/171/>).

¹⁵⁹ O termo *Documentação* surge em 4 cursos apenas e sempre associada ao termo *Biblioteconomia*. Número bastante reduzido comparando com os 41 cursos só com a designação de Biblioteconomia e os 18 ostentando apenas o termo Arquivologia, listagens retiradas do sistema em linha do Ministério de Educação brasileiro, URL: <http://emec.mec.gov.br/>, acessado a 26-8-2016.

No que respeita ao relacionamento da CI com a Biblioteconomia (Bib) a Arquivística (Arq), a análise das versões da TAC mostrou-se ilustrativo, do que se infere ser um conflito entre uma procura de identidade da área enquanto um todo e uma tentativa de afirmação por parte daquelas áreas disciplinares específicas. De acordo com as cinco versões da TAC (ver Tabela 12) a CI apresenta uma metamorfose representativa de cinco diferentes visões da mesma como, de seguida, se apresenta. Em 1976 a CI é uma *subárea* da Comunicação que tem como especialidades a Bib e a Arq. Seis anos mais tarde, em 1982, deixa a dependência da Comunicação constituindo, conjuntamente com a Bib e a Arq, uma *área* por direito próprio. Após dois anos, em 1984, a CI continua a ser uma *área*, mas a Bib e a Arq tornam-se *subáreas* da mesma. Por fim, na proposta de 2005, a Bib e a Arq passam novamente a *áreas* a par da CI, mas desta vez cada uma das três constitui-se como uma *área* independente. Como foi possível verificar o relacionamento atravessou todo um espectro de relações possíveis tendo como denominador comum o facto das duas disciplinas. a Biblioteconomia e a Arquivística, manterem o mesmo tipo de relação entre si.

Estudo de caso

No que concerne às conclusões relativas aos resultados do estudo de caso, efetuado no presente trabalho, estas serão apresentadas, relacionando-as, sempre que possível, com as hipóteses formuladas.

Em relação à distribuição da amostra pelo período considerado (2005 a 2015), concluiu-se que o número de trabalhos relativamente à Web Semântico se apresentou irregular, deste modo, não foi possível estabelecer um padrão, seja no sentido crescente, como o esperado, seja em sentido inverso (ponto 4.3.2). Também contrariando o esperado, foi no ano de 2013 que se registaram mais trabalhos (3 DM e 1 TD), não se tendo verificado uma relação entre a falta de visibilidade do sucesso da Web Semântica (descrito no ponto 1.3.2) e o interesse manifestado em estudar este assunto.

No que respeita à distribuição da amostra pelos dois países, conclui-se existir um desfazamento maior que o esperado entre ambos, dado todos os trabalhos incluídos na mesma serem produzidos no Brasil (ponto 5.1). Infere-se, deste facto, a existência de uma relação entre os contextos de origem da CI, europeu, para Portugal, e americano, no caso do Brasil, e a ligação à Informática/Ciência da Computação (ver ponto 2.3.1), área dominante no que diz respeito ao tratamento do assunto em questão (a WS)¹⁶⁰. Intui-se, ainda, que o número de cursos desta área e respetivo número de alunos matriculados em pós-graduação possam ser considerados um fator para explicar esta discrepância.

Relativamente à questão central do estudo de caso, os pontos de confluência entre as duas áreas, a CI e a WS, além da identificação dos três pontos já apresentados (i. a presença de uma contribuição mútua entre as duas áreas; ii. a própria WS enquanto tema de estudo; e iii. a complementaridade de conhecimentos das duas áreas), foi possível determinar a frequência com que os mesmos surgem nos resumos analisados, concluindo-se assim sobre a relevância do primeiro ponto de confluência que se destaca, de forma significativa (62%), dos restantes (Figura 27).

¹⁶⁰ Inferência que poderá estar relacionada com a ligação encontrada entre as percentagens obtidas pelos cursos incluídos na área Eletrónica e Automação (grupo C) em Portugal (31%) e a soma das percentagens do grupo C e A no Brasil (32%), relativamente ao total de trabalhos recuperados de cada país (ver ponto 5.1).

Foi, igualmente, possível inferir uma coerência entre as datas dos dois trabalhos (2015) e a presença do ponto de confluência iii. (complementaridade de conhecimentos das duas áreas), no sentido do seu uso efetivo no próprio trabalho onde é referido. Assim como, da coerência entre a única ocorrência encontrada no elemento do resumo *introdução*, (neste caso do ponto i. a presença de uma contribuição mútua), e a data, 2014, do mesmo trabalho. Estas situações permitem concluir que a confluência entre as duas áreas se alterou. Se no início se procurava os pontos de confluência em termos de contribuição entre as duas áreas, a partir de 2014 essa contribuição recíproca torna-se uma premissa ao mesmo tempo que se consolida as potencialidades do uso conjunto dos conhecimentos das mesmas (ponto 5.2.1).

Conclui-se, ainda, a existência de uma relação entre a abordagem efetuada à WS e as datas de realização dos trabalhos assim como as tipologias associadas aos mesmos (ponto 5.2). Os trabalhos anteriores a 2010 apresentam uma abordagem mais generalista à WS relativamente aos posteriores a essa data. Assim como as teses de doutoramento apresentam uma especificidade maior relativamente às dissertações de mestrado, nomeadamente não apresentando, as primeiras, a WS como objeto de estudo (ponto 5.2.2).

Confirmou-se, também, existir uma proporcionalidade direta entre o número de pontos de confluência encontrado nos resumos e o grau de proximidade determinado nos objetivos (ponto 5.2.3). No entanto, as inferências que se poderão retirar da mesma carecem de um ajuste ao modelo de racionalidade usado, além de outras considerações relativas às limitações do presente estudo que serão tratadas no respetivo ponto.

Dos resultados obtidos infere-se uma situação não conclusiva relativa ao desajuste entre a percentagem de ocorrências de pontos de confluência, entre a WS e a CI, no elemento do resumo *objetivos* (38%) em oposição às mesmas percentagens nos elementos *resultados* (14%) e *conclusões* (19%)¹⁶¹. Estes resultados implicam uma diferença, que se afigura demasiado elevada, entre os objetivos e os resultados e as conclusões (ponto 5.2.1). Desta situação, infere-se a possibilidade de os resumos não espelharem convenientemente os respetivos trabalhos¹⁶², situação que lança pistas para uma futura investigação que será abordada no respetivo ponto das conclusões.

Em termos do perfil académico dos orientadores e coorientadores, conclui-se que a amostra não é representativa do carácter interdisciplinar da WS enquanto tema trabalhado pelas áreas da Ciência da Computação e da Ciência da Informação, atendendo à conjugação de dois fatores: somente um docente apresenta formação conjunta nas duas áreas e só se registaram duas coorientações, sendo que apenas uma delas patenteia a contribuição disciplinar esperada. Ainda em relação à contribuição disciplinar, conclui-se que esta não apresenta o padrão espectável, que seria o contributo adicional de um docente da área da CC num trabalho de CI. Dos dois casos de coorientação, na DM é o coorientado que apresenta a formação em CI e na TD ambos docentes responsáveis pela orientação têm formação na CC e não na CI (ponto 5.3).

¹⁶¹ A diferença nas percentagens obtidas pelo elemento *objetivos* é de 16 pontos percentuais relativamente ao elemento *conclusões* e 31 para o elemento *resultados*.

¹⁶² Na generalidade dos resumos os objetivos específicos não se encontram expressos na sua totalidade, ou em parte. Acresce, ainda relativamente aos elementos internos do resumo, que nos trabalhos *dm.10.2011* *dm.14.2013* não foi possível determinar os resultados dos mesmos a partir dos respetivos resumos.

Confirma-se o carácter interdisciplinar da formação dos docentes da CI, apenas 2, em 20, apresentam formação exclusivamente nesta área. Apesar de se ter observado que dos 18 orientadores, 13 possuem formação específica em CI, contudo, entende-se que os 5 que não a apresentam constituem, ainda assim, um número significativo dado que se tratam de estudos que exigem dos orientadores um conhecimento muito específico da matéria. Deste modo, entende-se pertinente a questão de Brokes sobre os docentes dos cursos de CI (ponto 2.3.2). Idêntico cenário se observa na situação inversa¹⁶³.

Ainda relativamente aos três pontos de confluência detetados, um deles, a WS enquanto tema interdisciplinar (CI-CC), merece algumas considerações finais. Considerações merecidas uma vez que um dos objetos de estudo do presente trabalho, a WS, é, simultaneamente, resultado do mesmo. Do trabalho efetuado conclui-se que a WS, enquanto ponto de confluência entre a CI e a CC, não apresenta grande expressividade. Esta falta de expressividade verifica-se em três situações: nos resumos dos trabalhos analisados, apresenta uma representatividade de 29% (embora surja associada a 4 dos 5 elementos estruturais do mesmo)¹⁶⁴; na formação dos docentes orientadores a qual não é representativa da ligação interdisciplinar e, por fim, na pesquisa bibliográfica constatou-se um número pouco significativo de referências à WS enquanto tema interdisciplinar por investigadores que não pertencem à área da CI.

Perante este contexto poder-se-á inferir que a necessidade colocada em 2009 por um dos autores dos trabalhos analisados continua atual: “existe a necessidade de que o conhecimento gerado na CI, sobre a Web Semântica ultrapasse os limites da própria área.” (Carvalho, 2009, p. 6). Apesar desta situação e das poucas referências de autores que não pertencem à área da CI, a relação não é totalmente unilateral. Eventos como *O Estado da Arte* realizados pela USP¹⁶⁵ são excelentes pontos de intercâmbio disciplinar. Em um desses eventos, realizado em 2014, Renata Wassermann, investigadora da área da Computação, expressa da seguinte forma a relação entre as duas áreas, no que concerne à Web Semântica:

A ideia não é decorar essa linguagem OWL e toda essa tecnologia, essas ferramentas, na verdade nada disso importa, o que importa são os conceitos que se usa. (...) No nível do conceito não importa como isso vai ser representado (...) e essa é a parte que eu acho que vocês [CI] sabem fazer melhor que a gente [CC], (Wassermann, 2014, pt. 1:03:40).

Partindo da ideia expressa por Wassermann entende-se a Web Semântica como um campo de direito próprio, embora não exclusivo, da Ciência da Informação na medida em que antes da representação encontra-se a conceptualização.

Limitações do estudo

O presente estudo apresenta limitações de duas ordens: aquelas que decorrem dos limites estabelecidos dentro de uma metodologia que se pretendeu precisa e explícita, e as que emergiram no decorrer da aplicação desta mesma metodologia ao campo de ação.

¹⁶³ Da mesma forma que se considera importante a formação específica em CI para orientar um trabalho de mestrado ou doutoramento na área, numa abordagem ao tema da WS, incluída nesses mesmos trabalhos, se considera relevante a formação na área da informática, característica ausente em 13 orientações.

¹⁶⁴ Está presente nos *objetivos, metodologia, resultados e conclusões* do resumo e não na *introdução* do mesmo.

¹⁶⁵ Série de eventos promovidos pelo Departamento de Biblioteconomia e Documentação da Escola de Comunicações e Artes da Universidade de São Paulo.

Relativamente às primeiras, as mesmas foram devidamente justificadas ao longo do trabalho, pelo que relativamente a estas apenas será abordada uma, aquela que diz respeito à utilização do único termo de pesquisa (*web semântica*), dado a sua relação direta com os resultados do estudo. Embora essa opção tenha limitado a quantidade de trabalhos incluídos na amostra, a análise da mesma revelou um padrão, em termos de aproximação ao tema do estudo, que legitima a opção tomada (ponto 5.2). Um alargamento do âmbito da pesquisa a temas associados aumentaria potencialmente a quantidade de trabalhos sem referências a pontos de confluência entre as duas áreas.

Quanto às limitações que emergiram no decorrer da aplicação da metodologia usada, ao campo de ação, há a referir dois tipos. O primeiro relativo ao modelo utilizado na comparação entre a análise dos resumos e a dos objetivos (ponto 5.2.3). Embora o referido modelo tenha sido capaz de revelar a existência da relação entre as duas variáveis, apresentou limitações que impediram outras inferências. Entre os ajustamentos necessários foi possível detetar a necessidade de um enriquecimento qualitativo do eixo Y. Eixo marcado por uma escala exclusivamente quantitativa limitadora de um maior refinamento no posicionamento dos trabalhos analisados.

O segundo tipo de limitações encontra-se relacionado com as fontes de obtenção de dados (repositórios RCAAP e OASISbr). Além da questão abordada no respetivo ponto (4.1), sobre a natureza dinâmica da inserção de registos nos referidos portais, outras limitações surgiram tais como: os registos duplicados (ponto 4.3.1) e as discrepâncias entre os termos que constam nos campos *assunto* dos repositórios relativamente às palavras-chave escolhidas pelos autores dos respetivos trabalhos (ponto 5.2.3).

Relativamente aos registos duplicados, essa limitação dificultou a recolha uma vez que se encontrou desajustes relativos às datas apresentadas nos registos e as manifestadas nos respetivos documentos. Essas diferenças na datação também se observaram nos registos duplicados, quer quando estes pertenciam ao mesmo repositório quer quando pertenciam a cada um dos dois repositórios (RCAAP e OASISbr). Analisando esses casos foi possível apurar que os mesmos provinham de três universidades: UFG, UFRN e UNICAMP. Apresenta-se como razão plausível para estas discrepâncias o facto de a colheita da informação relativa ao ano ser efetuada em campos diferentes, uma vez que cada registo apresenta mais que um campo de data¹⁶⁶. Além da duplicação de registos com diferentes datas foram, também, encontrados no repositório da UNICAMP, Dissertações de Mestrado duplicadas como sendo Teses de Doutoramento.

As discrepâncias descritas ocorrem com maior frequência no RCAAP, sendo neste portal que se encontram os casos das três universidades enquanto no OASISbr apenas se observam os casos dos duplicados da UNICAMP. No portal português, os trabalhos das universidades UFG e UFRN, devido ao facto de a data do seu registo não se encontrar correta, estes não são recuperados no ano a que diz respeito os respetivos trabalhos. Sendo, essa, uma possível explicação para a ausência de 40 registos que apenas foram recuperados no OASISbr.

¹⁶⁶ No modo de visualização “registo completo” observa-se, na grande maioria dos repositórios de origem, quatro campos com valores “data” em cada registo, com diferentes designações tais como: i. *dateDefense*; ii. *date.issued*; iii. *date.available*; iii. *date.accessioned*; iv. *Date of availability*; v. *Date of Access*. O padrão observado aponta para uma maior coerência entre o ano presente nos trabalhos e o que consta nos respetivos registos quando associado aos campos i e ii.

Em relação à diferença entre as palavras-chave que se encontram registadas nos resumos dos documentos e aquelas que se encontram no campo *assunto* dos repositórios das Instituições de Ensino Superior (rIES), é possível inferir que a mesma esteja relacionada com os critérios usados pelo indexador. Já a possível explicação para o desfaseamento entre os termos que constam nos campos *assunto* dos rIES e os que se encontram nos repositórios agregadores (RCAAP e OASISbr) não se afigura tão linear. No entanto, pelo analisado é possível avançar com uma eventual explicação similar à questão das datas, ou seja, a existência de mais de um campo, nos rIES, nos quais se registam as “palavras-chave”. Como exemplo, pode apontar-se o caso do repositório da UFMG que apresenta dois campos (*Subject* e *Palavra-chave*) com diferentes termos, sendo que o OASISbr apresenta, no seu campo *Assunto*, unicamente os termos que constam no campo *Subject* do rIES da UFMG. Outro caso, que se apresenta como uma incompatibilidade na recolha de dados, tem a ver com o rIES da UNB. Os registos desse repositório não apresentam, no OASISbr, qualquer termo no campo *Assunto*.

Estas questões técnicas dos repositórios foram as grandes limitações ao estudo efetuado, com implicações diretas no número de trabalhos da amostra, (22 trabalhos, quando este número poderia ser maior, não se tivesse observado este tipo de limitações). Todavia, os resultados apresentados encontram-se circunscritos à amostra analisada não sendo, nesse sentido, posta em causa a sua validade. Acresce, ainda, que, enquanto estudo exploratório, considera-se os padrões encontrados indicadores úteis para estudos mais aprofundados.

Por fim, é conveniente referir que as limitações encontradas, durante a preparação do estudo, em vários rIES¹⁶⁷, bem como a disparidade de funcionalidades apresentadas pelos mesmos, torna inviável uma uniformização de critérios de recolha pretendida, justificando a opção da utilização dos portais considerados (RCAAP e OASISbr).

Contribuições para área da CI e linhas de futuras investigações

Considerando o trabalho desenvolvido, apontam-se algumas contribuições do mesmo para a área da CI, assim como possíveis linhas para futuras investigações.

Contribuições para área da CI

No que concerne às contribuições, estas são encaradas dentro da “descoberta de perspectivas teóricas novas do ponto de vista do investigador que efectuou o trabalho, ainda que estas sejam amplamente conhecidas noutros contextos.” (Quivy & Campenhoudt, 1992, pt. 246). Neste contexto aponta-se as seguintes:

- a) A clarificação de conceitos relativos à Web Semântica, uma vez que a discriminação dos mesmos contribui para a ausência de “ruído” nas comunicações;
- b) A simplificação do complexo conceito de Ciência da Informação através da construção de modelos;

¹⁶⁷ Como exemplo aponta-se: a impossibilidade ou limitação na pesquisa em campos específicos (PUCCAMP, UC, UFBA, UFMG, UL); a impossibilidade de definir um período temporal (UNB, UFPE, UEL); a impossibilidade de seleccionar a tipologia (UFSC) e, ainda, a limitação do acesso ao próprio repositório (UFF). Os testes foram efetuados na fase inicial de preparação do estudo em outubro e novembro de 2015. Pelos testes efetuados destaca-se pela positiva o rIES da UNESP/MAR pelas funcionalidades que o mesmo apresenta.

- c) Uma interpretação das complexas relações entre as disciplinas Biblioteconomia, Arquivística e Documentação e a Ciência da Informação no contexto dos cursos de ensino superior;
- d) Alguns indicadores relativos à relação entre a Ciência da Informação e a Informática/Ciência da Computação tendo como ponto de convergência a Web Semântica.

Linhas de futuras investigações

Como linhas de investigação indicam-se quatro dentro de outras possíveis:

- a) Um aprofundamento das questões relacionadas com os repositórios, nomeadamente no que diz respeito à ausência de registos ou de informação associada aos mesmos, assim como à sua duplicação;
- b) Um alargamento do estudo efetuado às formações académicas dos orientadores, não só a um maior número como também à formação dos próprios autores dos trabalhos, no sentido de verificar a questão levantada sobre a interdisciplinaridade;
- c) Uma aplicação do modelo de análise usado, devidamente enriquecido, a uma maior amostra, alargando a análise às secções *resultados* e *conclusões* (fora dos resumos);
- d) Um estudo diacrónico da utilização dos termos *web semântica*, *web of data* e *web of linked data* (ou simplesmente *linked data*) em elementos chave de documentos de carácter científico (título, palavras-chave e/ou resumos).

Referências Bibliográficas

- Almeida, M. B. (2013). Revisiting ontologies: A necessary clarification. *Journal of the American Society for Information Science and Technology*, 64(8), 1532–2890. <http://doi.org/10.1002/asi.22861>
- Almeida, M. B., Souza, R. R., & Fonseca, F. (2011). Semantics in the Semantic Web: A critical evaluation. *Knowledge Organization*, 38(3), 187–203. <http://doi.org/10.1186/2041-1480-2-S1-S1>
- Araújo, C. A. Á. (2003). A ciência da informação como ciência social. *Ciência da Informação*, 32(3), 21–27. Obtido de <http://www.scielo.br/pdf/ci/v32n3/19020.pdf>
- Araújo, C. A. Á. (2013). O que é Ciência da Informação? *Informação & Informação*, 19(1), 1–30. <http://doi.org/10.5433/1981-8920.2014v19n1p01>
- Bardin, L. (2011). *Análise de conteúdo*. (L. A. R. e A. Pinheiro, Trad.). São Paulo: Almedina.
- Barreto, A. de A. (2008). Uma quase história da ciência da informação. *DataGramaZero*, 9(2), 1–18. Obtido de http://eprints.rclis.org/17637/1/DataGramaZero - Revista de Ciência da Informação - Artigo 01_Aldo.pdf
- Belkin, N. J., & Robertson, S. E. (1976). Information Science and the Phenomenon of Information. *Journal of the American Society for Information Science*, (August), 197–204. <http://doi.org/10.1002/asi.4630270402>
- Berners-Lee, T. (1998, Outubro 14). A roadmap to the Semantic Web. Obtido 21 de Março de 2016, de <https://www.w3.org/DesignIssues/Semantic>
- Berners-Lee, T. (2006). The World Wide Web: Past, present and future (frequently asked questions). Obtido 16 de Março de 2016, de <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. Obtido de <https://youtu.be/KY5skobffk0>
- Berners-Lee, T. (2010). Long Live the Web. *Scientific American*, 22(December), 80–85. Obtido de https://jblomo.github.io/webarch253/slides/Long_Live_the_Web.pdf
- Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H. F., & Secret, A. (1994). The World-Wide Web. *Communications of the ACM*, 37(8), 76–82. <http://doi.org/10.1145/179606.179671>
- Berners-Lee, T., & Fischetti, M. (2000). *Weaving the Web: The original design and ultimate destiny of the world wide web by its inventor*. HarperCollins.
- Berners-Lee, T., Hall, W., Hendler, J. A., O'hara, K., Shadbolt, N., & Weitzner, D. J. (2006). A Framework for Web Science. *Foundations and Trends in Web Science*, 1(1), 1–130. <http://doi.org/10.1561/1800000001>
- Berners-Lee, T., & Hendler, J. A. (2001). Publishing on the Semantic Web. *Nature*, 410(April), 1023–1024. Obtido de <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. <http://doi.org/10.1007/978-3-642-29923-0>
- Bicalho, L. M. (2009). Interações disciplinares da pesquisa na Ciência da Informação. Em *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 235–250). Coimbra: Imprensa da Universidade de Coimbra.
- 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. Em I. E. Anagnostopoulos, M. Bieliková, P. Mylonas, & N. Tsapatsoulis (Eds.), *Semantic Hyper/Multimedia Adaptation: Schemes and Applications* (pp. 319–360). Berlin: Springer Berlin Heidelberg. http://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. Obtido de <http://eprints.soton.ac.uk/271285/>
- Borko, H. (1968). Information Science: What Is It? *American Documentation*, 19(1), 3–5. Obtido de <http://cdigital.uv.mx/bitstream/123456789/6699/2/Borko.pdf>
- Briet, S. (1951). *Qu'est-ce que la documentation?* Paris: Editions Documentaires. Obtido de <http://martinetl.free.fr/suzannebriet/questcequeladocumentation/briet.pdf>
- Brookes, B. C. (1980). The foundations of information science: Part I. Philosophical aspects. *Journal of Information Science*, 2(3–4), 125–133. <http://doi.org/10.1177/016555158000200302>
- Cambridge Semantics. (2016). Many Names for the Semantic Web. Obtido de

- <http://www.cambridgesemantics.com/semantic-university/many-names-semantic-web>
- Capurro, R. (2003, Novembro 10). Epistemologia e Ciência da Informação. (A. M. R. Cabral, E. W. Dias, I. Paim, L. M. M. Dumont, M. P. Aun, & M. E. N. Borges, Trans.) *V Encontro Nacional de Pesquisa em Ciência da Informação*. Belo Horizonte. Obtido de http://www.capurro.de/enancib_p.htm
- Capurro, R., & Hjørland, B. (2007). O conceito de informação. (A. M. P. Cardoso, M. da G. A. Ferreira, & M. A. de Azevedo, Trans.) *Perspectivas em Ciência da Informação*. <http://doi.org/10.1590/S1413-99362007000100012>
- Caro-castro, C. (2010). Las clasificaciones bibliográficas: De los estantes a la web. *Tabula*, 13, 11–23. Obtido de <http://hdl.handle.net/10366/121954>
- Carvalho, R. A. de. (2009). *Perspectivas na web semântica para a Ciência da Informação*. Pontifícia Universidade Católica de Campinas. Obtido de <http://tede.bibliotecadigital.puc-campinas.edu.br:8080/jspui/handle/tede/825>
- CNPq, CAPES, & FINEP. (2005). Nova tabela das áreas do conhecimento: Versão preliminar proposta para discussão. <http://doi.org/10.1590/S0037-86822005000400022>
- Codina, L. (2003). La web semántica: una visión crítica. *El Profesional de la Información*. <http://doi.org/10.1076/epri.12.2.149.15480>
- Costa, M. P., & Leite, F. C. L. (2016). Open access in the world and Latin America: A review since the Budapest Open Access Initiative. *Transinformacao*, 28(1), 33–45. <http://doi.org/10.1590/2318-08892016002800003>
- Costa, S. M. S., Kuramoto, H., & Leite, F. C. L. (2013). Acesso Aberto no Brasil: aspetos históricos, ações institucionais e panorama atual. Em E. Rodrigues, A. Swan, & A. A. Baptista (Eds.), *Uma Década de Acesso Aberto na UMinho e no Mundo* (pp. 133–150). Braga: Universidade do Minho, Serviços de Documentação. Obtido de https://repositorium.sdum.uminho.pt/bitstream/1822/26144/3/RepositoriumUM_10anos.pdf
- Feigenbaum, L. (2011). The Semantic Web Landscape: A practical introduction [Video]. Em *Cambridge Semantic Web Meetup*. Cambridge Semantics. Obtido de <http://www.cambridgesemantics.com/semantic-university/introduction-to-the-semantic-web>
- Fernandes, W. R., & Cedón, B. V. (2009). Ciência da Informação e Interdisciplinaridade: Análise das áreas de conhecimento correlatas. Em M. M. Borges & E. S. Casado (Eds.), *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 113–127). Coimbra: Imprensa da Universidade de Coimbra.
- Galvão, M. C. B. (1998). Construção de conceitos no campo da ciência da informação. *Ciência da Informação*, 27(1), 46–52. <http://doi.org/10.1590/S0100-19651998000100006>
- Garcia, J. C. R. (2002). Conferências do Georgia Institute of Technology e a Ciência da Informação: «de volta para o futuro». *Informação & Sociedade: Estudos*, 12(1). Obtido de <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/153>
- Gracioso, L. de S., & Saldanha, G. S. (2011). *Ciência da Informação e Filosofia da Linguagem: da pragmática informacional à web pragmática*. Araraquara: Junqueira&Marin.
- Hendler, J. A. (2011a). Semantic Web: «10 year update» [video]. Obtido de <http://www.vestforsk.no/video/wims/wims11-jahendler-keynote.ogv>
- Hendler, J. A. (2011b). «Why the Semantic Web will never work» (note the quotes) [Video]. Obtido de <https://youtu.be/oKiXpO2rbJM>
- Hendler, J. A. (2014). The Semantic Web: The Inside Story [video]. Em *Summer School in cognitive Science: Web Science and the Mind*. Montréal: Institut des sciences cognitives, UQAM. Obtido de <https://youtu.be/3Ap5FsvxjTQ>
- 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. Obtido de <http://www.w3.org/2002/07/swint>
- Hendler, J. A., Shadbolt, N., Hall, W., Berners-Lee, T., & Weitzner, D. J. (2008). Web science: An interdisciplinary approach to understanding the Web. *Communications of the Acm*, 51(7), 60–69. <http://doi.org/10.1145/1364782.1364798>
- 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. <http://doi.org/10.1002/asi>
- Isotani, S., & Bittencourt, I. I. (2015). Dados Abertos Conetados. Obtido 9 de Agosto de 2016, de

- <http://ceweb.br/livros/dados-abertos-conectados/>
- Laaff, M. (2011, Julho 22). Internet visionary Paul Otlet: Networked knowledge, decades before Google. (E. Ornstein, Trad.) *SPIEGEL ONLINE - International*. Obtido de <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, Trad.). Brasília: Briquet de Lemos/Livros.
- López Yepes, J. (1995). *La Documentación como disciplina: Teoría e historia* (2.ª ed.). Pamplona: Ediciones Universidade de Navarra.
- Marchiori, M. (2003). The Semantic Web Made Easy. Obtido 15 de Maio de 2016, de <https://www.w3.org/RDF/Metalog/docs/sw-easy>
- Marcondes, C. H. (2012). Em busca de uma semântica do digital, ou «As they may think». *PontoDeAcesso*, 6. Obtido de <http://www.portalseer.ufba.br/index.php/revistaici/article/viewArticle/6103>
- Mikhailov, A. I., & Giljarevskij, R. S. (1971). *An Introductory Course on Informatics/Documentation*. International Federation for Documentation. Obtido de <http://eric.ed.gov/?id=ED060875>
- Ministério das Actividades Económicas e do Trabalho. Portaria nº 256/2005 de 16 de Março: Classificação Nacional das Áreas de Educação e Formação, Diário da República — I Série-B, nº 53 2281–2313 (2005). Portugal. Obtido de <http://certifica.dgert.msess.pt/legislacao/cnaef-classificacao-nacional-de-areas-de-educacao-e-formacao.aspx>
- Otlet, P. (1934). *Traité de documentation: Le livre sur le livre: Théorie et pratique*. Bruxelles: Ediciones Mundaneum Palais Mondial. Obtido de http://lib.ugent.be/fulltxt/handle/1854/5612/Traite_de_documentation_ocr.pdf
- Parreiras, F. silva. (2012). *Semantic Web and Model-driven Engineering*. Wiley. Piscataway: John Wiley & Sons, Inc.
- Patel-Schneider, P. F., & Horrocks, I. (2007). A comparison of two modelling paradigms in the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(4). Obtido de <http://www.websemanticsjournal.org/index.php/ps/issue/view/21>
- Pinheiro, L. V. R. (1997). *A Ciência da Informação entre sombra e luz: Domínio epistemológico e campo interdisciplinar*. Universidade Federal do Rio de Janeiro. Obtido de <http://ridi.ibict.br/handle/123456789/35>
- Pinheiro, L. V. R. (2009). Configurações disciplinares e interdisciplinares da Ciência da Informação no ensino e pesquisa no Brasil. Em *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 99–111). Coimbra: Imprensa da Universidade de Coimbra.
- Pombo, O. (2002). Da classificação dos seres à classificação dos saberes. Lisboa: Instituto de Educação, Universidade de Lisboa. Obtido de <http://www.educ.fc.ul.pt/hyper/resources/opombo-classificacao.pdf>
- Quivy, R., & Campenhoutd, L. Van. (1992). *Manual de investigação em ciências sociais*. (J. M. Marques, M. A. Mendes, & M. Carvalho, Trads.) (6.ª ed.). Lisboa: Gradiva.
- Ray, K. (2010). Web 3.0 [Video]. Obtido de <https://vimeo.com/11529540>
- Ribeiro, F. (2006). Um modelo formativo em Ciência da Informação, de feição europeia e adequado a Bolonha: o caso da Universidade do Porto. *Cadernos de Biblioteconomia Arquivística e Documentação*, 1, 16–27. Obtido de http://www.apbad.pt/CadernosBAD/Caderno12006/CadBAD1_06_FernandaRibeiro.pdf
- Robredo, J. (2010). Ciência da informação e Web semântica: Linhas convergentes ou linhas paralelas? Em *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento* (pp. 12–47). Brasília: IBICT. Obtido de <http://repositorio.unb.br/handle/10482/7952>
- Rojas, M. Á. R., & Domínguez, S. E. C. (2009). La Archivística y las Disciplinas Informativas Documentales: Retos y Cuestionamientos Epistemológicos. Em *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 203–212). Coimbra: Imprensa da Universidade de Coimbra.
- Rousseau, J.-Y., & Couture, C. (1998). *Os fundamentos da disciplina arquivística*. Lisboa: Publicações Dom Quixote.
- Sales, R. de, & Murguia, E. I. (2015). Determinações políticas na produção científica da Ciência da Informação do Brasil: Impacto da Tabela de Áreas de Conhecimento (TAC) do CNPq. *Scire*, 21(1), 27–34. Obtido de <http://ibersid.eu/ojs/index.php/scire/article/view/4250/3802>
- Sandhaus, E. (2011). Entities, Relationships, and Semantics: the State of Structured Search [video]. Em D. Tunkelang (Ed.), *Strata New York 2011*. Obtido de

- <https://youtu.be/vr1blOJxXfQ?list=PLF8EX6mxQGp7IT3bhPRHuTYjtlyKue0IO>
- Saracevic, T. (1996). Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1(1), 41–62. Obtido de <http://www.brapci.inf.br/article.php?dd0=0000003224&dd90=bc40e71f02>
- Saracevic, T. (2009). Information Science. Em M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences* (pp. 2570–2585). New York: Taylor & Francis. <http://doi.org/10.1081/E-ELIS3-120043704>
- Setzer, V. W. (2015, Maio 25). Dado, Informação, Conhecimento e Competência. Obtido 14 de Maio de 2016, de <https://www.ime.usp.br/~vwsetzer/dado-info.html>
- Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. *IEEE Intelligent Systems*, 21(3), 96–101. <http://doi.org/10.1109/MIS.2006.62>
- Shera, J. H., & Cleveland, D. B. (1977). History and foundations of information-science. Em M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 12, pp. 249–275). New York: Knowledge Industry Publications Inc.
- 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. <http://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* (2.^a ed.). Porto: Edições Afrontamento.
- Simões, M. da G. (2008). *Da abstração à complexidade formal: relações conceptuais num tesouro*. Coimbra: Almedina.
- Simões, M. da G. (2011). *Classificações bibliográficas: percurso de uma teoria*. Coimbra: Almedina.
- Sletten, B. (2014, Janeiro 13). Keep On Keeping On. Dataversity, Smart Data. Obtido de <http://www.dataversity.net/keep-on-keepin-on/>
- Souza, R. R., Almeida, M. B., & Baracho, R. M. A. (2013). Ciência da Informação em transformação: Big Data, Nuvens, Redes Sociais e Web Semântica. *Ciência da Informação*, 40(2), 159–173. Obtido de <http://revista.ibict.br/index.php/ciinf/article/view/2283/1906>
- Souza, R. R., & Alvarenga, L. (2004). A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, 33(1), 132–141. Obtido de <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/50/50>
- Souza, R. R., Tudhope, D., & Almeida, M. B. (2012). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, 39(3), 179–192. Obtido de <http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=75340225&site=ehost-live&scope=site>
- Souza, E. D. de, & Dias, E. J. W. (2009). A epistemologia interdisciplinar na Ciência da Informação: Implicações da formação do campo científico. Em *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 129–142). Coimbra: Imprensa da Universidade de Coimbra.
- Souza, R. F. de, & Stumpf, I. R. C. (2009). Ciência da Informação como área do conhecimento: Abordagem no contexto da pesquisa e da Pós-Graduação no Brasil. *Perspectivas em Ciência da Informação*, 14(spe), 41–58. <http://doi.org/10.1590/S1413-99362009000400004>
- Stumpf, I. R. C. (2009). A Ciência da Informação no Brasil através de seus Programas de Pós-Graduação. Em *A Ciência da Informação criadora de conhecimento, vol.1, IV Encontro Ibérico EDIBCIC 2009* (pp. 167–176). Coimbra: Imprensa da Universidade de Coimbra.
- W3C - World Wide Web Consortium. (2013a). W3C Data Activity: Building the Web of Data. Obtido 18 de Outubro de 2015, de <http://www.w3.org/2013/data/>
- W3C - World Wide Web Consortium. (2013b, Dezembro 11). W3C Semantic Web Activity Homepage. Obtido 27 de Janeiro de 2016, de <https://www.w3.org/2001/sw/>
- W3C - World Wide Web Consortium. (2014, Janeiro 28). Semantic Web Wiki. Obtido 21 de Março de 2016, de https://www.w3.org/2001/sw/wiki/Main_Page
- W3C - World Wide Web Consortium. (2015a). W3C Standards: Linked Data. Obtido 1 de Maio de 2016, de <http://www.w3.org/standards/semanticweb/data>
- W3C - World Wide Web Consortium. (2015b). W3C Standards: Semantic Web. Obtido 21 de Março de 2016, de <https://www.w3.org/standards/semanticweb/>

- Wassermann, R. (2014). O Estado da Arte: Ontologia e Representação do Conhecimento [video]. [video file], São Paulo: Departamento de Biblioteconomia e Documentação da USP/ECA, IPTV USP. Obtido de <http://iptv.usp.br/portal/video.action?idItem=25631>
- Wersig, G., & Neveling, U. (1975). The phenomena of interest to Information Science. *Information Scientist*, 9(4). Obtido de <http://sigir.org/files/museum/pub-13/18.pdf>
- Yin, R. K. (2001). *Estudo De Caso - Planejamento E Metodos*. (D. Grassi, Trad.) (2.^a ed.). São Paulo: Bookman.
- 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. <http://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. <http://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. <http://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. <http://doi.org/10.1002/asi.20505>

Índice de Figuras

<i>Figura 1.</i> Relação dos objetivos específicos com objetivo geral a partir do ponto-chave inicial do mesmo.	5
<i>Figura 2.</i> Esquema do relacionamento das etapas da análise concetual efetuada, com descrição dos respetivos focos da mesma.....	6
<i>Figura 3.</i> Distribuição temporal das unidades de registo consideradas na dimensão de análise <i>Caraterísticas descritivas</i> de acordo com a data do documento de proveniência.	15
<i>Figura 4.</i> Distribuição temporal das unidades de registo consideradas na dimensão de análise <i>Base estrutural</i> de acordo com a data do documento de proveniência. As dimensões dos signos utilizados para representar as <i>Especificações/padrões</i> variam de acordo com a quantidade de elementos mencionados nas respetivas unidades de registo.	18
<i>Figura 5.</i> Três diapositivos do tutorial sobre <i>Linked Data</i> apresentado na 23ª conferência internacional da <i>World Wide Web</i> em abril de 2014. Fabien Gandon, <i>Semantic Web and Linked Data or how to link data and schemas on the web: A W3C tutorial at WWW2014</i> , http://www.slideshare.net/fabien_gandon/semantic-web-and-linked-data , (acedido a 4-5-2016).	24
<i>Figura 6.</i> Projeção isométrica da “pilha tecnológica” da Web Semântica, com destaque das especificações associadas a <i>Linked Data</i> . Adaptação de Ioannis Parapontis,, https://ioannis.parapontis.com/2012/12/07/semantic-web-technology-stacks/ , do original de Benjamin Nowack, http://bnode.org/blog/2009/07/08/the-semantic-web-not-a-piece-of-cake , (URLs acedidos em 8-6-2016).....	25
<i>Figura 7.</i> Evolução da <i>Linked Open Data Cloud</i> , de maio de 2007 (à esquerda) a agosto de 2014 (à direita). Richard Cyganiak e Anja Jentzsch, 2014, http://lod-cloud.net/ (acedido a 17-7-2016).	26
<i>Figura 8.</i> Perspetiva da “torre da web semântica” vista pelo “eixo P” (de Perceção e Pessoas). Massimo Marchiori, https://www.w3.org/RDF/Metalog/docs/sw-easy , (URL acedido a 12-6-2016).....	27
<i>Figura 9.</i> Distribuição das 15 obras selecionadas relativamente aos três momentos e respetivos eventos descritos.....	34
<i>Figura 10.</i> Esquema da abordagem dos fenómenos D-I-C no contexto da CI elaborado a partir da análise de Chain Zins (2007c, pp. 487–489).....	36
<i>Figura 11.</i> Esquema dos modelos concetuais para a definição de CI, identificados por Chain Zins, elaborado a partir da descrição dos mesmos (Zins, 2007b, pp. 340–350).....	38
<i>Figura 12.</i> Mapeamento das relações entre os componentes apresentados no modelo de Chain Zins para o mapa de conhecimento da CI, elaborado a partir do quadro: “FIG. 2. Knowledge map of information science” (Zins, 2007d, p. 529)	39
<i>Figura 13.</i> Esquema de relacionamento dos das características associadas às categorias principais do domínio Conhecimento Específico do Mapa do Conhecimento da CI, elaborado a partir do quadro: “FIG. 2. Knowledge map of information science” (Zins, 2007d, p. 529).....	40
<i>Figura 14.</i> O mapa da Ciência da Informação segundo Yves-François Le Coadic (1996, p. 24).	49
<i>Figura 15.</i> Diagrama do campo da Ciência da Informação segundo Armando Malheiro da Silva e Fernanda Ribeiro, (2008, p. 80).....	51
<i>Figura 16.</i> Distribuição temporal das referências às disciplinas mencionadas por três ou mais autores incluídos na Tabela 8 nas respetivas obras aí descritas.....	52
<i>Figura 17.</i> Distribuição da amostra pelas Instituições de Ensino Superior aí representadas.	73
<i>Figura 18.</i> Distribuição da amostra pelo período considerado.....	75
<i>Figura 19.</i> Relação dos critérios usados na análise do termo Web Semântica nos objetivos dos 22 trabalhos que constituem a amostra.	78

<i>Figura 20.</i> Distribuição dos 223 trabalhos recuperados (DM e TD), de acordo com o país de origem (49 de Portugal e 174 do Brasil), pelos grupos estabelecidos no presente trabalho para organização das áreas de educação/conhecimento em uso pelas entidades DGEEC e CAPES.	80
<i>Figura 21.</i> Distribuição dos 223 trabalhos recuperados pelos grupos estabelecidos no presente trabalho para organização das Áreas de Educação/Conhecimento e dos 41 cursos com designações diferentes (representados nos 223 trabalhos) pelos grupos estabelecidos no presente trabalho para organização das Áreas de Educação/Conhecimento.....	81
<i>Figura 22.</i> Distribuição das referências a pontos de convergência entre a WS e a CI pela amostra (22 trabalhos) e subamostra (13 trabalhos).....	82
<i>Figura 23.</i> Diagrama de relacionamento entre a subclasse <i>contributo da WS para a CI</i> (a) e respetivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respetivos documentos de origem, identificados pela referência da Tabela 16.	83
<i>Figura 24.</i> Diagrama de relacionamento entre a subclasse <i>contributo da CI para a WS</i> (b) e respetivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respetivos documentos de origem, identificados pela referência da Tabela 16.	84
<i>Figura 25.</i> Diagrama de relacionamento entre a subclasse <i>WS como tema na fronteira entre CI e CC</i> (c), e respetivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respetivos documentos de origem (DM/ TD, identificados pela referência da Tabela 16).....	85
<i>Figura 26.</i> Diagrama de relacionamento entre a subclasse <i>uso conjunto de conhecimentos da CI e da WS</i> (d) e respetivas unidades de registo (em tom laranja), com as subclasses dos elementos estruturais do resumo (ligações a vermelho). As unidades de registo estão conectadas (a cinza) aos respetivos documentos de origem, identificados pela referência da Tabela 16.....	85
<i>Figura 27.</i> Distribuição percentual das 21 unidades de registo pelas duas classes de análise: <i>pontos de confluência</i> e <i>elementos do resumo</i>	86
<i>Figura 28.</i> Distribuição das unidades de registo pela linha de tempo, de acordo com o ano dos respetivos documentos. Visão integrada das duas classes: <i>pontos de confluência</i> e <i>elementos do resumo</i>	87
<i>Figura 29.</i> Distribuição da percentagem dos 22 (100%) trabalhos da amostra pelos graus de proximidade relativa à WS enquanto tema dos mesmos.	87
<i>Figura 30.</i> Distribuição dos 22 trabalhos pelos graus de proximidade ao longo do período considerado.	88
<i>Figura 31.</i> Comparação entre os valores: frequência de pontos de confluência nos resumos – grau de proximidade nos objetivos, determinados pelas respetivas análises aos 22 trabalhos.	89
<i>Figura 32.</i> Representação dos processos esperados para: a inserção do termo Web Semântica no campo <i>assunto</i> dos registos das DM e TD, nos repositórios RCAAP e OASISbr, e para a posterior recuperação dos respetivos trabalhos.....	90
<i>Figura 33.</i> Distribuição dos orientadores e coorientadores pelas universidades associadas às dissertações de mestrado e teses de doutoramento que constituem a amostra.	91
<i>Figura 34.</i> Distribuição da formação académica dos 20 docentes, responsáveis pelas orientações das DM/TD da amostra, de acordo com os 4 grupos de áreas: A- CI; B- Informática/Computação; C- Eletrónica/Automação; D- outros cursos. Apresentação por: pós-doutoramento e doutoramento (pD/D); mestrado e especializações (M/Esp); licenciatura e bacharelato (L/B) e soma dos 3 estágios.....	92
<i>Figura 35.</i> Mapa das relações entre a formação académica dos orientadores e coorientadores e os respetivos trabalhos que constituem a amostra orientados.	93

Índice de Tabelas

Tabela 1.Referências documentais usadas na análise concetual da “Web Semântica”	12
Tabela 2.Análise do conceito Web Semântica: Grupos e respetivas unidades de registo consideradas na dimensão de análise 1- Características descritivas	14
Tabela 3.Análise do conceito Web Semântica: Grupo 2a (Procedimentos) e respetivas unidades de registo consideradas na dimensão de análise 2- Base estrutural	16
Tabela 4.Análise do conceito Web Semântica: Grupo 2b (Arquitetura) e respetivas unidades de registo consideradas na dimensão de análise 2- Base estrutural	17
Tabela 5.Análise do conceito Web Semântica: Unidades de registo, dentro da dimensão de análise Funções operativas, que apontam como objetivo geral da WS a utilização da Web como uma base de dados global e não apenas um repositório de documentos.....	18
Tabela 6.Análise do conceito Web Semântica: Unidades de registo, dentro da dimensão de análise Funções operativas, que apontam funções esperadas da WS relacionadas com o processamento automático e/ou destinado aos “utilizadores finais” do sistema computacional.....	19
Tabela 7.Exemplo das relações contraditórias de sinonímia e de pertença entre os conceitos Semantic Web, Web of Data e Linked Data encontradas numa mesma fonte	22
Tabela 8.Obras selecionadas para servirem de base à abordagem ao conceito de Ciência da Informação ...	33
Tabela 9.Distribuição dos fenómenos Dado (D), Informação (I) e Conhecimento (C) nos domínios universal e subjetivo de acordo com os modelos identificados por Chain Zins	35
Tabela 10.Disciplinas cuja relação com a Ciência da Informação é referida por três ou mais autores incluídos na Tabela 8 nas respetivas obras aí descritas	52
Tabela 11.Cursos do 2º e 3º Ciclos do Ensino Superior (mestrado e doutoramento) acreditados pela A3ES e em funcionamento em Portugal no ano letivo 2015-2016, listados na área de educação e formação “Biblioteconomia, Arquivo e Documentação (BAD)” da CNAEF e/ou considerados no presente trabalho como sendo de Ciência da Informação	57
Tabela 12.A Ciência da Informação nas versões da “Tabela de Áreas do Conhecimento” do CNPq.	59
Tabela 13.Cursos de pós-graduação (mestrado e doutorado) em funcionamento no Brasil no ano letivo 2015-2016, listados dentro da “Área do Conhecimento” Ciência da Informação no website da CAPES (confrontação entre as listagens de 1-11-2015 e de 1-12-2015)	61
Tabela 14.Listagem de termos com três ou mais ocorrências nas fontes de WS e com uma ou mais ocorrências nos cursos de CI.....	70
Tabela 15.Número total de ocorrências dos termos web semântica e ontologias nos cursos listados na Tabela 11 (Portugal) e na Tabela 13 (Brasil) onde foram encontradas menções aos termos da Tabela 14.	71
Tabela 16.Identificação dos trabalhos que compõem a amostra do presente estudo de caso	74
Tabela 17.Distribuição dos 223 trabalhos, de acordo com o curso que representam, pelas Áreas de Educação e Formação (Portugal) ou pelas Áreas Básicas ou do Conhecimento (Brasil) onde os respetivos cursos estão classificados.....	79

Lista de apêndices

- Apêndice 1 Lista dos 17 artigos reunidos sob o título “O Estatuto Epistemológico da Ciência da Informação” que constam na obra “A Ciência da Informação criadora de conhecimento – volume I” (2009), coordenada por Maria Manuel Borges e Elias Sanz Casado.
- Apêndice 2 Confrontação contextual e funcional entre o Repositório Científico de Acesso Aberto de Portugal (RCAAP) e o Portal Brasileiro de Acesso Aberto à Produção Científica (OASISbr).
- Apêndice 3a Folha de Recolha de Dados geral, para registo de informações relativas aos trabalhos recuperados (apresenta-se apenas alguns registos a título exemplificativo dado a quantidade de registos (251) e extensão dos elementos recolhidos).
- Apêndice 3b Folha de Recolha de Dados específica, para registo de informação proveniente de trabalhos de CI com acréscimo de dados complementares.
- Apêndice 4 Lista dos 14 termos recolhidos relativos à Web Semântica e confrontação entre as fontes usadas para a recolha (glossários de WS e programas, ementas e conteúdos das disciplinas de cursos/programas de CI).
- Apêndice 5 Esquema dos registos que constituem o universo e a amostra do presente trabalho e sua relação com os respetivos repositórios onde foram recuperados.
- Apêndice 6 Distribuição dos cursos e respetiva quantidade de trabalhos, correspondentes aos 223 registos recuperados nos portais RCAAP e OASISbr, pelos quatro grupos de áreas concebidos.
- Apêndice 7 Resumos e objetivos, geral e específicos, dos 22 trabalhos que constituem a amostra.
- Apêndice 8 Quadro de análise dos objetivos dos 22 trabalhos que constituem a amostra.

Apêndice 1

Artigos reunidos sob o título “O Estatuto Epistemológico da Ciência da Informação”, utilizados na seleção de obras referida no ponto 2.1.2. (Artigos incluídos na obra: Borges, M. M. & Casado, E. S., [Coords.]. (2009). *A Ciência da Informação criadora de conhecimento* (Vol. I). Coimbra: Imprensa da Universidade de Coimbra.)

Autoria	Artigo
Maria Izabel Moreira Arruda	<i>Biblioteconomia ou Ciência da Informação?</i>
Carlos Henrique Marcondes	<i>Fundamentos evolutivos e culturais da Ciência da Informação</i>
Cristina Dotta Ortega	<i>Surgimento e consolidação da Documentação: subsídios para compreensão da história da Ciência da Informação no Brasil</i>
Gustavo Freitas, Armando Malheiro da Silva	<i>Identificação de paradigmas nos Programas de Pós-graduação em Ciência da Informação oferecidos em Portugal e no Brasil</i>
Gustavo Silva Saldanha	<i>A complexidade e o conhecimento nos séculos XIII e XX: uma reflexão epistemológica</i>
Lena Vania Ribeiro Pinheiro	<i>Configurações Disciplinares e Interdisciplinares da Ciência da Informação no Ensino e Pesquisa</i>
Wesley Rodrigo Fernandes, Beatriz Valadares Cendón	<i>Ciência da Informação e Interdisciplinaridade: análise das áreas de conhecimento correlatas</i>
Edivanio Duarte de Souza, Eduardo José Wense Dias	<i>A Epistemologia Interdisciplinar na Ciência da Informação: Implicações na Formação do Campo Científico</i>
Isabel Villaseñor Rodríguez, Juan Antonio Gómez García	<i>Una forma posible de Interdisciplinariedad: La Documentación Iusfilosófica</i>
Renato Rocha Souza, Maurício Barcellos Almeida	<i>Representação do conhecimento: identidade ou esvaziamento da Ciência da informação?</i>
Ida Regina C. Stumpf	<i>A Ciência da Informação no Brasil através de seus Programas de Pós-graduação</i>
Leilah Santiago Bufrem, Aline Elis Arboit, Juliana Lazzaroto de Freitas	<i>Configuração epistemológica da Ciência da Informação na literatura periódica brasileira por meio de análise de citações (1972-2008)</i>
Miguel Ángel Rendón Rojas, Silvana Elisa Cruz Domínguez	<i>La Ciencia de la Información comprendida como un Sistema de Información y Documentación: la aplicación del paradigma emergente</i>
Ana Cristina Albuquerque	<i>Teoria da classificação e documentos fotográficos em Arquivos, Bibliotecas e Museus</i>
Eliany Alvarenga de Araújo, Marcos Antonio Alexandre Bezerra	<i>A Ética no contexto epistemológico da Ciência da Informação</i>
Lucinéia Maria Bicalho	<i>Interações disciplinares da pesquisa na Ciência da Informação</i>

Nota. Artigos listados pela ordem com que surgem na obra em epígrafe.

Apêndice 2

Confrontação contextual e funcional entre o Repositório Científico de Acesso Aberto de Portugal (RCAAP) e o Portal Brasileiro de Acesso Aberto à Produção Científica (OASISbr)

RCAAP	OASISbr		
<i>Início de atividade</i>			
2008, dezembro.	2006.		
<i>Enquadramento normativo local</i>			
<u>Decreto-lei nº 115/2013</u> Obriga ao depósito de uma cópia digital de todas as dissertações de mestrado e teses de doutoramento num repositório integrado na rede do RCAAP, no prazo máximo de 60 dias após a conceção dos respetivos graus (pontos 1 e 5 do artigo 50º).	<u>Projeto de Lei 1120/2007</u> Obriga as instituições públicas de ensino superior a construir os repositórios institucionais para depósito do inteiro teor da produção técnico-científica do corpo discente e docente.		
<u>Portaria nº 285/2015</u> Alarga o prazo para colocação no RCAAP dos trabalhos referentes aos graus de Mestre e de Doutor, conferidos entre agosto de 2013 e setembro de 2015, para, respetivamente, 31 de dezembro de 2016 e 31 de julho de 2016 (artigos 4º e 5º).	<u>Projeto de Lei do Senado 387/2011</u> Estabelece a obrigatoriedade das instituições de ensino superior e das unidades de pesquisa de caráter público elaborar os seus repositórios institucionais com a produção técnico-científica de seus pesquisadores, professores e alunos de mestrado, doutorado e pós-doutorado.		
<i>Software de suporte</i>			
Plataforma de pesquisa de código aberto <i>Apache Solr</i> .	Portal <i>LA Referencia</i> que tem como motor de base de dados o <i>software</i> de código aberto <i>PostgreSQ</i> .		
<i>Quantidade de campos de “pesquisa avançada” passíveis de serem usados simultaneamente</i>			
Três.	Número virtualmente ilimitado de grupos de pesquisa com um igual número “ilimitado” de campos cada.		
<i>Correspondência entre os campos de “pesquisa avançada” e os elementos das DM e TD^a</i>			
campo	elemento	campo	elemento
a) Título; b) Autor; c) Assunto; d) Descrição; e) Texto integral.	a) Título; b) Autor; c) Palavras-chave; d) Resumo; e) Todo documento	a) Título; b) Autor; c) Assunto; e) Ano da publicação; f) Recurso; f) Todos os campos.	a) Título; b) Autor; c) Palavras-chave; d) Ano da defesa; e) Sem correspondência; f) Todos os elementos com correspondência.
<i>Designação dos operadores booleanos entre campos de “pesquisa avançada”</i>			
E; OU; NÃO.		TODOS os termos; QUALQUER termo; NENHUM termo.	
<i>Filtros adicionais de refinamento prévio de pesquisas</i>			
Tipo de documento; Idioma; Tipo de acesso; Recurso; Data; Financiamento.		Tipo de documento; Idioma; Recursos; Ano da publicação; Rede.	
<i>Filtros adicionais de refinamento pós pesquisa</i>			
Autor; Assunto; Tipo de documento; Tipo de acesso; Data; Recurso.		Autor; Assunto; Tipo de documento; Tipo de acesso; Ano da publicação; Fonte; Rede; Idioma.	
<i>Filtros de ordenação dos resultados da pesquisa</i>			
Ordenar: Relevância ou Data; ordem: Descendente ou Ascendente.		Ordenar: Relevância; Data Descendente; Data Ascendente; Autor; Título.	

Nota. As correspondências descritas entre os campos de pesquisa e os elementos dos trabalhos foram esclarecidos via comunicação pessoal com as equipas de suporte dos respetivos repositórios.

^a DM: Dissertações de Mestrado; TD: Teses de Doutoramento.

Apêndice 3a

Folha de Recolha de Dados geral (exemplificação de alguns registros)

N.tbs	1.título	2.autoria	3.curso	4.instituição
1	Web semântica: uma análise focada no uso de metadados	Alves RCV	Ciência da Informação	UNESP/MAR- Universidade Estadual Paulista "Júlio de Mesquita Filho"
2	Hyperde: um framework e ambiente de desenvolvimento dirigido por ontologias para aplicações hipermídia	Nunes DA	Informática	Pontfícia Universidade Católica do Rio de Janeiro
3	Ontologias no suporte a portais semânticos	Jorge MA	Ciência da Informação	UFMG- Universidade Federal de Minas Gerais
4	Síntese de escalas utilizando semântica de passo máximo em redes de petri com temporização	--	--	--
5	Arquitetura para adaptação e meta-adaptação de sistemas hipermídia	Assis PSd	Informática	Pontfícia Universidade Católica do Rio de Janeiro
6	Web semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação	Ramalho RAS	Ciência da Informação	UNESP/MAR- Universidade Estadual Paulista "Júlio de Mesquita Filho"
7	Geração semi-automática de extratores de dados da web considerando contextos fracos	--	--	--
8	Uma arquitetura para sistemas tutores inteligentes apoiada por fundamentos de web semântica	Jacinto AdS	Engenharia Eletrônica e Computação	Instituto Tecnológico de Aeronáutica
(...)	(...)	(...)	(...)	(...)
251	Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática: um estudo crítico	Mota DAR	Ciência da Informação	USP- Universidade de São Paulo

5.ano	6.Tp	7.R	8.ci	9.País	10.variações	11. URL	12. observações (* tit.do duplic & « tiit.de recuperação qd <>s)	nP
2005	dm	OR	s	Br	=	http://repositorio.unesp.br/handle/11449/93690	* Web semântica: uma análise focada no uso de metadados /	1
2005	dm	OR	n	Br	=	http://www.maxwell.vrac.puc-rio.br/Busca_etds.php?strSecao=resultado&nrSeq=7617@1		2
2005	dm	O	s	Br	Portais semânticos	http://hdl.handle.net/1843/VALA-6KFNS2	não recuperado no RCAAP mesmo s/filtros temporais ou temáticos	3
2005	dm	OR	nn	Br	Utilizando semântica	--		4
2005	td	OR	n	Br	=	http://tede.ufam.edu.br/handle/tede/2938	« An architecture for adaptation and meta-adaptation in hypermedia systems	1
2006	dm	OR	s	Br	=	http://hdl.handle.net/11449/93709	* Web semântica : aspectos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação /	1
2006	dm	OR	nn	Br	Semi-automática	-		2
2006	dm	O	n	Br	=	http://www.bd.bibl.ita.br/tde_busca/arquivo.php?codArquivo=330	não recuperado no RCAAP mesmo s/filtros temporais ou temáticos	3
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
2015	td	O	s	Br	=	http://www.teses.usp.br/teses/disponiveis/27/27151/tde-27012016-135403/pt-br.php	não recuperado no RCAAP mesmo s/filtros temporais ou temáticos	2

Nota. No caso dos falsos positivos detetados (como os registros 4 e 7) alguns campos não foram preenchidos.

Apêndice 3b

Folha de Recolha de Dados específica: folha 1E [campos 1 a 9 de 15]

Nts	Ref.	1.título	2.autoria	3.ano	4.t	5.curso	6.IES	7.dep./esc.	8.Orientador	9.O-filiação
1	dm.01.2005	Web semântica: uma análise focada no uso de metadados	Rachel Alves	2005	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Plácida Leopoldina V. A. Costa Santos	Docente permanente do Programa de Pós-Graduação em CI da FFC/UNESP, na linha de pesquisa Informação e Tecnologia
3	dm.02.2005	Ontologias no suporte a portais semânticos	Marcelo Jorge	2005	dm	Ciência da Informação	UFMG	Escola de CI	Marcello Peixoto Bax	Professor Titular do Programa de Pós Graduação em CI da Escola de Ciência da Informação da UFMG
6	dm.03.2006	Web semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação	Rogério Ramalho	2006	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Silvana Aparecida B. G. Vidotti	Coordenadora do Programa de Pós-Graduação em Ciência da Informação da Unesp (2004 - 2011)
32	dm.04.2008	A construção do conhecimento científico: a web semântica como objeto de estudo	Cintia Pinheiro	2008	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Edberto Ferneda	professor do Departamento de CI da UNESP/Mar. Atua principalmente nas áreas de Indexação Automática e Recuperação de Informação.
40	dm.05.2008	Inserção da biblioteca digital de teses e dissertações no contexto da web semântica	Sinara Molossi	2008	dm	Ciência da Informação	UFSC	Centro de Ciências da Educação	Lígia Café	professor associado II da UFSC. Líder do Grupo de Pesquisa Núcleo de Estudos em Informação e Mediações Comunicacionais Contemporâneas (NEIMCOC)
46	dm.06.2008	Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da ciência da informação: uma abordagem do marcont initiative na era da web semântica	Fabiano Castro	2008	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Plácida Leopoldina V. A. Costa Santos	mesmo que dm.01.2005
72	dm.07.2009	Perspectivas na Web Semântica para a Ciência da Informação	Rodrigo Carvalho	2009	dm	Ciência da Informação	PUC CAMP	Centro de C.Humanas e Sociais Aplicadas	Mariângela Pisoni Zanaga	professor titular na Faculdade de Biblioteconomia da Pontifícia Universidade Católica de Campinas e no curso de pós-graduação em CI
98	td.01.2006	Autoria de documentos para a web semântica : um ambiente de produção de conhecimento baseado em ontologias	Edgard Oliveira	2006 (recup 2010)	td	Ciência da Informação	UNB	Fac.de Economia, Adm., Contab. e CI	Mamede Lima-Marques	Professor Titular da Universidade de Brasília, Faculdade de Ciência da Informação. Fundador e diretor do Centro de Pesquisa em Arquitetura da Informação - CPAI (FCI/UnB)
101	td.02.2010	Representação iterativa: um modelo para repositórios digitais	José Segundo	2010	td	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Silvana Aparecida B. G. Vidotti	mesmo que dm.03.2006
108	dm.09.2011	Estudo exploratório sobre compatibilização semântica devocabulário utilizando ontologia: fundamentação teórica-metodológica visando ao aperfeiçoamento do mhtx	Eliana Demarques	2011	dm	Ciência da Informação	UFMG	Escola de CI	Gercina Ângela Borémde Oliveira Lima	Professora associada da Escola de CI da UFMG
117	dm.10.2011	Registros de marcas da junta comercial do estado de pernambuco: um olhar semântico para a organização de um acervo histórico	Ana Araújo	2011	dm	Ciência da Informação	UFPE	Centro de Artes e Comunicação	Fábio Mascarenhas e Silva	Atualmente é professor adjunto da Universidade Federal de Pernambuco e docente permanente do Programa de Pós-graduação em Ciência da Informação da UFPE (PPGCI/UFPE)
133	dm.11.2012	Artigos científicos digitais na web: novas experiências para apresentação, acesso e leitura	Déborah Ambinder	2012	dm	Ciência da Informação	UFF	Istituto de Arte e Comunicação	Carlos Henrique Marcondes	Prof.Titular do Dep.de CI e do PPGCI da UFF e pesquisador 1D do Conselho Nacional de Desenvolvimento Científico e Tecnológico

Folha de Recolha de Dados específica: folha 1D [campos 10 a 15 de 15]

Ref.	10.O-pD/D	11.O-M/Esp	12.O-L/B	13.Coorient.	14.C-filiação	15.C-formação
dm.01.2005	em Semiótica e Lingüística Geral pela FFLCH/USP (1994) e livre-docente em Catalogação pela UNESP (2010)	mestre em Ciência da Informação pela PUC de Campinas (1983)	Bacharel em Biblioteconomia pela UNESP (1980)	n.a	n.a	n.a
dm.02.2005	em Ciência da Computação pela Universidade de Montpellier II, França.	DEA em Matemática e Computação e DESS em Informática pela Université d'Aix Marseille II, França	Bacharel em Ciência da Computação (PUC-MG)	n.a	n.a	n.a
dm.03.2006	em Educação - área de concentração Educação Brasileira - pela Faculdade de Filosofia e Ciências da UNESP (2001)	Mestre em Ciências - área de concentração - Ciências da Computação e Matemática Computacional - pelo Instituto de Ciências Matemáticas de São Carlos da USP (1993)	Licenciada em Matemática UNESP (1986). Especialista em Ciência da Computação pelo Instituto de Ciências Matemáticas de São Carlos da USP (1987)	Mariângela Spotti Lopes Fujita	Professora Titular do Departamento de Ciência da Informação da UNESP na linha de pesquisa "Produção e Organização da Informação" com enfoque na Indexação e Linguagens de Indexação	Dout.Ciências da Comunicação pela USP, 1992 e LDoc.em Análise Documentária e Linguagens Documentária Alfabéticas pela UNESP, 2003 Mestre em Ciências da Comunicação em 1986 pela USP Esp. em Curso de Documentação Científica IBICT, 1978 e em 1977 Graduação em Biblioteconomia.EBDSC (S.Carlos)
dm.04.2008	em Ciências da Comunicação pela USP, 2003. Pós-doutorado pela UFP, 2013	Mestre em Informática pela Universidade Federal da Paraíba (1997)	graduação em Processamento de Dados pela antiga Fundação Educacional de Bauru (1985)	n.a	n.a	n.a
dm.05.2008	em Lingüística - Université Laval Canadá (1999)	mestrado em Biblioteconomia e Documentação pela UnB (1988)	graduação em Biblioteconomia e Documentação pela UnB (1984)	n.a	n.a	n.a
dm.06.2008	mesmo que dm.01.2005	mesmo que dm.01.2005	mesmo que dm.01.2005	n.a	n.a	n.a
dm.07.2009	em Educação, na área de Educação, Ciência e Tecnologia pela Universidade Estadual de Campinas (2006)	mestre em Biblioteconomia, área de Planejamento e Administração de Sistemas de Informação (PUCCam, 1994)	bacharelado em Biblioteconomia pelaPUCCam (1973)	n.a	n.a	n.a
td.01.2006	em Ciência da Computação pela Université Toulouse III Paul Sabatier, França	Esp.em Ciência da Computação - IC/UNICAMP, DEA en Informatique e Doutorado	Eletrônica no Centro Federal de Educação Tecnológica de Minas Gerais; graduação em Filosofia pela PUCCam	Frank van Harmelen	Professor da Vrije Universiteit	PhD in 1989 for his research on meta-level reasoning The University of Edinburgh Mathematics and Computer Science, University of Amsterdam M.Sc. 1978 – 1983 Ciência da Computação: AI (KR, SW)
td.02.2010	mesmo que dm.03.2006	mesmo que dm.03.2006	mesmo que dm.03.2006	n.a	n.a	n.a
dm.09.2011	CI pela UFMG, 2004. Pós-doc.em Estudos sobre metodologias para o proc.de index. de docs textuais, pelo Dep.de Biblioteconomia y Documentación, Fac.de Humanidades, Comunicación y Documentación, Univ.Carlos III, Madrid (2014), e Pós-doc pela ECA da USP 2014	mestrado em Science in Library Service - Clark Atlanta University (1993)	graduação em Biblioteconomia pela Universidade Federal de Minas Gerais (1985)	n.a	n.a	n.a
dm.10.2011	em Ciência da Informação (Universidade de São Paulo - USP (2008)	Mestrado em Ciência da Informação. Pontifícia Universidade Católica de Campinas, PUC Campinas, 2003	Especialização em Informação Tecnológica. (Carga Horária: 365h) e graduação em Biblioteconomia, UFPE 1998 e 97	n.a	n.a	n.a
dm.11.2012	em Ciência da Informação pela Universidade Federal do Rio de Janeiro (1998)	mestrado em CI pela Universidade Federal do Rio de Janeiro (1992)	graduação em Arquitetura e Urbanismo pela Universidade Federal Fluminense (1978)	n.a	n.a	n.a

Folha de Recolha de Dados específica: folha 2E [campos 1 a 9 de 15]

Nts	Ref.	1.título	2.autoria	3.ano	4.t	5.curso	6.IES	7.dep./esc.	8.Orientador	9.O-filiação
134	dm.12 .2012	As dimensões semântica e pragmática da web e dos mecanismos de busca no ciberespaço	Ilza Andrade	2012	dm	Gestão da Informação	UEL	Centro de Educação, Comunicação e Artes	Silvana Drumond Monteiro	professora Associada da Un Est de Londrina, Departamento de CI, atuando na graduação e no mestrado acadêmico em CI
153	dm.13 .2012	Organização em sistema caótico: uso das tags para classificação da informação pelos usuários da rede social twitter	Amanda Sousa	2012	dm	Ciência da Informação	UFRJ/IBICT	Fac.de Adm.e Ciências Contábeis	Rosali Fernandez de Souza	pesquisadora titular do IBICT / MCT. Área: CI; subárea: organização e representação do conhecimento
181	dm.14 .2013	Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária	Francisco Gomes	2013	dm	Ciência da Informação	UFRJ/IBICT	Escola de Comunicação	Marcos do Couto Bezerra Cavalcanti	professor do Programa de Engenharia de Produção da COPPE/UFRJ. Coordenador do Crie (Centro de Referência em Inteligência Empresarial)
182	dm.15 .2013	As tecnologias da web semântica no domínio bibliográfico	Renata Silva	2013	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Plácida Leopoldina V. A. Costa Santos	mesmo que dm.01.2005
198	dm.16 .2013	Representação da informação dinâmica em ambientes digitais	Camila Ribeiro	2013	dm	Ciência da Informação	USP	Escola de Comunic.e Artes	Marcos Luiz Mucheroni	Ciência da Computação, com ênfase em Arquitetura de Sistemas de Computação (USP-ECA)
204	dm.08 .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	Maria Menezes	2010 (recup 2013)	dm	Ciência da Informação	UFBA	Instituto de CI	Zeny Duarte	Professora da Ufba, atual (3-2016) coordenadora do mestrado e doutorado em ciência da informação PPGCI/Ufba
213	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	Fernanda Monteiro	2013	td	Ciência da Informação	UNB	Fac.de Economia, Adm., Contab. e CI	Cláudio Gottschalg Duque	coord.do grupo de pesquisa "Research Expert Group for Intelligent Information in Multimodal Environment using Natural language Technologies and Ontologies" e coord. do Doutorado em CI UnB/UFES (DINTER), professor Adjunto da Faculdade de CI da U.de Brasília (FCI-UnB) e do PPG em CI (PPGCINF-UNB).
236	dm.17 .2014	Wikis semânticos: da web para a web semântica	Jaidier Ferreira	2014	dm	Ciência da Informação	UNESP/MAR	Fac.de Filosofia e Ciências	Plácida Leopoldina V. A. Costa Santos	mesmo que dm.01.2005
250	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	Julietti Andrade	2015	td	Ciência da Informação	USP	Escola de Comunic.e Artes	Nair Yumiko Kobashi	Professora livre-docente (Área: Análise documentária) da USP. Coordenadora adjunta da área Ciências Sociais Aplicadas I (CAPES),
251	td.05. 2015	Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática: um estudo crítico	Denyson Mota	2015	td	Ciência da Informação	USP	Escola de Comunic.e Artes	Marilda Lopes Ginez de Lara	professor livre-docente da USP e ministra aulas na Graduação e na Pós-Graduação A partir de 2010 é vice-coordenadora do PPG em CI da ECA

Folha de Recolha de Dados específica: folha 2D [campos 10 a 15 de 15]

Ref.	10.O-pD/D	11.O-M/Esp	12.O-L/B	13.Coorient.	14.C-filiação	15.C-formação
dm.12.2012	pós-doutorado em CI pela UFMG, doutorado em Comunicação e Semiótica pela PUCSP (2003)	mestrado em Ciência da Informação pela PUCCamp (1996)	graduação em Biblioteconomia pela Universidade Estadual de Londrina (1984).	n.a	n.a	n.a
dm.13.2012	realizado na Polytechnic of North London, conferido pelo Council for National Academic Awards (1984)	especialização em Documentação Científica (1969) e mestrado em Ciência da Informação (1973) pelo IBICT	graduação em Biblioteconomia e Documentação pela Universidade Santa Úrsula (1968)	n.a	n.a	n.a
dm.14.2013	em Informática - Université de Paris XI (Paris-Sud) (1993)	mestrado em Informática - Université de Paris XI (Paris-Sud) (1989)	graduação em Matemática pela Universidade Federal do Rio de Janeiro (1982)	n.a	n.a	n.a
dm.15.2013	mesmo que dm.01.2005	mesmo que dm.01.2005	mesmo que dm.01.2005	n.a	n.a	n.a
dm.16.2013	em Engenharia Elétrica pela Universidade de São Paulo (1996)	mestrado em Engenharia Mecânica pela Universidade de São Paulo (1988)	graduação em Ciência da Computação pela Universidade Federal de São Carlos (1980)	n.a	n.a	n.a
dm.08.2010	em Letras da UNova Lx1998 e pós-doc em Arquivologia Novas Tecnologias da Informação Uporto 2012	espec. em: restauration des photographies, Musée Carnavalet, 1993; arquivística internacional, Archives Nationales de France, 1990; arquivologia, Ufba, 1990; documentação científica, UFRJ e Ibict, 1983	Graduada em biblioteconomia e documentação, Ufba 1981	n.a	n.a	n.a
td.03.2013	em Linguística Computacional no Angewandte Sprachwissenschaft und Computerlinguist - Justus-Liebig-Universität Giessen (2004) e em Produção e Gestão da Informação no PPG em CI pela Escola de CI (2005)	mestrado em Psicolinguística no Programa de Pós-Graduação em Estudos Lingüísticos pela Faculdade de Letras da Universidade Federal de Minas Gerais (1998)	graduação em Letras habilitação em Português e Alemão pela Faculdade de Letras da Universidade Federal de Minas Gerais (1994)	n.a	n.a	n.a
dm.17.2014	mesmo que dm.01.2005	mesmo que dm.01.2005	mesmo que dm.01.2005	n.a	n.a	n.a
td.04.2015	em Ciências da Comunicação pela Universidade de São Paulo (1994). Realizou estágio de pesquisa na École des Hautes Études en Sciences Sociales, em 1991	mestre em Ciências da Comunicação pela Universidade de São Paulo (1988)	bacharel em Jornalismo pela Universidade de São Paulo (1980), bacharel em Biblioteconomia pela USP (1978)	n.a	n.a	n.a
td.05.2015	em Ciências da Comunicação pela Universidade de São Paulo (1999) e pós-doutorado pela Universidad Carlos III de Madrid	mestrado em Ciências da Comunicação pela Universidade de São Paulo (1993)	graduação em Curso de Biblioteconomia e Documentação pela Fundação Escola de Sociologia e Política de São Paulo (1973)	n.a	n.a	n.a

Apêndice 4

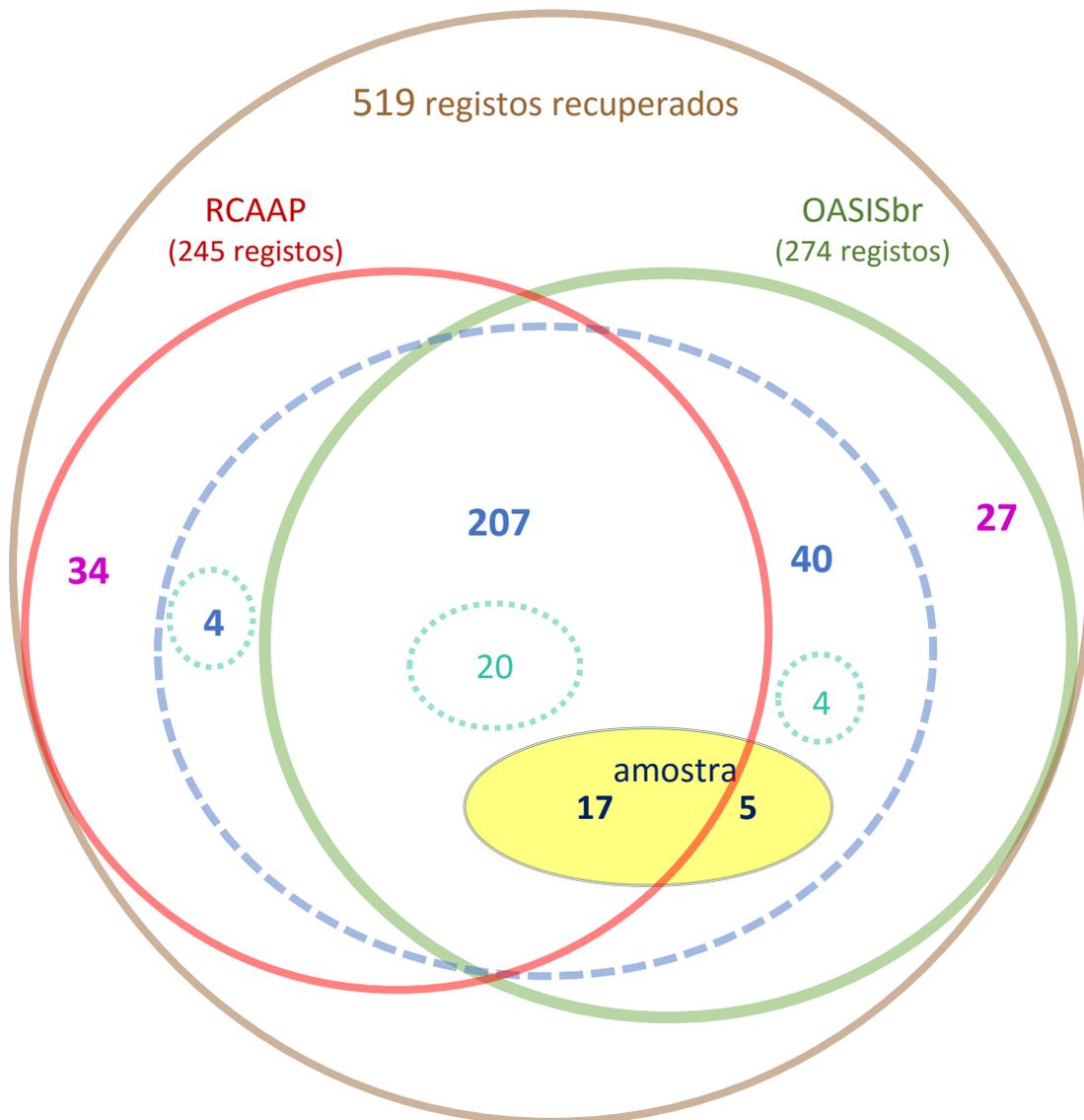
Lista dos 14 termos recolhidos relativos à Web Semântica e confrontação entre as fontes usadas para a recolha

termos		Class (aka: Kind)/(Defined/Anonymous) Class	GRDDL (Gleaning Resource Descriptions from Dialects of Languages)	Inferred hierarchy (vs.Asserted or Manual hierarchy)/Inference/Sistemas de Classificação Inferenciais/Inferência	Linked data/Open Data/Linked Open Data/dados abertos interligados	Ontology/Ontologias	OWL (Web Ontology Language)	Property (aka: Predicate)	RDF (Resource Description Framework)	RDFa (Resource Description Framework in attributes)	Reasoner (aka: semantic reasoner, reasoning engine, rules engine)/Automatização do	Relation/Relationships among those things/entidades e relações	Semantic Web/Web semântica	SKOS (Simple Knowledge Organisation System)	SPARQL (SPARQL Query Language for RDF)
fontes															
WS glossários	2011, W3C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2013, Base22	1		1		1	1	1			1	1	1		1
	2014, OSF	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2014, ws.org.br		1		1				1	1			1	1	
Cursos (refs. Tabelas 11 e 13)	Pt01.M				1								1		
	Pt02.M			2		4	2				1		2		
	Br01.MD				2	7	2		3		1		5	2	
	Br02.MD				2	4							3		
	Br04.MD					1							1		
	Br06.Ma					1									
	Br07.MD				2	5							3		
	Br08.MD				1	2							2		
	Br09.MD					2									
	Br11.MD					6						1			
Br15.Mp			1		1										
sommas	WS [A]	3	3	3	3	3	3	3	3	3	3	3	4	3	3
	CI [B]	0	0	2	5	10	2	0	1	0	2	1	7	1	0
	CI [C]	0	0	3	8	33	4	0	3	0	2	1	17	2	0
	[A]+[B]	3	3	5	8	13	5	3	4	3	5	4	11	4	3
	[A]+[C]	3	3	6	11	36	7	3	6	3	5	4	21	5	3

Nota. [A]: quantidade de ocorrências encontradas nos glossários/listas terminológicas; [B]: quantidade de ocorrências encontradas por curso, i.e., cada termo foi contabilizado uma única vez por curso, independentemente das vezes que ocorreram no mesmo curso; [C] quantidade de ocorrências encontradas nos programas, conteúdos ou ementas das disciplinas e linhas de pesquisa, todas as ocorrências foram contabilizadas.

Apêndice 5

Esquema dos registos que constituem o universo e a amostra do presente trabalho e sua relação com os respetivos repositórios onde foram recuperados



	RCAAP	Comuns aos dois repositórios	OASISbr	total
registos únicos	4	207	40	251
falsos positivos		20	4	28
registos duplicados	34	-	27	61
amostra	-	17	5	22

Apêndice 6

Distribuição dos cursos e respetiva quantidade de trabalhos, correspondentes aos 223 registos recuperados nos portais RCAAP e OASISbr, pelos quatro grupos de áreas concebidos

Designação dos cursos (Portugal)	Σ ·Pt	Designação dos cursos (Brasil)	Σ ·Br
<i>Grupo A: Portugal – 0 cursos; Brasil – 2 cursos</i>			
[s/representantes]		Ciência da Informação	21
		Gestão da Informação ^a	1
<i>Grupo B: Portugal – 9 cursos; Brasil – 7 cursos (total – 15 cursos)</i>			
Engenharia de Software	2	Ciência da Computação	57
Engenharia e Gestão de Sistemas de Informação	1	Ciências da Computação e Matemática Computacional	11
Engenharia Informática	14	Computação	6
Engenharia Informática e Computação	3	Computação Aplicada	3
Engenharia Informática e de Computadores	1	Informática	30
Informática	4	Informática Aplicada	3
Lógica Computacional	2	Sistemas e Computação	3
Tecnologias da Informação Aplicadas às Ciências Biológicas e Médicas	1		
Tecnologias e Sistemas de Informação	2		
<i>Grupo C: Portugal – 3 cursos; Brasil – 9 cursos (total – 12 cursos)</i>			
Engenharia de Computadores e Telemática	2	Engenharia de Computação	2
Engenharia Eletrotécnica e de Computadores	7	Engenharia e Gestão do Conhecimento	3
Tecnologias Geoespaciais	6	Engenharia Elétrica	11
		Engenharia Elétrica e Informática Industrial	2
		Engenharia Eletrônica e Computação	10
		Gestão de Informação e Comunicação em Saúde	1
		Gestão do Conhecimento e Tecnologias da Informação	1
		Modelagem Computacional	1
		Modelagem Computacional de Conhecimento	1
<i>Grupo D: Portugal – 4 cursos; Brasil – 7 cursos (total – 11 cursos)</i>			
Bioengenharia	1	Biologia Computacional e Sistemas	1
Gestão da Informação ^a	1	Design	1
Medicina Dentária	1	Engenharia Mecânica e de Materiais	1
Pedagogia do E-Learning	1	Gestão Empresarial	1
		Linguística e Língua Portuguesa	1
		Psicologia	1
		Tecnologia	1
<i>Soma total de cursos: 41 (Portugal – 16 cursos; Brasil – 25 cursos)</i>			

Nota. O número de registos recuperados referente a cada curso é apresentado na respetiva linha. Σ ·Pt: somas relativas aos trabalhos de cursos de Portugal; Σ ·Br: somas relativas aos trabalhos de cursos do Brasil.

^a Apesar da mesma designação, Gestão da Informação, o curso do Brasil está inserido no grupo A e o de Portugal no D, de acordo com as classificações usadas na CAPES e na DGEEC, respetivamente.

Apêndice 7

Resumos e Objetivos (gerais e específicos) dos 22 trabalhos que constituem a amostra

dm.01.2005 *Web semântica: uma análise focada no uso de metadados*

RESUMO: Atualmente a nossa sociedade, denominada sociedade da informação, vem sendo caracterizada pela valorização da informação, pelo uso cada vez maior de tecnologias de informação e comunicação e pelo crescimento exponencial dos recursos informacionais disponibilizados em diversos ambientes, principalmente na Web. Essa realidade trouxe algumas mudanças no acesso automatizado às informações. Se por um lado temos uma grande quantidade de recursos informacionais disponibilizados, por outro temos como consequência problemas relacionados à busca, localização, acesso e recuperação dessas informações em ambientes digitais. Nesse contexto, o problema que originou essa pesquisa está relacionado com a dificuldade na busca e na recuperação de recursos informacionais digitais na Web e a ausência de tratamento adequado para a representação informacional desses recursos. O maior desafio para a comunidade científica no momento está na identificação de padrões e métodos de representação da informação, ou seja, na construção de formas de representação do recurso informacional de maneira a proporcionar sua busca e recuperação de modo mais eficiente. Assim, a proposição apontada nesse trabalho como solução do problema refere-se ao estabelecimento da Web Semântica e a aplicação de padrões de metadados para a representação da informação, pois são consideradas como iniciativas importantes para proporcionar uma melhor estruturação e representação dos recursos informacionais em ambientes digitais. Com uma metodologia baseada na análise exploratória e descritiva do tema a partir da literatura disponível, apresenta-se uma análise da Web Semântica como uma nova proposta para organização dos recursos informacionais na Web e as ferramentas tecnológicas que permeiam sua construção, com enfoque no uso de metadados como elemento fundamental para proporcionar uma melhor representação dos recursos informacionais disponibilizados na Web e sua posterior recuperação. A proposta da Web Semântica é disponibilizar recursos informacionais melhor estruturados e representados, formando uma rede de informações conectadas que por meio de ferramentas tecnológicas, tais como: os agentes de software, a linguagem de marcação XML, arquitetura de metadados RDF, ontologias e, principalmente, padrões ou formatos e metadados. Como resultado pode-se destacar que a implementação da Web Semântica requer o trabalho conjunto das várias ferramentas tecnológicas estudadas e que proporcionará em pequena, média e grande escala a tão necessária estruturação e representação informacional dos recursos e consequentemente sua melhor recuperação. Além disso, **foi possível verificar que as tecnologias da Web Semântica convergem para a área de Ciência da Informação, estabelecendo uma estreita relação na questão da representação do conhecimento, principalmente com relação ao uso de metadados** que são considerados essenciais para se estabelecer uma boa representação dos recursos informacionais na rede. Sabendo que a representação da informação é necessária em qualquer ambiente para proporcionar uma recuperação mais eficiente, podemos considerar os metadados como ferramentas essenciais para estabelecer a representação dos recursos informacionais no ambiente da Web Semântica como instrumentos para a construção de uma rede de conhecimentos e recuperação da informação de modo mais eficiente.

OBJETIVO GERAL: Estudar os aspectos que permeiam a construção da **Web Semântica** em relação com a Ciência da Informação, bem como o uso e aplicação de metadados para a representação de recursos informacionais, já que são considerados fatores principais para a representação da informação na atualidade.

OBJETIVOS ESPECÍFICOS: a) Analisar as ferramentas tecnológicas que compõe a **Web Semântica**, destacando para a questão do uso das ontologias como responsáveis por garantir a semântica dos dados; b) Identificar as características da linguagem de marcação XML – eXtensible Markup Language, e sua importância para a estruturação dos recursos informacionais na **Web Semântica**; c) Analisar o uso da arquitetura de metadados, em especial a arquitetura RDF – Resource Description Framework, para o estabelecimento da interoperabilidade sintática, estrutural e semântica. d) Analisar as características, o uso e a aplicação dos diferentes tipos de metadados e seus formatos para o estabelecimento da representação de recursos informacionais e destacá-los como ferramentas essenciais para a construção da **Web Semântica**.

dm.02.2005 *Ontologias no suporte a portais semânticos*

RESUMO: Portais semânticos aparecem como uma evolução natural dos tradicionais portais web e podem ser entendidos como portais de informações que utilizam os padrões de representação de informação propostos pelo W3C para a chamada Web Semântica. As ontologias são a base desses novos portais web, propiciando melhor organização das informações, além de outras vantagens. O objetivo deste estudo é compreender como o uso de ontologias promove a organização da informação e agrega valor a esses portais. Para atingir este objetivo, o trabalho investiga as principais ferramentas semânticas baseadas em ontologias existentes na literatura científica e elabora uma ontologia que serve de sustentação para a proposição de um portal semântico a ser utilizado por uma comunidade universitária. Após uma pré-especificação do portal semântico, sua arquitetura funcional é sugerida e as funções semânticas são apresentadas e detalhadas através do uso da ontologia construída anteriormente. As ferramentas semânticas investigadas são utilizadas no contexto de especificação do portal semântico proposto e exemplos ilustrativos, baseados na ontologia elaborada, são extraídos. Com o trabalho espera-se contribuir para a discussão acerca dos benefícios que uma ontologia pode trazer para a organização de conceitos, promovendo a reutilização e o compartilhamento de informações em um determinado domínio do conhecimento.

OBJETIVO GERAL: Investigar e compreender, na prática, se o uso de uma ontologia, cuja construção também é proposta, apóia a organização da informação em portais (hipótese da pesquisa) e de que forma se dá tal apoio.

OBJETIVOS ESPECÍFICO: Determinar as principais vantagens e dificuldades do uso da ontologia em portais.

dm.03.2006 *Web semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação*

RESUMO: No âmbito da gestão de recursos informacionais os modelos e métodos de organização e recuperação de informações sempre estiveram condicionados às tecnologias utilizadas, de modo que com desenvolvimento e intensificação da utilização das tecnologias digitais uma nova gama de possibilidades vem sendo incorporada aos processos de produção, armazenamento, representação e recuperação de informações, atingindo um estágio em que os modelos clássicos de organização e recuperação de informações precisam ser (re)pensados sob diferentes perspectivas, pois os mesmos não parecem ser capazes de solucionar os

problemas identificados no ambiente Web, evidenciando a necessidade de desenvolvimento de novas tecnologias que permitam otimizar a recuperação de informações em ambientes digitais. Nesse sentido, os estudos relacionados ao projeto Web Semântica vêm destacando-se como uma nova perspectiva no desenvolvimento de tecnologias que possibilitem um aumento na qualidade e relevância das informações recuperadas, a partir do desenvolvimento de instrumentos que permitam descrever formalmente, em um formato que possa ser processado por máquinas, os aspectos semânticos inerentes aos recursos informacionais, contribuindo para a identificação e contextualização das informações disponíveis no ambiente Web. Deste modo, **a proposição deste trabalho é a realização de um estudo teórico e metodológico de caráter interdisciplinar acerca do projeto Web Semântica, buscando favorecer a desmistificação dos conceitos e tecnologias subjacentes e avaliar em que medida a área de Ciência da Informação pode contribuir para sua concretização, ressaltando os possíveis reflexos destas novas abordagens tecnológicas em seu corpus teórico.** Assim, apresenta-se um levantamento bibliográfico acerca do desenvolvimento da Internet e os principais conceitos e tecnologias relacionados ao projeto Web Semântica, contextualizando-os a partir dos aportes teóricos da área de Ciência da Informação. **Verificou-se que os conceitos e tecnologias subjacentes ao projeto Web Semântica podem ser considerados como uma renovação ou desdobramento dos métodos convencionais de representação, organização e recuperação de informações, apontando a possibilidade de contribuições da área de Ciência da Informação para o desenvolvimento do projeto Web Semântica,** devido ao seu embasamento teórico referente a formas de representação e as práticas profissionais identificadas em seu campo de atuação, **sendo evidente o caráter interdisciplinar que delineia o corpus teórico dos estudos relacionados ao projeto Web Semântica, apresentando-se como um campo fértil para pesquisas e indicando a possibilidade de desenvolvimento de novos métodos de organização e recuperação de informações, construídos por meio de esforços interdisciplinares que favoreçam a junção da fundamentação teórica inerente à área de Ciência da Informação com as novas tecnologias emergentes da área de Ciência da Computação.**

OBJETIVO GERAL: Contribuir para uma “desmistificação” dos conceitos e tecnologias subjacentes ao projeto **Web Semântica**, com o intuito de contribuir para um melhor entendimento das tecnologias relacionados.

OBJETIVOS ESPECÍFICOS: a) Identificando pontos de convergência entre as novas abordagens computacionais e os métodos de análise, síntese, organização e representação de recursos informacionais utilizados tradicionalmente no âmbito da área de Ciência da Informação; b) Indicando possibilidades de desenvolvimento de novos métodos de organização e recuperação de informações, que favoreçam a junção da fundamentação teórica inerente à área de Ciência da Informação com as novas tecnologias emergentes da área de Ciência da Computação.

dm.04.2008 A construção do conhecimento científico: a web semântica como objeto de estudo

RESUMO: 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. 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 e nas dissertações e teses defendidas sobre Web Semântica e suas tecnologias. 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.**

OBJETIVO GERAL: Refletir sobre a participação e integração da Ciência da Informação e da Ciência da Computação na construção e desenvolvimento do conhecimento científico sobre **Web Semântica**.

OBJETIVOS ESPECÍFICOS: a) Identificar a produção científica em **Web Semântica** disponível nas principais iniciativas brasileiras de bancos e bases de dados científicos em relação as matrizes disciplinares, canais de comunicação, colaboração científica; b) Identificar os autores mais citados nas dissertações e teses sobre o tema a fim de mapear as influências recebidas no seu desenvolvimento científico; c) Identificar a presença da Ciência da Informação nos canais de comunicação da Ciência da Computação e vice-versa; d) Identificar a colaboração científica da Ciência da Informação e da Computação através de co-autorias.

dm.05.2008 Inserção da biblioteca digital de teses e dissertações no contexto da web semântica

RESUMO: A problemática da explosão informacional detectada em meados dos anos 40 aliada as constatações das deficiências dos sistemas convencionais de consulta motivaram a aplicação das novas tecnologias da informação e comunicação no desenvolvimento de modernos modelos e padrões de representação e recuperação da informação. Dentre eles, encontra-se a Web Semântica que associa significado à informação, permitindo o trabalho cooperativo entre computadores e pessoas. **Com o intuito de contribuir para o desenvolvimento desta área no contexto brasileiro, esta pesquisa tem por objetivo analisar as tecnologias de recuperação da informação sob a perspectiva da Web Semântica para adaptação e aperfeiçoamento do método de recuperação atualmente adotado pela Biblioteca de Teses e Dissertações (BDTD) do Instituto Brasileiro de Informação e Tecnologia (IBICT).** Do ponto de vista metodológico, este estudo analisa a estrutura da BDTD, bem como os metadados Título, Assunto e Resumo, a fim de extrair os termos semanticamente relevantes desses campos e desenvolver uma ontologia para um domínio específico. Para o desenvolvimento da ontologia foram utilizados a ontoKEM como ferramenta de apoio a elaboração do projeto e a documentação do trabalho e o Protégé para a sua construção. Foi ainda adotado o arquivo OWL como meio de integrar essas duas visões no desenvolvimento (o projeto e a construção). Os resultados obtidos indicam que estudos de ontologia conduzem à interoperabilidade entre os padrões existentes e essa nova proposta contextualizada pela Web Semântica e suas tecnologias. No que diz respeito ao ambiente da BDTD, foi identificado que além da ontologia desenvolvida e de uma programação nos seus agentes de busca, a marcação semântica é peça importante nesse cenário. Como conclusão, são propostas melhorias para as técnicas de representação e recuperação da informação, utilizadas pela BDTD, respeitando as peculiaridades de sua estrutura. Por fim, sugere-se como trabalhos futuros a realização de um estudo posterior acerca da possibilidade dos metadados atenderem a função da

marcação semântica.

OBJETIVO GERAL: O trabalho tem como objetivo principal analisar as tecnologias de recuperação da informação sob a perspectiva da **Web Semântica** para adaptação e aperfeiçoamento do método de recuperação da informação atualmente adotado pela BDTD.

OBJETIVOS ESPECÍFICOS: a) Caracterizar no contexto da recuperação da informação, a proposta da **Web Semântica**; b) Caracterizar a ontologia como ferramenta da **Web Semântica**, responsável por expressar e permitir o processamento da semântica dos dados, destacando suas características, classificações, vantagens, desvantagens e principalmente alguns critérios que devem ser utilizados no desenvolvimento da mesma; c) Verificar a estrutura atual da BDTD e métodos para recuperação da informação, focalizando os padrões utilizados para representação de dados e metadados para recuperação da informação; d) Elaborar parte da ontologia para o domínio correspondente à intersecção entre a recuperação da informação e a **Web Semântica**, contemplado pela BDTD; e) Sugerir a integração do modelo proposto pela **Web Semântica** ao cenário atual da BDTD; f) Propor recomendações de melhorias na recuperação da informação da BDTD, fundamentadas nas tecnologias da **Web Semântica**, estudadas durante a pesquisa.

dm.06.2008 Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da CI...

RESUMO: A relevância tanto da Web quanto das bibliotecas digitais para os diversos ramos da ciência, tem impulsionado pesquisadores e comunidades científicas a buscar soluções de integração, intercâmbio e entendimento semântico sobre os conteúdos que nelas circulam. Verifica-se, no cenário atual, uma forte tendência para a disponibilização de conteúdos digitais e a falta de uso de padrões adequados para a representação e a descrição desses recursos em ambientes digitais, o que é apontado pela literatura como uma grande preocupação para a comunidade científica e focado neste trabalho enquanto problema da pesquisa. Nesse contexto, a proposição desta pesquisa é verificar na literatura científica a aplicabilidade e a funcionalidade das tecnologias e instrumentos disponíveis como padrões de representação e descrição de conteúdo e forma, no âmbito das bibliotecas digitais, com o objetivo de identificar e de caracterizar o uso e a aplicação do MarcOnt enquanto ferramenta para a construção de formas de representação da informação para bibliotecas digitais na atualidade, destacando-o como instrumento para descrição dos aspectos semânticos dos conteúdos digitais nestas ambiências. Como metodologia, o presente trabalho caracteriza-se por ser uma pesquisa de análise exploratória e descritiva do tema, com o qual é possível obter subsídios para um maior entendimento e compreensão mais detalhados do MarcOnt e as tecnologias que estão abarcadas no âmbito das bibliotecas digitais para a construção Web Semântica. Como resultado, **procurou-se demonstrar a aplicação da ferramenta MarcOnt em um ambiente informacional específico, a Biblioteca Digital Semântica JeromeDL, para analisar nos fundamentos da Catalogação as formas de representação dos recursos informacionais.** O que nos levou a definir categorias para análise futura do sistema em comparação com o padrão de metadados MARC21.

OBJETIVO GERAL: Identificação e estudo de métodos e ferramentas tecnológicas disponíveis para a representação e a descrição de recursos informacionais bem como de seu conteúdo, no contexto das bibliotecas digitais e da Web [atendendo aos cenários da **Web Semântica** e da Web 2.0], na tentativa de identificar o embasamento conceitual subjacente às ferramentas para a padronização da descrição dos aspectos de forma e do conteúdo dos recursos informacionais nestas ambiências, na abordagem da Ciência da Informação.

OBJETIVOS ESPECÍFICOS: a) Identificar e caracterizar o MarcOnt Initiative, como um padrão de representação da informação (forma e conteúdo) em bibliotecas digitais e na Web na atualidade; b) Justificar a adoção da ferramenta MarcOnt como possibilidade de proporcionar a interoperabilidade semântica entre bibliotecas digitais e outros sistemas de informação.

dm.07.2009 Perspectivas na web semântica para a Ciência da Informação

RESUMO: A Web se tornou o principal veículo de armazenamento e disseminação da informação, mas há a necessidade de que a informação nesse ambiente seja tratada de forma eficiente. **O objetivo da pesquisa é buscar perspectivas na Web Semântica para a Ciência da Informação**, através do mapeamento na literatura das propostas básicas dessa tecnologia e do desenvolvimento já alcançado e também da análise das teses e dissertações sobre o tema apresentado aos Programas de Pósgraduação em Ciência da Informação no Brasil. O método utilizado foi o dialético, com abordagem exploratória e descritiva, o procedimento técnico foi o bibliográfico e as técnicas para a estruturação dos dados foram a bibliometria e a análise de conteúdo. **Foram estudados os seguintes movimentos e relações: - do surgimento da Web até o estado atual da Web Semântica (mudança qualitativa); - Web atual e Web Semântica (movimento, mudança qualitativa); - Ciência da Informação (através da revisão de literatura) e Web semântica (relações); e - Ciência da Informação (em teses e dissertações do Brasil) e Web Semântica (relações).** O movimento da Web atual à Web Semântica foi significativo, com objetivos voltados a possibilitar inferências complexas por parte das máquinas, mas deixando a simplicidade da arquitetura da Web atual, a Web Semântica se desenvolveu pouco, não teve movimento qualitativo significativo. **Quanto à CI, sua arquitetura sustenta uma relação. A literatura da CI, no que diz respeito à recuperação e representação da informação e representação do conhecimento, possui recursos que podem colaborar para o desenvolvimento da Web Semântica** (como os tesouros, os metadados etc.). **Outro ponto que favorece essa aproximação é a relação interdisciplinar da área com a Ciência da Computação, com a Inteligência Artificial e com a Linguística (Terminologia).** Mas existe a necessidade de que o conhecimento gerado na CI, sobre a Web Semântica ultrapasse os limites da própria área.

OBJETIVO GERAL: Indicar perspectivas na **WS** para a CI, principalmente no que diz respeito à organização e à RI.

OBJETIVOS ESPECÍFICOS: a) 1. Identificar pontos de desenvolvimento da **WS**, principalmente conceituais, a partir de uma literatura representativa e pertinente sobre o assunto, escolhidas conforme critérios estabelecidos pela análise de conteúdo (AC). b) Indicar como a **WS** se integra à literatura da CI no Brasil, através das teses e das dissertações dos programas de Pós-Graduação em CI.

dm.08.2010 Web semântica e o governo eletrônico brasileiro...

RESUMO: A Web semântica e o Governo eletrônico brasileiro constituem a temática deste trabalho, que teve como objetivo analisar se a associação do padrão de metadados a ser adotado pelo Governo eletrônico brasileiro (e-PMG) à lista de assuntos do Governo

(LAG/VCGE) – taxonomia para navegação – permitirá a representação dos recursos informacionais governamentais para a recuperação da informação. O estudo partiu de revisão histórica e conceitual acerca da classificação do conhecimento e das iniciativas de filósofos imbuídos na tarefa de organizar as áreas dos saberes. Nessa perspectiva, referencia a teoria da classificação facetada, listando produções de cânones responsáveis pela definição de princípios e postulados e pela aplicação do método de classificar o conhecimento a partir da organização, descrição, representação de conteúdo e indexação. A seguir, realizou-se pesquisa sobre a Web semântica e as tecnologias inovadoras, com destaque nos metadados, por se tratar de conjunto de elementos com semântica padronizada na descrição de informações eletrônicas. A investigação empírica se deu através de um estudo de caso, associado a um estudo documental, numa amostra composta pelo segmento organização e intercâmbio de informação, do Padrão de Interoperabilidade do Governo Eletrônico. Foi utilizada a análise documental de manuais técnicos, legislação e procedimentos, e a realização de entrevista com coordenador do segmento pesquisado. A análise de dados foi realizada a partir das abordagens qualitativas e quantitativas, de forma a analisar a estrutura da taxonomia para navegação e as relações de similaridade e complementaridade entre o e-PMG e o Dublin Core. Os resultados obtidos indicaram que a associação entre o e-PMG e a LAG/VCGE trará benefícios significativos para a descrição e representação da informação do portal do Governo eletrônico brasileiro para a localização e recuperação da informação, uma vez que o vocabulário controlado possibilitará a uniformização dos termos a serem utilizados do elemento assunto do e-PMG. Ainda como resultados, foi apontado que a estrutura hierárquica e a relação entre os termos da LAG/VCGE mantêm coerência com o referencial teórico adotado, no que se refere ao dinamismo de atualização dos termos, ao controle do vocabulário e ao estabelecimento das relações semânticas. Deduz-se que a associação do e-PMG à LAG/VCGE, possibilitará a descrição dos recursos informacionais com termos previamente definidos e de acordo com a área ou domínio específico da administração pública federal. E, mais ainda, através das relações de similaridade e complementaridade, com o DC, identificou-se os elementos específicos e-PMG e os qualificadores de refinamento e de padrões utilizados. Por fim, verificou-se que o e-PMG apresentou maior número de qualificadores e que a sua adoção possibilitará maior detalhamento e especificidade na descrição dos recursos informacionais do Governo, assim como também o aumento da forma significativa de representação da informação a partir do conteúdo.

OBJETIVO GERAL: Analisar se a associação entre esse padrão de metadado (e-PMG) [adotado pelo Governo eletrônico brasileiro] à taxonomia para navegação (LAG/VCGE) permitirá a representação para a recuperação da informação.

OBJETIVOS ESPECÍFICOS: a) 1. Analisar a estrutura da taxonomia para navegação adotada pelo e-gov; b) Apresentar os elementos gerais e específicos do padrão de metadado criado para o contexto do e-gov; c) Analisar as relações de similaridade e complementaridade existentes entre o e-PMG e o padrão Dublin Core, padrão internacional tomado como referência; d) Identificar os pontos de aproximação entre o e-PMG e o Dublin Core; e) Verificar se a LAG/VCGE segue as orientações teóricas para construção de taxonomia.

dm.09.2011 *Estudo exploratório sobre compatibilização semântica de vocabulário utilizando ontologia ...*

RESUMO: Estudo exploratório sobre compatibilização semântica de vocabulário, utilizando uma ontologia de domínio no MHTX1, procurando refletir sobre as práticas de leitura e escrita de documentos publicados em bibliotecas digitais. Nesses ambientes, a ambiguidade do vocabulário é apontada como um dos aspectos cruciais para a compatibilização semântica de bases heterogêneas. A integração de modelos conceituais é uma das técnicas mais usadas para compartilhar informações com entendimento comum entre significados conceituais terminológicos. O MHTX favorece a criação de um ambiente interativo. A ideia é analisar a representação de diferentes publicações em um único mapa como um sistema de conceitos integrados, tendo como base a fundamentação teórica-metodológica sobre compatibilização de vocabulário, visando o aperfeiçoamento do MHTX. Propõe-se uma nova técnica de modelagem informacional de objetos digitais, considerando os mesmos princípios de classificação facetada. Esta técnica permitiu usar as funcionalidades do conceito para explicitar relações de equivalência em um único mapa, indo ao encontro dos pressupostos deste estudo que considera o conceito um metadado descritivo, administrativo e estrutural. Para isso, tanto a indexação quanto a recuperação deve ser baseada em uma terminologia consensual para o estabelecimento de categorias formais diretamente ligadas ao uso da linguagem. A combinação dessas categorias representadas em facetas de uma ontologia de domínio mostrou conexões semânticas em que as características expressas nos objetos hipertextuais estão presentes na definição e nas relações entre conceitos. Os resultados confirmam a viabilidade de inserir uma ontologia de domínio no MHTX, utilizando a modelagem proposta com mecanismos automatizados de indexação para conduzir compatibilização semântica do vocabulário.

OBJETIVO GERAL: Aperfeiçoamento do MHTX, tendo como base a fundamentação teórica-metodológica sobre compatibilização semântica de vocabulário, utilizando uma ontologia de domínio.

OBJETIVOS ESPECÍFICOS: a) Compreender as características e componentes básicos comuns presentes em uma ontologia e um sistema de conceitos através da bibliografia da área; b) Estudar mecanismos e ferramentas propostos pelo W3C para compatibilização e interoperabilidade semântica de metadados digitais, agregando aos subsídios teóricos e metodológicos da CI com processos de construção de ontologias, através de um estudo exploratório; c) Contribuir nas pesquisas da área de RI, bibliotecas digitais e ontologias.

dm.10.2011 *Registros de marcas da junta comercial do estado de Pernambuco ...*

RESUMO: A produção intelectual no Brasil é resguardada, dentre outras formas, através do registro de marca. Durante o período de 1875 a 1924, o registro de marca em Pernambuco foi realizado na Junta Comercial do Estado de Pernambuco (JUCEPE), antigo Tribunal do Comércio. Dessa atividade administrativa realizada pela Junta, resultou um manancial de documentos que foram digitalizados, cujo papel fundamental não está mais vinculado a resguardar juridicamente a propriedade das marcas, mas atuam como elementos representativos de preservação da cultura e memória social. A presente pesquisa traz uma discussão teórico-reflexiva proporcionada por temas referentes à organização do acervo digital dos registros de marcas da JUCEPE. A respeito dos aspectos metodológicos, configura-se como uma pesquisa qualitativa de caráter exploratório e investigativo e utiliza o levantamento documental como técnica de pesquisa, cujo objetivo é avaliar a necessidade/importância da descrição semântica das informações do acervo digital de registros de marcas da Junta Comercial do Estado de Pernambuco a partir das necessidades potenciais dos usuários. Diante disso, **procura discutir temas relacionados à Web semântica, principalmente metadados e ontologias, sob a**

ótica da Ciência da Informação, e verifica as suas contribuições para as reflexões teóricas sobre a descrição semântica do acervo de marcas registradas da JUCEPE em ambiente digital. Visa também apresentar e analisar os registros de marcas da Junta Comercial como fontes de informação histórica e tecnológica, caracterizar as necessidades informacionais dos usuários de marcas registradas e, por fim, fazer uma análise comparativa entre dois sistemas de recuperação de marcas registradas de caráter histórico, a Base de Marcas do Arquivo Histórico da Oficina Espanhola de Patentes e Marcas (OEPM) e o site da Fundação Joaquim Nabuco (FUNDAJ), dentro do contexto da Ciência da Informação.

OBJETIVO GERAL: Contribuir avaliando a importância da descrição semântica das informações do acervo digital de registros de marcas da Junta Comercial do Estado de Pernambuco a partir das necessidades potenciais dos usuários.

OBJETIVOS ESPECÍFICOS: a) [Discutir] temas relacionados à **Web Semântica**, sob a ótica da Ciência da Informação, a partir de revisão de literatura especializada [e verificar] quais suas contribuições para a descrição semântica da tipologia documental em questão. b) [Apresentar e analisar] o acervo digital de marcas registradas da JUCEPE como fonte de informação tecnológica e histórica, priorizando os documentos que contivessem imagem. c) [Caracterizar] as necessidades informacionais dos usuários de marcas registradas. d) [Analisar] comparativamente dois sistemas digitais de recuperação de registros de marcas históricas - a Base de Marcas do Arquivo Histórico da Oficina Espanhola de Patentes e Marcas (OEPM) e o site da Fundação Joaquim Nabuco (FUNDAJ) -, focando suas metodologias de organização da informação e do conhecimento.

dm.11.2012 *Artigos científicos digitais na web: novas experiências para apresentação, acesso e leitura*

RESUMO: O padrão de publicação científica da atualidade é o periódico científico eletrônico. Com a chegada da internet e, principalmente, com a chegada das publicações eletrônicas, a humanidade passou a dispor de muita informação ao mesmo tempo, o que ocasionou uma grande “explosão informacional”. Hoje, a informação é produzida em um ritmo que excede as habilidades humanas. É determinante neste cenário, mobilizar o computador para tratar e processar o conteúdo das informações disponíveis no ambiente Web. Mesmo com a facilidade de acesso ao texto completo dos artigos de periódicos científicos, através de bibliotecas digitais, repositórios digitais, o formato utilizado no meio eletrônico ainda é textual legível somente por pessoas, o que impossibilita o seu processamento semântico por programas. As páginas da Web foram construídas com semânticas locais, e este fato, se constitui como o maior obstáculo para integrar seus conteúdos. Pensar em organizar o caos informacional disponível na Web se tornou imperativo para possibilitar novas formas de acesso à informação digital. A Web 2.0 e Web 3.0 (Web Semântica) se configuram como novas propostas para alcance desses objetivos. A Web Semântica propõe incorporar sentido às informações de maneira que as máquinas possam compreender a linguagem humana, ou seja, fornecer estruturas e dar significado ao conteúdo das páginas Web; e a Web 2.0, além de facilitar a comunicação interpessoal e compartilhar informações, se destacando também pela colaboração científica, incentivando os periódicos científicos tradicionais a adotarem ferramentas colaborativas como os blogs em seus websites. Várias experiências estão sendo desenvolvidas atualmente no sentido de utilizar as tecnologias da Web 2.0 e Web Semântica em publicações acadêmicas eletrônicas. A proposta desta pesquisa é identificar projetos e experiências inovadoras de periódicos científicos que utilizam as tecnologias da Web Semântica e Web 2.0 para fornecer acesso direto ao conteúdo semântico dos artigos científicos digitais e ampliar o potencial de compreensão e recuperação do conteúdo semântico e de interação entre autores e leitores de artigos científicos digitais na Web. O desenvolvimento desta pesquisa está fundamentado nas bases da Ciência da Informação, em especial na Comunicação Científica, dando ênfase à evolução do periódico científico como canal privilegiado deste meio de comunicação e na Ciência da Computação, ao que diz respeito às tecnologias da Web Semântica e Web 2.0. Trata-se de uma pesquisa de natureza documental na Web, bibliográfica, aplicada, qualitativa exploratória e descritiva, que utiliza o método comparativo encontrado no estudo das Ciências Sociais, para a exploração dos fenômenos, identificação das características comuns e diferenças existentes nas dezesseis experiências analisadas por este estudo. Dentre os resultados destas análises, constata-se que o tradicional modelo de artigo científico impresso já não atende muitas das novas necessidades dos pesquisadores e não utiliza efetivamente as potencialidades oferecidas pelas novas Tecnologias de Informação e Comunicação para ampliar a comunicação científica. O grande percentual das experiências analisado está voltado para a área da Saúde, o que reflete o aspecto pioneiro da área Biomédica. As experiências propõem um passo adiante para a questão da recuperação e processamento semânticos de conteúdos em ambientes digitais. Ou seja, vão além do modelo do artigo impresso, lido exclusivamente por pessoas utilizando as tecnologias semânticas para que estes possam ser “inteligíveis” também por programas. Existe um uso efetivo de tecnologias da Web 2.0, com vistas a facilitar o relacionamento do pesquisador no ambiente digital, cujas métricas baseadas nestas atividades podem informar medidas mais rápidas de impacto, complementando as métricas tradicionais de citação, esboçando assim, um novo cenário para a cointermetria com o uso destas tecnologias.

OBJETIVO GERAL: Identificar experiências inovadoras de apresentação, leitura e acesso aos artigos científicos publicados na Web, que ultrapassem o formato textual e os tradicionais mecanismos de busca.

OBJETIVOS ESPECÍFICOS: a) Destacar as características principais dessas experiências; b) Apontar as ferramentas e técnicas utilizadas pelas mesmas; c) Distinguir se a experiência é operacional ou se trata de uma proposta teórica ou um protótipo.

dm.12.2012 *As dimensões semântica e pragmática da web e dos mecanismos de busca no ciberespaço*

RESUMO: O ciberespaço é um espaço semântico/semiótico, desterritorializado, em constante modificação e a Web, seu principal constructo, tem crescido de forma vertiginosa ao ponto de ser fracionada para fins de estudo, dividindo-se em Web 1.0 ou Sintática, Web 2.0 ou Social, Web 3.0 ou Web Semântica, Web Pragmática etc. Assim, a Web Semântica (WS) é uma representação do conhecimento que tem na sua estrutura as tecnologias para atribuir semântica baseada nas linguagens de programação de modo geral. Porém, a semântica da WS é do tipo formal, mais ligada à sintaxe e a lógica, diferente da semântica da Linguística. A Web Pragmática está sendo construída a partir das experiências da Web 2.0 ou Social, de modo que as pessoas estão contribuindo para sua construção, fazendo o upload e uso da linguagem dentro de um contexto. Contudo, somente uma pequena parte da Web está representada por metadados, ontologias etc e, a outra parte, que é a grande maioria, ainda tem o problema da taxonomia do conhecimento e da multiplicidade dos signos. Dessa forma, os mecanismos de busca como tecnologias da informação fazem a indexação automática dos conteúdos da Web; no entanto, tradicionalmente, utilizam palavras-chave ou descrições textuais para processarem a busca. A grande riqueza da indexação automática são as múltiplas sintaxes, não obstante, o aspecto semântico é necessário para atribuir sentido a query, ou seja, a questão de busca. Nesse sentido, para entender as dimensões da Web e dos mecanismos de busca, utilizamos os conceitos de semântica e pragmática da Linguística e da Filosofia, relacionando o conceito de semântica ao sentido e o conceito de pragmática ao contexto de

uso da linguagem. Trata-se de uma pesquisa teórico-informal que tem por objetivo estudar os mecanismos de busca que operam com semântica e a busca contextual. Para tanto, adotou-se a pesquisa documental com abordagem qualitativa e a análise documental como método e técnica tanto para construção do corpus teórico quanto para identificação, seleção, definição do corpus de análise e exemplificação da busca realizada pelos mecanismos de busca que operam com semântica. Foram analisadas a forma de organização (indexação) e o processo de busca dos mecanismos de busca selecionados. Os resultados demonstraram que a semântica e a pragmática são interdependentes quando se trata do estudo dos mecanismos de busca, porque não dá para desvincular a indexação e a busca, uma vez que o mecanismo é o interpretante da enunciação da busca e o leitor, o interpretante dos resultados. Os mecanismos de busca têm utilizado da colaboração dos leitores tanto na geração do índice quanto na definição de padrões de busca para que os resultados sejam obtidos em um contexto pragmático. Constatou-se que a pesquisa pode contribuir para a categorização ou tipologia dos mecanismos de busca, tanto que sugerimos a readigração da categoria “apresentação dos resultados”. Por fim, acredita-se que o estudo possibilitará aos bibliotecários o conhecimento das funcionalidades dos mecanismos de busca, uma vez que a tendência atual das interfaces de busca dos catálogos online (OPACs) é facilitar o acesso às coleções em um ambiente similar ao dos sites de busca da Web.

OBJETIVO GERAL: Estudar os mecanismos de busca que operam com semântica e a busca contextual.

OBJETIVOS ESPECÍFICOS: a) Identificar e selecionar os mecanismos de busca que operam com semântica; b) Definir um corpus e analisá-lo a partir do constructo teórico; c) Exemplificar como os mecanismos estão processando a busca (pragmática).

dm.13.2012 *As dimensões semântica e pragmática da web e dos mecanismos de busca no ciberespaço*

RESUMO: Aborda temas relacionados à classificação e organização do conhecimento, folksonomia, recuperação da informação e às teorias do caos e da complexidade em sistemas, com o objetivo de constituir o campo conceitual necessário para conduzir a investigação. Trata a classificação desde o seu caráter inato, passando pelos estudos filosóficos iniciados na antiguidade, chegando até a abordagem científica da concepção das classificações bibliográficas e das classificações que vem sendo utilizadas para a organização da informação na era da Web 2.0. Num segundo momento, a discussão teórica se desenvolve nas teorias do caos e da complexidade em sistemas e suas relações com as redes sociais na Internet, em particular, com a rede social Twitter que é uma rede de mensagens postadas pelos usuários do sistema, limitadas a 140 caracteres, sem o agrupamento dos usuários por comunidades ou grupos de interesse, como é comum em outras redes sociais. A pesquisa exploratória e empírica evidencia o uso e as finalidades das hashtags no Twitter, destacando o papel da colaboração na escolha dos termos. Conclui que a colaboração é essencial para tornar a recuperação da informação mais efetiva, além de atuar na convergência de saberes.

OBJETIVO GERAL: Investigar a classificação colaborativa da informação nas redes sociais da Internet através do uso das tags pelos usuários, considerando a presença de elementos de caos e de complexidade nesses sistemas.

OBJETIVOS ESPECÍFICOS: a) Investigar a dinâmica da disseminação da informação e da interação entre os usuários na rede social Twitter para identificar elementos das teorias do caos e da complexidade em sistemas; b) Identificar a lógica do uso das tags para a classificação da informação no Twitter estabelecendo relação com o desenvolvimento de sites e ferramentas de apoio, destacando a característica colaborativa desse tipo de classificação.

dm.14.2013 *Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária*

RESUMO: No âmbito das organizações, a gestão do conhecimento, com seus processos de aquisição, armazenamento e compartilhamento, se baseia em recursos informacionais destinados para o uso direto pelos seres humanos. Porém, estes recursos não são combinantes a ponto de serem usados ou reutilizados para produção de novos conhecimentos emergentes ou inovação. Diante do atual cenário de crescente aumento no volume de informações registradas e do desafio de se compreender cada vez mais e melhor a complexidade dos fenômenos, é preciso contar com alguma forma de memória organizacional integrada e integradora, de maneira atender as necessidades dos usuários em diferentes perspectivas, situações, dimensões, condições e contextos. Por outro lado, a tecnologia da Web Semântica manifesta características que possibilitam o desenvolvimento de uma arquitetura de repositório semântico que atenda essas demandas. Neste sentido, esta dissertação tem como objetivo delinear uma arquitetura de repositório semântico que integre os dados corporativos, podendo servir de referência para implementação em instituição de pesquisa agropecuária.

OBJETIVO GERAL: Avaliar o papel da arquitetura de repositório de dados com **Web Semântica** para a gestão do conhecimento corporativo no âmbito da organização de pesquisa científica no campo da agropecuária.

OBJETIVOS ESPECÍFICOS: a) Estudar a natureza do conhecimento científico na perspectiva do processo de geração, retenção e compartilhamento do conhecimento, dentro do contexto das crescentes demandas de pesquisa e suas necessidades de visão complexa, holística, e sistêmica; b) Mapear elementos que caracterizam uma arquitetura de repositório semântico de dados com a finalidade de promover o compartilhamento de conhecimento semanticamente estruturado em nível de dados, possibilitando o processamento de inferências, e facilitar a operação e intercomunicação entre sistemas de informação a partir do estabelecimento de uma terminologia comum no âmbito de uma organização de pesquisa agropecuária; c) Adequar o modelo genérico da arquitetura às especificidades da organização de pesquisa científica no campo da agropecuária, Embrapa.

dm.15.2013 *As tecnologias da web semântica no domínio bibliográfico*

RESUMO: A proposta de uma Web Semântica surgiu como uma alternativa que possibilitaria a interpretação das informações por máquinas, permitindo assim maior qualidade nas buscas e resultados mais relevantes aos usuários. A Web Semântica pode ser utilizada atualmente apenas em domínios restritos, como em sites de comércio eletrônico, devido à dificuldade de representar ontologicamente toda a Web. **Objetiva-se verificar como os conceitos, tecnologias, arquiteturas de metadados utilizados pela Web Semântica podem contribuir à construção, modelagem e arquitetura de metadados de catálogos bibliográficos, tomando por base os conceitos definidos no modelo conceitual desenvolvido para a representação do universo bibliográfico denominado Functional Requirements for Bibliographic Records (FRBR)**, além de explanar sobre a utilização do modelo conceitual como recurso ontológico. A proposta se pauta no estudo de arquiteturas de metadados semânticas, de modo a identificar suas características, funções e estruturas, além de estudar o modelo BIBFRAME (Bibliographic Framework), que se configura como a iniciativa mais recente acerca da implementação de tecnologias da Web à área da Biblioteconomia e Ciência da Informação. Esta pesquisa caracteriza-se por seu caráter teórico-exploratório e foi desenvolvida mediante análise e revisão de literatura sobre seus temas. Os resultados apresentam as principais arquiteturas de metadados utilizadas no contexto da Web Semântica e uma abordagem sobre ontologias, interoperabilidade em sistemas de informação, modelagem de catálogos online, além da apresentação do modelo BIBFRAME, com base em sua importância para a catalogação.

OBJETIVO GERAL: Identificar as contribuições que os conceitos e tecnologias utilizados pela **Web Semântica** podem oferecer à área da Ciência da Informação, mais especificamente ao desenvolvimento, modelagem e arquitetura de metadados (entendidas nesta pesquisa como padrões de metadados e linguagens para representação de ontologias) em catálogos online, tendo por base os

conceitos definidos nos Functional Requirements for Bibliographic Records (FRBR) e o framework bibliográfico BIBFRAME.

OBJETIVOS ESPECÍFICOS: a) Identificar os principais aspectos da **Web Semântica**, destacando suas características e conceitos que a fundamentam; b) Realizar um estudo acerca das arquiteturas de metadados identificadas (RDF; RDFS; SHOE; OIL; DAML; DAML-OIL; OWL); c) Apresentar o framework bibliográfico BIBFRAME (BIBFRAME Linked Data Model), tendo como base as tecnologias da **Web Semântica**.

dm.16.2013 Representação da informação dinâmica em ambientes digitais

RESUMO: Este trabalho é um estudo exploratório interdisciplinar, pois converge de duas áreas não pertencentes à mesma classe acadêmica, **Ciência da Informação (CI) e Ciência da Computação**. O objetivo é, além de estudar a representação no ambiente virtual, encontrar uma forma de representar a informação não textual (multimídia) que atenda essas "novas necessidades" e possibilidades que a Web Semântica requer no desenvolvimento de contextos com uso do XML. Conforme a complexidade dos documentos multimodais que envolvem textos, vídeos e imagens descritos em mais de um formato, a opção para a interoperabilidade da descrição foi representar o contexto destes documentos com uso de ontologia. Através de uma metodologia de pesquisa qualitativa de análise exploratória e descritiva, apresentam-se ontologias que permitam que esta descrição feita em padrões convencionais, mas interoperáveis, de formatos de descrição, e que possam atingir um conjunto de objetos multimodais. A descrição desta ontologia, em dois formatos interoperáveis, MARC21 e Dublin Core, foi criada utilizando o software Protégé; e para validação da ontologia, foram feitas 3 aplicações práticas com vídeos acadêmicos (uma aula, um trabalho de conclusão de curso e uma defesa de dissertação de mestrado), que possuem imagens retiradas dos slideshows e compostas num documento final. O resultado alcançado é uma representação dinâmica de vídeo, que faça as relações com os outros objetos que a vídeo trás além da interoperabilidade dos formatos de descrição, tais como: Dublin Core e MARC21.

OBJETIVO GERAL: Encontrar formas de representação que atenda às "novas necessidades" e possibilidades que a **Web Semântica** requer no desenvolvimento de contextos com uso do XML, e assim, compreender melhor as variáveis organizacionais e socioculturais que interferem as necessidades de informação dos usuários destes meios.

OBJETIVOS ESPECÍFICOS: a) [Analisar] os métodos de representação com uso de metadados que mostrem a informação dinâmica a partir dos formatos atuais tanto de representação descritiva como da **Web Semântica**; b) Apresentar ontologias que permitam que esta descrição possa atingir um conjunto de objetos multimodais

dm.17.2014 Wikis semânticos: da web para a web semântica

RESUMO: Com o desenvolvimento das Tecnologias de Informação e Comunicação, a Ciência da Informação precisou repensar a postura tradicional de gerenciamento informacional. O hipertexto, advindo principalmente por meio do ambiente Web, elevou ainda mais a complexidade do tratamento informacional. A popularização da Internet fez com que a Web se tornasse mais interativa e colaborativa tornando comum a utilização de sistemas wiki para o gerenciamento informacional. Idealizada por Tim Berners-Lee, surge a iniciativa da Web Semântica, na qual as máquinas são capazes de analisar os dados presentes na rede. Nesse contexto aparecem os wikis semânticos, caracterizados por fazerem uso das tecnologias da Web Semântica. Diante desse cenário, **considera-se que a Ciência da Informação, que já se preocupa com o desenvolvimento da Web e da Web Semântica, deve também se preocupar com os wikis semânticos**. Assim, por meio de uma pesquisa descritiva e exploratória, objetivou-se explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de descrição, de recuperação e de intercâmbio de informações apoiadas em tecnologias da Web Semântica, visando a favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais. Após se apresentar sobre as raízes históricas da Web Semântica, são destacados os padrões de representação, codificação, descrição, relação e consulta de dados estruturados (URI, XML, RDF, RDFS, OWL e SPARQL) que, junto a outras tecnologias, formam a base da Web Semântica e apoiam o funcionamento dos wikis semânticos. Os wikis semânticos são apresentados e definidos como sistemas wiki que se utilizam de tecnologias da Web Semântica para incorporar conhecimento formalizado, conteúdo, estruturas e links em suas páginas. Após essa etapa, são descritas as principais atividades de descrição, de recuperação e de intercâmbio de informações no Semantic MediaWiki, o wiki semântico mais utilizado até o momento. Como considerações finais, entende-se que os wikis semânticos favorecem o entendimento, a discussão e o uso de tecnologias da Web Semântica em ambientes informacionais digitais.

OBJETIVO GERAL: Explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de representação, de recuperação e de intercâmbio de informações apoiadas em tecnologias da **Web Semântica** para, assim, favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais.

OBJETIVOS ESPECÍFICOS: a) Apresentar brevemente conceitos de **Web Semântica**, wiki e wiki semântico para fornecer o entendimento básico da terminologia utilizada na pesquisa; b) Compreender os padrões de codificação, de representação, de relação e de consulta de dados estruturados relacionados ao desenvolvimento da **Web Semântica** (URI, XML, RDF, RDFS, OWL e SPARQL); c) Investigar as principais iniciativas que englobam as tecnologias wiki com as tecnologias da **Web Semântica**; d) Descrever as principais contribuições das tecnologias da **Web Semântica** aplicadas aos wikis para uma possível utilização dessas tecnologias em ambientes informacionais digitais.

td.01.2006 Autoria de documentos para a web semântica...

RESUMO: Nesta tese, apresentamos o momento da autoria de documentos como situação potencial para a representação do conhecimento e a conseqüente publicação de documentos. Sistemas baseados em ontologia para a representação da informação e autoria na Web Semântica estão surgindo e provendo soluções como anotação semântica, extração da informação, marcação ontológica, etc. Ao analisar esses sistemas, descobrimos as características que eles deveriam ter, a fim de que produzam e suportem documentos em conformidade com as recomendações da Web Semântica e conseqüentemente usufruam de suas potencialidades. Apresentamos aqui uma análise de ferramentas baseadas em ontologias, a partir de um critério de análise específico desenvolvido para identificar as características, requisitos e conformidade com as recomendações existentes, identificadas na revisão de literatura. O resultado da análise revelou que nesse novo paradigma de produção, os autores deverão desempenhar tarefas de estruturação do conhecimento com os quais eles ainda não estão familiarizados. Identificamos a necessidade de abordar essas questões no âmbito da Ciência da Informação, buscando dar apoio a autores no desenvolvimento de conteúdos bem estruturados e representados, a partir do potencial que têm as atuais tecnologias baseadas em ontologias. Após tratar dessas questões, propomos com esse trabalho um conjunto de requisitos e um modelo de arquitetura de ambiente de autoria baseado em ontologia para a Web Semântica, composto por módulos que sugerem novas funções e integram tecnologias correntes, fazendo uso de possibilidades existentes e propondo o desenvolvimento de novos requisitos.

OBJETIVO GERAL: Criar uma especificação de alto nível de um novo modelo de arquitetura de sistemas de informação para ambientes de autoria de documentos baseados em ontologia para a **Web Semântica**.

OBJETIVOS ESPECÍFICOS: a) Levantar critérios e modelos de análise de sistemas de autoria, de avaliação de ontologias, de conformidade com o W3C e de análise e avaliação de software; b) Realizar uma análise de requisitos de sistemas de autoria baseados em ontologia com base em critérios levantados; c) Especificar requisitos e modelar a arquitetura de um novo ambiente de autoria baseado em ontologias para a **Web Semântica**.

td.02.2010 *Representação iterativa: um modelo para repositórios digitais*

RESUMO: A recuperação da informação tem sido muito discutida e abordada dentro da Ciência da Informação nos últimos anos, principalmente depois da explosão informacional gerada pela Internet. A busca por informação de qualidade e compatível com a necessidade do usuário tem sido tratada como obsessão, atualmente. A utilização da Internet indicou novos modelos de armazenamento de informações, como os repositórios digitais, que têm sido utilizados em ambientes acadêmicos e de pesquisa como principal forma de autoarquivar e, principalmente, de disseminar informação, porém com uma estrutura de informação que sugere melhor descrição dos recursos do que a própria Web e indica uma melhor recuperação da informação nestes ambientes. Os repositórios ainda não estão aptos a recuperar informação de forma semântica e contextualizada. Os novos paradigmas de Internet sugerem utilização dos recursos de Web 2.0 e também de Web 3.0, permitindo, respectivamente, interatividade e também estrutura de informação semântica. Desta forma **o objetivo desta pesquisa é melhorar o processo de recuperação da informação, apresentando uma proposta de modelo estrutural no contexto da Web Semântica, abordando o uso de recursos da Web 2.0 e Web 3.0 em repositórios digitais, que permita recuperação semântica da informação, através da construção de uma camada de informação chamada Representação Iterativa**. Através do modelo sugerido e proposto — Representação Iterativa — será possível adequar os repositórios digitais para que utilizem Folksonomia e também vocabulário controlado de domínio, de forma a gerar uma camada de informação iterativa, que possibilite retroalimentação da informação, além de recuperação semântica da informação, através do modelo estrutural desenhado para repositórios. O modelo sugerido resultou na efetivação da tese de que através da Representação Iterativa é possível estabelecer um processo de recuperação semântica da informação em repositórios digitais.

OBJETIVO GERAL: Melhorar o processo de recuperação da informação, apresentando uma proposta de modelo estrutural no contexto da **Web Semântica**, abordando o uso de recursos da Web 2.0 e Web 3.0 em repositórios digitais, que permita recuperação semântica da informação, através da construção de uma camada de informação chamada Representação Iterativa.

OBJETIVOS ESPECÍFICOS: a) Estabelecer uma modelagem complementar de banco de dados que possa caracterizar o uso dos conceitos de Folksonomia em repositórios digitais; b) Aplicar uma metodologia de sugestão e descrição de tags, sugerindo a utilização de uma representação amparada em Folksonomia e Estruturas de Representação do Conhecimento, que se denomina Folksonomia Assistida; c) Construir um corpus de informação baseado em uma ontologia de domínio e ampliada e reciclada com a experiência do usuário através do uso da Folksonomia Assistida, criando uma estrutura nomeada Representação Iterativa; d) Utilizar o modelo construído, Representação Iterativa, no processo de recuperação da informação em repositórios digitais, através da elaboração de uma rede baseada na teoria dos grafos, de forma que seja possível recuperar informações com caráter semântico.

td.03.2013 *Web semântica e repositórios digitais educacionais na área de saúde...*

RESUMO: Apresenta modelagem baseada em Web Semântica para repositórios digitais educacionais na área de saúde, que disponibilizam objetos de aprendizagem com aspectos educacionais descritos em formulários de metadados. Foram identificadas características dos objetos de aprendizagem, com foco no objetivo de aprendizagem; agregadas tecnologias da Web Semântica à arquitetura desses repositórios e definida estrutura de informação adequada. Isso viabilizou inferências automáticas e contribuiu com o refinamento dos resultados de busca do usuário, aprimorando a recuperação da informação. Aborda repositórios digitais educacionais, as características e peculiaridades que influenciam sua utilização; Web Semântica e suas tecnologias para estrutura e sintaxe, lógica e semântica; e também são feitas considerações sobre metadados utilizados na descrição da informação dos objetos de aprendizagem. Como resultado, obtem-se modelagem que considera o processo de descrição e recuperação de objetos de aprendizagem, segundo seus aspectos educacionais; informações sobre a implementação; quadro com metadados LOM e Dublin Core correlacionados; e a simulação do resultado de busca em repositório digital educacional na área de saúde, com a implementação da modelagem proposta.

OBJETIVO GERAL: Propor modelagem baseada em **Web Semântica**, para repositórios digitais educacionais na área de saúde, com foco no objetivo de aprendizagem.

OBJETIVOS ESPECÍFICOS: a) Identificar as características dos objetos de aprendizagem disponíveis nos repositórios digitais educacionais na área de saúde; b) Verificar tecnologias da **Web Semântica** adequadas para a estruturação da informação descritiva dos objetos de aprendizagem dos repositórios digitais educacionais na área de saúde; c) Definir modelagem que contribua para o refinamento do resultado da busca dos objetos de aprendizagem disponíveis em repositórios digitais educacionais na área de saúde.

td.04.2015 *Interoperabilidade e mapeamentos entre sistemas de organização do conhecimento...*

RESUMO: Esta pesquisa apresenta o desenvolvimento de método de busca e recuperação de informações em bases de dados especializadas para produção do conhecimento científico na área da Saúde, com ênfase na Saúde Baseada em Evidências. Recorremos, neste trabalho, a diferentes metodologias considerando as especificidades de cada etapa: pesquisa exploratória, método hipotético dedutivo e estudo de caso empírico qualitativo. **Mobilizamos os fundamentos teórico-metodológicos da Ciência da Informação e da Saúde nos domínios da Organização e Recuperação da Informação e do Conhecimento, Web Semântica, Saúde Baseada em Evidências e Metodologia Científica**, assim como realizamos dois experimentos: estudo de caso em Ortopedia e Traumatologia no sentido de identificar e estabelecer critérios para busca, recuperação, organização e seleção de informações de modo que possam integrar parte da metodologia de trabalhos científicos na área da Saúde; e análise dos tipos de busca e recuperação e dos mapeamentos entre Sistemas de Organização do Conhecimento (SOC) propostos no Metatesauro no escopo da Unified Medical Language System (UMLS) da US National Library of Medicine e no BioPortal da National Center for Biomedical Ontology, ambos na área biomédica. O UMLS disponibiliza acesso a 151 SOC, e o BioPortal, um conjunto de 302 ontologias. Apresentam-se propostas para construção de estratégias de busca com uso de Sistemas de Organização do Conhecimento mapeados e interoperados, bem como para realização de pesquisas bibliográficas para elaboração de trabalhos científicos na área da Saúde.

OBJETIVO GERAL: A construção de método dirigido aos profissionais da informação atuantes na área da Saúde, que oriente a identificação e seleção de termos e conceitos para elaboração e aplicação de estratégias de busca sensibilizadas, a recuperação, a seleção dos trabalhos recuperados, assim como a organização dos resultados que contribua na explicitação da metodologia utilizada na produção do conhecimento científico na Saúde Baseada em Evidências e na Saúde em Geral.

OBJETIVOS ESPECÍFICOS: a) Identificar tipos de busca e recuperação de informações em uma amostra das bases de dados especializadas na área da Saúde. b) Estabelecer critérios para identificação e seleção de termos, conceitos e termos equivalentes

para construção de estratégias de busca sensibilizadas. c) Estabelecer formas de organização dos resultados de busca de modo que os pesquisadores tenham acesso aos trabalhos recuperados, selecionados, excluídos e repetidos. d) Estabelecer critérios para busca, recuperação, seleção e organização dos resultados de busca com base nos achados na literatura da Ciência da Informação e da Saúde no que se refere à: análise, construção e uso de linguagens, indexação, classificação, construção de filtros de pesquisa, bem como nos resultados do estudo de caso realizado em Ortopedia e Traumatologia: Artroplastias e Próteses de Joelho.

td.05.2015 Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática...

RESUMO: Estudo comparativo das propostas da Web Semântica e da Web Pragmática, com base nas teorias da organização e recuperação de informações, com o objetivo de propor uma modelagem de representação de dados, que inclua contextos, para melhorar a qualidade dos processos informacionais. Dirige-se o foco da pesquisa aos problemas da representação e recuperação de informações em acervos bibliográficos, principalmente com o uso de RDF e Topic Maps. Para isso, são apresentadas as origens da Web Semântica e da Web Pragmática e os conceitos fundamentais relacionados a esse ambiente, tais como linguagem, representação e recuperação da informação e do conhecimento; termos e terminologia; semântica e pragmática. A metodologia do trabalho foi construída a partir da análise e discussão dos conceitos fundamentais, a modelagem em Diagrama Entidade-Relacionamento e a codificação em RDF e XTM para representar contextos, aplicados a um repositório de textos científicos. Com base na análise crítica das propostas de organização e recuperação de informação das Web Semântica e Web Pragmática, apresenta-se uma proposta de representação de informações com contexto que poderá contribuir para melhorar a relevância dos resultados da recuperação de informações na WWW. Os contextos representados são: Citações, Domínio de Origem do Documento, Domínio da Palavra-Chave, Áreas de Formação do Indivíduo, Áreas de Publicação do Indivíduo, Áreas de Publicação da Revista e Interesses do Indivíduo. A pesquisa permitiu observar que há limites para introduzir contextos em sistemas de informação e compreender, também, que termos como semântica e pragmática requerem abordagem crítica. De fato, a operacionalização de conceitos semânticos e pragmáticos ainda está longe de ser realidade nos sistemas de informação contemporâneos no contexto da www. É importante ressaltar que **a presente pesquisa tem cunho interdisciplinar por abordar problemas discutidos tanto na Ciência da Computação como na Ciência da Informação.** Esta abordagem interdisciplinar ocorre, em primeiro lugar, porque o objeto de pesquisa, embora tenha origem na Ciência da Computação, requer as teorias e métodos da representação da informação estudados na Ciência da Informação para ser desenvolvido de forma adequada.

OBJETIVO GERAL: Propor uma abordagem teórico-metodológica para representação de contextos em acervos informacionais digitais, de modo a avançar em direção à Web Pragmática.

OBJETIVOS ESPECÍFICOS: a) Estudar criticamente os conceitos gerais que guiam a representação e recuperação de informações e de conhecimentos; b) Estudar os conceitos fundamentais da **Web Semântica** e Web Pragmática e discuti-los sob a ótica da Ciência da Informação; c) Estudar os possíveis contextos que podem ser representados em acervos informacionais digitais e em bases de dados de periódicos científicos; d) Explorar alternativas de modelagem de representação de contextos, em acervos digitais, utilizando os Topic Maps e RDF.

Nota. As unidades de registo consideradas encontram-se destacadas a negrito. Foram usadas duas limitações semânticas diferentes, nos resumos foi considerado o *tema* e nos objetivos o termo *Web Semântica*.

Apêndice 8

Quadro de análise dos objetivos dos 22 trabalhos que constituem a amostra

Objetivos	Verbo(s) de ação	Objeto(s) da ação	Elementos adicionais de contexto	Presença do termo WS	Trabs. (ref)
<i>Grau de Proximidade: 0 (zero)</i>					
Geral	Indicar	perspectivas na WS p/a CI.	[n.a.]	Sim	
Específicos	a) Identificar b) Indicar (como)	a) pontos de desenvolvimento da WS b) a WS se integra à literatura da CI no Brasil	a) principalmente conceituais. b) nas teses e dissertações dos PPG em CI.	2 de 2 100%	dm.07. 2008
Geral	Estudar	os aspectos que permeiam a construção da WS c/a CI.	[n.a.]	Sim	
Específicos	a) e c) Analisar b) Identificar	a) as ferramentas tecnológicas da WS. b) as características da linguagem XML c) os diferentes tipos de metadados	a) [n.a.] b) e sua importância p/a WS. c) enquanto ferramenta p/a WS.	3 de 4 75%	dm.01. 2005
Geral	Identificar	as contribuições da WS p/a CI	no desenvolvimento de metadados (especificamente).	Sim	dm.15.
Específicos	a) Identificar b) Apresentar	a) os principais aspectos da WS. b) o <i>framework</i> bibliográfico BIBFRAME	a) [n.a.] b) tendo como base as tecnologias da WS.	2 de 3 67%	2013
Geral	Refletir (sobre)	a participação e integração da CI e da CC na construção e desenvolvimento do conhecimento científico sobre a WS.	[n.a.]	Sim	
Específicos	a) e b) Identificar	a) a produção científica em WS b) os autores mais citados nas dissertações e teses sobre a WS	a) nas principais iniciativas brasileiras de BD's científicos. b) [n.r.]	2 de 4 50%	dm.04. 2008
Geral	Contribuir (p/uma "desmistificação" de)	os conceitos e tecnologias subjacentes ao projeto WS.	[n.a.]	Sim	dm.03. 2006
Específicos	[não apresenta objetivos específicos]				
<i>Grau de Proximidade: -1 (menos um)</i>					
Geral	Explorar, apresentar e descrever	as características dos wikis semânticos	apoiadas em tecnologias da WS.	Sim	
Específicos	a) Apresentar (brevemente) b) Compreender c) Investigar d) Descrever	a) conceitos de WS, wiki e wiki semântico. b) os padrões de codificação, de representação, de relação e de consulta de dados estruturados c) as principais iniciativas que englobam as tecnologias wiki d) as principais contribuições das tecnologias da WS	a) [n.a.] b) relacionados ao desenvolvimento da WS. c) c/as tecnologias da WS. d) aplicadas aos wikis.	4 de 4 100%	dm.17. 2014
Geral	Analisar	as tecnologias de recuperação da informação (RI)	sob a perspectiva da WS.	Sim	
Específicos	a) e b) Caracterizar c) Elaborar (parte de) d) Sugerir (a integração de)	a) a proposta da WS b) a ontologia c) a ontologia de domínio (interseção entre RI e WS) d) o modelo proposto pela WS	a) no contexto da RI. b) como ferramenta da WS. c) contemplado pela BDTD. d) ao cenário atual da BDTD.	4 de 6 67%	dm.05. 2008
Geral	Encontrar	formas de representação possíveis/ requeridas pela WS	p/o desenvolvimento de contextos c/uso do XML.	Sim	dm.16.
Específicos	Analisar	os métodos de representação da informação dinâmica, c/uso de metadados	a partir dos formatos de representação descritiva e da WS.	1 de 2 50%	2013
<i>Grau de Proximidade: -2 (menos dois)</i>					
Geral	Criar	uma especificação de alto nível de um novo modelo de arquitetura de sistemas de informação	p/ambientes de autoria de documentos baseados em ontologia p/a WS	Sim	td.01.2
Específicos	Especificar (requisitos) e modelar	a) a arquitetura de um novo ambiente de autoria	baseado em ontologias p/a WS	1 de 3 33%	006
Geral	Propor	modelagem p/repositórios digitais educacionais na área de saúde	baseada em WS	Sim	td.03.2
Específicos	Verificar	a) tecnologias da WS	adequadas p/a estruturação da informação descritiva dos objetos de aprendizagem	1 de 3 33%	013
Geral	Criar	uma proposta de modelo estrutural em repositórios digitais (Representação Iterativa)	no contexto da WS.	Sim	td.02.2
Específicos	-	-	-	0%	010
Geral	Avaliar	o papel da arquitetura de repositório de dados c/WS	p/a gestão do conhecimento corporativo/científico no campo da agropecuária.	Sim	dm.14.
Específicos	-	-	-	0%	2013
<i>Grau de Proximidade: -3 (menos três)</i>					
Geral	Identificar e estudar	os métodos e ferramentas tecnológicas disponíveis p/a representação e a descrição de recursos informacionais	no contexto das bibliotecas digitais e da Web, de acordo c/os novos cenários, tais como, WS, Web 2.0.	Sim	dm.06. 2008
Específicos	-	-	-	0%	
Geral	-	-	-	Não	dm.10.
Específicos	Discutir	temas relacionados à WS	sob a ótica da CI.	1 de 4 25%	2011
Geral	-	-	-	Não	td.05.2
Específicos	Estudar	os conceitos fundamentais da WS e da Web Pragmática.	-	1 de 4 25%	015
<i>Grau de Proximidade: -4 (menos quatro)</i>					
dm.02.2005; dm.08.2010; dm.09.2011; dm.11.2012; dm.12.2012; dm.13.2012; td.04.2015					

Estes trabalhos não apresentam o termo *Web Semântica* nos objetivos, seja no Geral seja nos Específicos.