

# A Web Semântica nas Empresas: Proposta de um Modelo de Implementação

**Reinaldo Manuel Lopes Dias Ferreira**

**Tese de Doutoramento em Informática**

Orientação: Professora Doutora Maria Isabel Calapez Cabrita Leal Seruca  
abril, 2015

c



UNIVERSIDADE PORTUCALENSE



# Agradecimentos

Não teria sido possível terminar esta etapa da minha vida sem a preciosa contribuição de várias pessoas, a quem dedico o trabalho realizado. Desde logo, aos meus pais, Isabel e Reinaldo, que me deram espaço para investigar desde cedo, permitindo que a curiosidade pelos mais diversos assuntos desse lugar a várias realizações. À minha companheira, Paula, que me percebe melhor que ninguém e que alimenta essa curiosidade das mais diversas formas, frequentemente em prejuízo de outras oportunidades. À minha filha, Bárbara, que me inspira de novo e me tira qualquer dúvida sobre a matéria da qual os sonhos são feitos. As pessoas mais interessantes são as que observam o passado, disfrutam o presente e constroem o futuro.

Ao Lino Oliveira, o meu parceiro de musculação intelectual, pelo exemplo com distinção que decidi seguir. À Professora Filomena Lopes e ao Professor Fernando Moreira, por me entusiasmarem em avançar com o Doutoramento. À Professora Isabel Seruca, pelas orientações preciosas e contribuições ponderadas, essenciais para a realização do trabalho ao longo dos últimos anos. Ao João Paulo Peixoto, pelo último impulso neste percurso. As contrariedades são por vezes amplificadas pela incapacidade de distinguir entre a aparência e a essência dos problemas.

Aos meus colegas, às empresas em geral e aos meus clientes em particular, a minha perseverança seria inútil sem a vossa atenção e reforço positivo. Inovar não é um impulso, é um processo.

A muitos outros que também, de forma direta ou indireta, consciente ou fortuita, participaram neste pequeno contributo para o conhecimento. Desejo que os negócios, as organizações e, em particular, o Estado, se concentrem na geração de conhecimento, criação de valor e promoção da liberdade: o caminho alternativo que tem sido seguido termina invariavelmente com a destruição.

A todos: o contexto depende mais do utilizador que do ambiente.

Reinaldo Ferreira



# A Web Semântica nas Empresas:

## Proposta de um Modelo de Implementação

### Resumo

Na geração atual da Web coexistem a tecnologia que lhe deu origem e gerou a Web dos documentos, a evolução que a tornou ubíqua e criou a Web social, e a tecnologia que poderá gerar a Web dos dados, a Web Semântica.

As oportunidades de investigação na Ciência da Web são inúmeras. Em particular, a implementação da Web Semântica nas Empresas é ainda um tema pouco explorado e encontra-se numa fase precoce de implementação.

Este trabalho apresenta a relevância do estudo da Web Semântica nas Empresas e propõe um modelo para facilitar e guiar a implementação desta tecnologia nas organizações. O modelo permite, numa primeira fase, analisar e determinar o grau de adoção da Web Semântica nas Empresas e, numa fase complementar, servir como guia, propondo um conjunto de recomendações e sistematizando o caminho a seguir para uma adoção mais rápida e intensiva.

Utilizando-se a metodologia de Investigação-Ação, delineou-se um trabalho de intervenção num conjunto de empresas em que se analisa a utilização atual da tecnologia, se propõem ações que promovem a sua implementação e se avaliam progressivamente os resultados obtidos.

**Palavras-chave:** Web Semântica, *Linked Data*, modelo de implementação



# The Enterprise Semantic Web: Proposal for an Implementation Model

## Abstract

The current Web generation embraces the original technology that generated the Web of documents, the evolution that turned it ubiquitous and created the social Web, and the technology that can generate the Web of data, the Semantic Web.

There are plenty of research opportunities within Web Science. In particular, the implementation of Semantic Web in enterprises is still a relatively unexplored subject and at an early stage of implementation.

This thesis presents the relevance of the study of Semantic Web in enterprises and proposes a model to facilitate and guide the implementation of this technology in organizations. The model allows, in a first stage, to analyze and determine the degree of adoption of the Semantic Web in enterprises and, in a complementary stage, is used as a guide, proposing a set of recommendations and systematizing the way forward to a faster and more intensive adoption.

Using the methodology of Research-Action, an intervention in a group of enterprises was outlined in which the current use of the technology is analysed, actions, to promote its implementation are proposed and results obtained are progressively evaluated.

**Keywords:** Semantic Web, Linked Data, implementation model



# Índice

1	Introdução .....	1
1.1	Enquadramento .....	1
1.2	Objetivos .....	2
1.3	Metodologia.....	3
1.4	Contribuição Científica.....	3
1.5	Estrutura da Tese .....	4
2	A Ciência da Web .....	7
2.1	A Internet .....	7
2.2	Web 1.0.....	8
2.2.1	O nascimento da <i>World Wide Web</i> .....	8
2.2.2	A primeira geração da Web .....	9
2.2.3	A cauda longa.....	10
2.3	Web 2.0.....	12
2.4	Web 3.0.....	14
2.4.1	A geração atual da Web.....	17
2.4.2	A Internet das coisas.....	18
2.4.3	Dados Universais.....	21
2.4.4	A economia da Web.....	23
2.4.5	A Ciência da Web .....	24
2.4.6	Oportunidades de Investigação na Ciência da Web .....	27
	Ciência dos Computadores .....	29
	Matemática.....	30

Direito .....	30
Sociologia.....	31
Economia .....	33
3 A Web Semântica .....	35
3.1 Conceito de Web Semântica .....	35
3.2 Tecnologias da Web Semântica.....	36
3.2.1 HTML.....	36
3.2.2 RDF.....	38
3.2.3 OWL .....	45
3.2.4 SPARQL.....	45
3.2.5 Tendências de evolução .....	46
3.3 Aplicações da Web Semântica.....	47
3.3.1 Exemplos de adoção da Web Semântica.....	48
3.3.2 O caso do schema.org.....	50
3.3.3 O caso do Google Knowledge Graph .....	51
3.3.4 O caso do Wolfram Alpha .....	57
3.4 Utilização da Web Semântica .....	59
3.4.1 Atividades do utilizador na Web.....	60
3.5 As Empresas e a Web .....	64
3.5.1 Caracterização das empresas .....	65
3.5.2 A Web Semântica nas Empresas.....	67
4 O Modelo e-swim.....	71
4.1 Preâmbulo .....	71
4.2 Desenvolvimento do Modelo .....	73
4.2.1 Adoção tecnológica .....	74

4.2.2	Proveniência.....	77
4.2.3	Acessibilidade dos dados .....	78
4.2.4	Atividades na Web .....	79
4.3	Proposta de Investigação - Ação.....	81
5	Utilização do Modelo .....	85
5.1	Seleção das Empresas .....	85
5.1.1	Projeto Raio X.....	85
5.1.2	Divulgação dos Resultados .....	86
5.1.3	Seleção da amostra .....	87
5.1.4	Protocolo de Participação .....	88
5.2	Estrutura da Intervenção .....	89
5.3	Diagnóstico inicial .....	91
5.3.1	Caracterização das Empresas.....	91
5.3.2	Evidências recolhidas .....	94
5.3.3	Primeiras ações implementadas.....	95
5.4	Avaliação da evolução das empresas .....	96
5.4.1	Ferramenta de Avaliação .....	96
5.4.2	A avaliação inicial .....	107
5.4.3	Reflexões sobre o trabalho realizado até à data .....	110
5.5	Elaboração de Recomendações .....	111
	Responsabilidade do Webmaster (Dimensão: Adoção) .....	113
	Colaboração Tecnológica (Dimensão Adoção) .....	114
	Estratégia de <i>e-business</i> (Dimensão: Adoção).....	114
	Propriedade na Web (Dimensão: Adoção) .....	115
	Importação de Dados (Dimensão: Proveniência) .....	115

Integração de Sistemas (Dimensão: Proveniência) .....	116
Acessibilidade de Conteúdos (Dimensão: Acessibilidade) .....	116
Adoção de Ontologias (Dimensão: Acessibilidade) .....	117
Design Adaptativo (Dimensão: Acessibilidade) .....	117
Disponibilização de Dados (Dimensão: Acessibilidade) .....	118
Tipos de Web Site (Dimensão: Acessibilidade) .....	118
Informação Estruturada da Empresa (Dimensão: Atividades) .....	119
5.5.1 Outras ações não sistematizadas .....	119
5.6 Avaliação final .....	123
5.6.1 Adoção e partilha de Recomendações .....	124
5.6.2 Análise da Intervenção .....	134
5.6.3 Avaliação Final .....	135
5.6.4 Aplicações e Atividades .....	141
5.6.5 Fatores de Sucesso .....	146
5.6.6 Barreiras identificadas .....	147
6 Conclusão .....	149
6.1 Síntese do trabalho realizado .....	149
6.2 Análise da contribuição científica .....	149
6.3 Trabalho Futuro .....	153
6.3.1 O futuro do Modelo e-swim .....	153
6.3.2 Oportunidades de Investigação .....	154
Bibliografia .....	159
Anexos .....	171
Anexo 1: Investigadores .....	173
Anexo 2: Web Science Subject Categorization (PT) .....	179

Anexo 3: Protocolo .....	185
Anexo 4: Empresas.....	189
Anexo 5: Ferramenta de Avaliação.....	193
Anexo 6: Recomendações.....	197
Anexo 7: Webmaster - o papel e o valor de um especialista na Presença Web de uma Empresa .....	237
Anexo 8: Aplicações e Atividades .....	247



# Lista de Abreviaturas e Siglas

AJAX: *Asynchronous JavaScript and XML*

CAE: Código de Atividade Económica

CSS: *Cascading Style Sheets*

CSV: *Comma Separated Values*

DHTML: *Dynamic HTML*

DOM: *Document Object Model*

e-swim: *Enterprise Semantic Web Implementation Model*

FOAF: *Friend Of A Friend*

GIF: *Graphics Interchange Format*

GPS: *Global Positioning System*

JPEG: *Joint Photographic Experts Group*

HTML: *HyperText Markup Language*

HTTP: *HyperText Transfer Protocol*

IANA: *Internet Assigned Numbers Authority*

ICANN: *Internet Corporation for Assigned Names and Numbers*

IP: *Internet Protocol*

OWL: acrónimo ajustado de *Web Ontology Language*

RDF: *Resource Description Framework*

SPARQL: acrónimo recursivo de *Sparql Protocol And Rdf Query Language*

SW: *Semantic Web*, Web Semântica

TCP/IP: *Transmission Control Protocol / Internet Protocol*

URI : *Universal Resource Identifier*

URL: *Universal Resource Locator*

XML: *eXtended Markup Language*

W3C: *World Wide Web Consortium*

WHATWG: *Web Hypertext Application Technology Working Group*

WSSC: *Web Science Subject Categorization*

# Índice de Figuras

Figura 2.1 Conceito de vendas "Long Tail" (Anderson, 2008) .....	11
Figura 2.2: Gerações da Web.....	17
Figura 2.3: Comparação entre as Gerações Web e o Ciclo de Vida do Produto .....	18
Figura 2.4: Linked Data – fases para a disponibilização de dados de forma universal (Berners-Lee, 2006a) .....	22
Figura 2.5. Classificação hierárquica de tópicos associados à Ciência da Web.....	27
Figura 3.1: Modelo de Dados - Indivíduos .....	39
Figura 3.2: Modelo de Dados - Publicações.....	39
Figura 3.3: Modelo de Dados - Publicações, tabela única.....	40
Figura 3.4: Modelo de Dados - Docentes e Publicações.....	41
Figura 3.5: Modelo de Dados - Docentes e Publicações - tabela única.....	41
Figura 3.6: Modelo de Dados - Docentes e Publicações - tabela única com URIs...	43
Figura 3.7: Modelo de Dados - Docentes e Publicações - RDF .....	43
Figura 3.8 Exemplo de RDFa para caracterização de um indivíduo (W3C, 2015c)..	44
Figura 3.9 schema.org no formato JSON-LD para o contacto de uma empresa (Google, 2015a) .....	44
Figura 3.10: Tecnologias da Web Semântica .....	47
Figura 3.11 Google Knowledge Graph resultante da pesquisa "porto" .....	52
Figura 3.12 Google Knowledge Graph resultante da pesquisa "porto pontos" .....	52
Figura 3.13 Resultado da pesquisa "a teoria" no Google .....	53
Figura 3.14 Resultado da pesquisa "sonae" no Google .....	54
Figura 3.15 Resultado da pesquisa "sonae" no Google Maps.....	54
Figura 3.16 Resultado da pesquisa "sonae sgps sa" no Google .....	55

Figura 3.17 Pesquisa no Wolfram Alpha para "mortgage \$150,000, 6.5%, 30 years" .....	56
Figura 4.1: Modelo de Implementação da Web Semântica nas Empresas – e-swim .....	74
Figura 4.2. A dimensão Adoção tecnológica .....	76
Figura 4.3. A dimensão Proveniência .....	78
Figura 4.4. A dimensão acessibilidade dos dados (adaptado de (Berners-Lee 2006)) .....	79
Figura 4.5. Caracterização de Tarefas e Atividades segundo a sua Conceptualização e Diversificação (adaptada de (Benjamins et al., 2011)).....	80
Figura 5.1. Web Site do Projeto Raio X, tal como disponibilizado em 23 de Outubro de 2014.....	87
Figura 5.2. Distribuição das empresas por setor de atividade .....	91
Figura 5.3. Distribuição das empresas por Secção da Classificação de Atividades Económicas.....	92
Figura 5.4. Distribuição das empresas por Dimensão .....	92
Figura 5.5. Ferramenta de Avaliação.....	106
Figura 5.6. Avaliação Global: variável e-swim .....	108
Figura 5.7. Avaliação por Dimensão .....	109
Figura 5.8. Avaliação Global por Empresa.....	109
Figura 5.9. Recomendações adotadas por Empresa .....	125
Figura 5.10. Recomendações por Empresa vs. Avaliação Inicial .....	125
Figura 5.11. Frequência de adoção das Recomendações .....	127
Figura 5.12. Avaliação Global: variável e-swim .....	135
Figura 5.13. Avaliação por Dimensão .....	136
Figura 5.14. Avaliação Global por Empresa.....	137

Figura 5.15. Recomendações implementadas por Empresa .....	138
Figura 5.16. Frequência de implementação das Recomendações .....	139
Figura 5.17. Recomendações do modelo e-swim.....	140
Figura 5.18. Tipos de Aplicações identificados no âmbito do projeto .....	141
Figura 5.19. Tipologias de Aplicação Web identificadas no âmbito do projeto....	142
Figura 5.20. Atividades observadas no âmbito do projeto .....	143
Figura 5.21. Atividades observadas no âmbito do projeto, gráfico ordenado.....	144
Figura 5.22. Caracterização de Tarefas e Atividades segundo a sua Conceptualização e Diversificação: uma nova proposta.....	146



# Índice de Tabelas

Tabela 3.1 Exemplo de FOAF .....	45
Tabela 3.2 Exemplo de SPARQL (Feigenbaum, 2009).....	46
Tabela 4.1. Web Semântica: características da Tecnologia e requisitos na Organização para a respetiva adoção .....	71
Tabela 5.1. Lista de Sessões de Divulgação do Projeto Raio X .....	88
Tabela 5.2. Estrutura da Intervenção .....	90
Tabela 5.3. Distribuição das empresas em função da sua dimensão, comparativa entre participantes no projeto e o universo de Empresas Portuguesas e da União Europeia (fonte: (Ecorys & European Commission, 2012; Eurostat, 2012)) .....	93
Tabela 5.4. Estrutura resumida da Ferramenta de Avaliação .....	98
Tabela 5.5 Dimensões e Recomendações do Modelo e-swim .....	112



# 1 Introdução

## 1.1 Enquadramento

Em cerca de duas décadas e meia, a Web evoluiu de uma ideia promissora de um projeto de laboratório e transformou-se num ambiente ubíquo e universal, a que se pode aceder praticamente em qualquer lugar e em que se encontra informação sobre os mais diversos assuntos. A Web dos Conteúdos evoluiu para a Web das Pessoas e dos conteúdos gerados pelo utilizador e deu lugar à Web Social (Fuchs et al., 2010). Mas a Web baseada nos documentos é ainda uma versão pouco estruturada da Web. Com a introdução do acesso aos dados, à sua estrutura e ao seu significado, pode-se ambicionar um ambiente Web mais integrado, robusto e com maior utilidade prática. As próximas duas décadas poderão proporcionar o tempo necessário à disseminação da Web dos Dados, sendo essa a expectativa da atividade científica atual.

A introdução da Web Semântica implica a transformação progressiva da Web baseada nos documentos, em que as hiperligações entre documentos constituíram a essência da primeira geração da Web, na Web baseada nos dados, em que a criação de hiperligações entre dados ou informação dará origem a uma base de dados distribuída à escala da Web (Heath & Bizer, 2011). Esta base de dados potencialmente de grandes dimensões está ainda pouco disseminada (Herman, 2012), embora seja possível encontrar exemplos em número suficiente para avaliar esse potencial, uma vez que alguns esforços de investigação em conjunto com algumas políticas governamentais para a publicação de dados produziram já resultados muito satisfatórios (Bizer, Heath, Idehen, & Berners-Lee, 2008). Apesar destas iniciativas, a utilização da Web Semântica nas Empresas é ainda um tema pouco explorado (Ahmed & Gerhard, 2010; Kuhn, 2010). Entre as várias razões para um reduzido grau de adoção podem referir-se as dificuldades em homogeneizar e validar as fontes de dados, definir regras e fronteiras ao conhecimento que permitam relacionar os dados uniformemente, analisar exemplos demasiado complexos, disponibilizar capacidade tecnológica a baixo custo para viabilizar a sua implementação, encontrar ferramentas de desenvolvimento, recrutar profissionais com

experiência, encontrar referências de sucesso e adotar a mudança de paradigma na modelização, projeto e programação (Ahmed & Gerhard, 2010; Kang et al., 2008; Kuhn, 2010; Pollock & Oracle, 2008).

Tendo em conta estas dificuldades, este trabalho propõe um Modelo de Implementação da Web Semântica nas Empresas com dois propósitos fundamentais. O primeiro propósito é o de facilitar a introdução da tecnologia em organizações com motivações e características diversas. O segundo propósito é o de avaliar o impacto da introdução da tecnologia nas aplicações utilizadas nessas organizações e nas tarefas desempenhadas pelos utilizadores, através da medição do progresso da adoção dessa tecnologia.

## 1.2 Objetivos

O principal objetivo desta tese é a proposta de um Modelo de Implementação da Web Semântica nas Empresas. Este modelo permitirá, numa primeira fase, analisar e determinar o grau de adoção da Web Semântica nas Empresas e, numa fase complementar, servir como guia, propondo um conjunto de recomendações e sistematizando o caminho a seguir para uma adoção mais rápida e intensiva. O modelo funcionará, assim, como um meio facilitador e de promoção da implementação da tecnologia nas Empresas.

O presente trabalho pretende ainda dar resposta a uma questão de investigação central:

- *Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?*

Foram, assim, estabelecidos os seguintes objetivos que, sendo complementares, permitirão a concretização do trabalho proposto:

- Avaliar o grau de adoção da tecnologia Web Semântica nas empresas;
- Desenvolver um Modelo de Implementação da Web Semântica em ambiente empresarial que facilite e sistematize o processo, evidenciando as suas etapas principais;

- Analisar os resultados da aplicação do Modelo de Implementação da Web Semântica e extrair conclusões.

### 1.3 Metodologia

A metodologia adotada para o trabalho de investigação apresentado nesta tese foi a Investigação-Ação (*Action Research*). A escolha desta metodologia teve em conta a necessidade de idealizar uma intervenção junto das empresas que permitisse uma avaliação adequada do grau de adoção da tecnologia Web Semântica e que permitisse simultaneamente uma intervenção por parte do autor, de forma a poder propor alterações à situação verificada que conduzissem a uma nova avaliação.

Assim, tendo em conta a evidência empírica do reduzido grau de adoção da Web Semântica na maior parte das empresas, considerou-se pertinente optar por uma metodologia de intervenção e implementação que permitisse alterar esse grau de adoção para que este pudesse ser posteriormente melhor observável. A estrutura da intervenção baseou-se em dois ciclos de Investigação-Ação interligados, a realizar em cada empresa.

### 1.4 Contribuição Científica

O trabalho reportado nesta tese baseia-se no estado da arte de evolução tecnológica da Web e nos constrangimentos atuais de implementação da tecnologia Web Semântica e apresenta um Modelo de Implementação da Web Semântica nas Empresas que promove e facilita a adoção da tecnologia. Assim, delineou-se um trabalho de intervenção nas empresas em que se analisa a utilização atual da tecnologia, se propõem ações que promovem a sua implementação e se avaliam progressivamente os resultados obtidos.

A proposta de investigação e o progresso do trabalho foram partilhados com a comunidade científica em conferências nacionais e internacionais. Até ao momento, o trabalho apresentado resultou em 4 artigos científicos em que esta tese se baseou:

Ferreira, R., & Seruca, I. (2014). e-swim - Enterprise Semantic Web Implementation Model Towards a systematic approach to implement the Semantic Web in Enterprises. Em Slimane Hammoudi, L. A. Maciaszek, & J. Cordeiro (Eds.), *Proceedings*

*of the 16th International Conference on Enterprise Information Systems (ICEIS 2014)*, Lisbon, Portugal, 27-30 April 2014 (pp. 653–658). Lisbon.

Ferreira, R., & Seruca, I. (2013). Modelo de Implementação da Web Semântica nas Empresas. Em Á. Rocha, L. P. Reis, M. P. Cota, M. Painho, & M. C. Neto (Eds.), *Atas da 8a Conferência Ibérica de Sistemas e Tecnologias de Informação (CISTI'2013)*, Lisboa, Portugal, 19-22 Junho 2013 (pp. 615–620). Lisbon: AISTI.

Ferreira, R., & Seruca, I. (2013). A Ciência da Web : oportunidades de investigação. Em António Guerreiro, P. R. da Silva, & R. Quaresma (Eds.), *Atas da 13a Conferência da Associação Portuguesa de Sistemas de Informação*, Évora, Portugal, 4-5 Outubro 2013 (pp. 13–31). Évora: Fundação Luis de Molina.

Ferreira, R. (2013). Applying Semantic Web technologies to City Tourism information. Em *1st International Conference - Porto as a Tourism Destination*, Porto, 26-28 September. Porto, Portugal: CEPES.

Como o trabalho descrito nesta tese resulta da evolução e do aprofundamento do trabalho descrito nos artigos publicados, estes não foram explicitamente citados, sendo alguns dos conteúdos incorporados no corpo da tese.

O resultado final desta tese consiste no Modelo de Implementação, acompanhado de uma Ferramenta de Avaliação e de um conjunto de Recomendações, que poderão servir de guia e orientação às organizações que desejem implementar a tecnologia.

## 1.5 Estrutura da Tese

Esta tese está organizada em seis capítulos que refletem o percurso do trabalho realizado.

No primeiro capítulo é feito um enquadramento geral ao tema a abordar e contextualizado o trabalho a desenvolver. É indicada a motivação do autor para a realização desta tese, os objetivos que se propôs a atingir, a metodologia a seguir para a concretização do trabalho e a contribuição científica esperada.

O Capítulo 2 inclui uma panorâmica da evolução da *World Wide Web*, caracterizando as várias gerações desde a sua criação até à geração atual, e fornece o

enquadramento científico para a sua evolução, destacando um corpo organizado de conhecimento e de oportunidades de investigação, designado por Ciência da Web.

No Capítulo 3 efetua-se uma revisão do conceito de Web Semântica, das tecnologias associadas e das suas aplicações, sendo destacados exemplos de adoção da Web Semântica em vários setores de atividade. É ainda dada uma panorâmica do estado atual do envolvimento das empresas com esta tecnologia e dos desafios a superar, abrindo caminho à contribuição deste trabalho.

No Capítulo 4 são apresentados a fundamentação teórica e o Modelo de Implementação da Web Semântica proposto neste trabalho.

O Capítulo 5 descreve a utilização do modelo no contexto da intervenção de investigação-ação realizada num conjunto de empresas e termina com a avaliação da intervenção realizada.

O último capítulo apresenta uma síntese do trabalho desenvolvido e uma análise à contribuição efetuada pelo trabalho. É revisitada a questão central de investigação colocada e são feitas considerações aos resultados obtidos. São ainda apresentadas sugestões de trabalho futuro que permitam dar continuidade ao trabalho desenvolvido.



## 2 A Ciência da Web

A World Wide Web transformou-se rapidamente num ambiente ubíquo e universal. No entanto, apesar da velocidade desta evolução, é possível identificar momentos, tecnologias e características que distinguem diferentes gerações da Web e que influenciam a direção da investigação científica, o ritmo de adoção tecnológica e, consequentemente, a sua visibilidade pública. Este capítulo endereça a identificação dessas gerações, descreve o grau de adoção da Web na atualidade e fornece o enquadramento científico para a sua evolução, destacando um corpo organizado de conhecimento e de oportunidades de investigação, designado por Ciência da Web.

### 2.1 A Internet

A ideia de interligar máquinas em rede remonta à década de 60 do século XX, no auge da Guerra Fria, e é frequentemente associada a ter sido impulsionada por pesquisas no âmbito militar, nomeadamente na necessidade, por parte dos Estados Unidos da América, de criar um sistema de comunicações que sobrevivesse a um possível ataque nuclear da URSS (Leiner et al., 1997). No seio da *Advanced Research Projects Agency* (ARPA) do Departamento de Defesa dos EUA, foi idealizada e encomendada a *Arpanet*, a primeira rede de computadores, cuja operação se iniciou em 1969.

Em 1981, com a publicação das especificações finais da família TCP/IP de protocolos, foram criadas as bases para uma rede de computadores de alcance mundial, uma rede à escala planetária ou global, constituída por redes locais interligadas por linhas de comunicação, criando o conceito de Internet. Em 1 de Janeiro de 1983, o TCP/IP passou a ser o único protocolo utilizado na *Arpanet*. No final da década de 80, vários países possuíam já redes próprias baseadas em TCP/IP. Em 1990, a *Arpanet* foi descontinuada (Leiner et al., 1997).

Na essência, esta rede é possível com a atribuição de um endereço único a cada dispositivo. O protocolo Internet, ou IP (*Internet Protocol*), na versão utilizada nas últimas décadas, IPv4, prevê a utilização de endereços individuais para cada dispositivo baseados em números de 32 bits, ou seja, o suporte de cerca de 4 mil milhões,  $4 \times 10^9$ , de endereços

distintos. Para colmatar a escassez de endereços previsível, com o aumento de dispositivos ligados e a extensão da Internet a dispositivos não tradicionais, a introdução do IPv6, em 1998, permitiu a utilização de endereços de 128 bits, ou seja, a cobertura de um número de endereços várias ordens de grandeza superior, cerca de  $3 \times 10^{38}$  endereços distintos.

Os problemas resultantes da necessidade de gerir a atribuição de endereços eram inicialmente confiados à *Internet Assigned Numbers Authority* (IANA), desde os anos 70 (IANA, 2015). Até 1982, os endereços IP eram as únicas representações disponíveis para identificar dispositivos ligados à rede, mas a introdução dos domínios e do *Domain Name System* (DNS) permitiram a criação de endereços mais significativos. Após a criação da *InterNIC* em 1992, esta foi fundida com a IANA em 1998, dando origem à *Internet Corporation for Assigned Names and Numbers* (ICANN), atualmente responsável pela gestão de endereços Internet (ICANN, 2015b).

Até 2020 perspectiva-se a extensão da utilização da Internet a outros planetas, com a implementação do projeto Internet Interplanetária (InterPlanetary Networking Special Interest Group, 2015). No âmbito desse projeto estão a ser testados novos protocolos de comunicação que permitam a comunicação entre a Terra, naves espaciais e outros planetas. Esses novos protocolos são necessários para se fazerem comunicações interplanetárias, uma vez que o atual protocolo TCP/IP, em que se baseia a Internet, depende de uma elevada disponibilidade das ligações, o que não acontece quando os corpos se movimentam a elevadas velocidades e distâncias, como no caso das naves e dos planetas. Os primeiros testes foram já realizados pela NASA com recurso à Estação Espacial Internacional (NASA, 2015).

## 2.2 Web 1.0

### 2.2.1 O nascimento da *World Wide Web*

Há 26 anos, em 1989, nascia no laboratório CERN (Centro Europeu para a Investigação Nuclear), na Suíça, a ideia que veio revolucionar a comunicação em rede e a vida de milhares de pessoas em todo o mundo. A *World Wide Web* foi a solução, na altura encontrada pelo cientista inglês Tim Berners-Lee, para permitir o acesso a um arquivo

comum através de computadores ligados entre si, criando as primeiras aplicações precursoras dos atuais *browsers* e servidores Web. No final de 1990, em apenas 4 meses e com um computador NeXT, a Web ganhou forma (Cailliau, 1995). Na realidade, em 1980, Berners-Lee já tinha apresentado um projeto, o ENQUIRE (Connolly, 2000), cujas premissas essenciais estão na base do conceito da *World Wide Web*. O projeto consistia num programa que permitia escrever texto e criar ligações entre partes distintas de um texto, uma ideia antecessora da invenção da Web criada no CERN. Mas foi talvez com a importante declaração do CERN, em 30 de Abril de 1993, considerando que a Tecnologia Web deveria ser livre e gratuita para todos, que a Web ganhou a forma que lhe permitiu a evolução até à atualidade (Connolly, 2000).

### 2.2.2 [A primeira geração da Web](#)

Os fundamentos básicos da *World Wide Web* referiam-se à teia de documentos e respetivas hiperligações construídos em HTML (*Hyper Text Markup Language*) (Connolly, 2000), publicados em servidores acessíveis a especialistas, destinados apenas a serem consultados pelo público em geral. A primeira geração da Web, ou usando a metáfora geracional habitualmente utilizada no software, a Web 1.0, baseava-se na capacidade de expor o conhecimento, sem pretensão a organizá-lo, deixando-o ao dispor do critério cognitivo individual (Fuchs et al., 2010). Era uma Web só de leitura, em que os utilizadores não podiam adicionar conteúdos, caracterizada por Web Sites maioritariamente estáticos, acedidos com larguras de banda de ordem de grandeza muito baixa, até 48 kbps.

O Web *browser* e o motor de pesquisa são as aplicações mais representativas desta geração, corporizando a plataforma que permitiu massificar a utilização da Web para além da utilização dada pelos especialistas. Por um lado, o Web *browser* permitiu o acesso aos documentos de forma uniforme, independente da tecnologia, da língua, do local de acesso ou do local de armazenamento. Por outro lado, o motor de pesquisa permitiu criar um índice da imensidão de documentos que já estavam disponíveis à escala global, pesquisável da forma mais natural, através da linguagem.

O precursor dos Web *browsers*, o Mosaic, foi criado em 1993, logo seguido por um dos *browsers* mais famosos, o Netscape Navigator, em 1994 e um dos mais utilizados, o

Microsoft Internet Explorer<sup>1</sup>, em 1995. A mudança de geração dos *Web browsers* inicia-se em 1998 com o Mozilla Firefox<sup>2</sup>.

O primeiro motor de pesquisa considerado como tal foi lançado em 1994 e estava acessível em WebCrawler.com. Seguiram-se-lhe em dimensão e importância, a nível internacional, o Altavista e o Yahoo<sup>3</sup> e, a nível nacional, o SAPO – Serviço de Apontadores Portugueses<sup>4</sup>, todos criados em 1995. A mudança de geração dos motores de pesquisa acontece também em 1998 com o aparecimento do Google. Um ano após ser lançado, o Google já respondia a cerca de 3.5 milhões de pesquisas diárias. Atualmente, responde a cerca de 3.5 mil milhões de pesquisas diárias (Internet Live Statistics, 2015).

### 2.2.3 A cauda longa

A exploração comercial da Web foi uma consequência natural da utilização do enorme potencial de comunicação disponibilizado por este serviço da Internet, permitindo a criação de novas empresas em todo o mundo e proporcionando a possibilidade de transformação dos modelos de negócio de inúmeras empresas. A primeira grande oportunidade comercial da Web traduziu-se no conceito de vendas *long tail* (cf. Figura 2.1). Este conceito está associado à seguinte constatação no processo de negócio de uma empresa.

Ao produzir um histograma do número de unidades vendidas por produto num dado segmento de produtos de massa, como por exemplo livros ou discos, verifica-se normalmente que os primeiros produtos, os mais vendidos, os *best-sellers*, apresentam vendas muito elevadas mesmo quando comparados com os que se lhes seguem, originando uma curva com uma descida acentuada. Como o número de produtos distintos em qualquer segmento tende para infinito e o número de unidades vendidas por produto tende para zero, a curva toma uma forma de cauda longa. Sendo difícil disponibilizar todos os produtos possíveis, o comércio optava por ter disponíveis os produtos mais vendidos e dispensava produtos cujo volume de vendas não fosse suficiente para justificar o investimento em inventário necessário.

---

<sup>1</sup> [windows.microsoft.com/en-us/internet-explorer/download-ie](http://windows.microsoft.com/en-us/internet-explorer/download-ie)

<sup>2</sup> [www.mozilla.org/pt-PT/firefox/new/](http://www.mozilla.org/pt-PT/firefox/new/)

<sup>3</sup> [search.yahoo.com/](http://search.yahoo.com/)

<sup>4</sup> [www.sapo.pt/](http://www.sapo.pt/)

# Long Tail

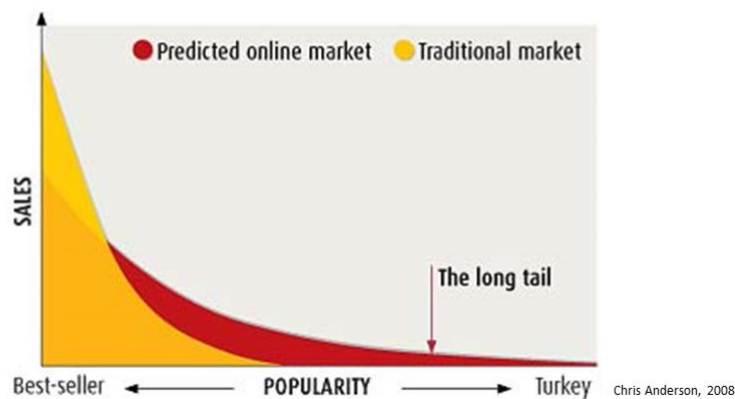


Figura 2.1 Conceito de vendas "Long Tail" (Anderson, 2008)

A ideia fundamental associada ao modelo de negócio “cauda longa” pressupõe ter disponível um número quase infinito de produtos, pelo que a venda desses produtos também tende quase para o infinito, dando origem a negócios de grande dimensão que exploram todas as possibilidades de venda, todos os segmentos, todos os nichos, ainda que o investimento em stocks seja partilhado com outros investidores de menor dimensão, todos suportados numa rede de informação integrada acessível na Web. O tema é largamente abordado numa obra de referência precisamente denominada “A Cauda Longa” (Anderson, 2007).

Grandes empresas “cauda longa” nasceram na primeira geração da Web e ainda se mantêm com enorme visibilidade, nomeadamente a Amazon<sup>1</sup> e a ebay<sup>2</sup>, ambas criadas em 1995, a Netflix<sup>3</sup>, criada em 1997, a Google<sup>4</sup> criada em 1998, a Audible<sup>5</sup>, criada em

---

<sup>1</sup> [www.amazon.co.uk](http://www.amazon.co.uk)

<sup>2</sup> [www.ebay.pt](http://www.ebay.pt)

<sup>3</sup> [www.netflix.com](http://www.netflix.com)

<sup>4</sup> [www.google.pt](http://www.google.pt)

<sup>5</sup> [www.audible.com](http://www.audible.com)

1999 e posteriormente adquirida pela Amazon, ou, mais recentemente, o serviço iTunes da Apple<sup>1</sup>, criado em 2003.

A Web 1.0 é frequentemente referida como a Web dos documentos (Bizer, Heath, & Berners-Lee, 2009), uma descrição que traduz não só o seu propósito inicial de publicar e interligar documentos como também o facto de ser constituída essencialmente por conteúdos estáticos, que os utilizadores não podiam alterar.

## 2.3 Web 2.0

A Web 2.0 situa-se em termos temporais nos anos compreendidos entre 2004 e 2008 e traduz essencialmente a geração das redes sociais, o conceito de comunidade, de livre adesão a interesses partilhados, uma Web Social como sociais são todas as interações assumidas entre indivíduos distintos, um passo na aproximação entre a Web e o mundo como um todo. Mas ainda assim, apenas uma Web de comunicação (Fuchs et al., 2010).

Sem pretensão de menosprezar o sucesso das redes sociais, por ventura a imagem mais atual da Internet, em muito devido a um fenómeno designado por Facebook<sup>2</sup>, rede social criada em 2004 e atualmente com mais de 1300 milhões de utilizadores em todo o mundo (Statista, 2015), esta geração da Web contribuiu apenas para a sua democratização ou mesmo liberalização, transformando-a num espaço de conteúdos gerados pelo utilizador, onde todos publicam para todos. É uma Web não só de leitura, como acontecia na geração anterior, mas também de escrita. Outros exemplos caracterizam bem esta geração, como o da rede social dedicada a profissionais, LinkedIn<sup>3</sup>, fundada em 2003, atualmente com mais de 300 milhões de utilizadores (Statista, 2015) ou a proliferação de blogs, que se estima serem acima de 200 milhões (Statista, 2011). Pode considerar-se que esta geração atingiu atualmente a sua maturidade.

No limiar da Web 3.0, uma Web de colaboração, de interações individuais que contribuem para um bem comum (Fuchs et al., 2010), talvez o conceito de *Wiki* pudesse ter sido o conceito que melhor poderia traduzir essa nova geração, mas apesar dos mais

---

<sup>1</sup> [www.itunes.com](http://www.itunes.com)

<sup>2</sup> [www.facebook.com](http://www.facebook.com)

<sup>3</sup> [www.Linkedin.pt](http://www.Linkedin.pt)

de 4 milhões de artigos em inglês e 800 mil em português da Wikipedia<sup>1</sup>, a enciclopédia fundada em 2001, a dificuldade em verificar a fonte de algumas informações mantém-se como um problema sensível, particularmente em assuntos que não são de interesse alargado a muitas pessoas. Assim, pode considerar-se que não traduz a melhor expressão da Web de colaboração e, por isso, o conceito de *Wiki* mantém-se como representativo da geração 2.0. No entanto, pode considerar-se que merece uma referência particular, como representativo da geração intermédia que poderia designar-se por geração 2.5, quer pelo papel fundamental que tem na rapidez de acesso a um nível de conhecimento por vezes superficial mas seguramente universal, quer pelo potencial de evolução que encerra e que poderá permitir uma evolução para a geração Web 3.0. A consulta da Wikipedia é por si uma experiência que justifica a sua referência fundamental na Web atual, sendo suficiente procurar alguns termos relevantes para verificar a sua abrangência. Um bom exemplo é o da própria definição de Web 2.0 (Wikipedia, 2015), que mostra que alguns temas podem ser abordados de forma completa mas simultaneamente discutível. Com o anúncio do projeto Wikidata<sup>2</sup>, a organização criadora da enciclopédia pretende desenvolver um acesso centralizado a dados, legível a pessoas e computadores, que poderá reposicioná-la facilmente numa geração mais avançada. Atualmente, este repositório contabiliza mais de 200 milhões de edições em mais de 13 milhões de entidades (The Wikimedia Foundation, 2015). Adicionalmente, o projeto DBPedia<sup>3</sup> toma como base a própria Wikipedia para extrair informação estruturada e a disponibilizar na Web e conta com 4.5 milhões de entidades classificadas que são acedidas mais de 3 milhões de vezes por dia (Lehmann et al., 2012).

Outros exemplos desta geração merecem destaque, como o do YouTube<sup>4</sup>, a rede de publicação de vídeo criada em 2005 e que se transformou no grande arquivo de vídeo da Internet, com mais de 2 mil milhões de vídeos vistos por dia (Internet Live Stats, 2015), adquirida pela Google em 2006, e o Twitter<sup>5</sup>, o serviço de mensagens curtas criado em

---

<sup>1</sup> [www.wikipedia.org](http://www.wikipedia.org)

<sup>2</sup> [www.wikidata.org](http://www.wikidata.org)

<sup>3</sup> [www.dbpedia.org](http://www.dbpedia.org)

<sup>4</sup> [www.youtube.com](http://www.youtube.com)

<sup>5</sup> [twitter.com](http://twitter.com)

2006 e que conta com mais de 300 milhões de utilizadores ativos (Internet Live Stats, 2015).

Os utilizadores passaram a dispor de largura de banda na ordem de 1 Mbps e o acesso à Internet através de dispositivos móveis começou a proliferar.

## 2.4 Web 3.0

A geração Web 3.0 é referida com a Web dos Dados, uma vez que, tirando partido das tecnologias da Web Semântica, pretende capitalizar tanto no significado dos dados como na sua integração (Bizer et al., 2009). A proliferação de dispositivos de utilização individual e de dispositivos mais ou menos autónomos, desde sensores a veículos inteligentes, e o papel preponderante do contexto da utilização levam a que a Web 3.0 possa também ser considerada a Web das Coisas (W3C, 2015f) ou a Web da Inteligência (Shroff, 2013).

A linguagem base da Web, o HTML, deu origem a várias evoluções. A primeira e fundamental com a introdução do XML (*eXtended Markup Language*), um formato definido para a criação de documentos estruturados que possam ser lidos por máquinas e que, em conjunto com um esquema gráfico, possam ser também lidos pelas pessoas. A introdução do XML permitiu disseminar um formato aberto, simples, de utilização generalizada, compatível com qualquer língua e adequado à representação de qualquer estrutura de dados (W3C, 2008b).

Outra evolução, quer do HTML, quer do XML, responde pelo acrónimo RDF (*Resource Description Framework*), uma estrutura que permite que os dados possam ser identificados, publicados, reutilizados e interligados com recurso aos mesmos endereços que se utilizam para identificar a localização de um documento. O RDF está, assim, na base da concretização da Web dos dados. As tecnologias de suporte à Web Semântica serão abordadas em mais detalhe no Capítulo 3.

Para que a Web 3.0 se concretize plenamente, espera-se a ajuda da Web Semântica e, novamente, de Tim Berners-Lee (Shannon, 2006). Os próximos anos serão testemunha do seu crescimento (Spivack, 2010). A Web 3.0 será para os dados o que foi a

Web 1.0 para os documentos, interligando e dando significado às mais diversas fontes de dados, disponibilizando-as a cada vez mais dispositivos.

No entanto, a importância e a viabilidade da Web Semântica é um tema onde não existe consenso absoluto. Por um lado, alguns autores salientam o crescimento verificado e o potencial futuro da Web dos dados, confiando na importância, na aplicabilidade e na disseminação do conceito (Bizer et al., 2008). Por outro lado, outros autores salientam a dificuldade em homogeneizar e validar as fontes de dados, em definir regras e fronteiras ao conhecimento que permitam relacionar os dados uniformemente e mesmo em disponibilizar capacidade tecnológica a baixo custo para viabilizar a sua implementação (Ahmed & Gerhard, 2010; Kuhn, 2010).

O desafio da Web Semântica é mais complexo do que quando abordado apenas na perspectiva tecnológica. A modelização conceptual prometida pela Web Semântica pode ser bem mais complicada do que aparenta (Kuhn, 2010). No entanto, poderá servir como ponte entre todo o tipo de informação, independentemente da sua fonte ou origem, um ponto de partida para agregar toda a informação (Hotho & Stumme, 2010).

O W3C considerou, no final de 2009, que a Web Semântica estaria no final da fase de adoção precoce (*early adopters phase*) (Herman, 2012), com algumas empresas a começar a explorar este recurso e nalguns casos a publicarem os seus próprios dados (Shadbolt, 2014), mas a realidade de utilização alargada da Web Semântica em processos de negócio poderá acontecer apenas até ao final da década (Benjamins et al., 2011). Salienta-se a opinião divergente entre investigadores e indústria, com os primeiros a estimar a adoção generalizada da tecnologia num prazo menor que os últimos (Cuel, Delteil, Louis, & Rizzi, 2004).

Um dos melhores exemplos de utilização potencial de dados disponíveis na Web é a abertura de dados à utilização pública proporcionada pela iniciativa “Data.gov.uk” do governo britânico, que atualmente apresenta mais de 20 mil conjuntos de dados sobre as mais variadas matérias (Data.gov.uk, 2015). Nesta iniciativa, não só os dados são disponibilizados em bruto como também são divulgadas aplicações criadas por empresas ou indivíduos para aceder facilmente a esses dados, como por exemplo uma aplicação que lista e localiza todas as escolas no Reino Unido, para que os pais possam encontrar a

melhor opção para os seus filhos. Deve referir-se também a iniciativa portuguesa congénere, denominada “dados.gov.pt”, com pretensões semelhantes mas com um número de catálogos de dados e aplicações ainda muito reduzido (Agência para a Modernização Administrativa, 2015).

A informação empresarial tem recebido particular atenção em termos mundiais, tendo em conta as necessidades de agregação de diversos dados económicos nacionais, regionais e globais. Uma empresa inglesa lançou-se na missão de disponibilizar um URL para cada empresa em todo o mundo. O resultado está disponível de forma aberta em “opencorporates.com”, por enquanto parcial, “limitado” à consulta de informação em mais de 80 milhões de empresas. Num relatório que avalia a disponibilidade dessa informação nos países de todo o mundo, destacam-se o Reino Unido, a República Checa, a Eslováquia e a Albânia (OpenCorporates, 2012).

Um exemplo relevante de utilização da Web Semântica é o do Wolfram Alpha<sup>1</sup>, que pode considerar-se um motor de pesquisa com base linguística e semântica. O sistema assume que o utilizador pretende obter determinados resultados, considerando a forma como é feita a pesquisa, sendo os resultados apresentados de forma hierarquizada, a partir de dados estruturados. Uma pesquisa demonstrativa desta capacidade pode ser feita introduzindo a expressão “*distance from porto to rome*”: o resultado devolvido assume que “*porto*” se refere à cidade do Porto, localizada em Portugal, que “*rome*” se refere à cidade de Roma, localizada em Itália, fornecendo como resposta 1760 km. O potencial da abordagem semântica pode ainda ser verificado com a apresentação de resultados adicionais, como a distância no sistema inglês (1094 milhas) ou o tempo aproximado de um voo direto de avião (2 horas). Finalmente pode verificar-se que reduzindo a pesquisa a “*porto rome*”, os resultados são entendidos primariamente como um pedido de comparação entre as duas cidades, pelo que primeiro é apresentada uma tabela comparativa e de seguida a distância entre as duas cidades, entre outra informação.

---

<sup>1</sup> [www.wolframalpha.com](http://www.wolframalpha.com)

### 2.4.1 A geração atual da Web

Por tudo isto, pode considerar-se que a Web está atualmente na sua geração 3.2, uma quantificação que reflete não só as realizações conseguidas com a disponibilização de dados abertos, para utilização individual ou para tratamento por outras aplicações, mas também o facto de a esmagadora maioria das aplicações estar ainda por explorar. No entanto, esta quantificação simplista, uma espécie de média estatística, esconde a dispersão que a Web apresenta realmente, pois é simples verificar que inúmeras presenças Web são presenças meramente de primeira geração, com disponibilização de conteúdos limitada, como também se podem encontrar aplicações da nova geração completamente funcionais, com recurso dinâmico a fontes de dados atualizadas.

Finalmente, é possível antecipar que no futuro surgirá a Web 4.0, provavelmente com a aplicação da Inteligência Artificial à utilização individual, servindo de suporte maciço e exclusivo a todo o software, como um *WebOS*, um Sistema Operativo que se confunde com a própria Web, onde dados e aplicações estão baseadas. Será provavelmente uma grande evolução a ocorrer nos próximos 10 anos e que transformará definitivamente a Web na rede omnipresente que se começa a detetar (Spivack, 2010).

## Gerações da Web

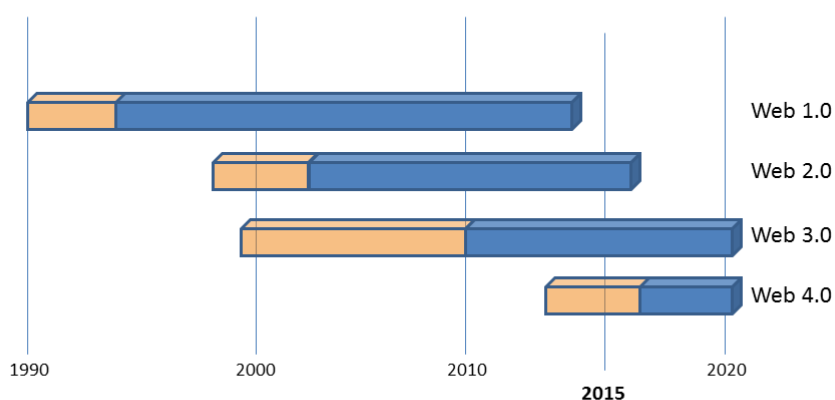


Figura 2.2: Gerações da Web

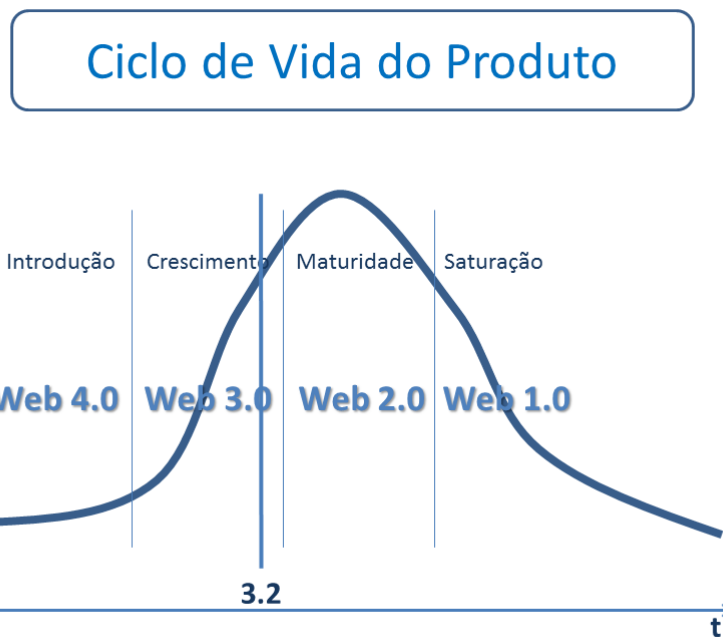


Figura 2.3: Comparação entre as Gerações Web e o Ciclo de Vida do Produto

A metáfora geracional utilizada permite enquadrar a evolução da Web de forma mais sucinta e entendê-la na perspetiva cronológica (cf. Figura 2.2) e económica.

A Figura 2.3 ilustra essa evolução e compara-a com o conceito de ciclo de vida do produto. Será consensual considerar que a Web 1.0 está atualmente plenamente implementada e relegada para segundo plano pela Web 2.0, pelo que é aceitável classificá-la principalmente na fase de Saturação. Refira-se que esta fase é muitas vezes vista como a fase do declínio, embora em muitos produtos se verifique a sua grande longevidade e portanto pareça mais correto falar apenas de Saturação, ou mais assertivamente, de disseminação plena. A Web 2.0, objeto de tanta atenção, encontra-se seguramente em fase de Maturidade, conforme já referido. A Web 3.0, ainda jovem mas com resultados visíveis, poderá considerar-se na fase de Crescimento. E, finalmente, a Web 4.0 poderá considerar-se na fase de Introdução, talvez mesmo precoce. A abcissa 3.2 coloca ênfase no estado atual da Web e salienta ainda o seu potencial de crescimento.

#### 2.4.2 [A Internet das coisas](#)

A evolução sofrida pela Web permite antever que o seu crescimento começa também a ser “invisível”, pois a “Internet das pessoas” começa a dar lugar à “Internet das

coisas” (*Internet of Things*), devido à introdução de dispositivos diversos com capacidade de se conectarem à internet e transmitirem dados, criando a Web Ubíqua (Janowicz & Hitzler, 2010b).

Atualmente acedem à Internet quase 3 bilhões de pessoas em todo o mundo, 40% da população mundial, número que chega aos 78% em países desenvolvidos. Com cerca de 125 telefones móveis ativos em cada 100 habitantes, com mais de 75% das pessoas e das casas com acesso à Internet (International Telecommunication Union, 2014), os Europeus experimentam já a omnipresença da Web. As duzentas mil mensagens escritas (SMS) enviadas por segundo em todo o mundo (International Telecommunications Union, 2010) caracterizam bem esta ideia de as pessoas estarem permanentemente ligadas. Embora de facto o envio destas mensagens não seja efetuado pela Internet, a quarta geração de comunicações móveis terá já por base a Internet: a especificação para a nova geração inclui uma rede baseada em *Internet Protocol* (IP), com velocidades de 100Mbps em utilização de elevada mobilidade (carro ou comboio, por exemplo) ou de 1Gbps em utilização de mobilidade reduzida e ainda a capacidade de ligação de proximidade, ponto a ponto (ITU, 2008). Atualmente, os operadores disponibilizaram já nalgumas regiões o que podem ser considerados equipamentos 3.9G, porque se aproximam muito desta especificação, embora não a satisfaçam completamente.

Em “A Internet das Coisas – um Plano de Ação para a Europa” (European Commission, 2009), a Comissão Europeia estabelece prioridades para a política europeia em relação ao desenvolvimento esperado da Internet. Como exemplos atuais são citados a utilização de telemóveis para identificar produtos a partir da imagem obtida pela câmara incorporada, a utilização de etiquetagem individualizada para rastreio (utilizando *Radio Frequency Identification*, RFID), a implementação de contadores inteligentes no fornecimento de energia elétrica para monitorização instantânea ou os “objetos inteligentes” utilizados no ciclo produtivo e de logística para aumentar a eficácia.

Para que não se torne definitiva a confusão entre os conceitos de Internet e Web que muitas vezes é feita, deve referir-se que a Web representa apenas cerca de 25% do tráfego na Internet, relegada ao segundo lugar pelo serviço P2P (*peer to peer*), serviço que suporta a transferência direta de dados entre dois computadores e que lidera com cerca de 50% do tráfego (Schulze & Mochalski, 2009). Apesar desse domínio ainda se

manter, nos últimos anos o peso do P2P tem diminuído, à custa do crescimento da própria Web e do *streaming* (Web sites como o YouTube são classificados como *streaming* e representam quase 10% do total do tráfego da Internet). O crescimento da Web é evidente, pela utilização mais frequente, pela utilização de mais largura de banda e pela proliferação da partilha de ficheiros, esta última a representar quase metade do tráfego Web, cerca de 10% do tráfego global na Internet.

O cenário que pode ser experimentado na atualidade é o de uma rede global de acesso móvel, com uma cobertura extensiva, um conjunto de aplicações de utilização generalizada, com acesso às mais variadas fontes de dados, com praticamente todos os indivíduos ligados. A única diferença entre o que pode ser experimentado hoje e o que se poderá experimentar daqui a poucos anos está na quantidade de dados a que é possível aceder e na forma como os mesmos serão utilizados. Daí a atualidade do termo “Internet das Coisas”, onde tudo pode estar ligado, onde os mais improváveis equipamentos comunicam permanentemente e geram ainda mais dados.

A informação transforma-se em conhecimento quando é interpretada por uma pessoa num determinado contexto (Nonaka, Toyama, & Konno, 2000). Para além do utilizador, do registo das suas ações, das suas preferências, o contexto inclui informação adicional sobre a localização, o ambiente, entre outras “coisas” relacionadas com o utilizador. A Internet das Coisas é atualmente constituída por cerca de 1200 milhões de dispositivos, com uma previsão de crescimento para 5400 milhões até 2020 (Verizon, 2015).

A combinação da Web Semântica com a Internet das Coisas fornece uma oportunidade para avançar na compreensão, gestão e utilização de fontes de dados provenientes de sensores, um tópico já a ser explorado mas com vários desafios por enfrentar (Corcho & García-Castro, 2010; Janowicz & Hitzler, 2010b). Pode antever-se o enorme crescimento da quantidade de dados utilizáveis que podem ser analisados quando a sua produção depender apenas de máquinas simples, como sensores de temperatura ou câmaras de vigilância, fornecerem um fluxo ilimitado de dados, com todas as consequências que daí se poderão extrair em termos de dimensão, qualidade ou capacidade de cálculo e armazenamento.

Reconhecendo a necessidade de abrir as restrições à utilização de tecnologia associada à Internet das Coisas ao potencial de utilização da Web, e em particular das Tecnologias da Web Semântica, na interligação de pessoas, locais, eventos, dispositivos, reais ou virtuais, o W3C criou em Janeiro de 2015 o *Web of Things Interest Group*, um projeto financiado pelo 7º Programa Quadro de Investigação e Desenvolvimento Tecnológico da União Europeia (W3C, 2015f).

### 2.4.3 Dados Universais

A premissa subjacente à Web Semântica é de permitir que o significado dos dados seja conhecido, identificando assim as “coisas” a que se referem. Por um lado, a Web das coisas tomará forma tirando partido da Web dos documentos, tal como é conhecida atualmente, que por sua vez tira partido da Web dos computadores, que por sua vez tira partido da rede de cabos (e ondas) (Berners-Lee, 2010). Por outro lado, os dados não têm que estar presos no interior dos documentos e podem formar ligações que sejam úteis sem o suporte documental, recorrendo, por exemplo, a um “*browser* de dados” e dando origem ao conceito de *Linked Data* (Berners-Lee, 2006a).

O caminho para a disponibilização dos dados de forma universal passa por várias fases, tendo Berners-Lee criado mesmo uma classificação com cinco estrelas que o caracteriza, apresentada na Figura 2.4.

A primeira fase visa disponibilizar os dados na Internet, para utilização livre. Tendo em conta a habitual confusão entre as diferentes autorizações de utilização, por livre entenda-se mesmo sem qualquer restrição. Alguns desafios se colocam nesta matéria, como sejam os da utilização indesejável dos dados, como na espionagem ou no terrorismo, ou da utilização excessiva, que implique sobrecarga dos sistemas.

A segunda fase consiste em disponibilizar esses dados de forma a que as máquinas os entendam. Um exemplo clássico será o de substituir a imagem de uma tabela, por exemplo no formato JPEG, por uma tabela no formato do Excel. Os desafios passarão então a ser estruturais, quer quanto ao entendimento da codificação da estrutura, quer quanto à própria qualidade da mesma.

## Linked Data: o caminho



Figura 2.4: Linked Data – fases para a disponibilização de dados de forma universal (Berners-Lee, 2006a)

A terceira fase será a de disponibilizar os mesmos dados numa estrutura não proprietária. Seguindo o mesmo exemplo, pode substituir-se a tabela no formato do Excel por uma estrutura de dados simples mas generalizada, como a CSV (*Comma Separated Values*).

A quarta fase permite adicionar a capacidade semântica, utilizando standards do W3C como o RDF e o SPARQL (*Sparql Protocol And Rdf Query Language*). Os dados publicados podem assim passar a ser referidos por terceiros, usando endereços que indicam não apenas a sua localização mas também a sua identidade.

Finalmente, a quinta fase consiste na ligação desses mesmos dados aos dados de terceiros, proporcionando contexto e validação. Cria-se, assim, neste último patamar, um cenário de dados interligados reciprocamente que corporizaria as ambições da Web Semântica.

Uma organização onde a criação de dados esteja fortemente ligada à sua partilha poderá ser designada “*Linked Data Enterprise*”. A utilização de vocabulários controlados ou ontologias está na base da criação de organizações com essa característica (Allemang, 2010).

Em paralelo, a abertura dos dados públicos aos cidadãos representa um avanço na transparência democrática possibilitado pela disponibilidade tecnológica. Destacam-se os esforços dos governos americano e britânico que levaram já à disponibilização alargada de fontes de dados com utilidade generalizada. A necessidade de explorar estas fontes de dados revela-se antes de mais como uma possibilidade de explorar fontes de riqueza (Koumenides, Salvadores, Alani, & Shadbolt, 2010). Os países abrem-se à colaboração, pelo que podem considerar-se no limiar da era das “*Wikinations*”, termo que tira partido da popularidade da Wikipedia e da obra *Wikinomics* (Tapscott & Williams, 2007), uma referência na literatura de gestão atual e que aborda tendências como a globalização e a utilização maciça da Web.

Ao chamar desafio ao percurso até à concretização da Web Semântica, considera-se a necessidade de investigar sobre a forma de integrar as diversas e fechadas fontes de informação existentes (Berners-Lee, Hendler, Hall, Shadbolt, & Weitzner, 2006).

#### 2.4.4 [A economia da Web](#)

Com frequência, nas mais diversas áreas, as dimensões científica, tecnológica e económica, entre outras, cruzam-se e dão origem a evoluções que não são consensuais. Uma teoria científica não tem necessariamente consequência em termos de produção de uma tecnologia, tal como uma tecnologia não é necessariamente adotada pela economia, entre outras possibilidades. A responsabilidade por essas evoluções é mesmo difícil de determinar, tais são as múltiplas influências entre as diferentes dimensões.

A dimensão económica da Web é, como em todas as áreas, determinante. Foi assim na geração Web 1.0, com a emergência das empresas “*cauda longa*”. Foi igualmente assim no crescimento da Web 2.0, com as redes sociais a serem exploradas na perspetiva da ligação das pessoas às empresas, às marcas e aos produtos. E será eventualmente assim com a Web 3.0, com as empresas a envolverem-se ativamente na sua construção.

Exemplos como a evolução da sistematização e utilização dos domínios salientam divergências e convergências entre as pretensões económicas e tecnológicas ou científicas. Se inicialmente os domínios de topo foram criados para identificar as orientações dos Web Sites e os países que representam, a realidade é que nos Estados

Unidos as empresas adotaram de forma hegemónica os domínios .com, .net e .org em detrimento dos mais específicos .us, .com.us, .net.us e .org.us. Não tardou muito para que se desenvolvessem outras práticas que subvertem os objetivos iniciais, entre as quais se destaca o domínio .tv que se deveria referir às Ilhas Tuvalu, no Pacífico, mas que em 2000 vendeu os direitos de utilização a uma empresa americana por 4 milhões de dólares anuais para os comercializar em projetos relacionados com televisão (Black, 2000). Seguiram-se o exemplo do agregador de hiperligações favoritas Delicious<sup>1</sup>, adquirido pela Yahoo<sup>2</sup> em 2005, que surgiu com o endereço del.icio.us, este curiosamente registado no país adequado. Mais recentemente, em 2012, é possível citar o exemplo da liberalização do domínio .pt promovida pela FCCN<sup>3</sup> em Portugal, que deu origem ao registo de quase 40000 endereços num só mês (dns.pt, 2015), aproximadamente o que se registava num ano completo. O endereço wra.pt é um dos exemplos de endereço registado com aproveitamento da linguística (*wra*pt significa embrulhado): o endereço foi utilizado por uma empresa americana dedicada a ofertas para o público masculino que o utiliza para criar a ilusão de embrulhar um Web Site<sup>4</sup>. Finalmente, ainda em 2012, a ICANN abriu concurso para leilão e criação de novos domínios, para registo de endereços a atribuir por empresas privadas, tendo concorrido 1154 empresas que apresentaram 1930 pedidos, num total de quase 400 milhões de dólares apenas em caucões (Ferreira, 2012). O processo ainda está em curso, tendo até ao momento sido contratada a criação de mais de 500 novos domínios de topo (ICANN, 2015a).

Estes exemplos ilustram bem o que é a economia da Web: rápida, volátil e imprevisível. A dimensão económica contribuirá também, seguramente, para o estudo da Web.

#### 2.4.5 [A Ciência da Web](#)

A breve história da Web não impede que se encontrem factos relevantes para o seu estudo e a sua evolução potencial convida a que se o faça, dando origem ao que se começou a designar por *Web Science*, Ciência da Web, uma ciência que resulta da

---

<sup>1</sup> del.icio.us ou delicious.com

<sup>2</sup> www.yahoo.com

<sup>3</sup> www.fccn.pt

<sup>4</sup> wra.pt

combinação de conteúdos, técnicas e competências em Matemática, Física, Ciência dos Computadores, Psicologia, Ecologia, Sociologia, Direito, Ciências Políticas e Economia, entre outras disciplinas (Shadbolt & Berners-Lee, 2008). A ciência é necessária para analisar e sintetizar e, frequentemente, a capacidade de analisar e perceber ultrapassa o potencial da criação pura. A Ciência da Web é certamente interdisciplinar, atribuindo-lhe o endereçamento de algumas das questões mais desafiantes e intrigantes do século XXI (Berners-Lee, Weitzner, et al., 2006). A compreensão dos temas emergentes requer frequentemente uma análise interdisciplinar das suas várias funcionalidades (O'Hara, Contractor, Hall, Hendler, & Shadbolt, 2013).

Na origem da Ciência da Web estiveram naturalmente trabalhos no âmbito da Ciência dos Computadores, que contribuíram para a sua génese, bem como a sua componente Matemática. Como exemplo porventura mais popular, o trabalho dos fundadores da Google teve origem em Stanford e deu origem ao motor de busca mais utilizado em todo o mundo (Brin & Page, 1998). De forma mais discreta, um americano de origem húngara tem desenvolvido a teoria das redes livres de escala (*scale-free networks*), explicando a sua formação através da expansão contínua e da adesão preferencial (Barabasi, 1999), o que permite compreender como as redes se desenvolvem, com aplicação às mais diversas áreas, incluindo a biologia, as redes sociais ou a Web como um todo.

Então o que pode um indivíduo fazer para participar neste desafio? A pergunta pertinente tem uma resposta tão vasta quanto as áreas que envolve, como já foi referido neste trabalho.

No documento “A Internet das Coisas – um Plano de Ação para a Europa” (European Commission, 2009), a Comissão Europeia refere, na Linha de ação 7, Investigação e desenvolvimento, que “continuará a financiar, no âmbito do sétimo Programa-Quadro, projetos de investigação no domínio da Internet das Coisas, privilegiando, por um lado, domínios tecnológicos importantes, como a microeletrónica, os componentes sem silício, as tecnologias de colheita de energia (*harvesting technologies*), as tecnologias ubíquas de determinação da posição, as redes de sistemas inteligentes que comunicam sem fios, a semântica, as tecnologias que, logo na conceção, prevejam a proteção da privacidade e a segurança e o software que rivalize com o

raciocínio humano, e, por outro, aplicações inéditas”. Neste mesmo documento, antecipa-se a iniciativa “*Future Internet*”, uma Parceria Público-Privada de integração dos esforços de Investigação e Desenvolvimento em Tecnologias de Informação e Comunicação em relação ao futuro da Internet. No âmbito desta iniciativa, mais de 150 projetos de investigação associaram-se no sentido de reforçar a competitividade europeia no mercado global. Este foco no investimento europeu continua no âmbito do programa Horizonte 2020, em particular da Agenda Digital para a Europa, considerando a evolução da Internet para um meio de serviços, coisas e infraestrutura que inclui aparelhos inteligentes que comunicam entre si, roupas que monitorizam a nossa saúde, carros que não colidem e tecnologias móveis e plataformas na nuvem que gerem os nossos negócios (European Commission, 2015).

A multidisciplinaridade da Web atual proporciona uma infinidade de caminhos de investigação, resultantes das múltiplas combinações de tópicos possíveis. Assim, existe um grande número de investigadores dedicados à área, em vários centros de investigação, em Universidades e Empresas, reunidos em conferências e encontros, proporcionando um ambiente de aprendizagem global, como seria de esperar, em especial tratando-se desta matéria. Num esforço para identificar alguns desses centros e investigadores, ainda que de forma não exaustiva, e de proporcionar a imersão do autor nessa rede, foi criada uma lista, no âmbito deste trabalho, onde se identificam indivíduos e locais de interesse para o estudo desta área, apresentada no Anexo 1. A concretização desse trabalho de *networking* enriquecerá o trabalho aqui proposto.

## *Web Science Subject Categorization*

- Geral (...)
- História da Web e Metodologia (...)
- Tecnologias Web (...)
- Análise da Web (...)
- A Sociedade da Web
  - Economia e Negócios (...)
  - Envolvimento Social e Ciências Sociais (...)
  - Envolvimento pessoal e Psicologia (...)
  - Filosofia (...)
  - Direito (...)
  - Política e Governança (...)
- Ensinando sobre a Web

Figura 2.5. Classificação hierárquica de tópicos associados à Ciência da Web

A Ciência da Web toma entretanto a sua primeira forma, denominada *Web Science Subject Categorization* (WSSC), um sistema de classificação destinado a facilitar a comunicação e a colaboração entre investigadores da área (Vafopoulos, 2011), através de uma organização temática das diferentes disciplinas. No Anexo 2 apresenta-se uma versão dessa categorização em Português, onde alguns termos foram propositadamente mantidos em Inglês para melhor compreensão geral. A Figura 2.5 exhibe o primeiro nível da classificação e o detalhe de segundo nível para o tópico “A Sociedade da Web”.

### 2.4.6 Oportunidades de Investigação na Ciência da Web

A identificação das áreas de investigação relevantes para a Ciência da Web é referida por vários autores, com destaque para Nigel Shadbolt (Shadbolt, 2008). O trabalho de vários investigadores contribui diariamente para responder a alguns aspetos das áreas identificadas e para especificar outros onde é necessária intervenção. A pesquisa realizada no âmbito deste trabalho identificou ainda aspetos adicionais onde essa intervenção seria desejável.

Nos últimos anos tem-se verificado uma abordagem generalizada dos investigadores à evolução da Web, nos seus mais diferentes aspetos. A abordagem foi tão generalizada que dificulta mesmo a análise da progressão da própria investigação, através

do contacto simultâneo e, por vezes, indistinto com momentos de descoberta científica, de discussão filosófica ou de promoção comercial. Pode mesmo considerar-se que se vive um momento de euforia superior ao que se verificou na viragem de século, referido habitualmente como o “boom das dot com”, pois como em todos os momentos de euforia, se sobrepõe a emoção aos factos e se distorce facilmente a realidade. Se é verdade que a história se repete e seria possível prever uma nova queda como a que aconteceu na viragem de século, com o conseqüente arrefecimento do interesse generalizado sobre a Internet, também é verdade que o momento económico catastrófico vivido a nível global é também uma repetição e que em nada se fica a dever à Internet e, bem pelo contrário, teve na origem a economia mais tradicional, com destaque para os sectores da banca e imobiliário. Outros factos parecem entretanto suportar uma importância crescente da Web como envolvente social e económica, entre os quais a revolução em vários países árabes, e, em menor escala, as revoltas populares em vários países europeus, como Inglaterra, Espanha e Portugal, em que a Web desempenhou um papel essencial no suporte à subversão. Precisamente o tipo de factos que colocam muitas perguntas e lançam um ruído generalizado que em nada facilita a tarefa daqueles que pretendem antes de mais identificar e responder às grandes questões científicas.

A dificuldade em identificar as áreas de investigação relevantes na Ciência da Web parece não surgir pela falta de questões pertinentes, mas sim pelo excesso de oportunidades. O esforço de identificação e avaliação de oportunidades deve assim ser orientado por uma estrutura de base científica. Foi essa perspectiva que regulou inicialmente este trabalho e que fundamenta a sua sequência.

As oportunidades de investigação apresentadas de seguida corporizam assim um panorama global da Ciência da Web, organizado por disciplinas, menos estanques do que talvez fosse desejável e que, por isso, não pretendem fechar os temas a alguma especulação, pois bem pelo contrário pretendem promover a discussão em torno dos mesmos.

## Ciência dos Computadores

A Ciência dos Computadores tem por objetivo estudar a informação e a computação, pelo que a sua relação com a Web Semântica é não apenas óbvia na generalidade como também o é em particular em relação ao foco nos dados que se pretendem analisar. No campo da Ciência dos Computadores são identificadas naturalmente inúmeras oportunidades de investigação, entre as quais:

- Quais são as características dos repositórios colaborativos de informação (por exemplo, Wikipedia)? (Shadbolt, 2008)
- Como contribuem os indivíduos para os repositórios colaborativos? (Shadbolt, 2008)
- Como suportar a inferência à escala da Web? (Shadbolt, 2008)
- Como suportar a confiança e a proveniência na Web? (Shadbolt, 2008)
- Como identificar partes não funcionais, atrofiadas ou redundantes da Web? (Shadbolt, 2008)

Ainda mais em concreto, na área específica da Web Semântica, foi possível identificar várias questões:

- Que aplicações podem evoluir com a Web Semântica? (Heath, 2010)
- Como pode a Web 2.0 ajudar utilizadores inexperientes na partilha do conhecimento? (Hotho & Stumme, 2010)
- Como pode uma pessoa identificar-se perante uma aplicação? (Aroyo & Houben, 2010)
- Como pode uma pessoa gerir as suas identidades em diversas aplicações? (Aroyo & Houben, 2010)
- Como pode uma aplicação obter a identidade de um utilizador de outra aplicação? (Aroyo & Houben, 2010)
- Como pode ser fornecida confiança e privacidade em mecanismos de identificação? (Aroyo & Houben, 2010)
- Qual o comportamento dos utilizadores e as consequências sociais e legais em sistemas de partilha de identificação? (Aroyo & Houben, 2010)

- Como podem as propriedades únicas dos utilizadores ser identificadas? (Aroyo & Houben, 2010)
- Como pode ser partilhado um vocabulário comum de propriedades de utilizadores? (Aroyo & Houben, 2010)
- Como podem ser partilhadas as escalas de valorização de propriedades de utilizadores? (Aroyo & Houben, 2010)
- Como podem ser anexadas às propriedades dos utilizadores noções de certeza e precisão? (Aroyo & Houben, 2010)
- Como é percebida pelos utilizadores uma abordagem aberta ao conhecimento dos utilizadores? (Aroyo & Houben, 2010)
- Como podem os utilizadores ter a oportunidade de inspecionar e corrigir o conhecimento dos utilizadores que uma aplicação mantém? (Aroyo & Houben, 2010)

### Matemática

A Matemática estuda a quantidade, o espaço, a estrutura e a mudança, pelo que a sua relação com a Web Semântica sai reforçada em particular. No campo específico da Matemática, foi possível encontrar várias possibilidades aplicadas:

- Como distinguir a Web permanente da Web efémera? (Shadbolt, 2008)
- Como usar representações de incerteza na Web? (Shadbolt, 2008)
- Qual a estrutura da Web e como representá-la? (Shadbolt, 2008)
- Como é afetada a estrutura aparente da Web em cada pesquisa realizada por um utilizador? (Shadbolt, 2008)
- Como representar e suportar o contexto da informação semântica?
- Como medir o nível de complexidade da Web? (Shadbolt, 2008)
- Como identificar tendências de fracionamento da Web? (Shadbolt & Berners-Lee, 2008)

### Direito

A sociedade torna-se mais livre e mais complexa, também por influência da tecnologia, e, conseqüentemente, o Direito terá que tornar-se mais vasto, antecipando

todos os possíveis cenários. A interação entre a lei e a Web lança também um grande desafio: trata-se de perceber em que medida uma influencia ou restringe a outra. Algumas possibilidades de trabalho no campo do Direito podem incluir:

- Como adaptar o direito de autor aos conteúdos digitais? (Shadbolt & Berners-Lee, 2008)
- Como definir os direitos de autor das obras digitais com contribuições individuais de reduzida dimensão? (Shadbolt & Berners-Lee, 2008)
- Como incorporar regras de utilização nos conteúdos? (Shadbolt & Berners-Lee, 2008)
- Qual o relacionamento entre as regras de utilização e a distribuição da informação? (Shadbolt & Berners-Lee, 2008)
- Qual o impacto do Direito na formação da Web? (Shadbolt & Berners-Lee, 2008)
- Deverá o Direito ser um catalisador ou um observador reativo da mudança económica, social e tecnológica? (Shadbolt & Berners-Lee, 2008)
- Qual a aplicação do direito de autor a conteúdos que tendem a ser crescentemente gerados por computador? (Shadbolt & Berners-Lee, 2008)
- Que tecnologias devem ser mantidas abertas e quais as consequências dessa decisão? (Shadbolt & Berners-Lee, 2008)
- Até que ponto serão os prestadores de serviços os únicos representantes da política pública de proteção quanto a conteúdos ilegais ou danosos? (Shadbolt & Berners-Lee, 2008)
- Quais os desafios de privacidade que se colocam a um ambiente de informação de complexidade crescente? (Shadbolt & Berners-Lee, 2008)
- Quais as novas formas de regulação emergentes? (Shadbolt, 2008)

### Sociologia

A Sociologia estuda a sociedade e é na sociedade que a Web tem encontrado o suporte para o seu desenvolvimento mais acelerado. No campo da Sociologia, a Web tem já provado que a sua influência é determinante, mas muitas possibilidades de estudo se mantêm disponíveis:

- Quais serão os métodos mais eficazes na determinação da confiança dos conteúdos? (Shadbolt & Berners-Lee, 2008)
- Qual a relação entre confiança nos indivíduos e nos seus avatares? (Shadbolt & Berners-Lee, 2008)
- Como realizar a integração digital de serviços sociais, médicos, financeiros ou educacionais? (Shadbolt & Berners-Lee, 2008)
- Como prevenir o roubo de identidade, perseguição, agressão ou espionagem com base digital? (Shadbolt & Berners-Lee, 2008)
- Como perceber o fenómeno social-tecnológico da Web? (Shadbolt, 2008)
- Quais as relações entre a estrutura e a utilização da Web? (Shadbolt, 2008)
- Quais os conceitos sociológicos e psicológicos a estudar para entender a forma como as pessoas usam a Web? (Shadbolt, 2008)
- Quais as implicações da Web na diversidade? (Shadbolt, 2008)
- Como se posiciona a Web em relação às redes de poder? (Shadbolt, 2008)
- Qual o poder da Web? (Shadbolt, 2008)
- Como pode ser promovida a inteligência coletiva de um ponto de vista tecnológico? (Shadbolt, 2008)
- Qual o efeito da língua, raça, género e outras diferenças na criação de uma inteligência coletiva? (Shadbolt, 2008)
- Quais são as razões socioeconómicas para a participação individual em empreendimentos coletivos? (Shadbolt, 2008)
- Quais são os mecanismos de identificação psicológica de um indivíduo com uma comunidade? (Shadbolt, 2008)
- Qual deve ser o papel das entidades públicas no processo de criação colaborativa? (Shadbolt, 2008)
- Como identificar padrões de comportamento na Web? (Shadbolt, 2008)
- Poderão a incerteza, a insegurança e a invasão de privacidade acabar com a Web? (Shadbolt, 2008)
- Poderão os utilizadores encontrar formas de se protegerem quanto à incapacidade da Web o fazer por eles? (Shadbolt, 2008)

- Como poderá mudar a Web assim que novas populações lhe acedam? (Shadbolt, 2008)
- Como Identificar normas sociais que podem promover a utilização da Web? (Shadbolt & Berners-Lee, 2008)
- Como utilizar a Web como fonte de informação para a aprendizagem? (Hall & O'Hara, 2008)

## Economia

A Economia, enquanto ciência social, desempenha um papel preponderante no desenvolvimento da Web, como já referido, e, naturalmente, apresenta muitas oportunidades de investigação:

- Quais as oportunidades económicas da Web 2.0? (Shadbolt, 2008)
- Quais as implicações económicas da produção colaborativa de conteúdos em redes *ad-hoc*? (Shadbolt, 2008)
- Quais as forças económicas que modelam a formação das redes sociais? (Shadbolt, 2008)
- Quais são as propriedades fundamentais das redes sociais? (Shadbolt, 2008)
- Qual a relação entre as estruturas social, económica e matemática da Web? (Shadbolt, 2008)
- Quais os incentivos comerciais proporcionados pela Web? (Shadbolt, 2008)
- Quais os incentivos industriais proporcionados pela Web? (Shadbolt, 2008)
- Será a Web propensa à concentração, a um domínio de poucos e grandes intervenientes? (Shadbolt, 2008)
- Quais serão as forças que permitem a coexistência entre grandes e pequenos intervenientes? (Shadbolt, 2008)
- Quais os argumentos económicos a favor e contra as plataformas abertas na Web? (Shadbolt, 2008)
- Deverá a política pública e económica influenciar a abertura de plataformas? (Shadbolt, 2008)
- Que mecanismos económicos e sociais poderão melhorar o desempenho da Web? (Shadbolt, 2008)

- Como pode a economia resolver problemas como os da pirataria, privacidade e identidade? (Shadbolt, 2008)
- Quais as razões socioeconómicas para a participação individual em empreendimentos coletivos? (Shadbolt, 2008)
- Qual deve ser o papel das entidades públicas no processo de criação colaborativa? (Shadbolt, 2008)
- Qual a relação entre a abertura da Web e a inovação? (Shadbolt, 2008)
- Serão abertura da informação e requisitos de segurança compatíveis? (Shadbolt, 2008)
- Qual o processo para avaliar um modelo de publicação em relação a um modelo de negócio proprietário? (Shadbolt, 2008)

São, assim, inúmeras as oportunidades de investigação, em áreas mais ou menos tecnológicas, proporcionando um espaço de colaboração interdisciplinar onde se esperam várias realizações. Com o objetivo de apoiar o desenvolvimento global da Ciência da Web, foi criada em 2006 a *Web Science Research Initiative*, mais tarde transformada no *Web Science Trust*<sup>1</sup>.

---

<sup>1</sup> [webscience.org](http://webscience.org)

## 3 A Web Semântica

Neste capítulo descreve-se o conceito de Web Semântica e identificam-se as tecnologias associadas. São também apresentadas aplicações da Web Semântica, sendo destacados exemplos de adoção desta tecnologia em vários setores de atividade. Finalmente é fornecida uma sistematização das atividades do utilizador na Web e dada uma panorâmica do estado atual do envolvimento das empresas com a tecnologia da Web Semântica e dos desafios a superar, abrindo caminho à contribuição deste trabalho.

### 3.1 Conceito de Web Semântica

O conceito de Web Semântica traduz a visão do W3C sobre uma Web de dados interligados (W3C, 2015e). Esta visão materializou-se com a criação de um conjunto de tecnologias que possibilitam que a Web clássica baseada em documentos possa evoluir para a Web dos dados.

Estes dados podem ter qualquer origem, desde que estejam acessíveis na Web, permitindo que as tecnologias da Web Semântica definam regras para o seu armazenamento. Assim, é possível considerar cada página Web como um local onde potencialmente podem ser armazenados registos da base de dados global, por exemplo, na página do Web Site de uma empresa, armazenar os dados que a caracterizam, como o telefone ou a morada da sede, entre outros.

Esses dados podem ser relativos a qualquer tema, pelo que as tecnologias da Web Semântica deverão ser capazes de os caracterizar de forma a que seja percebido o seu significado. Para esse efeito, cada registo de dados deve ser relacionado com um esquema que permita compatibilizar os vários registos, por exemplo, poder reunir os registos de várias empresas de forma a identificar todos os números de telefone de empresas.

Esses dados estarão distribuídos pela Web, pelo que será necessário identificar a sua localização e extrair a informação relevante em cada momento. Assim, é necessária

uma linguagem que permita elaborar e parametrizar os critérios dessa extração, como, por exemplo, aceder aos endereços e telefones de todas as empresas na região do Porto.

A Secção 3.2 aborda as tecnologias que permitem criar repositórios de dados, construir vocabulários e escrever regras para manusear esses dados.

## 3.2 Tecnologias da Web Semântica

A identificação e caracterização da Web Semântica na sua vertente tecnológica é essencial, pois tal como é comum confundir-se o conceito de Web com o de Internet, a tecnologia que suporta a Web Semântica também é frequentemente confundida com as suas realizações ou aplicações.

### 3.2.1 [HTML](#)

A linguagem base da Web, o HTML, sofreu uma evolução significativa. Desde a primeira versão criada por Tim Berners-Lee em 1990, várias atualizações lhe sucederam, nomeadamente a versão 2.0 em 1995, a versão 3.2 em 1997 e a versão 4.01 em 1998. Em 2000, o W3C introduz o XHTML 1.0, uma reformulação do HTML 4. Cada atualização adicionou mais funcionalidades à versão anterior e fez convergir o esforço de desenvolvimento de várias entidades, muitas delas envolvidas na produção de *Web browsers* com características inovadoras que, com alguma disseminação e demonstração do potencial, influenciaram a evolução da linguagem.

Com a introdução do XHTML 1.0, o W3C pretendeu reformular o HTML 4 com a finalidade de o substituir e criar espaço para o desenvolvimento da linguagem com base em extensões, o que permitiria uma evolução mais rápida. Em 2004, a Apple, Mozilla e Opera entraram em divergência com o W3C e acordaram no desenvolvimento paralelo do HTML, numa iniciativa conhecida como WHATWG (*Web Hypertext Application Technology Working Group*). Em 2006, o W3C indicou o seu interesse em participar no desenvolvimento do então emergente HTML 5 e desde 2007 começou a trabalhar em conjunto com o WHATWG (W3C, 2014a).

A utilização da nova versão do HTML, o HTML 5, está já bastante disseminada, apesar de existir em versão definitiva apenas desde outubro de 2014, pois a sua grande orientação para a vertente gráfica permite desenvolver aplicações ricas em multimédia e

movimento sem necessidade de recurso a *plug-ins*. O muito utilizado Adobe Flash está já a ser largamente substituído pelo HTML 5, com o exemplo máximo do YouTube que disponibiliza a alternativa de visualizar vídeos com base em HTML 5. O sucesso de vendas do iPad e do iPhone e a proliferação da plataforma Android para *tablets* e *smartphones* deu também origem ao desenvolvimento de inúmeras pequenas aplicações, as *apps*, que fazem também uso intensivo do HTML 5 e permitem o desenvolvimento de modelos de aplicação reutilizáveis para diferentes plataformas.

A aparência foi sempre uma preocupação na Web, obrigando não apenas à simples disponibilização de documentos onde o conteúdo é fundamental mas não suficiente, mas também à criação de documentos incluindo formatos e elementos multimédia que valorizam o conteúdo. A introdução de CSS (*Cascading Style Sheets*) em 1996 permitiu introduzir uma separação entre os elementos posicionadores e estruturais do documento HTML e o estilo com que estes devem ser exibidos. A evolução do standard CSS levou ao desenvolvimento do nível 3: as versões CSS são disponibilizadas em níveis, em que cada nível inclui as definições dos níveis anteriores, por isso justificando a utilização do termo *cascading* (W3C, 2015a).

Para além da aparência, o dinamismo foi também uma preocupação constante dos criadores de Web Sites e documentos. Desde as animações precoces baseadas em ficheiros de imagem no formato “gif animado” até à utilização intensiva do Flash, a saturação dos utilizadores criou espaço para a utilização de outro tipo de dinamismo, o conteúdo interativo. Utilizando a capacidade de interpretação de código JavaScript dos *browsers*, as ações do utilizador passam a poder ser tratadas e dar origem a alterações do conteúdo dentro do mesmo documento. Essas alterações podem ser apenas alterações de formato, usualmente recorrendo à alteração do HTML e dando origem ao DHTML (*Dynamic HTML*), ou alterações do conteúdo propriamente dito, usualmente recorrendo ao AJAX (*Asynchronous JavaScript and XML*). Neste último caso, a utilização de JavaScript permite que o conteúdo de um documento seja alterado com base em dados externos solicitados em formato XML sobre o protocolo HTTP. O potencial da combinação destas tecnologias é enorme, possibilitando a criação de páginas em que o conteúdo é reutilizado de outras páginas, recorrendo a *widgets* e criando uma espécie de manta de retalhos a que se começou a chamar *mash-up*, sinónimo de mistura, muito característica

da Web 2.0 (Clarkin & Holmes, 2007). Um dos exemplos mais comuns é a utilização de um *widget* do Google Maps no Web Site de uma empresa que assinala a sua localização, bem como, com recurso a uma hiperligação, iniciar a determinação de um itinerário que tenha esse local como destino, facilitando essa tarefa ao utilizador (Google, 2015c).

### 3.2.2 [RDF](#)

A génese da Web Semântica confunde-se com a da própria Web, pois a proposta original de Tim Berners-Lee que levou à criação da *World Wide Web* referia diferentes tipos de ligações entre diferentes tipos de documentos. O W3C publicou a primeira recomendação para a representação genérica de metadados em 1999, criando assim o RDF (*Resource Description Framework*) (W3C, 2004). Conforme referido na Secção 2.4, o RDF permite que os dados possam ser identificados, publicados, reutilizados e interligados com recurso aos mesmos endereços que se utilizam para identificar a localização de um documento e, por isso, está na base da concretização da Web dos dados.

Concretamente, o RDF é uma linguagem para representação de modelos de dados que utiliza afirmações constituídas por triplos: um sujeito, um predicado e um objeto. Tendo em conta a influência do RDF neste trabalho, é importante efetuar a clarificação dos seus fundamentos e a sua estrutura.

A organização de dados com base numa tabela, onde cada coluna representa um campo ou propriedade e cada linha representa um registo ou um elemento é amplamente reconhecida. Considere-se o exemplo apresentado na Figura 3.1, uma tabela de indivíduos que inclui os respetivos nomes e datas de nascimento.

# Indivíduos

Nome	Data de Nascimento
António	13/09/1948
Beatriz	05/07/1972
Camilo	21/02/1976
Daniela	10/10/1941
Eduardo	29/06/1966

Figura 3.1: Modelo de Dados - Indivíduos

# Publicações

Nome	Data Nascimento
António	13/09/1948
Beatriz	05/07/1972
Camilo	21/02/1976
Daniela	10/10/1941
Eduardo	29/06/1966



Nome	Documento	Data Publicação
António	"What about RDF?"	11/09/1998
Beatriz	"RDF triples make the Web"	15/02/2004
Beatriz	"One tale, two Webs"	26/04/2008
Daniela	"Mining the Global Data Space"	17/06/2002
Eduardo	"Understanding OWL"	19/11/2009
Eduardo	"SPARQLing: Water"	13/08/2011

Figura 3.2: Modelo de Dados - Publicações

Considere-se agora uma segunda tabela, apresentada na Figura 3.2, onde se identificam documentos publicados por esses indivíduos.

Estas representações são bastante comuns e intuitivas. O modelo Entidade-Relação de representação de dados tem por base estas representações, onde se apresenta a Entidade em causa para cada tabela e se definem as Relações entre as Entidades. Neste caso, um Indivíduo pode publicar vários documentos. O objetivo desta descrição não é obviamente o de tentar explicar todos os conceitos envolvidos no modelo, mas antes de salientar as características fundamentais que o distinguem do RDF, pelo que não se descrevem mais pormenores nem se discute porque não poderia um documento ser publicado por mais que um indivíduo.

Ao substituir esta estrutura por uma estrutura semelhante, podem-se antecipar várias consequências, conforme se pode verificar pela análise da Figura 3.3.

A informação disponível ficou agora menos intuitiva. Não apenas se tem agora mais dificuldade em analisar a informação disponível, como se torna mais difícil responder a perguntas simples, como por exemplo, quantos documentos foram publicados antes de 1 de Janeiro de 2000. O esforço computacional necessário para processar esta estrutura também aumentou na mesma proporção.

## Publicações

Entidade	Predicado	Objecto
António	nasceu em	13/09/1948
Beatriz	nasceu em	05/07/1972
Camilo	nasceu em	21/02/1976
Daniela	nasceu em	10/10/1941
Eduardo	nasceu em	29/06/1966
António	publicou	"What about RDF?"
Beatriz	Publicou	"RDF triples make the Web"
Beatriz	Publicou	"One tale, two Webs"
Daniela	Publicou	"Mining the Global Data Space"
Eduardo	Publicou	"Understanding OWL"
Eduardo	Publicou	"SPARQLing: Water"
"What about RDF?"	publicado em	11/09/1998
"RDF triples make the Web"	publicado em	15/02/2004
"One tale, two Webs"	publicado em	26/04/2008
"Mining the Global Data Space"	publicado em	17/06/2002
"Understanding OWL"	publicado em	19/11/2009
"SPARQLing: Water"	publicado em	13/08/2011

Figura 3.3: Modelo de Dados - Publicações, tabela única

## Docentes e Publicações

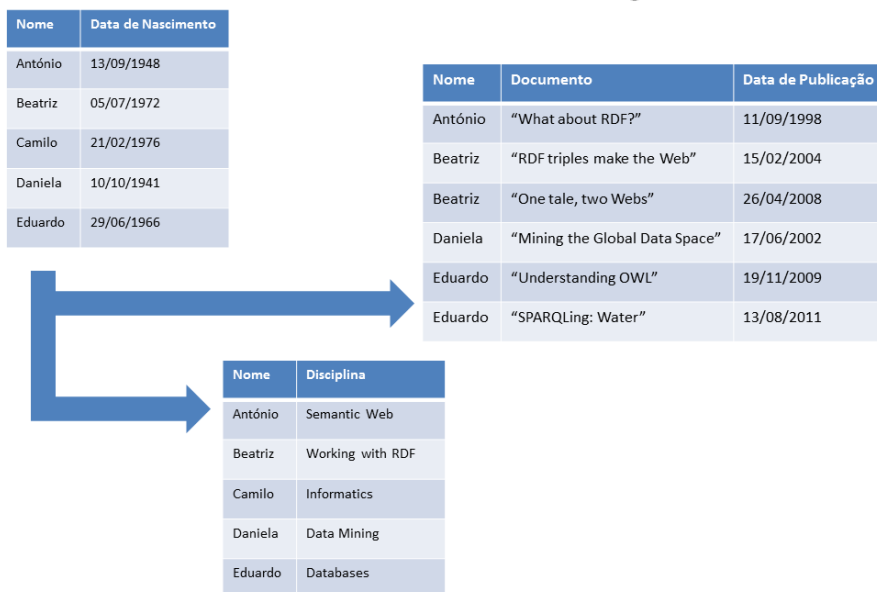


Figura 3.4: Modelo de Dados - Docentes e Publicações

## Docentes e Publicações

Entidade	Predicado	Objecto
António	nasceu em	13/09/1948
Beatriz	nasceu em	05/07/1972
Camilo	nasceu em	21/02/1976
Daniela	nasceu em	10/10/1941
Eduardo	nasceu em	29/06/1966
António	publicou	"What about RDF?"
Beatriz	Publicou	"RDF triples make the Web"
Beatriz	Publicou	"One tale, two Webs"
Daniela	Publicou	"Mining the Global Data Space"
Eduardo	Publicou	"Understanding OWL"
Eduardo	Publicou	"SPARQLing: Water"
"What about RDF?"	publicado em	11/09/1998
"RDF triples make the Web"	publicado em	15/02/2004
"One tale, two Webs"	publicado em	26/04/2008
"Mining the Global Data Space"	publicado em	17/06/2002
"Understanding OWL"	publicado em	19/11/2009
"SPARQLing: Water"	publicado em	13/08/2011
António	docente de	Semantic Web
Beatriz	docente de	Working with RDF
Camilo	docente de	Informatics
Daniela	docente de	Data Mining
Eduardo	docente de	Databases

Figura 3.5: Modelo de Dados - Docentes e Publicações - tabela única

No entanto, a estrutura tornou-se mais flexível. Observe-se na Figura 3.4 o que seria necessário fazer para adicionar as disciplinas lecionadas pelos indivíduos ao primeiro cenário.

Por outro lado, a Figura 3.5 apresenta o impacto da mesma alteração no segundo cenário. Agora, a estrutura permanece inalterada e a simples adição de informação à estrutura permite estender o modelo. Aparentemente, a simplicidade da forma garante-lhe flexibilidade mas parece tornar o conteúdo cada vez menos claro.

Entretanto, o compromisso entre esforço computacional e flexibilidade tem vindo a pender em favor do segundo lado, tendo em conta a verificação da projeção ditada pela Lei de Moore, pois mais capacidade de computação disponível leva a menos preocupação com a simplificação ou otimização. Por outro lado, a economia tem beneficiado da inovação e deu origem a um ambiente de elevada competitividade onde flexibilidade e velocidade assumiram um papel por vezes mais importante que robustez e confiança. Pode assim questionar-se a validade económica atual dos modelos de normalização de dados em comparação com a flexibilidade e promessa de universalidade da Web Semântica (Segaran, Evans, & Taylor, 2009).

Talvez a principal característica do RDF seja a identificação universal de todos os recursos por um *Universal Resource Identifier* (URI), fazendo uso generalizado da Web não apenas para obter informação acessível em *Universal Resource Locators* (URL), mas também para identificar qualquer coisa, seja ela possível de aceder eletronicamente ou não (Segaran et al., 2009; W3C & IETF, 2001; W3C, 2008a). No caso específico de *Linked Data*, são utilizados HTTP URI em detrimento de outras possibilidades, permitindo que qualquer cliente de HTTP possa obter uma descrição do recurso identificado pelo endereço (Heath & Bizer, 2011). Tal como no Modelo Entidade-Relação se teria identificado cada indivíduo por um identificador único, por exemplo, com um número inteiro diferente para cada indivíduo, no RDF espera-se confiar nos URI como forma de identificação universal. Assim, pode aceitar-se que o mesmo indivíduo tenha mais do que um URI, desde que cada um dos URI referencie os restantes como sendo indicadores do mesmo indivíduo, eliminando a duplicação e facilitando a extensibilidade.

## Docentes e Publicações

Entidade	Predicado	Objecto
http://example.org/people/antonio	tem nome	António
http://example.org/people/beatriz	tem nome	Beatriz
http://example.org/people/camilo	tem nome	Camilo
http://example.org/people/daniela	tem nome	Daniela
http://example.org/people/eduardo	tem nome	Eduardo
http://example.org/people/antonio	nasceu em	13/09/1948
http://example.org/people/beatriz	nasceu em	05/07/1972
http://example.org/people/camilo	nasceu em	21/02/1976
http://example.org/people/daniela	nasceu em	10/10/1941
http://example.org/people/eduardo	nasceu em	29/06/1966
http://example.org/documents/what_about_rdf	tem título	"What about RDF?"
http://example.org/documents/rdf_triples_make_the_web	tem título	"RDF triples make the Web"
http://example.org/documents/one_tale_two_webs	tem título	"One tale, two Webs"
http://example.org/documents/minig_the_global_data_space	tem título	"Mining the Global Data Space"
http://example.org/documents/understanding_owl	tem título	"Understanding OWL"
http://example.org/documents/sparqling_water	tem título	"SPARQLing: Water"
http://example.org/people/antonio	publicou	http://example.org/documents/what_about_rdf
http://example.org/people/beatriz	Publicou	http://example.org/documents/rdf_triples_make_the_web
http://example.org/people/beatriz	Publicou	http://example.org/documents/one_tale_two_webs
http://example.org/people/daniela	Publicou	http://example.org/documents/minig_the_global_data_space
http://example.org/people/eduardo	Publicou	http://example.org/documents/understanding_owl
http://example.org/people/eduardo	Publicou	http://example.org/documents/sparqling_water
http://example.org/documents/what_about_rdf	publicado em	11/09/1998

Figura 3.6: Modelo de Dados - Docentes e Publicações - tabela única com URIs

## Docentes e Publicações

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:example="http://example.org/rdf/example/1.0">
  <rdf:Description rdf:about="http://example.org/people/antonio">
    <rdf:type rdf:resource="http://example.org/rdf/example/1.0/Person" />
    <example:name>António</example:name>
    <example:birthday>13/09/1948</example:birthday>
    <example:publication>http://example.org/documents/what_about_rdf
    </example:publication>
    <example:lecturer>Semantic Web</example:lecturer>
  </rdf:Description>
  (...)
  <rdf:Description rdf:about="http://example.org/documents/what_about_rdf">
    <rdf:type rdf:resource="http://example.org/rdf/example/1.0/Document" />
    <example:title>What about RDF?</example:title>
    <example:published>11/09/1998</example:published>
  </rdf:Description>
  (...)
</rdf:RDF>
```

Figura 3.7: Modelo de Dados - Docentes e Publicações - RDF

O exemplo corrente passaria a ter então um formato um pouco distinto, conforme apresentado na Figura 3.6, em versão incompleta para melhor visualização.

```

<p vocab="http://schema.org/" typeof="Person">
  My name is
  <span property="name">Manu Sporny</span>
  and you can give me a ring via
  <span property="telephone">1-800-555-0199</span>
  or visit
  <a property="url" href="http://manu.sporny.org/">my homepage</a>.
</p>

```

Figura 3.8 Exemplo de RDFa para caracterização de um indivíduo (W3C, 2015c)

```

<script type="application/ld+json">
{ "@context" : "http://schema.org",
  "@type" : "Organization",
  "url" : "http://www.your-company-site.com",
  "contactPoint" : [
    { "@type" : "ContactPoint",
      "telephone" : "+1-401-555-1212",
      "contactType" : "customer service"
    } ] }
</script>

```

Figura 3.9 schema.org no formato JSON-LD para o contacto de uma empresa (Google, 2015a)

Representando o mesmo exemplo em RDF, obter-se-ia uma estrutura como a apresentada na Figura 3.7, onde se ilustra apenas 20% da estrutura para melhor legibilidade.

A representação do RDF pode tomar vários formatos, e, neste caso, o RDF/XML (W3C, 2014d) apresentado na Figura 3.7 corresponde à sintaxe mais representativa. Outros formatos como RDFa (W3C, 2015d), JSON/LD (W3C, 2014b), N-Triples (W3C, 2014c), N3 (W3C, 2011) ou Turtle (W3C, 2015b) permitem a representação do mesmo modelo de dados. Não se trata de um assunto crítico no âmbito deste trabalho, uma vez que todas as representações são facilmente tratadas com uma biblioteca específica e a transformação de uma representação noutra é um exercício dominado na maior parte das linguagens de programação.

Com frequência, é necessário recorrer à inserção de metadados em páginas Web, permitindo que o mesmo documento inclua a parte legível ao indivíduo e a parte legível às máquinas, o que conduz à introdução do RDF em páginas HTML. Nesses casos destacam-se os formatos RDFa, (*RDF in attributes*), e JSON-LD. A principal diferença entre ambos reside na forma como interferem na construção do HTML. No caso do RDFa, o HTML DOM (*Document Object Model*) é utilizado como base para a inserção de metadados, recorrendo a atributos das *tags* HTML, como exemplificado na Figura 3.8.

Esta representação é semelhante à representação microdata (W3C, 2013), uma representação alternativa mais simples, também ela um standard do W3C. No caso do JSON-LD, os metadados são introduzidos numa tag script e completamente isolados do restante HTML, como exemplificado na Figura 3.9.

### 3.2.3 OWL

O esforço de universalização intrínseco à criação deste espaço único de dados globais envolve um não menor esforço de interpretação de todos os dados, na procura e formalização do seu verdadeiro significado. As entidades, o seu significado e as suas relações são representados na forma de ontologias, um termo importado da Filosofia que traduz a caracterização da realidade como um todo. A OWL (*Web Ontology Language*) é uma linguagem baseada em RDF que define classes e propriedades de representação, que graças à sua extensibilidade permitirá a representação de quase tudo (W3C, 2012).

Um exemplo da aplicação de RDF e OWL é o FOAF (*Friend of a Friend*), um vocabulário que descreve pessoas e, principalmente, permite indicar ligações entre elas (Brickley & Miller, 2015). Iniciado no ano 2000, pode ser considerado o primeiro projeto baseado na Web Social e um pioneiro na Web Semântica. Na Tabela 3.1 apresenta-se a utilização de FOAF para caracterizar um indivíduo e os seus contactos numa página HTML, com recurso a RDFa.

```
<div vocab="http://xmlns.com/foaf/0.1/" typeof="Person"><p>
  <p>
    <span property="name">Alice Birpemswick</span>,
    Email: <a property="mbox" href="mailto:alice@example.com">alice@example.com</a>,
    Phone: <a property="phone" href="tel:+1-617-555-7332">+1 617.555.7332</a>
  </p>
</div>
```

Tabela 3.1 Exemplo de FOAF

### 3.2.4 SPARQL

Tal como as hiperligações na Web clássica ligam documentos num espaço global de informação, as ligações definidas no RDF ligam dados num espaço único de dados globais (Heath & Bizer, 2011). Para poder interrogar esse espaço global e obter respostas

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX type: <http://dbpedia.org/class/yago/>
PREFIX prop: <http://dbpedia.org/property/>
SELECT ?country_name ?population
WHERE {
  ?country a type:LandlockedCountries ;
    rdfs:label ?country_name ;
    prop:populationEstimate ?population .
  FILTER (?population > 15000000) .
}

```

Tabela 3.2 Exemplo de SPARQL (Feigenbaum, 2009)

adequadas, é necessário definir uma linguagem apropriada. Tal como o SQL se tornou quase standard para bases de dados relacionais, o SPARQL (*Sparql Protocol And Rdf Query Language*) (W3C, 2008c) assume o protagonismo na extração de dados de RDF e segue inclusive uma sintaxe muito semelhante ao SQL.

O exemplo apresentado na Tabela 3.2 indica a sintaxe para interrogar a DBPedia e obter uma lista de países sem fronteiras com o mar e com mais de 15 milhões de habitantes.

### 3.2.5 Tendências de evolução

A Web das páginas e dos documentos está a transformar-se na Web das aplicações. A exploração das capacidades do HTML e dos *browsers* que o interpretam permite desenvolver soluções muito completas baseadas em páginas disponibilizadas na Web que rivalizam em funcionalidade com o seu equivalente mais tradicional, a aplicação cliente. Com vantagens e desvantagens que se podem identificar em ambas as abordagens, um argumento de peso emerge: o aumento exponencial do número de aparelhos que utilizam a Web, das suas capacidades integradas, como câmaras, giroscópios ou GPS (*Global Positioning System*), e das oportunidades para a sua utilização, particularmente tendo em conta o acesso móvel, contribui decisivamente para o potencial de utilização da Web.

A Figura 3.10 resume as tecnologias referidas na Secção 3.2 e a os seus inter-relacionamentos numa “pilha tecnológica” (*technology stack*), em que cada camada se suporta nas camadas inferiores com que está relacionada. A base da pilha representa a

# Web Semântica - Tecnologias

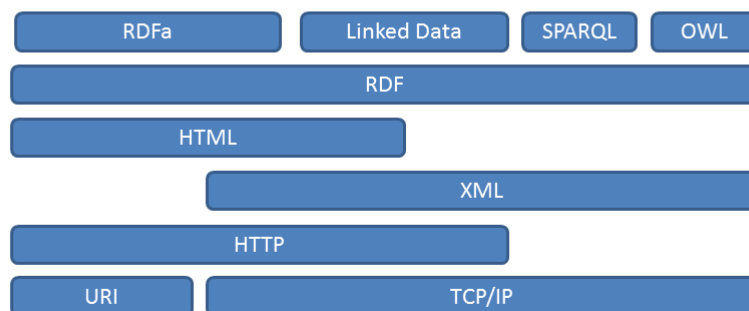


Figura 3.10: Tecnologias da Web Semântica

Internet e a base tecnológica da Web (W3C, 2010). Algumas barras não se estendem totalmente na horizontal representando a possível independência de algumas tecnologias sobre as restantes. O topo da pilha inclui as tecnologias específicas da Web Semântica abordadas nesta secção. Adicionalmente, deveria ainda incluir áreas onde a tecnologia é ainda emergente, como, por exemplo, nas áreas da confiança (validação das fontes de dados) e segurança (controlo da publicação ou utilização dos dados) (Berners-Lee, 2006b).

## 3.3 Aplicações da Web Semântica

A tecnologia da Web Semântica poderá transformar o software empresarial, dar origem a novos modelos de negócio e até reduzir custos em áreas como a integração de dados, gestão de dados centrais e gestão da informação empresarial (Pollock & Oracle, 2008).

Redes empresariais internas baseadas em *Linked Data* poderão reduzir significativamente os custos de integração de informação, entre outros benefícios, como por exemplo na integração da informação de produtos, fornecedores, materiais, legislação, mercado, finanças ou outras informações entre fontes internas e externas

(Janowicz & Hitzler, 2010a). A adoção de ontologias pode responder a questões essenciais na modelização de informação empresarial (Antunes, Bakhshandeh, Mayer, Borbinha, & Caetano, 2014).

No entanto, a adoção crescente das tecnologias da Web Semântica e os princípios de *Linked Data* fomentam a pergunta sobre que aplicações podem ser desenvolvidas para utilizar este potencial. A resposta poderá ser obtida com a identificação das áreas onde estas tecnologias e princípios podem contribuir de forma única (Heath, 2010).

### 3.3.1 Exemplos de adoção da Web Semântica

Atualmente podem encontrar-se exemplos de adoção da tecnologia Web Semântica nas organizações que caracterizam bem o seu potencial. Nos exemplos apresentados de seguida podem identificar-se aplicações de vários tipos, em setores de atividade diversos:

- Europeana (*Content Management, Cultura*)<sup>1</sup>  
Disponibilização de dados no formato *Linked Data* para serem utilizados livremente por profissionais e empresas dos setores da cultura e das indústrias criativas;
- BBC Digital Public Space (*Content Management, Media*)<sup>2</sup>  
Disponibilização de arquivos públicos de forma a serem utilizados na forma de catálogo de recursos;
- Biogen Idec (*Supply Chain Management, Farmacêutico*)<sup>3</sup>  
Integração de informação de Supply Chain Management, proporcionando maior visibilidade operacional aos fornecedores e obtendo informação mais atualizada;
- UCB (*R&D, Farmacêutico*) (Dataversity, 2009)  
Acesso a informação para projetos de investigação bio farmacêutica;
- Best Buy (*e-commerce, Retalho*) (MacManus, 2010)

---

<sup>1</sup> <http://pro.europeana.eu/linked-open-data>

<sup>2</sup> [http://www.bbc.co.uk/blogs/bbcinternet/2011/10/digital\\_public\\_space\\_idea.html](http://www.bbc.co.uk/blogs/bbcinternet/2011/10/digital_public_space_idea.html)

<sup>3</sup> <http://www.americanlaboratory.com/914-Application-Notes/35621-Pharmaceutical-Supply-Chain-Management-Using-Semantic-Web-Technology-Case-Study/>

Informação sobre 450,000 produtos na Web, para que possa ser indexada por motores de pesquisa e Web Sites de agregação de preços;

- Bankinter (*e-banking*, Financeiro) (Fensel, Facca, Simperl, & Toma, 2011)  
Empresa que opera nos mercados financeiros, utiliza a tecnologia semântica para modelizar operações e apoiar a decisão de investimento dos seus clientes;
- University of Plymouth, Talis (*Resource List Management*, Educação) (Clarke & Greig, 2009)  
Gestão de referências bibliográficas de disciplinas de Licenciatura e gestão de leitura e notas de alunos;
- Liverpool John Moores University (*Tagging*, Educação) (Morris, 2012)  
Marcação de clips de vídeo e relacionamento com coreografias de dança com o objetivo de suportar a aprendizagem;
- CRUZAR, Zaragoza (*Route Planning*, Turismo) (Fernández & Campos, 2008)  
Criação de rotas turísticas personalizadas em função das preferências do turista;
- iLAW, Ministry of Justice of Korea (*Information Management*, Direito) (Jung et al., 2010)  
Desenvolvimento de uma ontologia e de um sistema de indexação e pesquisa para legislação de vários países;
- FAO, United Nations (*Search*, Agricultura) (Salokhe, Sini, & Keizer, 2007)  
Desenvolvimento de uma ontologia e de um sistema de indexação e pesquisa para produtos agrícolas;
- NATAS, TATA (*Natural Language Processing*, Software) (Anantaram, 2007)  
Interface em linguagem natural para tratamento de instruções sobre aplicações de gestão;
- Audi (*Integração*, Automóvel) (Syldatke, Chen, Angele, Nierlich, & Ullrich, 2008)  
Integração e análise de dados de monitorização de componentes eletrónicos.

Alguns destes exemplos têm mais projeção que outros, mas a realidade é que sendo uma tecnologia recente, a Web Semântica ainda não deu origem a um número elevado de projetos que a simbolizem de forma significativa. No entanto, alguns projetos conquistaram já um espaço relevante junto do mercado. Pode observar-se que uma grande parte dos exemplos é proveniente de alguns setores da economia onde o conteúdo ou o conhecimento são parte fundamental do seu negócio nuclear. Na primeira das situações, o conteúdo é fundamental em setores como os Media, a Cultura ou mesmo o Direito, na segunda, o conhecimento é fundamental em setores como o Farmacêutico, a Educação ou o Turismo.

### 3.3.2 [O caso do schema.org](#)

O schema.org<sup>1</sup> resulta de uma iniciativa lançada em Junho de 2011 por Google, Bing e Yahoo, mais tarde acompanhados pelo Yandex, destinada a criar um conjunto comum de esquemas de marcação de dados para utilização generalizada em Web Sites. Esta iniciativa suporta os objetivos dos motores de pesquisa no sentido de adicionar à indexação de todas as páginas Web, informação adicional sobre o significado do seu conteúdo.

Entre mais de 600 tipos de Entidade, as mais utilizadas são:

- Trabalho criativo: livro, filme, música, receita, série...
- Objeto embebido: áudio, imagem, vídeo
- Evento
- Organização
- Pessoa
- Local, Negócio Local, Restaurante...
- Produto, Oferta...
- Revisão, Avaliação

A adição dos metadados com base em schema.org a uma página Web pode ser feita com recurso a três formatos. Os formatos RDFa e microdata permitem inserir atributos no HTML que caracterizam o seu conteúdo. O formato JSON-LD (cf. Figura 3.9)

---

<sup>1</sup> [www.schema.org](http://www.schema.org)

permite inserir um bloco “script” em qualquer local da página, tornando-se assim mais simples de gerar, embora possa gerar divergências entre os dados inseridos e o texto visível para o utilizador.

De acordo com uma comunicação de agosto de 2014 de Ramanathan V. Guha, um dos responsáveis pela criação do schema.org, mais de 20% dos Web Sites utilizam já este esquema, num total de mais de 15 mil milhões de entidades identificadas (Zaino, 2014). As plataformas de Gestão de Conteúdos mais populares, como por exemplo o Wordpress e o Joomla, disponibilizam *plug-ins* para marcação de dados com estes esquemas. A iniciativa teve também repercussão nos trabalhos do W3C, com a *Semantic Web Interest Group Web Schemas Task Force*<sup>1</sup> a concentrar-se na análise de propostas de extensão do schema.org. A iniciativa tem assim o potencial de se tornar uma referência principal nesta área, ou mesmo parte do lote de tecnologias standard da Web Semântica.

### 3.3.3 [O caso do Google Knowledge Graph](#)

Aproximadamente um ano após a criação do schema.org, a Google lança o *Knowledge Graph*<sup>2</sup>, traduzido para português como Gráfico do Conhecimento, apesar de se tratar efetivamente de um Grafo de Conhecimento, uma representação de conceitos interligados como definido na Teoria dos Grafos (Carlson, 2015). A ideia subjacente ao produto da Google não é nova e pretende fornecer resultados com base semântica, originando o que se começou a designar por *Semantic Search*. Na prática, o utilizador recebe como habitualmente os resultados da sua pesquisa na forma de descrições e hiperligações para as páginas que o algoritmo de pesquisa considera mais relevantes, mas também, complementarmente, a informação resumida sobre a entidade eventualmente detetada nessa pesquisa.

---

<sup>1</sup> [www.w3.org/wiki/WebSchemas](http://www.w3.org/wiki/WebSchemas)

<sup>2</sup> [www.google.com/insidesearch/features/search/knowledge.html](http://www.google.com/insidesearch/features/search/knowledge.html)

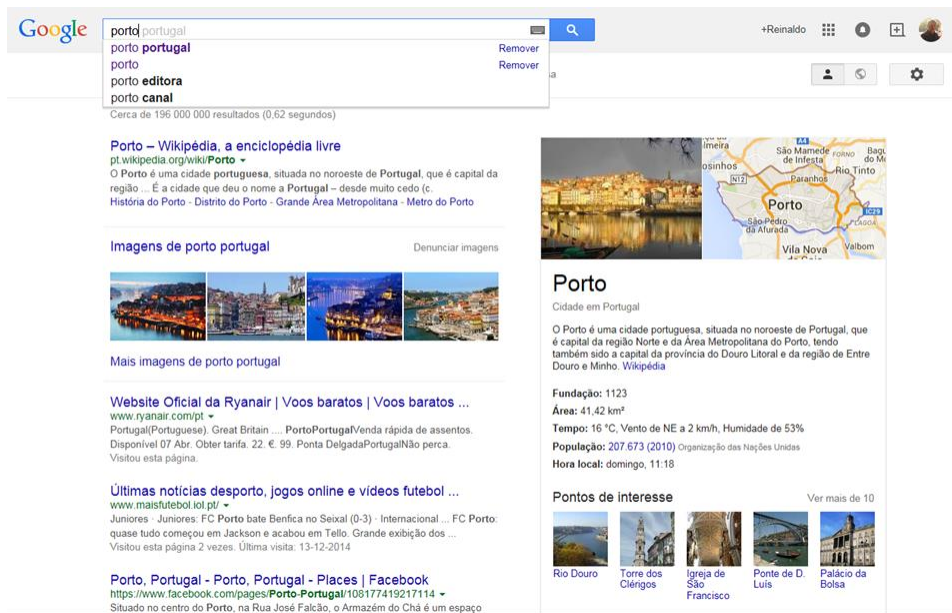


Figura 3.11 Google Knowledge Graph resultante da pesquisa "porto"

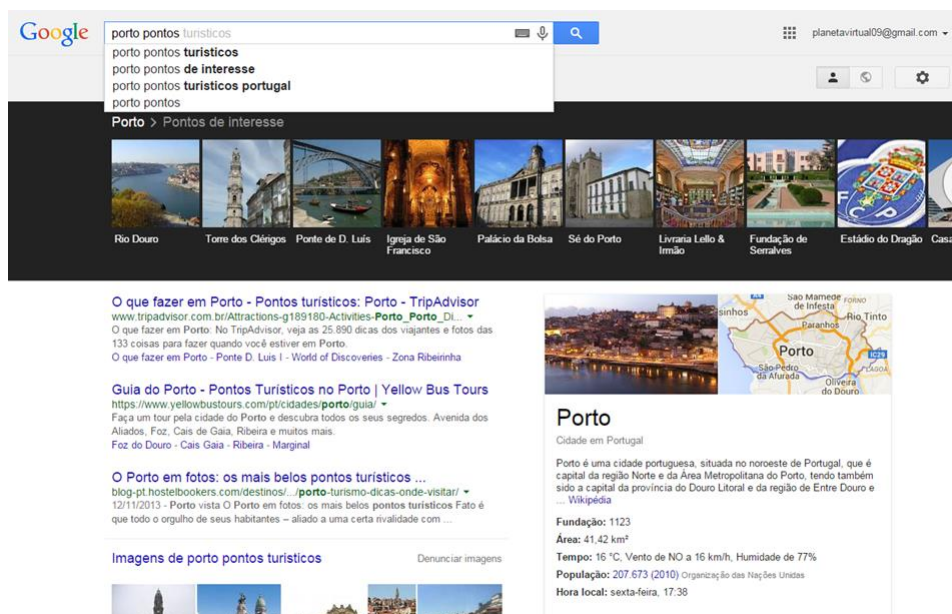


Figura 3.12 Google Knowledge Graph resultante da pesquisa "porto pontos "

A Figura 3.11 ilustra esses dois tipos de resultados para o caso de uma pesquisa efetuada com a palavra “Porto”: na coluna da esquerda são apresentadas as páginas e imagens consideradas mais relevantes para a pesquisa; na coluna da direita, a entidade “Porto, Cidade em Portugal” é identificada como sendo um resultado muito provável. Para além da informação resumida sobre a cidade do Porto, o resultado apresenta também alguns Pontos de Interesse, permitindo que o utilizador navegue entre as entidades, ao longo do Grafo definido.

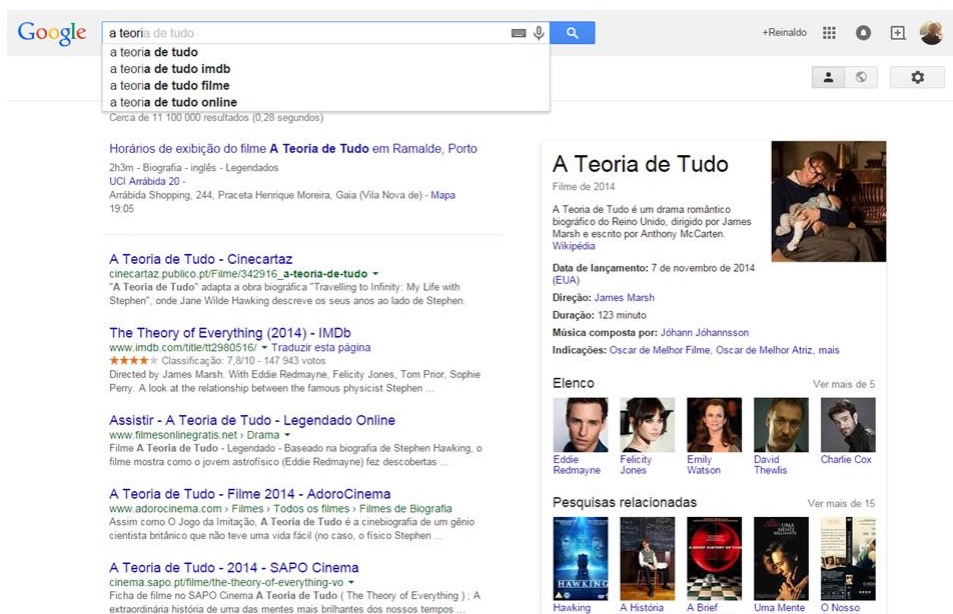


Figura 3.13 Resultado da pesquisa "a teoria" no Google

Por último, conforme ilustrado na Figura 3.12, quando os resultados de uma pesquisa são múltiplos, podem ser apresentados com recurso a uma “roda” que facilita a navegação através de um número elevado de itens.

Muitas considerações poderiam ser feitas sobre esta funcionalidade, mas deve referir-se uma em particular. Caso a pesquisa não seja totalmente conducente a um resultado único, como muitas vezes não o será, apesar das centenas de critérios tidos em conta (Google, 2015d), como a localização do utilizador, o seu histórico de pesquisa ou a hora do dia, o resultado do Google Knowledge Graph sugere resultados alternativos. A título de exemplo, sugere-se uma experiência com a pesquisa “*europa*”, pois tendo em conta o utilizador que a faz, os resultados apresentados primária e alternativamente são ou o continente com esse nome em língua inglesa ou a banda rock de quem todos conhecem apenas uma música.

A Figura 3.13 apresenta uma extensão da utilização da tecnologia para além do Google Knowledge Graph, no caso sugerindo locais para o utilizador ir ao cinema: assumindo que ao pesquisar “a teoria” o utilizador se está a referir ao filme “A Teoria de Tudo”, que está localizado em “Ramalde, Porto” e considerando a hora atual, o motor de pesquisa apresenta uma sugestão da próxima sessão a que o utilizador poderá assistir. Naturalmente, este resultado está condicionado aos cinemas que disponibilizam na Web metadados relativos aos filmes e respetivos horários de exibição.

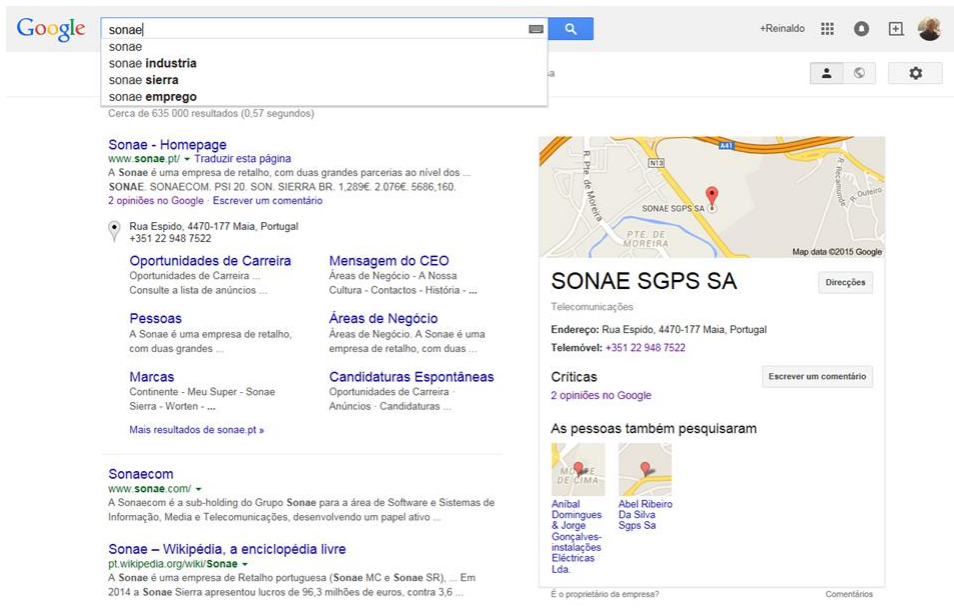


Figura 3.14 Resultado da pesquisa "sonae" no Google

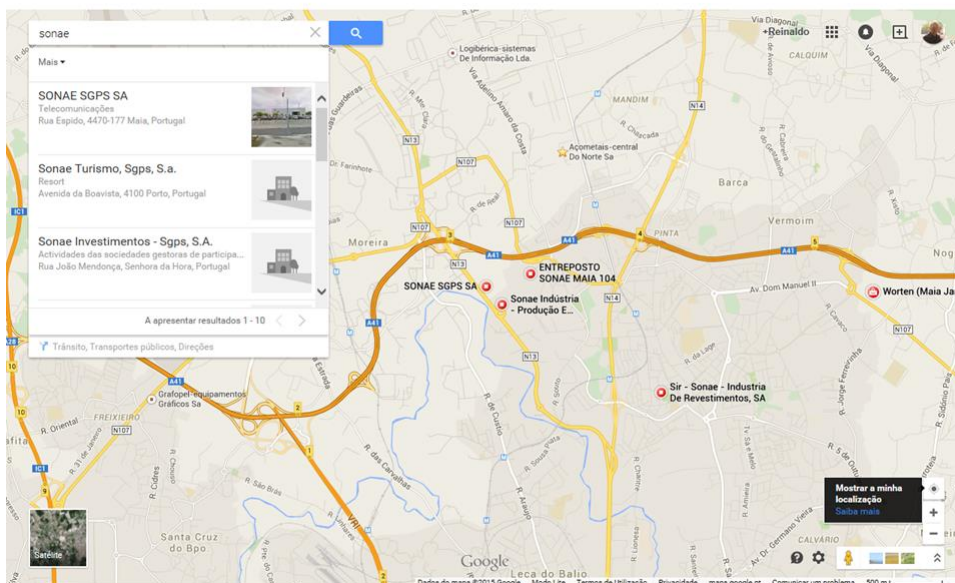


Figura 3.15 Resultado da pesquisa "sonae" no Google Maps

Os motores de pesquisa começaram também a ser usados para tarefas muito simples que libertam os utilizadores de algumas obrigações, como por exemplo procurar o número de telefone de uma empresa.

A Figura 3.14 apresenta um resultado semelhante ao obtido com o Google Knowledge Graph. No momento da publicação do trabalho, os resultados relativos à informação das empresas são provenientes de várias fontes. Entre as primeiras fontes a ser utilizadas pela Google estão bases de dados de empresas, carregadas sem o

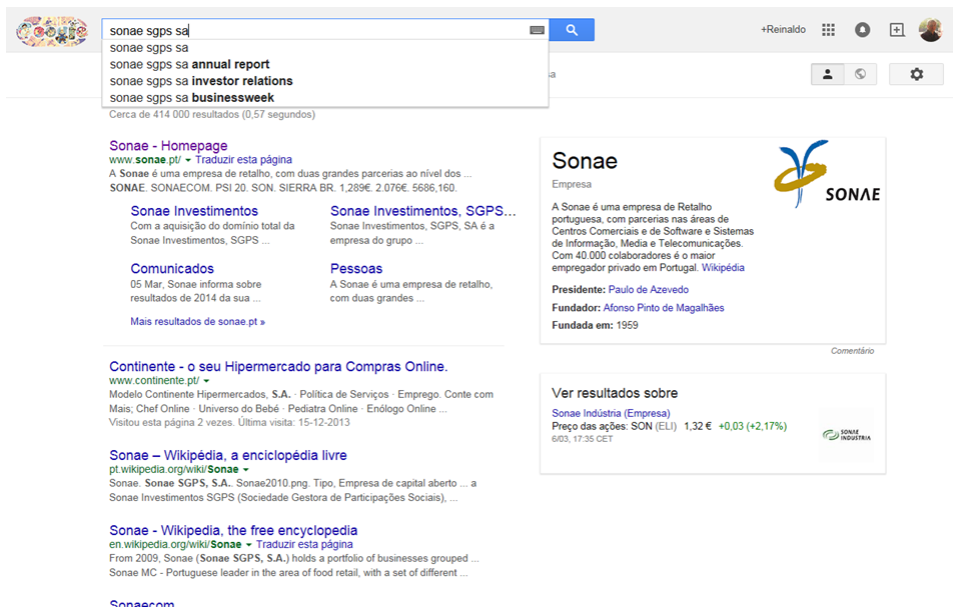


Figura 3.16 Resultado da pesquisa "sonae sgps sa" no Google

conhecimento expresso das mesmas. O exemplo ilustrado na figura é um desses casos, perceptível pela frase apresentada “É o proprietário desta empresa?”, que convida o utilizador a resgatar esta informação e, depois de validada a sua autoridade para tal, poder editá-la. Se o fizer, o utilizador é convidado a criar e gerir uma página da empresa no Google+, a rede social da Google.

Outra fonte para a informação apresentada numa pesquisa é o Google Maps, o serviço de cartografia que apresenta uma camada de dados sobreposta para complementar os resultados de uma pesquisa efetuada, como por exemplo de um restaurante nas proximidades do utilizador. A Figura 3.15 apresenta o resultado obtido para a mesma pesquisa realizada.

Adicionalmente, os resultados podem ser fornecidos ao Google através da inclusão de metadados nos Web Sites das empresas. A Figura 3.16 apresenta um resultado semelhante ao obtido com essa técnica, embora não seja possível validar com certeza a fonte dos dados apresentados.

Os exemplos aqui utilizados ilustram os desafios e ambiguidades da Web Semântica. Por um lado, a resposta fornecida de forma inequívoca e imediata concretiza o objetivo do utilizador; por outro lado, qualquer ambiguidade pode conduzir o utilizador a pontos muito distantes do resultado pretendido. Para que o resultado da pesquisa seja o desejado contribuem a destreza do utilizador em formular a pesquisa, a capacidade do

**WolframAlpha** computational knowledge engine

mortgage \$150,000, 6.5%, 30 years

Examples Random

Assuming fixed rate mortgage | Use adjustable rate mortgage instead  
 Assuming "\$150,000" refers to loan amount | Use sale amount or down payment instead  
 Assuming "6.5%" refers to annual percentage rate | Use down payment instead

Also include: points | interest-only period | tax rate | balloon payment

Share this page:  
 f t+ g+ t  
 More ▾

**WOLFRAM**  
 is hiring!  
 Find out more >>

Take Wolfram|Alpha anywhere...

Input information:

fixed rate mortgage	
loan amount	\$150 000 (US dollars)
annual percentage rate	6.5%
loan period	30 years

Monthly payments:

monthly payment	\$948
effective interest rate	6.697%

Mortgage totals:

principal paid	\$150 000
total interest paid	\$191 317
total payments	\$341 317

Payments and balances:

Payments table:

year	monthly payment	ending balance	yearly principal paid	yearly interest paid
1	\$948	\$148 323	\$1677	\$9701
2	\$948	\$146 535	\$1789	\$9588
3	\$948	\$144 626	\$1909	\$9469
4	\$948	\$142 589	\$2037	\$9341
5	\$948	\$140 416	\$2173	\$9204
6	\$948	\$138 098	\$2318	\$9059
7	\$948	\$135 624	\$2474	\$8904
8	\$948	\$132 985	\$2639	\$8738
9	\$948	\$130 169	\$2816	\$8561
10	\$948	\$127 164	\$3005	\$8373
11	\$948	\$123 958	\$3206	\$8171
12	\$948	\$120 538	\$3421	\$7957
13	\$948	\$116 888	\$3650	\$7727
14	\$948	\$112 994	\$3894	\$7483
15	\$948	\$108 839	\$4155	\$7222
16	\$948	\$104 405	\$4433	\$6944
17	\$948	\$99 675	\$4730	\$6647
18	\$948	\$94 628	\$5047	\$6330
19	\$948	\$89 243	\$5385	\$5992
20	\$948	\$83 498	\$5746	\$5632
21	\$948	\$77 368	\$6130	\$5247

Figura 3.17 Pesquisa no Wolfram Alpha para "mortgage \$150,000, 6.5%, 30 years"

algoritmo em interpretar os vários fatores relevantes e o esforço dos autores em caracterizar os dados publicados. Focando-nos neste último fator e no exemplo apresentado, o desafio dos autores é atualmente de entender a multiplicidade de plataformas em que a sua intervenção é pedida. Como referido no início desta secção, à data da publicação deste trabalho, os resultados relativos à informação das empresas são provenientes de várias fontes, incluindo bases de dados publicadas, páginas do Google+, inserções no Google Maps ou metadados provenientes de páginas Web. Com a introdução do Google My Business<sup>1</sup>, a Google pretende fazer convergir esse esforço, pelo menos no que diz respeito aos seus produtos. Sabendo que essa convergência tem como fonte principal a publicação de dados estruturados baseados em schema.org nos Web Sites oficiais das empresas (Google, 2015b), é também de esperar que essa possibilidade seja suportada por mais motores de pesquisa, para empresas de todo o mundo.

#### 3.3.4 [O caso do Wolfram Alpha](#)

No mercado desde maio de 2009, o Wolfram Alpha<sup>2</sup> disponibiliza respostas com base em fontes de dados estruturados, pelo que pode ser considerado um motor de pesquisa semântica, embora seja descrito pelos seus criadores como um motor de conhecimento computacional, uma vez que inclui funcionalidades de cálculo e processamento de dados que ultrapassam o mero fornecimento de respostas.

Disponível apenas em língua inglesa, o motor de pesquisa aceita perguntas em linguagem natural, embora o utilizador consiga facilmente utilizar uma linguagem simplificada que proporciona resultados com mais confiança. Os resultados são também dirigidos à melhor interpretação possível para as potenciais ambiguidades.

A Figura 3.17 ilustra os pressupostos assumidos para o caso do exemplo apresentado, nomeadamente que se trata de um pedido relativo a uma amortização com taxa fixa em alternativa a uma amortização com taxa variável, a um montante de empréstimo em alternativa a um montante de compra ou de um pagamento inicial e a uma taxa de juro anual em alternativa ao pagamento inicial. Caso um dos pressupostos esteja errado, o utilizador é convidado a alterá-lo com um simples clique. Os resultados

---

<sup>1</sup> [www.google.com/business/](http://www.google.com/business/)

<sup>2</sup> [www.wolframalpha.com](http://www.wolframalpha.com)

apresentados incluem o valor da amortização mensal, os valores acumulados, os gráficos de capital e juros, prestações e capitais e ainda uma tabela de pagamentos completa (visível na imagem apenas parcialmente devido à sua extensão).

A extensão dos dados disponíveis é tal que não é viável apresentar mais exemplos, podendo a plataforma ser utilizada em diversas áreas, nomeadamente Matemática, Línguas, Unidades e Medidas, Estatística, História, Química, Física, Finanças, Cultura, Música, Geografia, Engenharia, entre outras. Os resultados das seguintes pesquisas poderão também ser avaliados, para melhor entendimento do potencial do motor de pesquisa:

- ascii 99
- how many web sites
- 7 day weather forecast for Porto
- orbital period of the moon
- sulphuric acid
- reynolds number
- Aristotle
- How many religions are in Portugal?
- energy consumption Europe with Russia vs North America
- fish consumption portugal, spain
- vehicles in use in Europe
- barcode Portugal
- azores, madeira, canary islands
- star wars vs star trek
- roger moore, sean connery, daniel craig
- oscar best picture 1966
- what is love
- the answer to life, the universe and everything
- how many

### 3.4 Utilização da Web Semântica

O desígnio de suportar uma base de dados à escala global poderia parecer suficiente para entusiasmar a generalidade das pessoas a adotarem a Web Semântica. No entanto, na realidade, a tecnologia parece ser uma alternativa de maior risco às tecnologias mais convencionais, com ferramentas de desenvolvimento escassas, competências difíceis de recrutar, poucas referências de sucesso e a necessidade de uma verdadeira mudança de paradigma na modelização, projeto e programação (Pollock & Oracle, 2008). Para além disso, o entusiasmo precoce pela tecnologia está também limitado pela falta de resposta a questões essenciais. Enquadra-se nessa limitação a questão colocada por Tom Heath sobre que aplicações podem evoluir com a Web Semântica (Heath, 2010), já identificada na Secção 2.4.6 como oportunidade de investigação.

Os Sistemas de Informação têm sido classificados de acordo com várias abordagens, umas mais generalistas, outras com finalidades mais específicas (Lopes, Morais, & Carvalho, 2005). Apesar do elevado número de esforços realizados, a questão da classificação está longe de estar esgotada. Assim, será importante encontrar uma classificação que tenha por base a utilização da Web como ambiente de trabalho, em alternativa ou complementaridade às classificações mais tradicionais.

A maioria das metáforas utilizadas no desenvolvimento de software de HCI (*Human Computer Interface*) traduzem mais rapidamente o mundo físico que o mundo matemático. Os computadores e dispositivos móveis utilizados atualmente mostram documentos, pastas, botões, menus, tampos de secretária, caixotes de lixo, caixas de entrada e outros objetos do mundo físico adaptados ao imaginário digital. A Web também está organizada em locais, ainda que virtuais, páginas, ainda que de dimensão variável, e conteúdo, ainda que digital. Mas introduz o primeiro conceito que altera toda a realidade física, a hiperligação. Parece ser um único passo de esforço reduzido que permite ligar quaisquer dois documentos, mas é também um passo de energia quase infinita que permite ligar os menos prováveis ou mais distantes elementos de conteúdo. Parece representar uma rede comparável às ligações neuronais que configuram todo o

nosso conhecimento. Estas imagens traduzem o potencial da Web, suficientemente demonstrado na atualidade.

Adicionalmente, o crescimento das redes móveis e da utilização de aplicações específicas que suportam e promovem a mobilidade, está a alterar a forma como são utilizados os computadores e dispositivos móveis. Um exemplo disso é a utilização de serviços de armazenamento de ficheiros na Web, disponíveis para vários sistemas operativos, em alternativa ou complementaridade aos suportes físicos, como as recentemente populares *pen drives*. Com a possibilidade de usar uma simples página Web como alvo de uma operação de arrasto de um documento, possibilitada agora pelo HTML 5, as páginas Web passam a ser também um local de armazenamento de fácil utilização. Estas mudanças provocarão inevitavelmente alterações na forma como são arquitetados e desenvolvidos os diferentes tipos de sistemas de informação.

#### 3.4.1 Atividades do utilizador na Web

Poderá a Web não ter qualquer efeito sobre a forma como vivemos, agimos, trabalhamos ou nos relacionamos? As oportunidades de investigação referidas na Secção 2.4.6 parecem antecipar uma resposta negativa. Assim, parece recomendável que se estudem as aplicações tomando como base a própria Web e a forma como é utilizada.

Existem vários estudos conhecidos relativos à utilização da Web. Algumas conclusões apontaram para o facto de o utilizador pesquisar a Web em busca de informação, navegação ou transação (Broder, 2002). A procura, a colheita, a observação, a transação, a comunicação ou a manutenção foram também identificados como atividades (Sellen, Murphy, & Shaw, 2002). É, no entanto, o trabalho de sistematização de Tom Heath que fornece esclarecimentos adicionais sobre esta análise, completando-a com a concentração nas atividades que têm uma finalidade ou propósito, em detrimento das que se referem a meios ou métodos, criando assim uma taxonomia de tarefas.

Com base na sistematização realizada por (Heath, 2010) e estendida neste trabalho, as principais atividades do utilizador desempenhadas com o suporte da Web, podem organizar-se nas seguintes categorias:

- 1) Localização

A procura de um objeto que se presume que exista, como um artigo de jornal, uma fotografia, um número de telefone ou informação sobre um livro.

Esta tarefa é muito frequente, uma vez que a Web se transformou na fonte principal de informação generalizada para todos os fins. Assim, quando se pretende simplesmente obter um número de telefone de uma empresa, é mais provável que se pesquise o nome da empresa na Web em geral, tentando encontrar o Web Site da empresa e, de seguida, encontrar o respetivo número de telefone, do que eventualmente utilizar um serviço que liste os números de telefone das empresas. Ou que se procure um artigo no Web Site de um jornal, em alternativa a procurá-lo na edição impressa, página a página.

Palavras-chave: *locate, look, find*

## 2) Exploração

Recolher informação sobre um determinado assunto ou entidade, com o propósito de o entender ou conhecer, como na exploração de uma teoria para entender os seus fundamentos ou na recolha de informação de uma empresa antes de uma entrevista de emprego.

A dimensão da Web tornou-a na mais importante fonte de conhecimento da atualidade, em praticamente todos os domínios. Tornou-se assim comum utilizar a Web para explorar várias fontes de informação, recolher vários documentos, imagens ou conceitos.

Palavras chave: *explore, gather, research*

## 3) Navegação

Percorrer fontes sem um objetivo em particular, com o propósito de encontrar elementos interessantes, como ao seguir ligações num *blog* que pareçam interessantes ou ao seguir ligações para as notícias mais lidas do dia.

Para além de fonte de informação, a Web está a tornar-se no mais importante meio de entretenimento. É possível assim encontrar diversas fontes que permitem utilizar a Web como se tratasse de um recurso infinito.

Palavras-chave: *graze, browse, follow*

#### 4) Monitorização

Verificar regular ou repetidamente fontes com a intenção de detetar a ocorrência e a natureza de alterações, como na verificação da existência de mensagens novas numa conta de correio eletrónico ou na análise de um fórum de discussão à espera de novas ideias.

A Web é não apenas uma fonte imensa de informação, como é também muito rápida a disseminá-la. A combinação da intervenção generalizada de todos os utilizadores com a extensão e ubiquidade da Web fazem com que, quanto maior for a novidade, maior será a velocidade com que a Web a distribui.

Palavras-chave: *monitor, check, detect*

#### 5) Partilha

Tornar um objeto ou pedaço de informação disponível a terceiros, como na partilha de fotos das férias com um colega ou no carregamento de um artigo de jornal para um Web Site pessoal.

A partilha ou colaboração é um dos aspetos mais disruptivos do efeito da Web na sociedade. A Web proporciona um meio único de colaboração maciça que, com as ferramentas apropriadas e alguma inovação, transforma setores inteiros da economia.

Palavras-chave: *share, distribute, collaborate*

#### 6) Notificação

Informar terceiros de um evento ou uma alteração no estado de uma entidade, como enviar uma mensagem a amigos para os informar de um acontecimento ou bloquear a sua disponibilidade na aplicação de mensagens instantâneas.

A noção de estado, muito usada na Informática, ganhou novo fôlego pela utilização de uma rede de comunicação disponível em permanência, possibilitando, por exemplo, comunicar automaticamente que nos afastamos do computador, através da utilização do atributo “estado” de uma aplicação de mensagens instantâneas.

Palavras-chave: *notify, state, inform, communicate*

#### 7) Afirmação

Produzir afirmações ou opiniões sem nenhuma utilização concreta, como na escrita da crítica de um filme ou na indicação num Web Site que se possui um determinado livro.

A disponibilidade de acesso individual a uma rede de comunicação de massas abriu espaço para que a opinião de qualquer um possa chegar a mais pessoas, da mesma forma que a opinião de cada vez mais pessoas chega a cada indivíduo.

Palavras-chave: *assert, opinion, suggestion*

#### 8) Discussão

Trocar conhecimento e opiniões sobre um determinado assunto com terceiros, como discordar de uma entrada anterior num fórum de discussão ou alterar o conteúdo de uma página de um *Wiki*.

A disponibilidade de acesso individual, em conjunto com um espaço adequado, é também responsável pelo aumento da discussão, do confronto de ideias, da possibilidade de exteriorizar a discordância e, eventualmente, encontrar acordos, consensos ou conhecimento.

Palavras-chave: *discuss, comment, respond*

#### 9) Avaliação

Determinar se uma informação é verdadeira ou analisar várias opções para determinar a melhor, como na escolha do cinema em que se pretende ver um determinado filme ou na comparação de fontes de informação para decisão sobre a mais correta.

Desde a comparação de preços para um dado produto à determinação da bomba de gasolina mais próxima, a disponibilidade de informação de livre acesso diminuiu as barreiras de entrada, mas aumentou a concorrência dos mercados.

Palavras-chave: *evaluate, assess, analyze*

### 10) Combinação

Combinar com terceiros de forma a que algo possa ser possível, como na marcação de uma estadia num hotel disponível para uma dada data ou na colocação de uma encomenda de um produto que existe em stock.

Com a informação disponível, o indivíduo está mais capacitado para tomar decisões mais complexas, aumentando o leque de possibilidades e reduzindo o risco de erro.

Palavras-chave: *arrange, combine, negotiate*

### 11) Transação

Transferir dinheiro ou crédito entre partes, como no pagamento de um serviço através do banco ou na marcação e utilização de um dia de férias disponível.

Tratando-se de um meio de comunicação digital, apenas o que pode ser digitalizado pode ser transportado. O dinheiro eletrónico ganha um estatuto especial de destaque neste contexto, pois ainda se pensa nele como tangível, apesar de cada vez mais ser movimentado virtualmente, em particular utilizando a Web.

Palavras-chave: *transfer, pay, transact*

Este agrupamento de tarefas pode servir de base para o projeto de novas soluções baseadas na Web Semântica, que explorem o seu potencial e produzam resultados relevantes (Heath, 2010). A adoção desta abordagem permite trabalhar sobre um âmbito reduzido e objetivo de funcionalidades.

## 3.5 As Empresas e a Web

Considerando a evolução da Web nos últimos anos, nas suas diferentes fases, diferentes intervenientes desempenharam papéis bem identificados. Na primeira geração, a Web dos documentos, o papel dos autores dos documentos foi determinante, inundando a Web de conteúdos e criando a necessidade que fez do motor de pesquisa uma das suas aplicações de referência. Na segunda geração, a Web social, o papel dos utilizadores em geral foi determinante, como produtores e redistribuidores de conteúdos, criando espaço para o crescimento das redes sociais. Na terceira geração, a Web dos

dados, o papel das máquinas será fundamental, fornecendo informação caracterizada e integrada, criando eventualmente uma base de dados para um profundo desenvolvimento do conhecimento.

As empresas, enquanto organizações que procurando o lucro procuram também utilizar todo o potencial da tecnologia e captar a atenção da sociedade, estiveram presentes ou terão liderado mesmo o crescimento da Web. Na primeira geração, criando a sua própria presença institucional de comunicação pública maciça, fazendo uso intensivo do comércio eletrónico e desenvolvendo a base para o conhecimento organizacional. Na segunda geração, utilizando a rede para promover o envolvimento dos colaboradores, clientes e consumidores, implementando mecanismos de negócio digital que apoiam a interação entre indivíduos e máquinas. Terá já chegado a terceira geração da Web às empresas? E se chegou, estarão as empresas conscientes das oportunidades que lhes são apresentadas? O endereçamento destas perguntas contribuirá para a concretização da proposta de trabalho apresentada nesta tese.

### 3.5.1 Caracterização das empresas

A presença das empresas na Web apresenta características significativas e representativas. Segundo os dados do Eurostat (Eurostat, 2014) e dos relatórios “ICT usage in Enterprises 2010” (Giannakouris & Smihily, 2010) e “ICT usage in Enterprises 2011” (Giannakouris & Smihily, 2011), pode observar-se o seguinte:

- 95% das empresas europeias possuem acesso à Internet, valor que varia entre 94% para as pequenas empresas e 100% para as grandes empresas e que é comparável com 94% em Portugal;
- 69% das empresas europeias possuem um Web Site, valor que varia entre 66% para as pequenas empresas e 92% para as grandes empresas e que é consideravelmente superior aos 52% verificados em média em Portugal;
- 14% da faturação das empresas europeias provém do Comércio Eletrónico, variando entre os 25% da República Checa e os 5% da Itália e 1% de Chipre, com Portugal nos 11%, ligeiramente abaixo da média;

- 21% das empresas europeias possuem uma Extranet, valor que varia entre 18% para as pequenas empresas e 52% para as grandes empresas<sup>1</sup>;
- 40% das empresas europeias partilham informação digital internamente, valor que varia entre os 37% para as pequenas empresas e os 72% para as grandes empresas;
- 19% das empresas partilham informação digital com fornecedores ou clientes, valor que varia entre os 17% para as pequenas empresas e os 40% para as grandes empresas;
- 51% das empresas usam transmissão de dados digitais entre empresas, valor que varia entre os 63% no setor da Reparação de Computadores e Equipamentos, 60% nas Agências e Operadores de Viagens, 60% no setor da Informação e Comunicação e 33% no Alojamento e Restauração.

Com base nesta informação, podem tecer-se algumas considerações:

- O acesso à Web nas Empresas é generalizado e quase absoluto;
- A disponibilidade de um Web Site é ainda uma diferenciação negativa de Portugal em relação à Europa;
- A diferença de utilização das diferentes aplicações entre pequenas e grandes empresas é considerável;
- As diferenças entre setores poderão não ser significativas e apenas um reflexo das características específicas dos seus modelos de negócio.

Tendo em conta o exposto, as conclusões deste trabalho podem ser afetadas tendo em conta a localização geográfica das empresas no plano internacional. Sendo conhecido ainda que Portugal se distingue da Europa de uma forma geral pela menor utilização de tecnologia nas empresas, os resultados deste trabalho podem apresentar desvios, quer porque o interesse e a experiência das empresas sejam mais reduzidos, quer porque a oportunidade de desenvolvimento seja maior.

---

<sup>1</sup> Dados por país não disponíveis para todos os indicadores

Adicionalmente, tendo em conta que a dimensão das empresas é um fator relevante que as distingue na utilização de tecnologia, mais uma vez os resultados podem apresentar desvios com origem nas mesmas causas.

Finalmente, também diferentes setores podem permitir a identificação de mais benefícios que outros, até porque as diferentes áreas de negócio apresentam características que as distinguem, podendo, por isso, influenciar igualmente os resultados.

Estas considerações, embora pertinentes, parecem ser de menor importância quando comparadas com a possibilidade das empresas, em geral, poderem tirar partido de uma tecnologia emergente. Tratando-se a Web Semântica de uma oportunidade demonstrada, não parece adequado limitar o trabalho em função do que distingue as empresas e, pelo contrário, parece mais sensato impulsioná-lo através do foco naquilo que as torna semelhantes. Em conclusão, a localização, a dimensão e o setor de negócio das empresas podem ser interessantes, mas desempenharão nesta fase do trabalho apenas o papel de parâmetros de caracterização da amostra.

Finalmente, deve referir-se que o esforço meritório de alguns investigadores no sentido de promover a utilização da Web dos dados foi ambiciosamente dirigido às nações (Koumenides et al., 2010). Meritório e ambicioso, exigindo esforços vários de influência, relações públicas, investigação, comunicação, desenvolvimento ou manutenção, entre outros. O enorme potencial dos dados públicos é acompanhado também de vários riscos. Questões como o desempenho computacional, interligação de dados, qualidade dos dados e interação com o utilizador são importantes (Janowicz & Hitzler, 2010a). Estas questões colocadas no ambiente dos dados públicos e cruzadas com questões de privacidade e direitos de utilização podem tornar-se críticas. As possibilidades de investigação nesta direção estão longe de estar esgotadas e podem fornecer orientações a seguir no âmbito deste trabalho.

### 3.5.2 A Web Semântica nas Empresas

A abordagem idealista à universalidade da interligação de todos os dados na Web faz com que não existam muitos exemplos de aplicação com sucesso da Web Semântica nas Empresas (Kang et al., 2008). Às dificuldades já referidas na Secção 2.4 acrescentam-

se preocupações quanto à capacidade dos utilizadores em geral, dos decisores de negócio e mesmo dos programadores em entenderem a tecnologia e quantificarem o seu valor para o negócio. Ainda segundo os mesmos autores, a maioria dos exemplos são demasiado complexos, incluindo hierarquias conceptuais e modelos matemáticos e de raciocínio, tentando que as máquinas entendam completamente os dados como as pessoas o fazem.

Entretanto, diversos intervenientes adotaram vocabulários distintos, muitas vezes influenciados por disciplinas, culturas ou nacionalidades, dando origem à necessidade e dificuldade de compatibilizar vocabulários, detetar entidades referidas por múltiplos identificadores ou detetar entidades distintas referidas pelo mesmo identificador. A tão popularizada atribuição de etiquetas (*tags*) está sujeita a inúmeras ambiguidades, tornando impossível saber se “jaguar” é uma referência a um automóvel, a um modelo de avião ou a um animal. Nesse sentido, de um ponto de vista prático, mais esforço deveria ser colocado no desenvolvimento de métodos, ferramentas e práticas que possibilitem a prevenção em substituição da reparação (Hyvönen, 2010).

É ainda importante referir que os desafios da adoção tecnológica frequentemente não dependem da qualidade da investigação efetuada mas sim de modelos de negócio de sucesso e da perceção da complexidade e dos desafios da adoção tecnológica nas empresas (Dimitrov, 2014).

Tendo em conta a natureza distinta das organizações, faz sentido que a abordagem empresarial seja distinta da académica. As empresas procurarão resultados mais práticos, de curto-prazo e com elevada probabilidade de concretização e sucesso, enquanto a investigação académica procura tipicamente resultados teóricos, mais ambiciosos e de médio-prazo. No entanto, demasiada divergência pode causar uma fratura, em que a empresa se foca em questões de pequena dimensão e a academia em aspetos complexos que nunca serão utilizados (Cardoso, Miller, Su, & Pollock, 2008), sendo aconselhável um equilíbrio entre as duas abordagens.

Para além disso, a tecnologia semântica atingiu um ponto de inflexão em que deixou de constituir interesse exclusivo por parte do meio académico e se tornou num tópico de interesse para a indústria (Dau, 2010).

Este trabalho assume assim, a partir deste passo, a necessidade de conhecer o grau de implementação da Web Semântica nas Empresas e de determinar como pode a sua implementação ser promovida e facilitada. Trata-se de uma necessidade que traduz duas dimensões de trabalho, uma orientada para quantificar a adoção das tecnologias e outra orientada para qualificar a forma como essa adoção pode acontecer. A satisfação dessa necessidade será traduzida na contribuição de investigação apresentada no Capítulo 4.



## 4 O Modelo e-swim

É consensual na comunidade científica e na indústria que a tecnologia Web Semântica tem potencial para solucionar problemas relevantes, nomeadamente na modelização de informação empresarial. No entanto, são também reconhecidas limitações na sua implementação nas organizações, conforme descrito na Secção 2.4. Este enquadramento fundamenta a contribuição de investigação proposta neste trabalho.

Este capítulo descreve os pressupostos e a fundamentação teórica que levaram à proposta de um Modelo de Implementação da Web Semântica nas Empresas e termina com a Proposta de Investigação que conduzirá à sua validação.

### 4.1 Preâmbulo

No trabalho apresentado até ao momento, foram identificadas características da tecnologia da Web Semântica e requisitos, ao nível da organização, para a adoção da tecnologia que são sintetizados na Tabela 4.1.

Estas características e requisitos devem, assim, ser considerados no contexto da contribuição efetuada neste trabalho, que deverá garantir a respetiva identificação, implementação ou avaliação. Poder-se-á afirmar que, o suporte destes aspetos favorecerá a adoção da tecnologia Web Semântica nas organizações.

<b>Tecnologia Web Semântica</b>	<b>Requisitos na Organização</b>
Tecnologia acessível e pouco dispendiosa. Homogeneização e validação das fontes de dados. Simplificação da análise de modelos complexos. Disponibilização de ferramentas de desenvolvimento. Modelização, projeto e programação orientadas para os dados distribuídos.	Liderança com conhecimento da existência da tecnologia. Existência de profissionais qualificados para a implementação. Definição de regras e fronteiras ao conhecimento com vista à sua uniformização. Disponibilização de casos de sucesso para partilha.

**Tabela 4.1. Web Semântica: características da Tecnologia e requisitos na Organização para a respetiva adoção**

As interrogações quanto à adoção da tecnologia nas organizações recolhidas ao longo do trabalho até ao momento apontaram numa direção quantitativa (“Qual o grau de implementação...”) mas também numa direção qualitativa (“Como pode ser implementada...”). Coloca-se agora a questão de fundo quanto ao objeto que pode corporizar uma proposta de trabalho, a questão “O quê?”, ou mais concretamente, a questão que pode ajudar a responder à pergunta “Como... e quanto...?”.

Surge, assim, a hipótese de desenvolvimento de um Modelo de Implementação da tecnologia. No contexto científico, um modelo pode ser apenas uma representação aproximada de um fenómeno concreto, uma representação de uma teoria que visa explicar um fenómeno, ou mesmo tomar ambas as formas. A construção do modelo pode ser baseada numa linguagem, na imagem ou socorrer-se de ambas, na tentativa de sintetizar uma explicação teórica parcial para uma realidade mais vasta, complexa e desconhecida (Frigg & Hartmann, 2012).

Relativamente ao conceito de implementação, este engloba a realização de uma ideia ou a concretização de um plano, focando-se na ação e na mudança. Nas Ciências de Computadores, a implementação refere-se à realização material de um artefacto tecnológico (Turner, 2013). No entanto, a indústria das Tecnologias de Informação adota uma definição mais vasta de implementação, associando o conceito ao processo de execução de mudanças planeadas nas estratégias e aplicações de negócio, habitualmente incluindo a fase de aquisição, instalação, configuração e teste dos sistemas, considerando não apenas a colocação em funcionamento de artefactos tecnológicos específicos como também outros aspetos do processo, nomeadamente a análise de requisitos, a identificação de melhores práticas e o envolvimento dos utilizadores (Marakas & O’Brien, 2008).

Para além disso, o responsável pela implementação assume o papel de agente catalisador da mudança, através da interação que potencia com a organização e os respetivos métodos, negociando com todos os intervenientes a aceitação das consequências associadas à mudança (Laudon & Laudon, 2015).

Conclui-se, assim, que um “Modelo de Implementação” poderá ser a conceptualização adequada ao problema de investigação identificado no Capítulo 1, uma vez que endereça os seguintes aspetos:

- Fornecer uma explicação para o problema que auxilia o utilizador não especialista a perceber a sua envolvente
- Criar uma visualização gráfica e linguística da solução que simplifica a análise e a determinação de ações
- Introduzir um agente de mudança a influenciar o resultado, como é o caso do autor neste projeto ou, no futuro, um colaborador da empresa no seu interior;
- Disponibilizar um conjunto de artefactos, por exemplo na forma de ferramentas, que suportem o processo de mudança.

## 4.2 Desenvolvimento do Modelo

Tendo em conta a diversidade de aspetos a considerar na adoção da Web Semântica nas Empresas, um Modelo de Implementação dessa tecnologia deverá considerar várias dimensões. Cada dimensão deverá cobrir a necessidade de satisfazer requisitos para a implementação da tecnologia, numa perspetiva discreta, mas deverá representar simultaneamente o percurso que possibilita a sua implementação, em diferentes estádios, numa perspetiva contínua.

O modelo proposto neste trabalho é designado por “e-swim”, um acrónimo para “*Enterprise Semantic Web Implementation Model*”, e inclui quatro dimensões: Adoção, Proveniência, Acessibilidade e Atividades. A Figura 4.1 apresenta uma ilustração do modelo, que evidencia a estrutura descrita. Nas subsecções seguintes descreve-se a fundamentação associada a cada uma das dimensões do modelo.

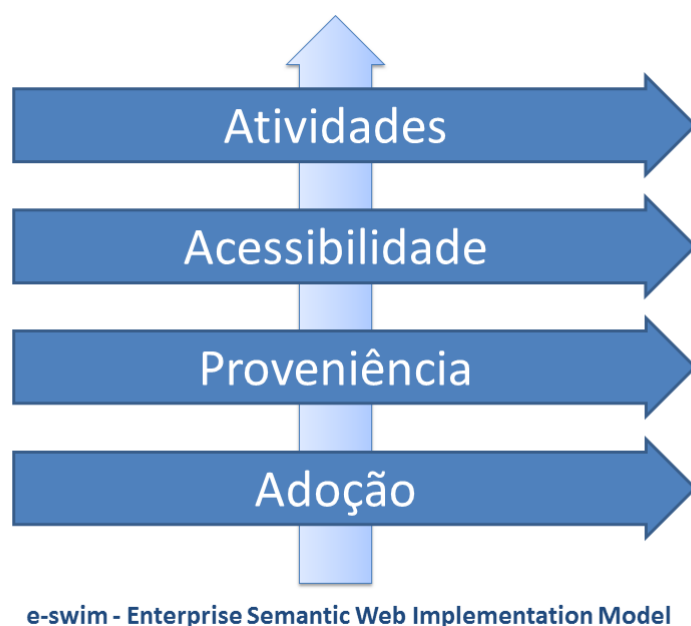


Figura 4.1: Modelo de Implementação da Web Semântica nas Empresas – e-swim

O formato do modelo foi estudado de entre várias alternativas para transmitir uma imagem composta de várias intenções. Em primeiro lugar, a utilização de dimensões quantitativas permite uma abordagem científica em detrimento de práticas de mercado com menor fundamentação. Essas dimensões constituem as 4 dimensões que compõem o modelo e a dimensão global do modelo que as suporta. A dimensão global surge na forma de um pilar, indicando um caminho ascendente que traduz a evolução prevista na implementação da tecnologia com a aplicação do modelo. As 4 dimensões foram posicionadas no modelo, segundo uma ordem ascendente de abordagem, começando com a “Adoção” na base pois constitui uma condição essencial à implementação, como será descrito posteriormente. O posicionamento das dimensões continua com a “Proveniência” e a “Acessibilidade”, dimensões de âmbito mais tecnológico, terminando com a dimensão “Atividades”, de carácter organizacional, que surge em último lugar, como que uma consequência das dimensões anteriores.

#### 4.2.1 Adoção tecnológica

Os principais modelos de adoção tecnológica com aplicação empresarial são o modelo da Difusão da Inovação e o modelo Tecnologia-Organização-Ambiente (Oliveira & Martins, 2011).

Segundo o modelo da Difusão da Inovação (Rogers, 1995), a inovação organizacional está essencialmente dependente das seguintes características:

- Liderança, tendo em conta a atitude em relação à mudança que é adotada pelos indivíduos que lideram a organização;
- Estrutura, considerando a organização das pessoas, a sua disponibilidade, interligação, mobilidade e conhecimento;
- Abertura, analisando a abertura da organização ao ambiente exterior.

Ainda de acordo com Rogers (1995), a inovação é adotada a velocidades diferentes, segundo uma distribuição normal do grau de adoção, sendo, a esse respeito, identificados cinco grupos:

- *Innovators*, o grupo com mais rápida adoção, 2.5% do total;
- *Early adopters*, o grupo com maior aceleração na adoção, 13.5% do total;
- *Early Majority*, o maior grupo em adoção crescente, 34% do total;
- *Late Majority*, o primeiro grupo com adoção a um ritmo decrescente, 34% do total;
- *Laggards*, o grupo com adoção mais tardia e residual, 16% do total.

Por sua vez, o modelo Tecnologia-Organização-Ambiente (Tornatzky & Fleischer, 1990) propõe três aspetos do contexto empresarial que influenciam a inovação tecnológica:

- Tecnologia, que se refere às tecnologias relevantes para a empresa;
- Organização, que se refere à estrutura organizacional e à sua atuação;
- Ambiente, que se refere ao espaço externo onde a empresa se insere, o setor onde atua, os seus concorrentes e a sua relação com o estado.

Estes modelos sobrepõem-se em grande parte mas complementam-se em detalhes que não parecem de desprezar. Assim, propõe-se uma abordagem que permita suportar conclusões referentes a ambos os modelos e que ofereça uma análise abrangente, aqui designada apenas por “Adoção tecnológica” e que inclui as seguintes características:

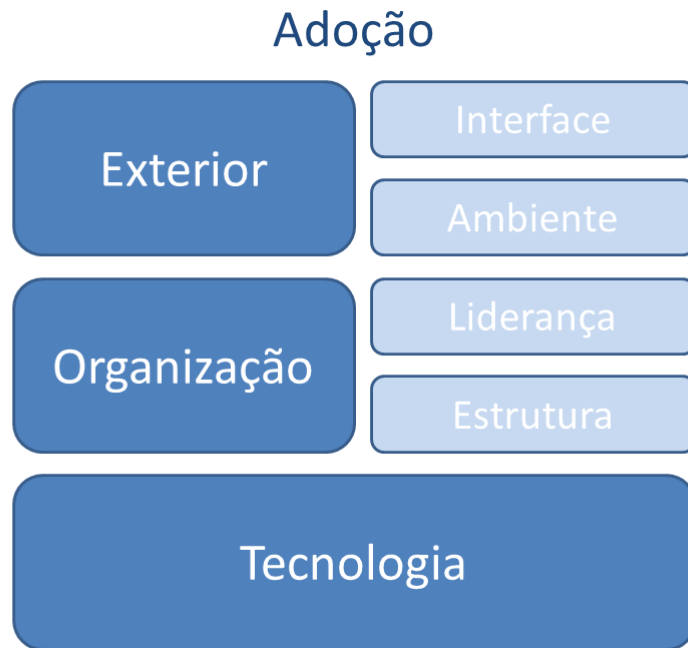


Figura 4.2. A dimensão Adoção tecnológica

- Tecnologia, que inclui as oportunidades de utilização da tecnologia pela organização, neste caso a Web Semântica e as suas aplicações;
- Organização, que se refere à formalização das relações internas, considerando ainda:
  - Liderança, que se refere à atitude em relação à mudança da gestão de topo;
  - Estrutura, que se refere às relações entre pessoas que compõem a organização;
- Exterior, que aborda a vertente externa da organização, no que diz respeito a:
  - Interface, que aborda a abertura da empresa ao exterior;
  - Ambiente, que aborda os intervenientes no espaço onde a empresa se situa.

Estas características e respetivos inter-relacionamentos são ilustrados na Figura 4.2.

Assim, a intervenção planeada nas empresas considera estas características e, através da observação e interrogação, tenta quantificar a posição de cada empresa no caminho da menor para a maior adoção tecnológica.

#### 4.2.2 Proveniência

Esta dimensão do Modelo de Implementação considera a proveniência dos dados como um fator determinante. Assim, a grande questão endereçada é a de identificar as fontes desses dados.

Por um lado, as diferentes Aplicações Web introduzem aspetos de utilização muito concretos que influenciarão a sua contribuição para a adoção da Web Semântica pelas Empresas, nomeadamente aplicações com tipologia de Web Site, Extranet, Intranet, *Web App* ou *Web Service*. Por outro lado, importa situar as fontes de dados no interior ou no exterior da empresa.

A inovação e a vantagem competitiva de uma empresa dependem em larga escala da capacidade desta de lidar com um fluxo permanente e crescente de informação. Consequentemente, os esforços de integração da informação têm de acompanhar esse crescimento. A utilização da Web Semântica nesse esforço de integração pode aumentar substancialmente o seu retorno, reduzindo os custos de integração e aumentando os benefícios decorrentes (Janowicz & Hitzler, 2010a).

A abertura dos dados públicos aos cidadãos representa um avanço na transparência democrática, permitido pela disponibilidade de tecnologia adequada. A esse respeito, destacam-se os esforços dos governos americano e britânico que levaram já à disponibilização alargada de fontes de dados com utilidade generalizada. A necessidade de explorar estas fontes de dados revela-se antes de mais como uma oportunidade de explorar fontes de riqueza (Koumenides et al., 2010).

A dimensão “Proveniência” deve, assim, focar-se na identificação das fontes de dados, internas e externas, e no seu fluxo de utilização, à imagem da tipologia distribuída da Web Semântica.

A Figura 4.3 apresenta um esquema representativo da dimensão “Proveniência”.

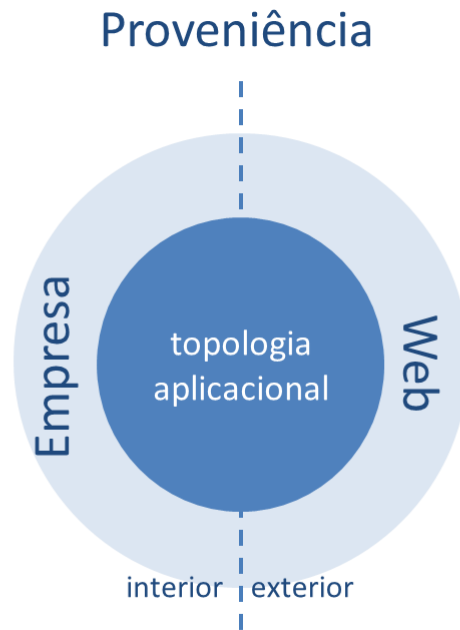


Figura 4.3. A dimensão Proveniência

A este respeito, a intervenção planeada tem como objetivo identificar fontes de dados, internas e externas, no contexto da utilização dos Sistemas de Informação Empresariais e quantificar a sua respetiva utilização.

#### 4.2.3 [Acessibilidade dos dados](#)

A perceção do utilizador comum quanto à acessibilidade dos dados na Web é largamente influenciada pela sua capacidade de entender esses dados sob diferentes formas. O caminho para a disponibilização dos dados de forma universal passa por várias fases, apresentadas e claramente identificadas por Tim Berners-Lee (Berners-Lee, 2006a) na forma de uma classificação com cinco estrelas ou cinco estados de evolução (cf. Figura 4.4) com a seguinte designação e significado para cada um dos estados:

- publicado: apenas disponibilizando dados na Web
- formatado: usando formatos proprietários
- aberto: usando formatos abertos
- semântico: através da utilização de standards semânticos
- ligado: com dados hiperligados

## Acessibilidade

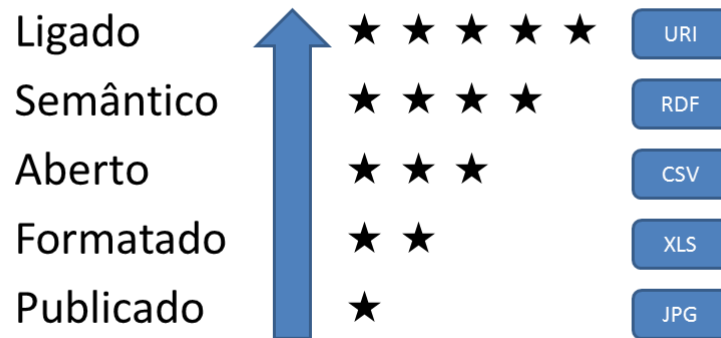


Figura 4.4. A dimensão acessibilidade dos dados (adaptado de (Berners-Lee 2006))

A intervenção planeada determinará o grau de abertura das fontes de dados e, conseqüentemente, quantificará a sua adequação aos standards da Web Semântica.

### 4.2.4 Atividades na Web

A Web Semântica constitui um espaço de utilização alargado que não se restringe à tecnologia standard e aberta disponível atualmente e inclui um largo espetro de soluções em várias áreas. Essas soluções incluirão formas de representação do conhecimento progressivamente mais capazes, que evoluem desde as formas mais simples de organização do conhecimento, como categorizações e taxonomias utilizadas na organização e pesquisa, para as formas mais complexas, baseadas em modelos semânticos que dão resposta a questões colocadas pelos utilizadores (Davis, 2008). A evolução prevista para essas soluções está mais relacionada com a utilização da tecnologia do que propriamente com a tecnologia de suporte.

Conforme referido na Secção 3.4.1, as principais atividades do utilizador desempenhadas com o suporte da Web podem ser sistematizadas e organizadas num conjunto de categorias.

A Figura 4.5 apresenta o espaço de implementação da tecnologia semântica, considerando a diversificação e conceptualização das tarefas e atividades desempenhadas.

Da análise da Figura 4.5 verifica-se que o grau de conceptualização e o grau de diversificação das tarefas desempenhadas influenciam a viabilidade de implementação da tecnologia. As tarefas simples e repetitivas são aquelas que mais naturalmente já são satisfeitas pelas aplicações de software tradicionais e as que menos poderão beneficiar da tecnologia Web Semântica. À medida que a conceptualização das tarefas aumenta, surgem oportunidades para esta tecnologia, mas no auge da conceptualização, quando as tarefas são dificilmente tipificadas e exigem a intervenção criativa de uma pessoa, o esforço para a sua implementação passa a ser superior a simplesmente executar as tarefas.

Por outro lado, sendo as tarefas pouco diversificadas as que mais facilmente são suportadas pelo software tradicional, no limite da diversificação dificilmente será rentável sistematizar as tarefas ao ponto de serem suportadas pela tecnologia semântica. Assim, o limiar ideal para a implementação da tecnologia semântica será aquele em que alguma



Figura 4.5. Caracterização de Tarefas e Atividades segundo a sua Conceptualização e Diversificação (adaptada de (Benjamins et al., 2011))

conceptualização ou diversificação tornam as tarefas demasiado complexas ou demasiado diversificadas para o software tradicional (Benjamins et al., 2011).

A intervenção planeada permitirá identificar as atividades dos utilizadores suportadas por tecnologia Web Semântica.

### 4.3 Proposta de Investigação - Ação

Com a proposta do Modelo de Implementação e-swim, esta contribuição inclui agora uma ferramenta que suportará a conceptualização do problema identificado e permitirá determinar o grau de adoção da Web Semântica nas Empresas, assim como promover e facilitar a sua implementação. Em concreto, determinou-se ser necessário atingir os seguintes objetivos:

- Avaliar o grau de adoção da tecnologia Web Semântica nas empresas;
- Desenvolver um Modelo de Implementação da Web Semântica em ambiente empresarial que facilite e sistematize o processo, evidenciando as suas etapas principais;
- Analisar os resultados da aplicação do Modelo de Implementação da Web Semântica e extrair conclusões.

A partir desta fase, será possível responder à questão de investigação central:

- ***Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?***

Na realização desta tese foi necessário identificar a metodologia de investigação que melhor enquadrasse o âmbito definido para o trabalho e os objetivos a atingir e que pudesse proporcionar resultados relevantes e com um grau aceitável de validação.

Em relação ao objeto principal de estudo, a Web Semântica, as suas tecnologias de base e as suas aplicações, foi necessário idealizar uma intervenção que permitisse uma avaliação adequada, junto dos indivíduos capazes de responder às exigências da análise. Uma vez que o estudo tinha como alvo as empresas em geral, foi necessário assegurar uma aproximação adequada ao ambiente empresarial, que permitisse analisar a utilização das tecnologias. Por outro lado, em relação às aplicações da tecnologia, perspetivou-se

ser necessário verificar em que atividades se enquadravam, tendo em conta as tarefas que os indivíduos desempenhavam em cada momento. Finalmente, tendo em conta a evidência empírica do grau de adoção da Web Semântica ser reduzido na maior parte das empresas, considerou-se pertinente optar por uma metodologia de intervenção e implementação que permitisse alterar esse grau de adoção para que este pudesse ser observável. Tendo em conta estes requisitos, a opção de metodologia recaiu sobre a Investigação – Ação (*Action Research*).

A Investigação - Ação visa o estudo sistemático e refletivo das ações e dos seus efeitos num contexto de trabalho. O investigador partilha a sua perspetiva com os interessados e observa o modo como estes modificam o seu comportamento. De seguida, o investigador partilha a sua análise e idealiza novas ações, numa sequência de ciclos que permite analisar as reações e otimizar os resultados da implementação (Riel, 2010).

A utilização de Investigação - Ação no âmbito deste trabalho permitirá abordar cenários de utilização experimental e recolher resultados que confirmarão ou não as hipóteses colocadas. Por um lado a Ação refere-se à mudança, melhoramento ou implementação num dado ambiente de trabalho. Por outro lado a Investigação refere-se à aprendizagem e compreensão (Dick, 2000). A Investigação - Ação contribui assim simultaneamente para os objetivos científicos e para resolver problemas práticos que se apresentam, através do estudo do sistema e da colaboração dos intervenientes na sua alteração segundo uma direção de interesse comum (O'Brien, 1998). A Investigação - Ação é uma metodologia científica uma vez que os efeitos das ações são observados através da examinação sistemática das evidências. Os resultados deste tipo de investigação são práticos, relevantes e podem contribuir para o desenvolvimento da teoria. A metodologia é, no entanto, diferente de outras formas de investigação pois preocupa-se menos com a universalidade da descoberta e coloca mais valor na sua relevância para o investigador e para os interessados (Riel, 2010).

Para que a Investigação - Ação produza resultados tanto relevantes como rigorosos, devem ser considerados os seguintes requisitos (Baskerville & Wood-Harper, 1996; Davison, Martinsons, & Kock, 2004):

- *Acordo*: Investigador e Empresa acordam num documento que esclarece o âmbito, compromisso, responsabilidades e objetivos da intervenção;
- *Ciclo*: a intervenção deverá ser baseada num ciclo iterativo Diagnóstico – Plano – Intervenção – Avaliação – Reflexão, que proporciona a coexistência da ação, análise e aprendizagem;
- *Teoria*: a intervenção deve ser baseada em teorias relevantes e significativas para o investigador e os interessados, que forneça suporte ao problema, à análise e às observações;
- *Mudança*: as ações implementadas devem introduzir mudanças, ser bem documentadas e avaliadas;
- *Reflexão*: a aprendizagem deve resultar da reflexão do investigador sobre os resultados obtidos e ser comunicada regularmente à comunidade interessada e à comunidade científica.

A Investigação - Ação enquanto metodologia de investigação oferece alguns desafios específicos, nomeadamente por envolver interesses externos à investigação que se pretendem endereçar com o cumprimento destes requisitos e com outras considerações a identificar com a execução do trabalho, tendo em conta a realidade que for sendo enfrentada.

Outras alternativas de metodologias foram equacionadas para a realização deste trabalho, tendo, no entanto, sido preteridas pelos motivos que a seguir se descrevem.

Uma metodologia de investigação que oferece igualmente contributo científico e apoio na resolução prática de problemas é a Investigação Aplicada sobre o Desenho (*Design Research*). No caso de se ter optado pela sua adoção, o objetivo do trabalho deveria ser orientado para a criação de um artefacto tecnológico em oposição ao foco na organização (Cole, Puroo, Rossi, & Sein, 2005), pelo que se revelou menos adequada ao trabalho proposto.

Finalmente, a hipotética adoção da metodologia de Caso de Estudo (*Case Study*) implicaria uma análise não interventiva (Baskerville & Wood-Harper, 1996) numa área onde se identificam poucos exemplos de sucesso concluídos, pelo que poderia colocar em causa o sucesso do trabalho por falta de oportunidades de observação.

O trabalho prossegue assim com uma intervenção de Investigação-Ação nas empresas que pretende responder à questão de investigação central “Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?”

## 5 Utilização do Modelo

A utilização no terreno do Modelo e-swim, apresentado no Capítulo 4, pressupõe a disponibilidade das organizações para a avaliação da implementação da tecnologia Web Semântica no seu ambiente. Com base nas considerações metodológicas apresentadas na Secção 4.3, a aproximação às empresas foi planeada e realizada em várias fases, nomeadamente:

- Iniciando-se com a seleção das empresas segundo critérios bem definidos
- Prosseguindo com a avaliação da sua situação inicial
- Recomendando ações que facilitem a sua evolução
- Terminando com a avaliação final resultante do efeito dessas ações

Este capítulo descreve esse percurso de aproximadamente dois anos numa sequência cronológica que permite apreciar o progresso conseguido.

### 5.1 Seleção das Empresas

Garantir a disponibilidade das empresas para a participação num Projeto de Investigação é uma tarefa que deve ser bem planeada para ter êxito, particularmente garantindo a procura de resultados de cariz mais prático, a obter no curto-prazo e com possibilidade de generalização. Para isso, consideraram-se algumas condições prévias para motivar as empresas à participação no projeto.

#### 5.1.1 [Projeto Raio X](#)

Os aparelhos de radiologia médica habitualmente designados por Raio X, em referência à radiação eletromagnética que lhes serve de base, permitem conseguir um efeito semelhante à visão do interior de um corpo sólido. Expressões como “um Raio X da empresa” ou “um Raio X das finanças da empresa” são habituais no meio empresarial, tal como expressões semelhantes o são noutros contextos.

A Web Semântica é geralmente invisível ao utilizador comum. Ainda que os resultados da utilização da tecnologia possam ser facilmente percebidos, como por

exemplo, no caso de uma pesquisa no Google, com a apresentação do Google Knowledge Graph nos resultados da pesquisa, de forma a permitir navegar pelos diferentes conceitos, o utilizador comum não percebe que este efeito é realmente o resultado de uma ação prévia de disponibilização de conteúdos estruturados. Mais ainda, esse mesmo utilizador comum, apesar de utilizar um Sistema de Informação empresarial com fins que lhe são muito mais próximos, como registar as transações realizadas ou gerir os colaboradores, ignora frequentemente que um modelo de dados estruturados suporta várias atividades dos utilizadores de forma a proporcionar os resultados a que se habituou, sejam eles analisar a evolução da sua faturação ou calcular os custos com pessoal.

A ideia de um Raio X que permitisse “mostrar” aquilo que normalmente não é visto foi assim considerada como potencialmente indutora de uma disponibilidade das empresas à participação no projeto, pelo que este foi simplesmente designado “Projeto Raio X”. O seu objetivo: permitir que as empresas vejam como a Web lhes permite desenvolver o seu negócio.

### 5.1.2 Divulgação dos Resultados

Apesar de, por definição, os resultados do projeto serem públicos, a divulgação pública poderia ser um fator condicionante da decisão de participação das empresas. Seria necessário encontrar então uma forma de participação que satisfizesse este impasse.

Optou-se assim pela criação de um Web Site (cf. Figura 5.1) de acesso privado, em que as empresas pudessem acompanhar a evolução do projeto e aceder aos seus resultados genéricos, com o compromisso de que os resultados de cada empresa fossem apenas do conhecimento da própria empresa. A análise da situação da empresa e da sua evolução ao longo do projeto seria, assim, possível sem que nenhuma situação particular pudesse ser causa de constrangimentos.

O desenvolvimento do Web Site com recurso à aplicação Google Sites<sup>1</sup> foi principalmente motivado, na perspetiva do utilizador, pela popularidade das contas

---

<sup>1</sup> sites.google.com

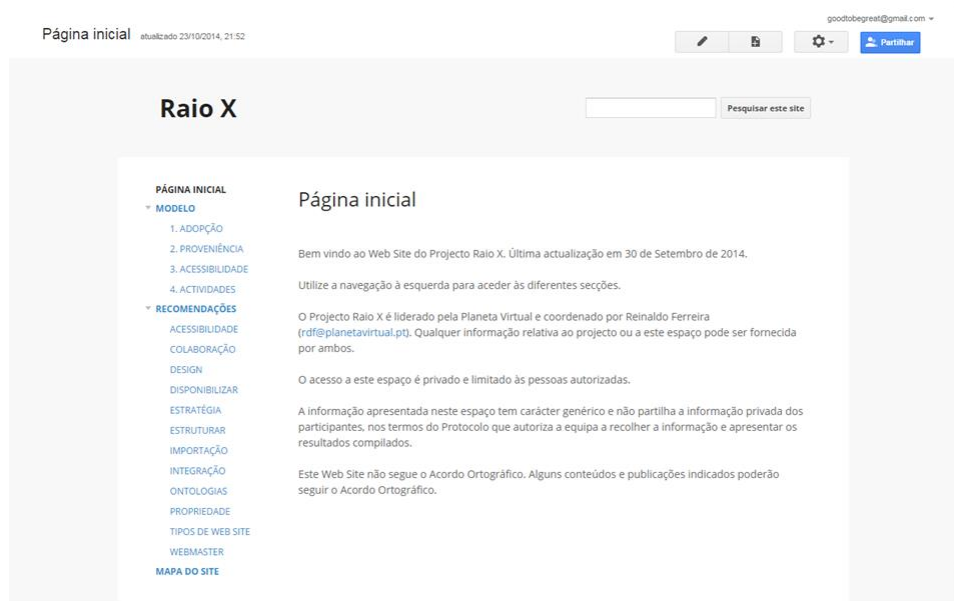


Figura 5.1. Web Site do Projeto Raio X, tal como disponibilizado em 23 de Outubro de 2014

Google, particularmente como consequência da utilização do serviço de correio eletrónico Gmail, e, na perspetiva do autor, pela facilidade e flexibilidade de atualização dos conteúdos.

### 5.1.3 Seleção da amostra

A divulgação do projeto foi efetuada de forma passiva, através da Web, com uma página de apresentação e convite à participação (Planeta Virtual, 2015), e de forma ativa, através de contactos do autor com empresas efetuados no âmbito da sua atividade profissional regular e através de sessões de divulgação do projeto às empresas, o que permitiu alargar o leque de potenciais participantes (cf. Tabela 5.1).

Por outro lado, a possibilidade de influenciar os resultados da utilização do modelo deveria ser acautelada, pelo que não foram consideradas elegíveis à participação no projeto as empresas em que existisse qualquer ligação formal, societária ou laboral, com a empresa promotora do projeto, a título individual e coletivo. Assim, assegurou-se que as decisões tomadas pelas empresas no âmbito do projeto foram totalmente independentes da influência do autor, no papel de consultor, e semelhantes às que resultariam de qualquer simples relação laboral ou comercial.

Data	Local	Empresas participantes
2013/02/15	Universidade Portucalense	5
2013/04/17	Universidade Fernando Pessoa	10
2013/07/16	Instituto de Soldadura e Qualidade	20
2013/09/26	Centro de Estudos da População, Economia e Sociedade	25
2013/11/11	Instituto de Estudos Superiores Financeiros e Fiscais	20
2013/12/10	Instituto de Soldadura e Qualidade	20

Tabela 5.1. Lista de Sessões de Divulgação do Projeto Raio X

#### 5.1.4 Protocolo de Participação

Com vista à formalização do acordo de participação das empresas na avaliação a efetuar, foi elaborado um Protocolo de Participação, que requer o envolvimento de três partes: a empresa, o autor e o patrocinador do projeto, a empresa Planeta Virtual.

A Planeta Virtual é uma empresa da área das Tecnologias da Informação fundada em 1997 por Reinaldo Ferreira, em conjunto com outros sócios, com vasta experiência no desenvolvimento de software, em particular de aplicações de gestão e e-business baseadas em tecnologia Web. A Planeta Virtual possui um Sistema de Gestão da Qualidade implementado de acordo com a norma ISO 9001, no domínio da comercialização, desenvolvimento, implementação e manutenção de software, e um Sistema de Gestão da Investigação, Desenvolvimento e Inovação implementado de acordo com a norma NP 4457. A empresa investe no desenvolvimento de produtos próprios e dos seus colaboradores, participando permanentemente em projetos conjuntos, nomeadamente com Instituições de Ensino Superior, entre as quais se destacam o Instituto Superior de Engenharia do Porto (ISEP), o Instituto de Estudos Superiores Financeiros e Fiscais (IESF) e a Universidade Portucalense (UPT).

Como patrocinadora do “Projeto Raio X”, a Planeta Virtual teve como objetivo compreender a utilização das tecnologias Web em ambiente empresarial, de forma a suportar o desenvolvimento futuro dos seus produtos e soluções para clientes. O seu contributo específico para o projeto consistiu na disponibilização de meios financeiros e materiais para a sua execução.

A minuta do Protocolo que foi elaborado com as empresas é apresentada no Anexo 3.

## 5.2 Estrutura da Intervenção

O trabalho a realizar em cada empresa baseia-se principalmente em dois ciclos de Investigação-Ação interligados, que garantem simultaneamente uma intervenção independente em cada empresa e a reaplicação das ações de Diagnóstico, Plano, Intervenção, Avaliação e Reflexão em todas as empresas.

De uma forma geral, a intervenção foi estruturada em dois ciclos com o objetivo de, em primeiro lugar, recolher evidências individualizadas para cada empresa e, em segundo lugar, generalizar as conclusões de forma a poderem ser reaplicadas em todas as empresas. Em particular, tendo em conta os objetivos científicos específicos, a intervenção ocorreu efetivamente em 4 grandes fases. A relação entre as atividades previstas nos dois ciclos de Investigação-Ação e as fases da intervenção é apresentada na Tabela 5.2.

Investigação-Ação	1º ciclo	2º ciclo
<b>Diagnóstico</b>	Realização de um diagnóstico inicial baseado em observação empírica	Adaptação do diagnóstico aos resultados das avaliações
<b>Plano</b>	Elaboração de um plano de intervenção adequado à realidade de cada empresa	Atualização do plano de intervenção
<b>Intervenção</b>	Início da intervenção	Conclusão da intervenção
<b>Avaliação</b>	Avaliação da intervenção Avaliação Inicial e Intermédia	Avaliação da intervenção Avaliação Final
<b>Reflexão</b>	Reflexões sobre a intervenção	Reflexões finais sobre a intervenção

**Tabela 5.2. Estrutura da Intervenção**

A análise da Tabela 5.2 permite as seguintes constatações. Em primeiro lugar, a tabela tem um carácter provisório, característico de uma ferramenta de investigação. Com efeito, o Modelo de Implementação que se pretende investigar será validado no terreno com a intervenção efetuada nas empresas participantes. Isto significa que, uma vez validado, o modelo poderá ser posteriormente reaplicado noutras empresas. Essa reaplicação terá essencialmente como base o resultado das atividades das Fases 3 e 4 apresentadas na tabela.

Em segundo lugar, existe alguma assimetria entre os dois ciclos e as fases que os constituem, que resulta do carácter experimentalista da intervenção, na sua busca de resultados que podem ou não confirmar as hipóteses colocadas.

As secções seguintes descrevem a evolução da intervenção nas suas diferentes fases, nomeadamente Diagnóstico inicial (Fase 1), Avaliação da evolução das empresas (Fase 2), Elaboração de Recomendações (Fase 3) e Avaliação Final (Fase 4), detalhando as atividades efetuadas em cada uma das fases.

## 5.3 Diagnóstico inicial

Esta secção descreve as empresas participantes e as evidências recolhidas inicialmente em termos de implementação da tecnologia, com vista ao fornecimento de algumas pistas para o desenvolvimento detalhado da intervenção.

### 5.3.1 Caracterização das Empresas

Como resultado do convite à participação das empresas nas condições descritas, o Projeto Raio X desenrolou-se com a participação de 19 empresas. A lista das empresas participantes é apresentada para referência no Anexo 4. Todos os resultados apresentados neste trabalho têm por base essa amostra.

Estão maioritariamente representadas no Projeto as empresas do setor dos serviços, que totalizam 11 das 19 empresas intervencionadas (cf. Figura 5.2).

Analisando a distribuição das empresas relativamente à Classificação de Atividades Económicas em vigor em Portugal, verifica-se que a maior participação é a da Classe M, atividades nas áreas da Gestão, Engenharia e Tecnologia, com 4 empresas em 19 (cf. Figura 5.3). Verifica-se ainda que a dispersão das empresas participantes por 9 classes distintas, de um total de 21, traduz uma representatividade multissetorial

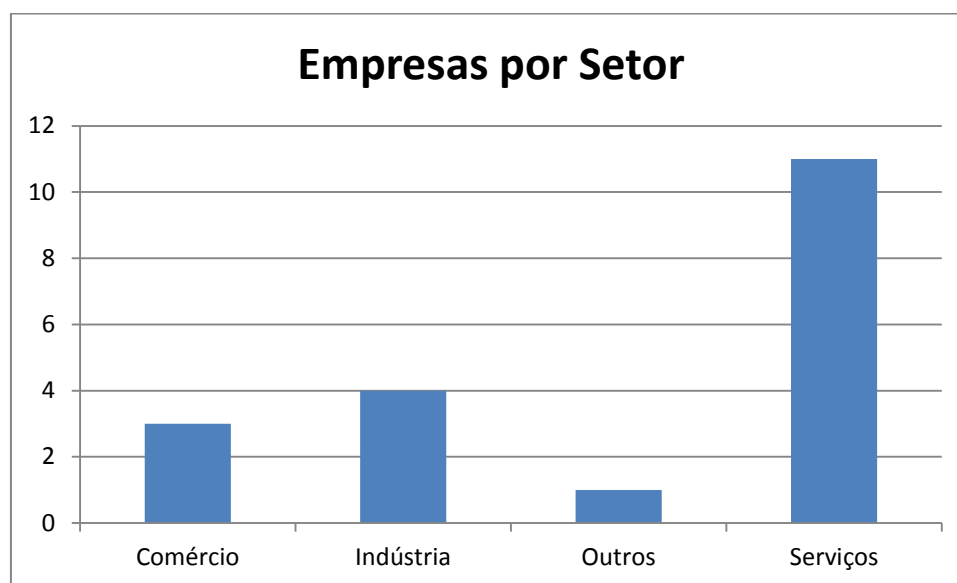


Figura 5.2. Distribuição das empresas por setor de atividade

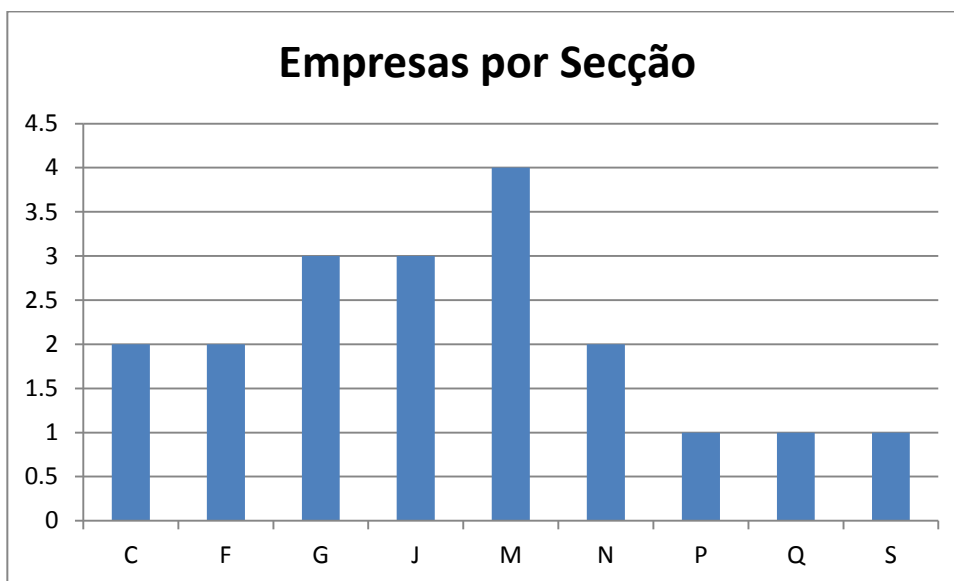


Figura 5.3. Distribuição das empresas por Secção da Classificação de Atividades Económicas

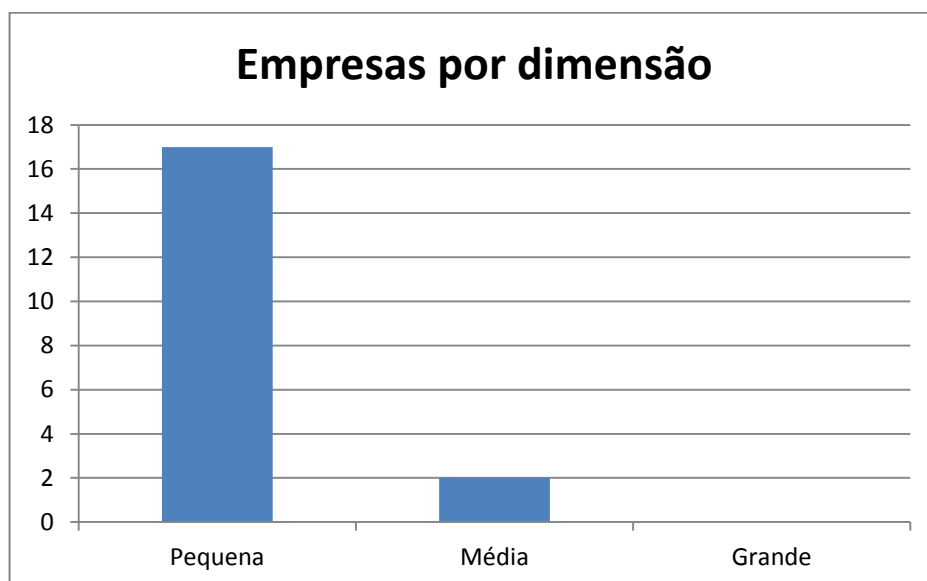


Figura 5.4. Distribuição das empresas por Dimensão

apreciável e desejável para os objetivos do projeto.

Em relação à representatividade da amostra relativamente à dimensão das empresas, de acordo com os critérios definidos pela Comissão Europeia, das 19 empresas participantes, 17 são classificadas como Pequenas Empresas e 2 como Médias Empresas (cf. Figura 5.4). Não participou no estudo qualquer empresa classificada como tendo Grande dimensão.

Dimensão	UE	PT	Estudo
Pequenas	98.7%	99.3%	90%
Médias	1.1%	0.6%	10%
Grandes	0.2%	0.1%	0.0%

**Tabela 5.3. Distribuição das empresas em função da sua dimensão, comparativa entre participantes no projeto e o universo de Empresas Portuguesas e da União Europeia (fonte: (Ecorys & European Commission, 2012; Eurostat, 2012))**

A Tabela 5.3 apresenta uma comparação entre a distribuição das empresas na União Europeia e em Portugal com a distribuição das empresas participantes no estudo, por dimensão.

Assim, as empresas participantes representam o universo de empresas de forma aproximada à realidade, com uma representação na ordem dos 90% para as Pequenas Empresas. No entanto, existe uma representação um pouco superior das Médias Empresas e uma representação inferior das Grandes Empresas. Porém, ambos os grupos só poderiam ser representados de forma mais aproximada se a amostra fosse muito superior.

Em relação à localização geográfica, das 19 empresas participantes, 16 tinham sede no Distrito do Porto e 3 no Distrito de Braga. A proximidade das empresas participantes constituía uma restrição significativa para o desenvolvimento do projeto dada a necessidade de intervir localmente.

Por último, no que diz respeito à forma de contacto utilizada para inclusão no Projeto, das 19 empresas participantes, a inclusão de 3 empresas adveio das sessões de divulgação realizadas, e a das restantes 16 empresas teve origem em convites pessoais.

Tal como já referido, a localização, a dimensão e o setor de negócio das empresas poderiam constituir variáveis interessantes para o estudo realizado, mas não sendo consideradas determinantes, desempenharão neste trabalho apenas parâmetros de caracterização da amostra.

A estimativa de tempo total necessário para a intervenção nas empresas era, nesta fase, de 40 horas por empresa, pelo que o tempo total de intervenção nas empresas se situa em 760 horas, cerca de 5 meses de trabalho a tempo inteiro.

### 5.3.2 Evidências recolhidas

Nesta secção são sistematizadas as evidências empíricas recolhidas no contacto inicial com as empresas e que servem de referência aos passos seguintes.

O contacto inicial com as empresas envolveu na totalidade dos casos a liderança das empresas, principalmente por necessidade de formalizar a sua participação. Na maioria dos casos, este contacto inicial envolveu também colaboradores das empresas com responsabilidade ou maior conhecimento na área das Tecnologias de Informação. As primeiras impressões foram recolhidas de forma não estruturada, com o objetivo de perceber a relação entre a organização, incluindo as suas prioridades, e as Tecnologias de Informação, quanto às oportunidades que apresentam. Apresenta-se de seguida o resumo dessas primeiras impressões.

#### **A tecnologia Web Semântica é desconhecida para praticamente todos os participantes**

Apenas duas empresas participantes no projeto conheciam a existência da tecnologia Web Semântica e dessas, em apenas uma, a tecnologia era conhecida pelos colaboradores, para além da liderança da empresa.

#### **A principal Aplicação Web é o Web Site da Empresa**

Quando questionadas sobre quais as Aplicações Web que a Empresa dispõe, a maioria indicou unicamente o Web Site da Empresa. Quatro empresas participantes não possuíam Web Site no início do projeto e admitiram de imediato que essa falha deveria ser colmatada no curto prazo.

A maioria das empresas referiu ainda que o Web Site não satisfaz as suas expectativas por estar desatualizado, seja nos conteúdos, seja no design gráfico ou estrutura.

Quando foi, entretanto, explorada a existência de outras aplicações nas Empresas, aproximadamente metade das empresas possuía Aplicações Web para além do Web Site, quase sempre uma Intranet.

### **A gestão da presença Web da Empresa é distribuída**

Nenhuma das empresas participantes demonstrou ter uma abordagem sistemática da Presença Web nas diversas vertentes. Apenas quatro empresas indicaram um “Webmaster”, um responsável pela propriedade, desenvolvimento e manutenção da Presença Web da empresa e apenas em uma delas esse responsável era um colaborador interno da empresa. As restantes empresas apenas indicaram a empresa ou profissional que desenvolve e/ou mantém o Web Site, sem outras responsabilidades adicionais.

#### **5.3.3 Primeiras ações implementadas**

Tentou-se com este projeto não colocar barreiras à ação das empresas e permitir que estas pudessem avançar com ações imediatas que considerassem pertinentes, ainda que eventualmente não contribuíssem diretamente para a análise a realizar. As ações seguintes foram registadas em algumas empresas e refletem uma vocação para o curto prazo que lhes é característica.

#### **Identificação do Webmaster**

Uma grande parte das empresas demonstrou uma preocupação imediata com as implicações da falta de atribuição de responsabilidades a um indivíduo que desempenhe o papel de Webmaster e iniciou rapidamente os contactos necessários à sua identificação.

#### **Desenvolver um Web Site**

Com a motivação resultante da participação no Projeto, pelo menos 8 empresas desencadearam de imediato o processo de criação ou alteração do Web Site para incorporar as primeiras alterações resultantes da aprendizagem realizada.

## **Incorporar metadados no Web Site**

Pelo menos 5 empresas procuraram de imediato efetuar a atualização do seu Web Site para incorporar metadados. No entanto, em alguns casos, essa ação não teve qualquer consequência prática imediata, quer porque o software utilizado não possuía uma funcionalidade adequada a essa incorporação, quer porque os profissionais que habitualmente atualizam o Web Site desconheciam a tecnologia a utilizar e adiaram a sua introdução.

## **5.4 Avaliação da evolução das empresas**

Esta secção descreve a ferramenta de avaliação e a avaliação inicial efetuada às empresas e sistematiza algumas reflexões resultantes da intervenção nas empresas.

### **5.4.1 Ferramenta de Avaliação**

O Modelo de Implementação proposto no Capítulo 4 apresenta quatro dimensões distintas e estruturadas. A sua utilização prática requer uma ferramenta de avaliação quantitativa que permita determinar o nível de implementação da tecnologia na empresa em momentos distintos e observar a evolução ocorrida. A ferramenta deverá poder ser utilizada em qualquer momento ao longo da vida da empresa, fornecendo indicações concretas quanto às possibilidades de manutenção e melhoria da implementação da tecnologia.

Na ferramenta desenvolvida, cada dimensão é avaliada separadamente, sendo subdividida em critérios e subcritérios definidos de acordo com a estrutura do Modelo de Implementação. A valorização de cada critério é calculada com base numa média aritmética simples da pontuação obtida nos subcritérios. A avaliação de cada dimensão e a avaliação global são também resultantes da média aritmética das avaliações obtidas respetivamente nos critérios e nas dimensões que a compõem. Foi considerada a possibilidade de introduzir uma ponderação na avaliação, sendo essa hipótese posteriormente abandonada por não se encontrarem fatores significativos para esse efeito, nomeadamente que pudessem introduzir mais confiança nos resultados do que ambiguidades resultantes de considerações adicionais.

A opção por escalas descritivas resultou da vontade de contribuir para resultados mais objetivos através da utilização de descrições complementares que fornecem opções mais concretas aos avaliadores. No âmbito deste projeto, o facto do avaliador ser o mesmo em todas as empresas participantes permitiu uma garantia prévia de homogeneidade de critérios. No entanto, a utilização futura da ferramenta por outros avaliadores será mais eficaz se estes tiverem acesso a descrições detalhadas que reduzam ao máximo a ambiguidade das respostas. Por outro lado, os interlocutores nas empresas mostraram alguma ansiedade quanto à representatividade de uma avaliação quantitativa na comparação de empresas distintas, pelo que esta descrição exaustiva permitiria reforçar a confiança na avaliação por parte dos participantes.

Em cada critério e subcritério é utilizada uma escala discreta de 0 a 5 pontos. Para maior homogeneidade nos resultados a obter, optou-se por considerar todos os valores calculados arredondados a uma casa decimal.

A ferramenta de avaliação foi construída progressivamente a partir das observações empíricas do início da intervenção. A primeira versão da ferramenta foi testada com recurso a 3 voluntários não envolvidos no projeto para determinar eventuais falhas ou ambiguidades. A versão final foi posteriormente utilizada em todos os momentos de avaliação, com todas as empresas.

A Tabela 5.4 apresenta a estrutura resumida da ferramenta de avaliação e do cálculo da avaliação dos subcritérios, critérios e dimensões.

Dimensão	Critério	Subcritério	Avaliação	Cálculo
	Tecnologia		<input type="checkbox"/>	
		Liderança	<input type="checkbox"/>	
		Estrutura	<input type="checkbox"/>	
	Organização		<input type="checkbox"/>	
		Ambiente	<input type="checkbox"/>	
		Interface	<input type="checkbox"/>	
	Exterior		<input type="checkbox"/>	
Adoção			<input type="checkbox"/>	
	Empresa		<input type="checkbox"/>	
	Web		<input type="checkbox"/>	
Proveniência			<input type="checkbox"/>	
Acessibilidade			<input type="checkbox"/>	
Atividades			<input type="checkbox"/>	
e-swim			<input type="checkbox"/>	

Tabela 5.4. Estrutura resumida da Ferramenta de Avaliação

De seguida apresenta-se a estrutura completa da ferramenta de Avaliação. Para maior clareza, apresenta-se uma avaliação fictícia realizada a título de exemplo.

### **Avaliação da dimensão “Adoção”**

A avaliação da dimensão Adoção, a mais complexa das 4 avaliações, processa-se de acordo com as tabelas seguintes.

	Subcritério	Critério	Dimensão
<b>Quais são as oportunidades de utilização da tecnologia?</b>		<b>Tecnologia</b>	<b>Adoção</b>
Não existe conhecimento da existência da tecnologia.		0	
A existência da tecnologia é conhecida mas não existem oportunidades identificadas.		1	
Para as oportunidades identificadas não existem planos de implementação concretos.		2	
Existem planos de implementação para a tecnologia em áreas específicas.		3	
A tecnologia é utilizada em aplicações específicas.		4	
A tecnologia é utilizada ou está planeado utilizá-la de forma alargada nas diferentes aplicações utilizadas.		5	

O critério “Tecnologia” é assim avaliado de forma simples, com a avaliação exemplo a obter 4 valores.

	Subcritério	Critério	Dimensão
<b>Qual a atitude da Liderança em relação à tecnologia?</b>	<b>Liderança</b>	<b>Organização</b>	<b>Adoção</b>
A liderança desconhece a existência da tecnologia.	0		
A liderança conhece a existência da tecnologia mas não conhece aplicação prática identificada.	1		
A liderança conhece aplicações práticas embora não conheça planos de implementação concretos.	2		
A liderança conhece e promove planos de implementação para a tecnologia em áreas específicas.	3		
A liderança conhece e promove a utilização da tecnologia em aplicações específicas.	4		
A liderança conhece e promove a tecnologia e planeia utilizá-la de forma alargada nas diferentes aplicações utilizadas.	5		

	Subcritério	Critério	Dimensão
<b>Qual a atitude da estrutura em relação à tecnologia?</b>	<b>Estrutura</b>	<b>Organização</b>	<b>Adoção</b>
A estrutura (estrutura organizacional, colaboradores, equipas, departamentos, ...) desconhece a existência da tecnologia.	0		
A estrutura conhece a existência da tecnologia mas não identifica aplicação prática.	1		
A estrutura conhece aplicações práticas embora não conheça planos de implementação concretos.	2		
A estrutura conhece planos de implementação para a tecnologia em áreas específicas.	3		
A estrutura conhece e participa na utilização da tecnologia em aplicações específicas.	4		
A estrutura conhece a tecnologia e utiliza-a de forma alargada nas diferentes aplicações.	5		

A média aritmética da avaliação obtida nos subcritérios “Liderança” e “Estrutura” resulta na avaliação do critério “Organização”. Assim, a avaliação exemplo, com os valores parciais de 3 e 3, obtém uma média de 3.0 valores, que será a valorização obtida para o critério “Organização”.

	Subcritério	Critério	Dimensão
		<b>Organização</b>	<b>Adoção</b>
		3.0	

	Subcritério	Critério	Dimensão
<b>Qual a atitude dos intervenientes no ambiente relevante?</b>	<b>Ambiente</b>	<b>Exterior</b>	<b>Adoção</b>
O ambiente (setor, subsetor, região, ...) ignora a existência da tecnologia.	0		
O ambiente conhece a existência da tecnologia mas não apresenta aplicações práticas.	1		
O ambiente conhece aplicações práticas e participa em planos de implementação concretos.	2		
O ambiente apresenta planos de implementação para a tecnologia em áreas específicas.	3		
O ambiente lidera a utilização da tecnologia em aplicações específicas.	4		
O ambiente lidera a utilização da tecnologia de forma alargada.	5		

	Subcritério	Critério	Dimensão
<b>Qual a atitude do interface em relação à tecnologia?</b>	<b>Interface</b>	<b>Exterior</b>	<b>Adoção</b>
O interface (parceiros, fornecedores, clientes, escolas, associações, ...) ignora a existência da tecnologia.	0		
O interface conhece a existência da tecnologia mas não apresenta aplicações práticas.	1		
O interface conhece aplicações práticas e participa em planos de implementação concretos.	2		
O interface apresenta planos de implementação para a tecnologia em áreas específicas.	3		
O interface lidera a utilização da tecnologia em aplicações específicas.	4		
O interface lidera a utilização da tecnologia de forma alargada.	5		

A média aritmética dos valores obtidos na avaliação dos subcritérios “Ambiente” e “Interface” resulta na avaliação do critério “Exterior”. Assim, a avaliação exemplo obtém no critério “Exterior” 2.0 valores, que corresponde à média dos valores 1 e 3 obtidos nos dois subcritérios.

	Subcritério	Critério	Dimensão
		<b>Exterior</b>	<b>Adoção</b>
		2.0	

Por último, a avaliação da dimensão “Adoção” é calculada com base na média aritmética dos valores obtidos nos critérios “Tecnologia”, “Organização” e “Exterior”. No exemplo, esses valores foram respectivamente de 4.0, 3.0 e 2.0 valores, pelo que a avaliação da dimensão resulta em 3.0 valores (9.0 valores no total, divididos por 3 critérios).

	Subcritério	Critério	Dimensão
			<b>Adoção</b>
			3.0

### **Avaliação da dimensão “Proveniência”**

A avaliação da dimensão Proveniência processa-se de acordo com as tabelas seguintes.

	Subcritério	Critério	Dimensão
<b>Quais são e como são utilizadas as fontes de dados internas?</b>		<b>Empresa</b>	<b>Proveniência</b>
Não são utilizadas fontes de dados na empresa (para além de ser conhecido que as mesmas existem como suporte aos Sistemas de Informação utilizados).		0	
Existem fontes de dados dispersas e esforços pontuais, não estruturados, para a sua utilização.		1	
Existem fontes de dados bem identificadas e utilizadas regularmente de forma pouco estruturada (p. ex. Excel, SharePoint)		2	
Algumas fontes de dados são utilizadas para satisfazer as necessidades de outros subsistemas, internos ou externos (p. ex. Extranet)		3	
São utilizadas várias fontes de dados, integradas com os subsistemas (p. ex. <i>Business Intelligence</i> )		4	
Existe uma estratégia global de utilização das várias fontes de dados de forma a maximizar a sua utilização da forma mais eficiente para a organização.		5	

	Subcritério	Critério	Dimensão
<b>Quais são e como são utilizadas as fontes de dados externas?</b>		<b>Web</b>	<b>Proveniência</b>
Não são utilizadas fontes de dados externas à empresa.		0	
Existem fontes de dados isoladas e esforços pontuais, não estruturados, para a sua utilização.		1	
Existem fontes de dados bem identificadas e utilizadas regularmente de forma pouco estruturada (p. ex. listagens de produtos)		2	
Algumas fontes de dados são utilizadas para satisfazer as necessidades de outros subsistemas, internos ou externos (p. ex. atualização de câmbios)		3	
São utilizadas várias fontes de dados externas, integradas com os subsistemas da organização (p. ex. <i>Supply Chain</i> )		4	
Existe uma estratégia global de utilização das várias fontes de dados externas de forma a maximizar a sua utilização da forma mais eficiente para a organização.		5	

A avaliação da dimensão “Proveniência” é calculada com base na média aritmética dos critérios “Empresa” e “Web”. Assim, a avaliação exemplo obtém nesta dimensão 3.0 valores, a média dos dois resultados de 3 valores obtidos nos 2 critérios.

	Subcritério	Critério	Dimensão
			<b>Proveniência</b>
			3.0

### **Avaliação da dimensão “Acessibilidade”**

A avaliação desta dimensão processa-se de acordo com a tabela seguinte.

	Subcritério	Critério	Dimensão
<b>Como são estruturadas as fontes de dados utilizadas?</b>			<b>Acessibilidade</b>
Não são utilizadas fontes de dados na empresa (para além de ser conhecido que as mesmas existem como suporte aos Sistemas de Informação utilizados).			0
As fontes de dados dão origem apenas a utilizações não estruturadas (p. ex. PDF, JPG).			1
As fontes de dados dão origem apenas a utilizações estruturadas de acordo com aplicações específicas (p. ex. XLS, SQL).			2
Algumas fontes de dados são disponibilizadas de forma aberta (p. ex. CSV, XML)			3
São utilizadas fontes de dados baseadas em tecnologia Web Semântica (p. ex. RDF, RDFa)			4
São utilizados URIs para identificar os dados de fontes abertas e baseadas em tecnologia Web Semântica.			5

A avaliação da dimensão “Acessibilidade” é resultante da avaliação de um único critério. No caso do exemplo, resulta numa avaliação de 4 valores.

#### **Avaliação da dimensão “Atividades”**

A avaliação da dimensão “Atividades” processa-se de acordo com a tabela seguinte.

	Subcritério	Critério	Dimensão
<b>Quais são as atividades suportadas na tecnologia?</b>			<b>Atividades</b>
Não são conhecidas atividades que utilizem ou sejam suportadas por tecnologia Web Semântica.			0
Apenas uma atividade foi identificada como utilizadora ou tendo como suporte a tecnologia.			1
Foram identificadas apenas atividades diversas ou realizadas por indivíduos isolados.			2
Foram identificadas pelo menos 4 atividades realizadas de forma pouco sistemática.			3
Foram identificadas pelo menos 6 atividades realizadas de forma sistemática por diversos indivíduos.			4
Foram identificadas 8 ou mais atividades com utilização sistemática da tecnologia por toda a organização.			5

A avaliação da dimensão “Atividades” é resultante da avaliação de um único critério. Assim, a avaliação exemplo resulta em 3 valores.

#### **Formulário de aplicação da ferramenta de avaliação**

A Figura 5.5 apresenta um formulário elaborado para facilitar a aplicação no terreno da ferramenta de avaliação. No Anexo 5 apresenta-se a ferramenta de uma forma completa e com maior resolução. O formulário de avaliação é disponibilizado no formato de folha de cálculo, incluindo as fórmulas necessárias à realização de todos os cálculos e alguma formatação condicional e validação que facilitem e validem o processo.

e-swim		Modelo de Implementação da Web Semântica nas Empresas						
Empresa: Exemplo								
Dimensão	Avaliação						Avaliação Inicial	Avaliação Final
	Subcritério	0	1	2	3	4	5	
Tecnologia			1			4		
	Não existe conhecimento da existência da tecnologia.	A existência da tecnologia é conhecida mas não existe aplicação prática identificada.	Para as aplicações práticas identificadas não existem planos de implementação concretos.	Existem planos de implementação para a tecnologia em áreas específicas.	A tecnologia é utilizada em aplicações específicas.	A tecnologia é utilizada e está planejado ou está planejado utiliza-la de forma alargada nas diferentes aplicações.	1.0	4.0
Liderança		0			3			
	A liderança desconhece a existência da tecnologia.	A liderança conhece a existência da tecnologia mas não conhece aplicação prática identificada.	A liderança conhece aplicações práticas embora não conheça planos de implementação concretos.	A liderança conhece e promove planos de implementação para a tecnologia em áreas específicas.	A liderança conhece e promove a utilização da tecnologia em aplicações específicas.	A liderança conhece e promove a tecnologia e planeia utilizá-la de forma alargada nas diferentes aplicações utilizadas.		
Estrutura			1		3			
	(estrutura organizacional, colaboradores, equipas, departamentos, ...) desconhece a existência da tecnologia.	A estrutura conhece a existência da tecnologia mas não identifica aplicação prática.	A estrutura conhece aplicações práticas embora não conheça planos de implementação concretos.	A estrutura conhece planos de implementação para a tecnologia em áreas específicas.	A estrutura conhece e participa na utilização da tecnologia em aplicações específicas.	A estrutura conhece e utiliza-a de forma alargada nas diferentes aplicações.		
Organização								
Ambiente			1					
O ambiente (setor, subsector, região, ...) ignora a existência da tecnologia.	conhece a existência da tecnologia mas não apresenta aplicações práticas.	conhece aplicações práticas e participa em planos de implementação concretos.	apresenta planos de implementação para a tecnologia em áreas específicas.	O ambiente lidera a utilização da tecnologia em aplicações específicas.	O ambiente lidera a utilização da tecnologia de forma alargada.			
Interface			1		3			
O interface (parceiros, fornecedores, clientes, escolas, associações, ...) ignora a existência da tecnologia.	O interface conhece a existência da tecnologia mas não apresenta aplicações práticas.	O interface conhece aplicações práticas e participa em planos de implementação concretos.	O interface apresenta planos de implementação para a tecnologia em áreas específicas.	O interface lidera a utilização da tecnologia em aplicações específicas.	O interface lidera a utilização da tecnologia de forma alargada.			
Exterior								
Adoção								
Empresa				2	3			
Não são utilizadas fontes de dados na empresa (para além de ser conhecido que as mesmas existem como suporte aos Sistemas de Informação).	Existem fontes de dados dispersas e esforços pontuais, não estruturados, para a sua utilização.	Existem fontes de dados bem identificadas e utilizadas regularmente de forma pouco estruturada (p. ex. Excel, SharePoint).	Algumas fontes de dados são utilizadas para satisfazer as necessidades de outros subsistemas, internos ou externos (p. ex. Extranet).	São utilizadas várias fontes de dados, integradas com os subsistemas (p. ex. Business Intelligence).	Existe uma estratégia global de utilização das várias fontes de dados de forma a maximizar a sua utilização de forma mais eficiente para a organização.	2.0	3.0	
Web		0			3			
Não são utilizadas fontes de dados externas à empresa.	Existem fontes de dados isoladas e esforços pontuais, não estruturados, para a sua utilização.	Existem fontes de dados bem identificadas e utilizadas regularmente de forma pouco estruturada (p. ex. listagens de produtos).	Algumas fontes de dados são utilizadas para satisfazer as necessidades de outros subsistemas, internos ou externos (p. ex. atualização de câmbio).	São utilizadas várias fontes de dados externas, integradas com os subsistemas da organização (p. ex. Supply Chain).	Existe uma estratégia global de utilização das várias fontes de dados externas de forma a maximizar a sua utilização de forma mais eficiente para a organização.	0.0	3.0	
Proveniência								
Acessibilidade			1			4		
Não são utilizadas fontes de dados na empresa (para além de ser conhecido que as mesmas existem como suporte aos Sistemas de Informação).	As fontes de dados têm origem apenas a utilizações não estruturadas (p. ex. PDF, JPG).	As fontes de dados têm origem apenas a utilizações estruturadas de acordo com aplicações específicas (p. ex. XLS, SQL).	Algumas fontes de dados são disponibilizadas de forma aberta (p. ex. CSV, XML).	São utilizadas fontes de dados baseadas em tecnologia Web Semântica (p. ex. RDF, R2FA).	São utilizados URLs para identificar os dados de fontes abertas e baseadas em tecnologia Web Semântica.	1.0	4.0	
Atividades					3			
Não são conhecidas atividades que utilizem ou sejam suportadas por tecnologia Web Semântica.	Apenas uma atividade foi identificada como utilizadora ou tendo como suporte a tecnologia.	Foram identificadas apenas atividades diversas ou realizadas por indivíduos isolados.	Foram identificadas pelo menos 4 atividades realizadas de forma pouco sistemática.	Foram realizadas 6 atividades realizadas de forma sistemática por diversos.	Foram identificadas 8 ou mais atividades com utilização sistemática da tecnologia por toda a organização.	0.0	3.0	
<b>Total</b>							0.7	3.3
<b>e-swim</b>							14	66

Figura 5.5. Ferramenta de Avaliação

A avaliação num determinado momento apenas permite conhecer a situação momentânea da empresa. Como o que se pretende é orientar a empresa para um percurso de implementação, a avaliação será efetuada em dois momentos distintos, permitindo conhecer a evolução da empresa como resultado da implementação das ações. Assim, são incluídas duas avaliações no formulário, denominadas “Avaliação Inicial” e “Avaliação Final”.

Percebeu-se, com a evolução do projeto, que a escala numérica de avaliação era adequada aos critérios de avaliação definidos mas resultava numa avaliação global pouco representativa. Por um lado, a implementação precoce verificada na maioria das empresas, que será objeto de análise detalhada a seguir, implicava naturalmente valores de avaliação baixos, e que simultaneamente pareciam estranhos (p. ex. 0.6). Por outro lado, a perceção da evolução que a empresa pode efetuar é pouco clara numa escala de 0 a 5 valores e é mais evidente quando se utiliza uma escala percentual (utilizando o mesmo exemplo, 12%). Assim, proporcionalizou-se a avaliação final multiplicando o resultado por 20, para que uma implementação nula corresponda a 0 pontos e uma implementação plena corresponda a 100 pontos. A variável que contém a avaliação final foi designada de “e-swim”, tal como o acrónimo do Modelo de Implementação, e é apresentada no final do formulário.

Assim, a avaliação apresentada a título de exemplo obtém nas 4 dimensões, respetivamente, 3.0, 3.0, 4.0 e 3.0 valores, ficando com um resultado de 3.3 valores, equivalente a uma avaliação e-swim de 66 pontos percentuais. O exemplo é ilustrado pela Avaliação Final da Figura 5.5.

#### 5.4.2 [A avaliação inicial](#)

No caso específico da utilização da ferramenta de avaliação no âmbito deste projeto, os momentos principais a considerar serão naturalmente o início e o final do projeto. No entanto, a necessidade de validar o Modelo e-swim, efetuar um acompanhamento mais próximo do trabalho e promover a motivação das empresas participantes aconselharam uma abordagem mais detalhada. Assim, foram considerados já nesta fase dois momentos de avaliação:

- Avaliação Inicial, referente ao momento imediatamente anterior ao início do projeto, uma vez que a própria decisão de adesão já influenciou positivamente alguns dos critérios de avaliação
- Avaliação Intermédia, referente à melhor previsão para o momento final do projeto, considerando as possibilidades e expetativas da empresa e do investigador

Foram assim efetuadas duas avaliações a todas as empresas participantes. A Avaliação Inicial coloca a variável e-swim no valor médio de 6 pontos percentuais, um valor que traduz fidedignamente a evidência empírica verificada. A Avaliação Intermédia para as empresas participantes atingiu o valor médio de 48 pontos percentuais (cf. Figura 5.6), indicando uma evolução expressiva que será seguidamente objeto de análise mais detalhada.

A análise da Figura 5.7 evidencia que essa evolução acontece de forma equitativa em todas as dimensões do modelo, sem nenhum destaque em particular.

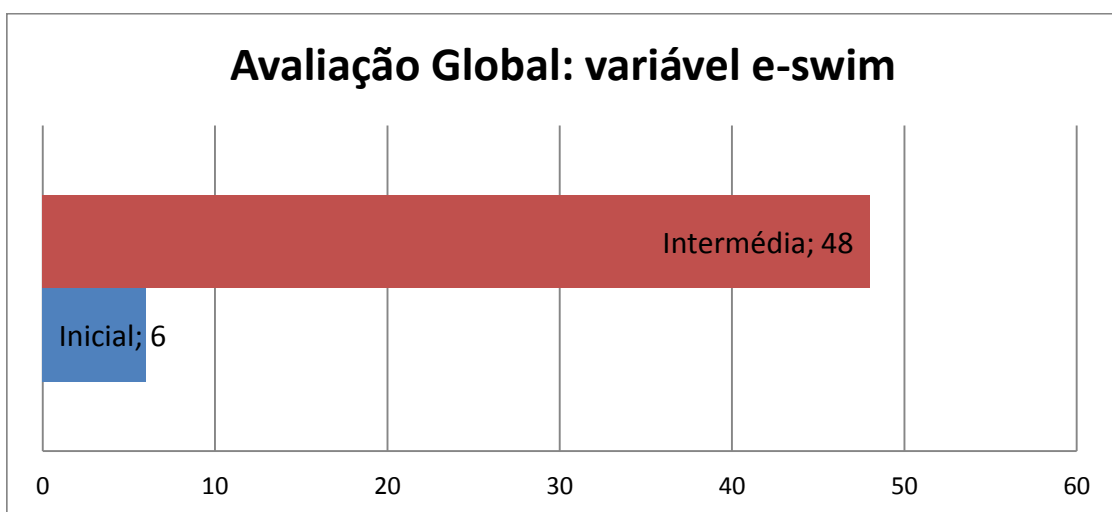


Figura 5.6. Avaliação Global: variável e-swim

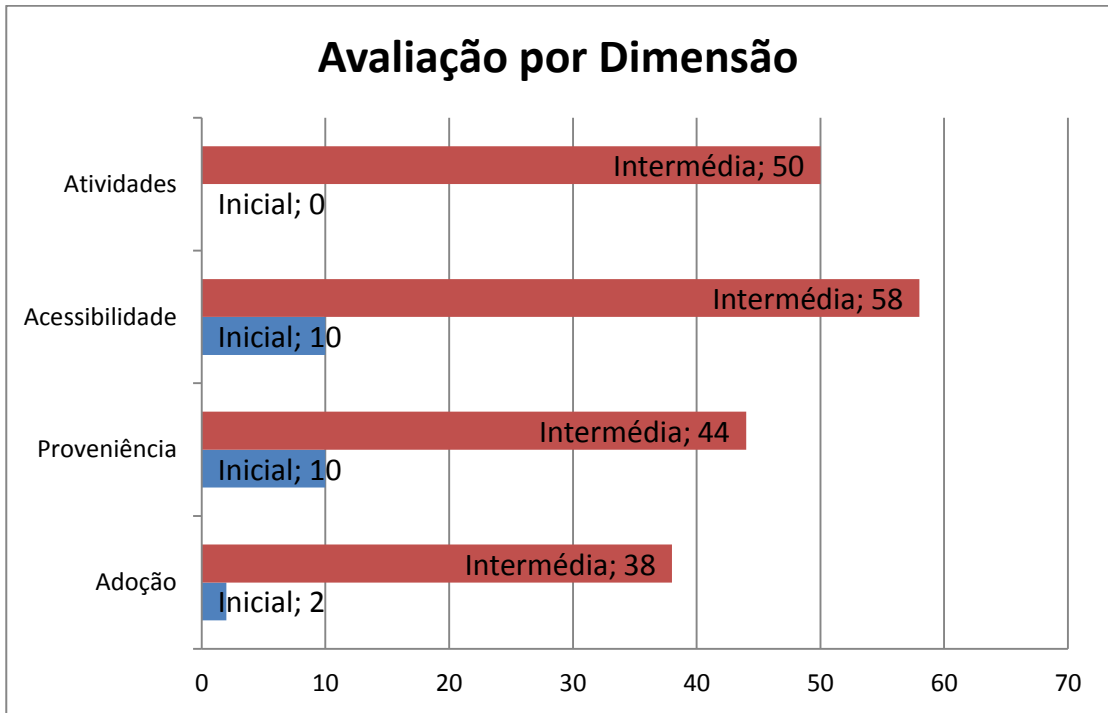


Figura 5.7. Avaliação por Dimensão

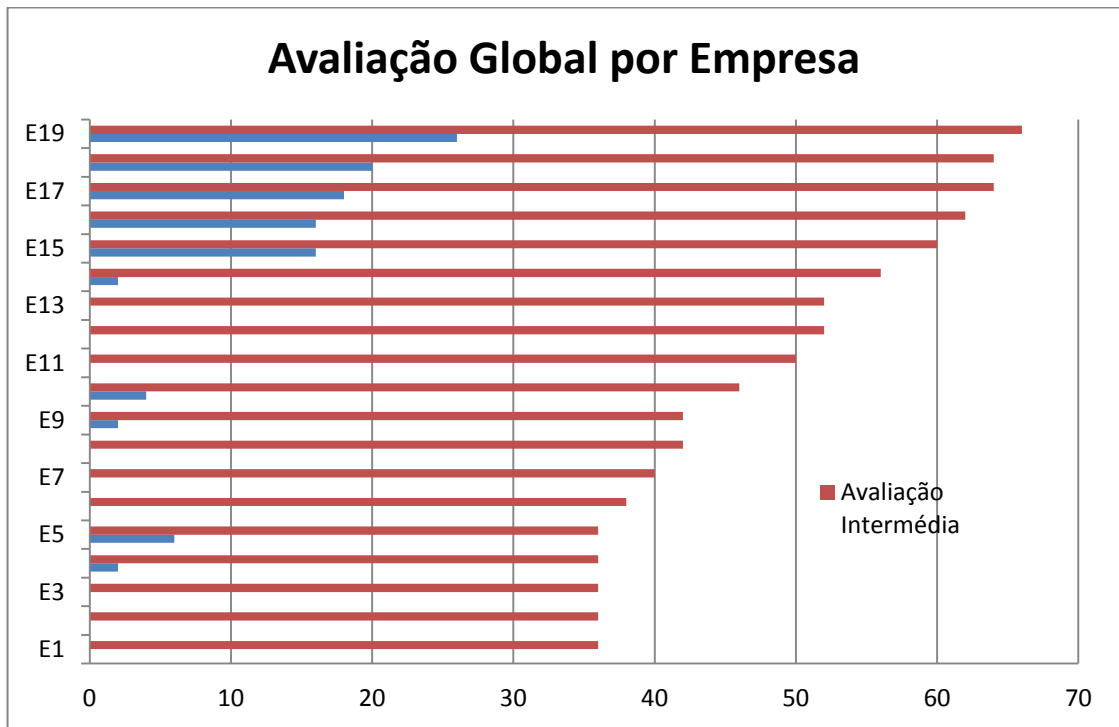


Figura 5.8. Avaliação Global por Empresa

Na dimensão Adoção, regista-se a menor das evoluções, 36 pontos, de 2 para 38 pontos percentuais, e na dimensão Atividades regista-se a maior, 50 pontos, de 0 para 50 pontos percentuais.

Em relação à evolução prevista para cada empresa, registou-se também uma evolução muito proporcional, sem destaques fundamentais. Essa evolução pode ser visualizada na Figura 5.8.

De uma forma geral, as evoluções menores prevêem-se nas empresas que partiram de avaliações mais baixas, com 5 casos que evoluem para 36 pontos percentuais, sendo que um deles parte de 6 pontos, outro de 2 pontos e três de apenas zero pontos. Por outro lado, as maiores evoluções prevêem-se numa empresa que partiu de 2 pontos e deverá evoluir para 56 pontos e em duas empresas que partiram de avaliações nulas para uma avaliação intermédia de 52 pontos.

Verifica-se ainda que a evolução da avaliação não afetou significativamente a dispersão de valores entre as diferentes empresas, que inicialmente se situava no intervalo de 0 a 26 pontos e passou a situar-se no intervalo de 36 a 64 pontos, uma amplitude de 28 pontos.

As Avaliações Inicial e Intermédia foram assim úteis, respetivamente, para quantificar a situação encontrada nas empresas e estimar a evolução possível para cada uma delas, até ao final do projeto. Uma vez que a Avaliação Intermédia não é mais que a melhor estimativa para a Avaliação Final, não será aconselhável retirar conclusões com base nesses resultados baseados em previsões.

#### 5.4.3 Reflexões sobre o trabalho realizado até à data

O primeiro ciclo de Investigação – Ação termina com a produção de um Relatório que sintetiza as reflexões das diferentes intervenções e contribui para o conhecimento do tema visado no projeto. Tendo em conta a experiência obtida em cada empresa, o relatório inclui reflexões generalizadas para a compreensão da intervenção como um todo, e que fornecerão considerações relevantes para a continuação do trabalho. O Relatório foi disponibilizado a todas as empresas e resumiu essencialmente o trabalho efetuado até ao momento no projeto, incluindo o resultado das Avaliações,

nomeadamente a baixa Avaliação Inicial e as expectativas espelhadas na Avaliação Intermédia.

Foi evidente que a decisão de participação no projeto foi desde logo resultado de um trabalho prévio de sensibilização da Liderança ou da Estrutura das empresas participantes, pelo que a evolução na dimensão Organização seria natural. No entanto, essa sensibilização deu lugar a ações concretas sugeridas ou tomadas pelos participantes em função das oportunidades que consideraram ser relevantes para o seu negócio. Algumas dessas ações surgiram como consequência natural do projeto, embora não tenham sido totalmente alinhadas com os seus objetivos, como, por exemplo, avançar para a construção de um Web Site corporativo sem considerar questões tecnológicas prévias. Porém, também estas ações constituíram parte da experiência realizada e contribuíram para as conclusões do projeto. Conjuntamente com sugestões dadas pelo autor constituíram a base para o trabalho que a partir daí se realizou, por vezes acarretando dificuldades para a investigação, que acompanhou ações executadas de forma assíncrona. Essas ações criaram frequentemente dificuldades no fecho de atividades previstas em algumas empresas, sem prejuízo de se terem iniciado também atividades adicionais noutras empresas. Apenas no início da fase seguinte foi possível estabilizar o plano de ação com o esforço de sistematização realizado.

## 5.5 Elaboração de Recomendações

Esta secção descreve a fase de sistematização das ações planeadas ou implementadas nas empresas com a criação de um conjunto de recomendações de aplicação generalizada, terminando com a sua disponibilização a todas as empresas.

As Recomendações desenvolvidas no âmbito do Projeto Raio X são apresentadas na Tabela 5.5, estando organizadas segundo as dimensões do Modelo de Implementação. Efetivamente, algumas Recomendações têm impacto em mais do que uma dimensão, mas a sua implementação terá um impacto mais forte na dimensão a que estão primariamente associadas.

Dimensão	Recomendação	Objetivo
Adoção	Responsabilidade do Webmaster	Identificar e definir responsabilidades do Gestor de Informação ou Webmaster na empresa.
	Colaboração Tecnológica	Aderir a esforços colaborativos de âmbito tecnológico.
	Estratégia de <i>e-business</i>	Definir e atualizar uma estratégia de <i>e-business</i> para a empresa.
	Propriedade na Web	Identificar o âmbito da propriedade da empresa em matéria de Web e Propriedade Intelectual.
Proveniência	Importação de Dados	Adotar fontes de dados externas relevantes para o negócio da empresa.
	Integração de Sistemas	Integrar fontes de dados e sistemas de informação com o objetivo de disponibilizar informação de negócio.
Acessibilidade	Acessibilidade de Conteúdos	Disponibilizar conteúdos Web compatíveis com as orientações de acessibilidade.
	Adoção de Ontologias	Adotar Ontologias relevantes para o sector de atuação da empresa.
	Design Adaptativo	Proporcionar uma experiência de utilizador compatível com os diferentes dispositivos e resoluções.
	Disponibilização de dados	Disponibilizar dados estruturados e metadados que possam ser automaticamente interpretados.
	Tipos de Web Site	Definir os tipos de Web Site e a sua relação com a estratégia da empresa.
Atividades	Informação Estruturada da Empresa	Identificar e classificar a empresa e os seus produtos e serviços através de ontologias.

Tabela 5.5 Dimensões e Recomendações do Modelo e-swim

As Recomendações desenvolvidas foram partilhadas com todas as empresas participantes através do Web Site do projeto, com o objetivo de reuplicar em ambientes distintos as potenciais oportunidades detetadas noutras empresas. Pretendeu-se, assim, que cada Recomendação constituísse uma estrutura que fosse simultaneamente pedagógica e orientada para a prática. Nesse sentido, as Recomendações foram estruturadas da seguinte forma:

- Nome: um nome significativo que traduz a ideia presente na Recomendação

- Objetivo: uma ação concreta que traduz a Recomendação da forma mais sucinta possível
- Resumo: a justificação fundamental para a Recomendação
- Desenvolvimento: um texto de apoio para o entendimento completo da Recomendação
- Ações recomendadas: as ações que na prática podem conduzir à implementação do teor da Recomendação
- Referências: indicações para leitura adicional e hiperligações para Web Sites úteis

Foram elaboradas no total 12 Recomendações, apresentadas de forma resumida no texto que se segue. O conteúdo integral das Recomendações é apresentado no Anexo 6.

### Responsabilidade do Webmaster (Dimensão: Adoção)

#### **Objetivo**

Identificar e definir responsabilidades do Gestor de Informação ou Webmaster na organização da empresa.

#### **Resumo**

O Webmaster deve assegurar a coordenação global da presença Web de uma Empresa, garantindo que todas as necessidades estão a ser satisfeitas, seja através das suas competências, seja através das competências dos membros da sua equipa interna e dos fornecedores de serviços. As empresas devem identificar quem desempenha essa função.

#### **Ações recomendadas**

- Identificar o Webmaster na organização
- Alinhar a atividade do Webmaster com a estratégia da empresa

## Colaboração Tecnológica (Dimensão Adoção)

### **Objetivo**

Aderir a esforços colaborativos de âmbito tecnológico.

### **Resumo**

O investimento em tecnologia de uma organização não tem de ser realizado em exclusivo pela mesma. Os vários esforços colaborativos disponíveis a todas as organizações permitem evoluir mais rapidamente, com menor custo e melhores resultados.

### **Ações recomendadas**

- Identificar e formalizar os grupos na organização onde a colaboração tecnológica pode ser desenvolvida
- Identificar e formalizar os interfaces da organização com o exterior onde a colaboração tecnológica pode ser desenvolvida

## Estratégia de *e-business* (Dimensão: Adoção)

### **Objetivo**

Definir e atualizar uma estratégia de *e-business* para a empresa.

### **Resumo**

As Tecnologias de Informação e a Internet são responsáveis por uma transformação radical dos negócios nas últimas décadas. Uma estratégia de *e-business* ou negócio eletrónico identifica a forma como a empresa pretende tirar partido dessa transformação.

### **Ações recomendadas**

- Identificar a estratégia de *e-business* da Empresa e avaliar o grau de conhecimento de toda a organização sobre a mesma
- Identificar as grandes opções da Empresa para o futuro e integrá-las com a evolução da estratégia definida.

## Propriedade na Web (Dimensão: Adoção)

### **Objetivo**

Identificar o âmbito da propriedade da empresa em matéria de Web e Propriedade Intelectual.

### **Resumo**

Os Ativos Intangíveis de uma empresa incluem as suas Marcas, os seus Domínios e também as suas Aplicações de Software, com destaque para aquelas das quais é proprietária integral, mas também o seu conhecimento e a forma como este é distribuído na Web.

### **Ações recomendadas**

- Identificar as diversas componentes da presença Web da organização e classificá-las quanto à sua importância para os resultados pretendidos.

## Importação de Dados (Dimensão: Proveniência)

### **Objetivo**

Adotar fontes de dados externas relevantes para o negócio da empresa.

### **Resumo**

A utilização da informação disponível e relevante para os negócios é uma necessidade das empresas que se pretendem manter competitivas. Essa informação é progressivamente mais vasta e encontra-se muitas vezes fora das fronteiras da empresa, em repositórios públicos ou associativos, pelo que a sua adoção exige uma atitude ativa.

### **Ações recomendadas**

- Identificar e utilizar fontes de dados externas na atividade regular da empresa

## Integração de Sistemas (Dimensão: Proveniência)

### **Objetivo**

Integrar fontes de dados e sistemas de informação com o objetivo de disponibilizar informação de negócio.

### **Resumo**

A integração de dados é uma das prioridades em muitas empresas, com o objetivo de reduzir a duplicação de tarefas e promover a qualidade da informação, disponibilizando informação de uma forma consistente e atualizada a toda a organização. Com a introdução da Internet não só as redes alargadas ganharam protagonismo sobre as redes locais como também possibilitaram novos cenários de aplicação.

### **Ações recomendadas**

- Definir as prioridades de integração de sistemas da empresa em função das necessidades estratégicas do negócio.

## Acessibilidade de Conteúdos (Dimensão: Acessibilidade)

### **Objetivo**

Disponibilizar conteúdos Web compatíveis com as orientações de acessibilidade.

### **Resumo**

A Web é uma rede de utilização livre suportada por tecnologias abertas. No entanto, as recomendações de utilização e em particular as de acessibilidade dos conteúdos e dos dados existem e devem ser respeitadas e adotadas.

### **Ações recomendadas**

- Avaliar a acessibilidade do(s) Web Site(s) da empresa
- Gerir metadados

## Adoção de Ontologias (Dimensão: Acessibilidade)

### **Objetivo**

Adotar Ontologias relevantes para o sector de atuação da empresa.

### **Resumo**

As particularidades de cada sector de atividade dão origem a conhecimento específico que pode ser relevante para o nível de competitividade das empresas. As ontologias podem desempenhar um papel fundamental na forma de sistematizar esse conhecimento.

### **Ações recomendadas**

- Identificar as áreas temáticas em que a organização se enquadra e quais as oportunidades de definição de significados relevantes para o seu negócio nuclear e áreas de suporte
- Identificar ontologias setoriais relacionadas com a atividade da organização

## Design Adaptativo (Dimensão: Acessibilidade)

### **Objetivo**

Proporcionar uma experiência de utilizador compatível com os diferentes dispositivos e resoluções.

### **Resumo**

A evolução das tecnologias de écrans planos e de comunicações móveis criaram um cenário de utilização frequente da Web com recurso a múltiplos dispositivos, com vantagens e condicionantes. O design adaptativo tem um papel fundamental na disponibilização de dados e informação para proporcionar uma experiência de utilizador adequada.

### **Ações recomendadas**

- Avaliar o design adaptativo do(s) Web Site(s) da empresa
- Avaliar a estrutura de dados e conteúdos disponibilizada pela empresa

## Disponibilização de Dados (Dimensão: Acessibilidade)

### **Objetivo**

Disponibilizar dados estruturados e metadados que possam ser automaticamente interpretados.

### **Resumo**

A forma como os dados são disponibilizados tem um impacto direto no potencial da sua reutilização, nomeadamente tornando possível uma interpretação sistematizada que conduza à sua reutilização sucessiva.

### **Ações recomendadas**

- Avaliar a orientação dos metadados do(s) Web Site(s) da empresa em termos da sua posterior utilização pelos motores de pesquisa
- Avaliar em que medida a disponibilização (controlada) ou publicação (livre) de dados estruturados contribui para os objetivos definidos para o(s) Web Site(s) da empresa
- Implementar a disponibilização de metadados estruturais no(s) Web Site(s) da empresa

## Tipos de Web Site (Dimensão: Acessibilidade)

### **Objetivo**

Definir os tipos de Web Site e a sua relação com a estratégia da empresa.

### **Resumo**

Os Web Sites podem ser classificados quanto à proveniência dos seus utilizadores e ao tipo de aplicações que os suportam. A escolha sobre os diferentes tipos de Web Site está diretamente ligada à estratégia da empresa. Pretende-se que as escolhas efetuadas ao longo do tempo estejam sempre associadas à estratégia da empresa em cada momento, fazendo com que os Web Sites sejam de facto uma peça que suporte a prossecução dos objetivos empresariais.

### **Ações recomendadas**

- Avaliar os objetivos do(s) Web Site(s) da empresa
- Avaliar o alinhamento estratégico dos Web Sites da empresa

### Informação Estruturada da Empresa (Dimensão: Atividades)

#### **Objetivo**

Identificar e classificar a empresa e os seus produtos e serviços através de ontologias.

#### **Resumo**

A utilização de ontologias para a classificação da informação de uma empresa e dos seus produtos facilita a sua interpretação por sistemas externos, com particular destaque para os indexadores da Web que alimentam os motores de pesquisa.

### **Ações recomendadas**

- Avaliar o schema.org do(s) Web Site(s) da empresa
- Avaliar a utilização de uma ontologia compatível com schema.org nos Sistemas de Informação da empresa

#### 5.5.1 Outras ações não sistematizadas

Algumas ações planeadas ou implementadas nas empresas foram consideradas de menor relevância para os objetivos do projeto, não obstante tenham relevância tecnológica ou de negócio, pelo que são aqui referidas apenas para memória e trabalho futuro.

#### **Alojamento**

Uma aplicação Web, como por exemplo um Web Site, necessita de ser instalada num servidor com ligação permanente à Internet para que possa ser acedida permanentemente. É comum designar o espaço ocupado por Alojamento, embora este possa existir sob diversas formas:

- Servidor próprio

- a empresa possui um Servidor Web que utiliza para alojar as suas aplicações
- comum no caso de Intranets
- com custos diretos e indiretos dificilmente justificáveis quando se trata de Web Sites públicos, exceto em caso de aplicações críticas ou de grande dimensão
- Servidor colocado
  - também designado por *Housing*
  - o servidor continua a ser propriedade da empresa mas é colocado num *Datacenter* com condições técnicas superiores
  - com custos diretos e indiretos elevados, uma opção pouco utilizada exceto em caso de aplicações críticas ou de grande dimensão
- Alojamento partilhado
  - também designado por *Hosting*
  - um prestador de serviços disponibiliza um espaço variável num dos seus servidores para alojamento da aplicação, tal como o faz a outros clientes
  - é a solução mais barata e mais frequente
  - com custos baixos, é a solução base aconselhável
- Alojamento virtual
  - também designado por *Cloud Hosting*
  - um prestador de serviços disponibiliza um espaço variável no seu cluster de servidores para alojamento da aplicação, tal como o faz a outros clientes
  - é a solução em franco crescimento na atualidade, um pouco mais dispendiosa que o Alojamento partilhado, de todas a mais flexível
- Alojamento dedicado
  - também designado por Servidor dedicado
  - um prestador de serviços disponibiliza um dos seus servidores para alojamento da aplicação, dedicado completamente ao cliente

- é uma solução acessível, mais cara que o Alojamento partilhado, mas com a vantagem de ser usada em exclusivo por uma empresa

A escolha da forma adequada de Alojamento de uma Aplicação Web deve ser feita tendo em conta o impacto económico e o impacto no desempenho da aplicação. Uma vez que não existe uma solução única ou ideal para esse efeito, as opções tomadas podem condicionar a evolução futura dos projetos. A título de exemplo, a disponibilização de um *Web Service* para integração de dados de produtos e fornecedores, poderá ser muito mais difícil ou mesmo impossível de implementar em algumas situações e profundamente dependente das condições comerciais de cada prestador de serviços.

### ***Naked Domain***

A utilização de Domínios surge principalmente para facilitar a memorização de endereços. Utilizando o *Domain Name System* (DNS), é possível associar um nome de domínio a um endereço IP e utilizá-lo como alternativa ao endereço numérico (por exemplo, aceder a planetavirtual.pt em alternativa a 151.236.50.21).

Para além da utilização do domínio, as melhores práticas consideram que se utilize um subdomínio para identificar as diferentes aplicações que podem ser acedidas nesse domínio. O caso mais comum é o do Web Site que é habitualmente identificado pelo prefixo www (por exemplo, www.planetavirtual.pt) e o do servidor de correio eletrónico que é habitualmente identificado pelo prefixo mail (por exemplo, mail.planetavirtual.pt). Outros exemplos poderiam incluir uma Extranet para clientes disponível no endereço clientes.planetavirtual.pt ou um *Web Service* para disponibilização de informação de produtos disponível em products.planetavirtual.pt.

Entretanto, a raiz do domínio, o endereço acessível sem qualquer prefixo (por exemplo planetavirtual.pt), é muitas vezes negligenciado e não tem qualquer utilidade. Até há pouco tempo essa prática era comum, mas ocorreram duas alterações significativas:

- a procura de endereços representativos de pequena dimensão levou a uma maior utilização de endereços Web sem qualquer prefixo (por exemplo goo.gl)

- alguns *browsers* e outras aplicações começaram a tentar interpretar os endereços Web e a fornecer hiperligações de forma automática, pelo que em muitos casos essa hiperligação equivale ao próprio domínio “raiz”

Assim, passou a ser necessário fornecer uma indicação concreta do que se pretende que aconteça quando um pedido é feito ao endereço Web sem qualquer prefixo, podendo consistir em uma das seguintes opções:

- o visitante deve ser redirecionado para o Web Site público, normalmente disponível no endereço com o prefixo *www*
- deve surgir um Web Site que é o principal Web Site público disponível nesse endereço, caso em que o acesso com o prefixo *www* deveria redirecionar o visitante para o endereço “raiz”
- deve surgir o mesmo Web Site, provavelmente utilizando uma entrada dupla no DNS em que ambos os endereços estão associados ao mesmo Web Site, prática pouco aconselhável em termos de *Search Engine Optimization*

Quando nenhuma destas opções é utilizada, o pedido de acesso a um endereço Web resulta num erro, que muitas vezes é confundido pelo utilizador com a situação do Web Site estar indisponível. Esses casos são habitualmente designados por *Naked Domain*.

### **Formação**

A utilização da Web pode constituir uma vantagem competitiva para uma Empresa em relação aos seus concorrentes diretos. Se a presença da Empresa fosse aquela que é obrigatória por lei para todas as empresas, então estar presente na Web seria apenas uma banalidade, obrigatória para todos, sem vantagem para nenhum.

Porque todas as empresas devem encarar a Internet como uma oportunidade de negócio? A razão é muito simples, explicada por Ann Mettler (Mettler, 2012), diretora executiva do Lisbon Council:

*“Um estudo recente da Comissão Europeia concluiu que 85% dos novos empregos na Europa são criados por pequenas e médias empresas (estas representam cerca de 99%*

*em número de empresas). Mais ainda, as empresas que fazem uso intensivo de tecnologias digitais crescem duas vezes mais, exportam duas vezes mais e criam duas vezes mais postos de trabalho que as empresas que são lentas na sua adoção. Para ser clara: não se trata apenas – nem sequer principalmente – das novas empresas de alta tecnologia, que tantas vezes são associadas à agenda europeia. Trata-se da utilização de tecnologia em todos os negócios e através de todos os sectores. Não é sem razão que, de acordo com o McKinsey Global Institute, 75% do valor económico criado pela Internet resulta de empresas tradicionais que utilizam tecnologia baseada na Web para fomentar a produtividade, reduzir os custos e alcançar novos clientes e mercados. (...) Desemprego jovem em número recorde? E que tal usar as capacidades da geração digital para enfrentar o défice previsto de 700,000 profissionais de Tecnologias de Informação em 2015?”*

A *World Wide Web* constitui um ambiente com tecnologias e aplicações diversas que evoluiu consideravelmente em pouco mais de duas décadas de existência. Os colaboradores das empresas questionam frequentemente qual a forma de utilizar a Web para que seja uma vantagem competitiva para o seu negócio. Por um lado, as competências das empresas, entenda-se, da sua direção e dos seus colaboradores, podem ser uma forte restrição à evolução dos seus negócios se não forem adequadas. Por outro lado, o perfil de competências de uma empresa pode ser melhorado através da contratação de novos colaboradores com competências adequadas, da promoção da colaboração interna bem como da formação aplicada.

A Formação nas empresas em temas como *e-business*, Marketing Digital, Web Semântica, Integração de Informação, entre outros, constitui uma necessidade que é necessário satisfazer no curto prazo. A procura e seleção de conteúdos adequados, por exemplo quanto à qualidade e língua, e formadores experientes, devem ser uma prioridade para todas as organizações.

## 5.6 Avaliação final

Esta secção descreve a fase que se inicia com a análise da adoção das Recomendações elaboradas, tendo implicado, por parte das empresas, uma

implementação das ações incluídas nessas recomendações, terminando com a Avaliação Final da intervenção nas empresas e uma reflexão sobre os resultados obtidos.

### 5.6.1 Adoção e partilha de Recomendações

O objetivo fundamental do segundo ciclo de Investigação-Ação era o de contribuir para a disseminação de práticas entre as diferentes empresas através da partilha de todas as Recomendações desenvolvidas e da análise da sua adequabilidade a cada caso particular. Se no primeiro ciclo a intervenção individual nas empresas permitiu analisar cada cenário de trabalho em detalhe e definir ações a aplicar, a parte da intervenção consubstanciada no segundo ciclo permitiu reaplicar práticas que forneceram resultados positivos.

A aplicação futura do Modelo de Implementação terá como apoio possível o conjunto de Recomendações entretanto elaboradas, pelo que a análise da sua reaplicação nas empresas participantes é fundamental para entender a sua viabilidade.

Foram elaboradas 12 Recomendações no total, apresentadas na Secção 5.5, que foram partilhadas pelas empresas participantes no projeto. Para cada empresa foi analisada a viabilidade de adoção de todas as Recomendações, tendo como pressupostos um prazo de implementação realista de alguns meses e um custo de implementação adequado à dimensão e recursos financeiros das empresas. Esta análise foi inicialmente superficial e foi ganhando corpo com os diferentes contatos com as empresas, quer presenciais, quer à distância. Finalmente, o conjunto de Recomendações a adotar no curto prazo, isto é, com possibilidade de implementação no decorrer do Projeto, foi sendo sucessivamente definido pelas diferentes empresas. Por exclusão, as Recomendações não adotadas foram consideradas inoportunas ou inviáveis para implementação no decorrer do projeto.

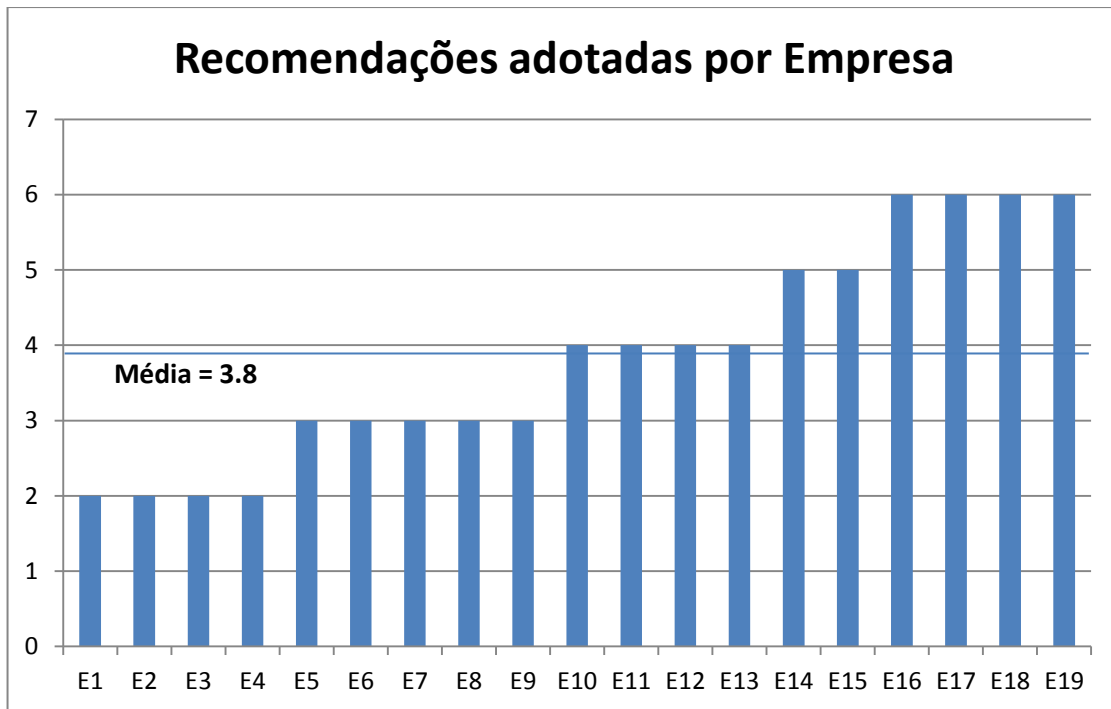


Figura 5.9. Recomendações adotadas por Empresa

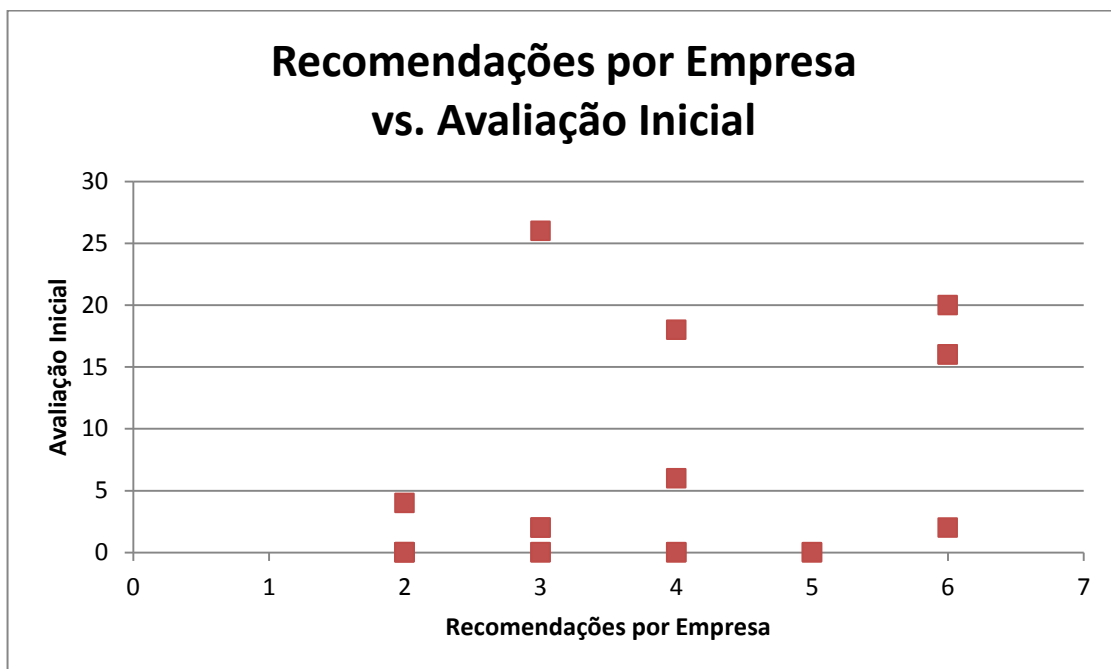


Figura 5.10. Recomendações por Empresa vs. Avaliação Inicial

Cada empresa adotou um mínimo de duas e um máximo de 6 Recomendações, com uma média de 3.8 adoções por empresa. A Figura 5.9 apresenta o número de Recomendações adotadas em cada empresa. Tendo em conta um prazo de

implementação relativamente flexível e os diferentes cenários experimentados em cada empresa, não é de estranhar alguma dispersão dos valores.

A Figura 5.10 ilustra a relação entre o número de Recomendações adotado em cada empresa e a Avaliação Inicial obtida pela mesma empresa. Da análise do gráfico, conclui-se que não existe uma relação evidente entre as duas variáveis.

A Figura 5.11 apresenta um gráfico relativo à frequência de adoção das diferentes Recomendações pelas empresas participantes. O interesse em cada uma das Recomendações por parte das empresas foi, como é visível, muito variável e dependente de vários fatores. A análise que se segue foca-se nos resultados obtidos com a adoção das Recomendações e na identificação desses fatores.

### **Responsabilidade do Webmaster adotada a 100%**

A Recomendação “Responsabilidade do Webmaster” foi adotada em 15 das 19 empresas participantes, 79% do total. As restantes 4 empresas não implementaram a recomendação apenas porque já a satisfaziam previamente, pelo que, implicitamente, a adoção desta recomendação foi total.

O principal fator que contribuiu para esta adoção generalizada foi o simples facto de a ação recomendada se traduzir apenas numa decisão, a identificação de um indivíduo que desempenhe o papel de Webmaster, sem mais implicações imediatas. Um dos fatores mais frequentes em todas as decisões de adoção foi o custo, que neste caso, genericamente, não está presente. Em relação a esta ação recomendada, as empresas participantes dividiram-se em dois grupos:

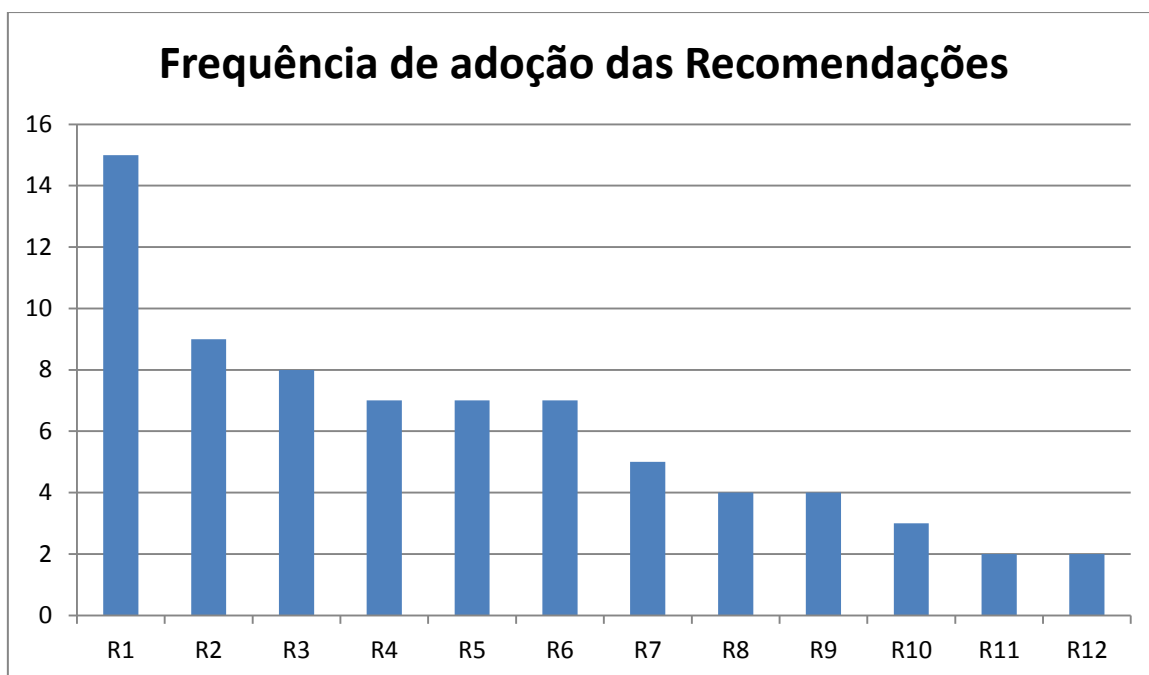


Figura 5.11. Frequência de adoção das Recomendações

- Webmaster como colaborador interno: formalmente, as empresas esperam a disponibilidade total do seu colaborador, pelo que confiam nele para satisfazer as necessidades identificadas, mesmo que na maior parte dos casos, tratando-se como é comum de um Técnico de Suporte Informático, a sua atuação cubra apenas uma pequena parte das funções esperadas, pelo que o custo associado ao colaborador foi previamente assumido
- Webmaster como colaborador externo: implicitamente, as empresas esperam que o seu prestador de serviços atue diligentemente na prossecução dos seus interesses, apresentando propostas de implementação que satisfaçam as necessidades mediante um custo associado

Em qualquer dos casos, as empresas presumem que o custo de implementação desta recomendação seja nulo. Comparativamente, a expectativa é equivalente a considerar que a Medicina seja exercida apenas pelos enfermeiros que aplicam os tratamentos e pelos farmacêuticos que fornecem os meios, desconsiderando na totalidade o trabalho dos médicos que analisam, exploram e prescrevem os tratamentos quando estes são necessários.

O antagonismo desta expectativa e a decisão de implementação generalizada de ações por uma grande parte das empresas levaram a que este tema fosse objeto de um estudo específico. Assim, em abril de 2014 o autor realizou um trabalho intitulado “Webmaster: o papel e o valor de um especialista na Presença Web de uma Empresa” cujo conteúdo integral é apresentado no Anexo 7.

Com esse trabalho pode concluir-se que as expectativas sobre as funções de um Webmaster são muito diversas e levarão naturalmente a que a função seja desempenhada por um conjunto de indivíduos com diversas especialidades, eventualmente liderados por um indivíduo com capacidades mais generalistas nas diversas áreas de conhecimento essenciais à Web da atualidade. Essa conclusão foi incorporada na Recomendação elaborada no âmbito deste projeto, que pode ser consultada em versão integral no Anexo 6.

O trabalho concluiu ainda que o valor percebido para a função Webmaster numa pequena empresa se situa em média nos €267 mensais. A maioria dos inquiridos valorizou mais um contrato com custos mensais constantes que vários serviços com custos parciais, aceitando implicitamente o valor da atuação do Webmaster mas desconsiderando a contribuição explícita que possa surgir de uma atuação pontual específica. Utilizando a mesma comparação com a Medicina, os inquiridos demonstram aparentemente a vontade de suportar o custo de um “plano médico” que inclua todos os custos regulares e recorrentes, incluindo a atuação do “clínico geral”, enquanto demonstram maior resistência em consultar o seu médico de forma regular e preventiva.

Sendo evidente que os valores encontrados foram claramente inferiores à possibilidade de contratar um profissional a tempo inteiro, concluiu-se que poderá ser necessário a uma pequena empresa recorrer exclusivamente a serviços externos que assegurem esses resultados. As empresas de maior dimensão, onde é viável a contratação de colaboradores a tempo inteiro para essas funções, poderão ainda ter que recorrer a serviços externos no sentido de completar as competências existentes no interior das suas organizações.

### **Metade das Recomendações foram adotadas por mais de um terço dos participantes**

Para além da Recomendação “Responsabilidade do Webmaster”, cinco Recomendações foram adotadas por mais de um terço das empresas. A Recomendação “Disponibilização de Dados” foi adotada por 9 das empresas participantes, 47% do total. A Recomendação “Acessibilidade de Conteúdos” foi adotada por 8 empresas (42% do total) e as Recomendações “Estratégia de *e-business*”, “Informação estruturada da Empresa” e “Tipos de Web Site” foram adotadas por 7 empresas (37% do total).

A este conjunto de recomendações com maior adoção por parte das empresas estão associados os seguintes fatores:

- dependência apenas da decisão da empresa
- orientação para a implementação imediata
- adoção imediata condicionada ao custo de implementação

### **Todas as Recomendações foram validadas por todos os participantes**

Uma das necessidades identificadas neste projeto foi a de validar que mesmo as Recomendações rejeitadas fazem algum sentido para o negócio das empresas, ainda que possam apenas ser implementadas no médio ou longo prazo. Tratando-se de um exercício especulativo e tendente a criar uma resposta positiva com menor convicção e, por isso, se traduzir em dados eventualmente questionáveis, optou-se por criar um cenário hipotético que facilitasse uma resposta com maior fundamentação.

Para a construção desse cenário contribuiu decisivamente a obrigatoriedade de utilização de programas de faturação certificados, implementada em Portugal durante o decurso do projeto. Esta obrigação fiscal a que todas as empresas, com poucas exceções, foram sujeitas, introduziu novidades tecnológicas na generalidade das empresas mas, principalmente, tornou claro que as Tecnologias de Informação e, em particular, a Internet, se cruzaram definitivamente no caminho dos negócios.

O cenário hipotético que foi elaborado tomou, assim, como ponto de partida o pressuposto real que as empresas comunicam à Autoridade Tributária e Aduaneira os dados da sua faturação e aproveitou alguns aspetos tecnológicos dessa comunicação,

nomeadamente alguns requisitos do SAF-T PT, e algumas considerações sobre uma eventual otimização de serviços já existentes na Administração Pública. Desta forma, as perguntas do tipo “*Ainda assim, no futuro, implementaria esta Recomendação na sua empresa?*” foram substituídas por uma única pergunta: “*No futuro, se este cenário fosse obrigatório, a sua empresa estaria preparada para o satisfazer?*”

O cenário hipotético é descrito no texto que se segue.

---

### *Comunicação do Ministério das Finanças*

#### *e-CAE (Código de Atividade Económica eletrónico)*

*As empresas passam a ser obrigadas a definir em concreto qual ou quais as atividades que desempenham através da caracterização dos seus produtos e serviços no sistema e-CAE. Assim, todas as empresas devem indicar um responsável pelo contacto com a Administração Pública, eventualmente substituindo os detentores de cargos sociais, para assegurar o fluxo constante de informação de produtos e serviços, de acordo com um conjunto de pressupostos.*

*Em primeiro lugar, todas as empresas necessitam de criar códigos de produto ou serviço que identifiquem qualquer produto ou serviço que seja faturado pela mesma. O grau de detalhe desses códigos é deixado ao critério da empresa, sendo que não é permitida a utilização do mesmo código para produtos ou serviços de natureza distinta. A título de exemplo refira-se que os códigos para a venda de um equipamento e a prestação de assistência técnica ao mesmo equipamento são obrigatoriamente distintos.*

*Em segundo lugar, esses códigos de produto ou serviço são comunicados eletronicamente ao serviço e-CAE para validação prévia dos Códigos de Atividade Económica em que a mesma atua. A criação de códigos associados a Códigos de Atividade Económica em que a empresa inicia a sua atuação fica sujeita a autorização prévia do serviço e-CAE. Esta autorização é fornecida eletronicamente e é vinculativa para efeitos fiscais e de atualização do registo de sociedades comerciais através do Sistema de Informação da Classificação Portuguesa das Atividades Económicas – SICAE. Esta*

*informação encontra-se disponível publicamente através da introdução do número fiscal de contribuinte, atualmente no serviço SICAE e posteriormente no serviço e-CAE ou em qualquer aplicação informática que possibilite a utilização dos serviços Web disponibilizados pelo e-CAE.*

*Em terceiro lugar, a faturação da empresa fica condicionada a serem assinalados os códigos de produto ou serviço respetivos para cada linha de faturação. Os códigos introduzidos são validados posteriormente no momento de envio dos ficheiros SAF-T PT respeitantes à faturação.*

*Finalmente, de forma a incentivar que as empresas promovam ou transacionem qualquer produto ou serviço de forma eletrónica, em qualquer plataforma, deverão disponibilizar na Internet os códigos de produto ou serviço correspondentes, seja através do serviço e-CAE seja através do seu próprio Web Site.*

*Mais se informa que a tecnologia Web adotada na implementação do e-CAE segue os standards internacionais do W3C aplicados a dados, nomeadamente utilizando RDF como linguagem base, e substituirá gradualmente as diferentes formas não interoperáveis de comunicação entre sistemas da Administração Pública e Empresas. Para reduzir o custo de implementação do e-CAE a nível nacional, as rotinas de integração e utilização de dados dos serviços Web disponibilizados foram implementadas nas linguagens mais comuns e são disponibilizadas em regime de código aberto para utilização conjunta entre a Administração Pública, as Empresas em geral e as empresas de software em particular. O e-CAE disponibiliza um ponto único de recolha de informação eletrónica a todas as empresas, que pode ser utilizado em exclusivo ou de forma integrada com o Web Site da empresa, nos casos em que exista, e que permitirá validar informação básica de negócios, contribuindo assim para o reforço da formalização da economia. O e-CAE disponibilizará ainda indicadores de referência extraídos dos dados acumulados de cada CAE às empresas que atuam com o mesmo CAE, proporcionando o acesso a informação relevante e atualizada para os negócios.*

---

A análise seguinte justifica a introdução de todas as Recomendações elaboradas, indicadas de forma intercalada com o texto apresentado previamente.

---

*Comunicação do Ministério das Finanças*

*e-CAE (Código de Atividade Económica eletrónico)*

*As empresas passam a ser obrigadas a definir em concreto qual ou quais as atividades que desempenham através da caracterização dos seus produtos e serviços no sistema e-CAE. Assim, todas as empresas devem indicar um responsável (“Responsabilidade do Webmaster”) pelo contacto com a Administração Pública, eventualmente substituindo os detentores de cargos sociais, para assegurar o fluxo constante de informação de produtos e serviços, de acordo com um conjunto de pressupostos.*

*Em primeiro lugar, todas as empresas necessitam de criar códigos de produto ou serviço (“Informação estruturada da empresa”) que identifiquem qualquer produto ou serviço que seja faturado pela mesma. O grau de detalhe desses códigos é deixado ao critério da empresa, sendo que não é permitida a utilização do mesmo código para produtos ou serviços de natureza distinta (“Adoção de Ontologias”). A título de exemplo refira-se que os códigos para a venda de um equipamento e a prestação de assistência técnica ao mesmo equipamento são obrigatoriamente distintos.*

*Em segundo lugar, esses códigos de produto ou serviço são comunicados eletronicamente ao serviço e-CAE (“Disponibilização de Dados”) para validação prévia dos Códigos de Atividade Económica em que a mesma atua. A criação de códigos associados a Códigos de Atividade Económica em que a empresa inicia a sua atuação fica sujeita a autorização prévia do serviço e-CAE. Esta autorização é fornecida eletronicamente (“Integração de Sistemas”) e é vinculativa para efeitos fiscais e de atualização do registo de sociedades comerciais através do Sistema de Informação da Classificação Portuguesa das Atividades Económicas – SICAE. Esta informação encontra-se disponível publicamente através da introdução do número fiscal de contribuinte, atualmente no serviço SICAE e*

*posteriormente no serviço e-CAE ou em qualquer aplicação informática que possibilite a utilização dos serviços Web disponibilizados pelo e-CAE.*

*Em terceiro lugar, a faturação da empresa fica condicionada a serem assinalados os códigos de produto ou serviço respetivos para cada linha de faturação. Os códigos introduzidos são validados posteriormente no momento de envio dos ficheiros SAF-T PT respeitantes à faturação.*

*Finalmente, de forma a incentivar que as empresas promovam ou transacionem qualquer produto ou serviço de forma eletrónica (“Estratégia de e-business”), em qualquer plataforma (“Design Adaptativo”), deverão disponibilizar na Internet os códigos de produto ou serviço correspondentes (“Acessibilidade de Conteúdos”), seja através do serviço e-CAE seja através do seu próprio Web Site.*

*Mais se informa que a tecnologia Web adotada na implementação do e-CAE segue os standards internacionais do W3C aplicados a dados, nomeadamente utilizando RDF como linguagem base, e substituirá gradualmente as diferentes formas não interoperáveis de comunicação entre sistemas da Administração Pública e Empresas. Para reduzir o custo de implementação do e-CAE a nível nacional, as rotinas de integração e utilização de dados dos serviços Web disponibilizados foram implementadas nas linguagens mais comuns e são disponibilizadas em regime de código aberto para utilização conjunta entre a Administração Pública, as Empresas em geral e as empresas de software em particular (“Colaboração Tecnológica”). O e-CAE disponibiliza um ponto único de recolha de informação eletrónica a todas as empresas, que pode ser utilizado em exclusivo ou de forma integrada com o Web Site da empresa, nos casos em que exista (“Tipos de Web Site”), e que permitirá validar informação básica de negócios, contribuindo assim para o reforço da formalização da economia. O e-CAE disponibilizará ainda indicadores de referência extraídos dos dados acumulados de cada CAE às empresas que atuam com o mesmo CAE, proporcionando o acesso a informação relevante e atualizada (“Importação de Dados”) para os negócios.*

---

Com base neste cenário hipotético criado propositadamente para validar a adequabilidade e verosimilidade de todas as Recomendações, obteve-se a aprovação do mesmo por parte de 100% das empresas participantes. Implicitamente, as empresas validaram o conteúdo da totalidade das Recomendações.

A título de curiosidade, a implementação deste cenário, perfeitamente possível por ser pouco diferente das funcionalidades atuais implementadas isoladamente, levaria todas as empresas portuguesas a obter uma classificação e-swim de 58 pontos, um valor um pouco acima da média de 46 pontos obtida na Avaliação Intermédia e apenas inferior à Avaliação Intermédia de três das 19 empresas participantes.

### 5.6.2 Análise da Intervenção

Um trabalho baseado na metodologia Investigação-Ação deverá continuar apenas enquanto não existirem soluções para o problema ou até que os resultados provenientes de ciclos adicionais sejam marginalmente irrelevantes. Considera-se nesta fase que as empresas participantes assimilaram o processo de trabalho, compreenderam o Modelo de Implementação e adotaram as Recomendações elaboradas. Por isso, um terceiro ciclo de Investigação-Ação não acrescentaria mais factos que permitam uma evolução da proposta de trabalho. Pelo contrário, a opinião partilhada pelas empresas indica a preferência pelo prolongamento do prazo para a implementação das ações e não para a procura de mais ações ou ações distintas. Por essa razão, concluiu-se que os dois ciclos de Investigação-Ação previstos foram suficientes para a realização do trabalho proposto.

Ainda a propósito do prazo para a implementação das ações, um dos aspetos mais críticos na execução do trabalho, poder-se-á especular que mais tempo disponível poderia influenciar os resultados obtidos, de certa forma satisfazendo essa vontade demonstrada pelas empresas. No entanto, o objetivo deste trabalho era determinar o grau de adoção da Web Semântica nas Empresas e promover e facilitar a sua implementação através da proposta de um modelo. Ora nesta fase, tendo sido proposto o modelo e medido o grau de adoção inicial e intermédio (a melhor previsão possível para o final), considera-se esse objetivo atingido. Por isso, a extensão no tempo do segundo ciclo de Investigação-Ação apenas faria com que o grau de adoção da tecnologia fosse um pouco maior, como natural e previsivelmente acontecerá após o final do projeto, caso as

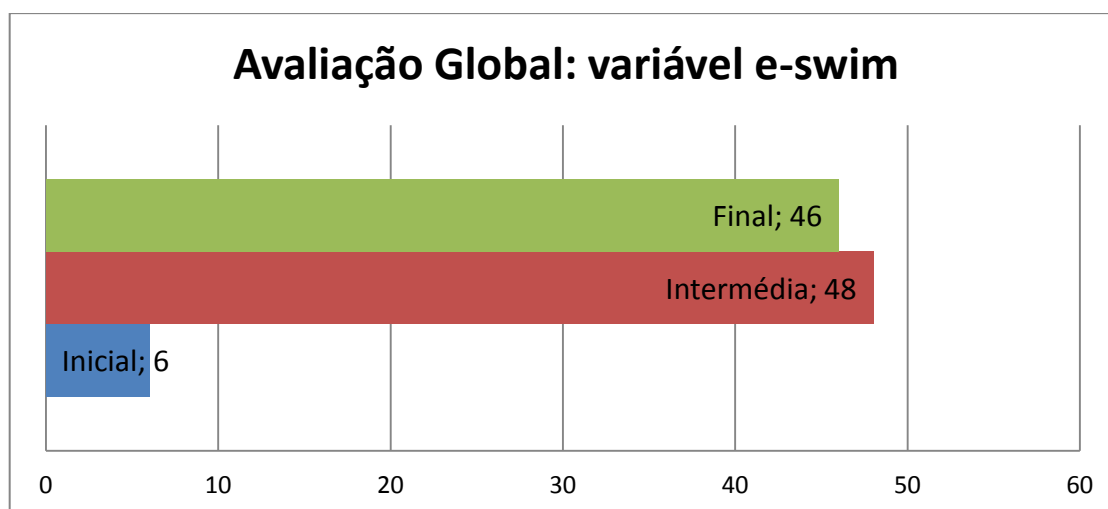


Figura 5.12. Avaliação Global: variável e-swim

empresas continuem a utilizar o modelo proposto. Assim, a decisão de terminar o projeto surgiu naturalmente e a Avaliação Final foi efetuada.

### 5.6.3 Avaliação Final

A Avaliação Final traduz a situação nas empresas relativa ao momento final do projeto, nomeadamente fevereiro de 2015. Uma vez que tinham sido efetuadas uma Avaliação Inicial, um retrato da empresa no momento imediatamente anterior ao início da intervenção, e uma Avaliação Intermédia, referente à melhor previsão para o momento final do projeto, ambas realizadas entre maio e setembro de 2014 nas diferentes empresas, apresentam-se de seguida os resultados das três avaliações.

A Avaliação final da variável e-swim situou-se em média nos 46 pontos percentuais para as empresas participantes, 2 pontos abaixo da Avaliação Intermédia que se tinha situado nos 48 pontos (cf. Figura 5.12), podendo concluir-se que a previsão efetuada aquando da Avaliação Intermédia teve uma precisão apreciável.

A análise da Figura 5.13 evidencia que as dimensões “Adoção”, “Proveniência” e “Acessibilidade” registaram uma Avaliação Final igual à Avaliação intermédia, tendo apenas a dimensão “Atividades” registado uma variação de 50 para 42 pontos percentuais.

Em relação à evolução em cada empresa, pode verificar-se pela análise da Figura 5.14, que a maioria das empresas (68%) registou uma Avaliação Final igual à Avaliação Intermédia, com 6 exceções que serão analisadas de seguida.

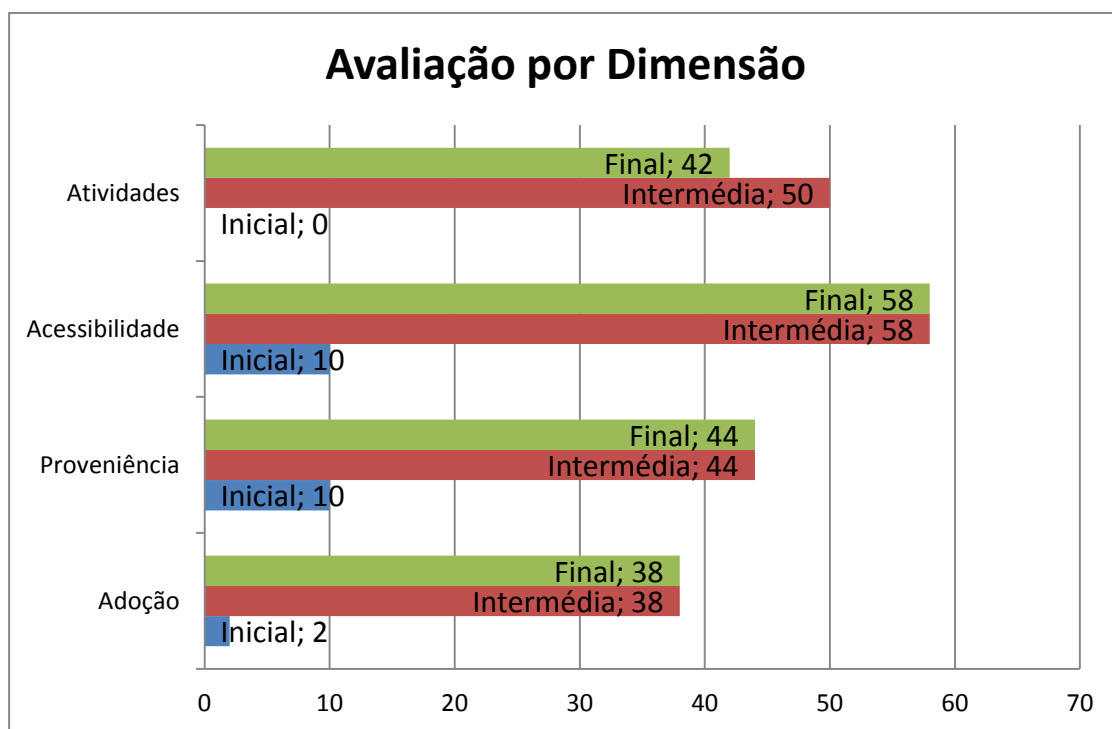


Figura 5.13. Avaliação por Dimensão

As principais diferenças entre a Avaliação Intermédia e a Avaliação Final tiveram duas origens distintas, cuja explicação é detalhada no texto que se segue.

A dimensão Atividades teve uma Avaliação Final inferior à estimada inicialmente em 5 empresas como resultado direto da avaliação quantitativa definida. Conforme descrito na Secção 5.4.1, a avaliação da dimensão Atividades resultou da avaliação de um único critério. Nesse critério a classificação de 3 pontos correspondia ao cumprimento da condição “foram identificadas pelo menos 4 atividades realizadas de forma pouco sistemática”. O facto de 5 empresas terem ficado classificadas apenas com 2 pontos, correspondentes à condição “foram identificadas apenas atividades diversas ou realizadas por indivíduos isolados”, ditou a redução da avaliação em relação ao previsto. Trata-se de uma diferença de pequena dimensão, quando comparada com a dimensão da amostra (26% das empresas) ou com o impacto na avaliação (8% na avaliação da dimensão, 2% na avaliação global). Poder-se ia argumentar com base em dados empíricos que esta diferença seria atenuada se a duração do projeto fosse um pouco mais extensa, eventualmente de mais dois meses. No entanto, considera-se que tendo em conta o impacto reduzido nos resultados finais, essas considerações não acrescentariam um valor relevante para o trabalho realizado.

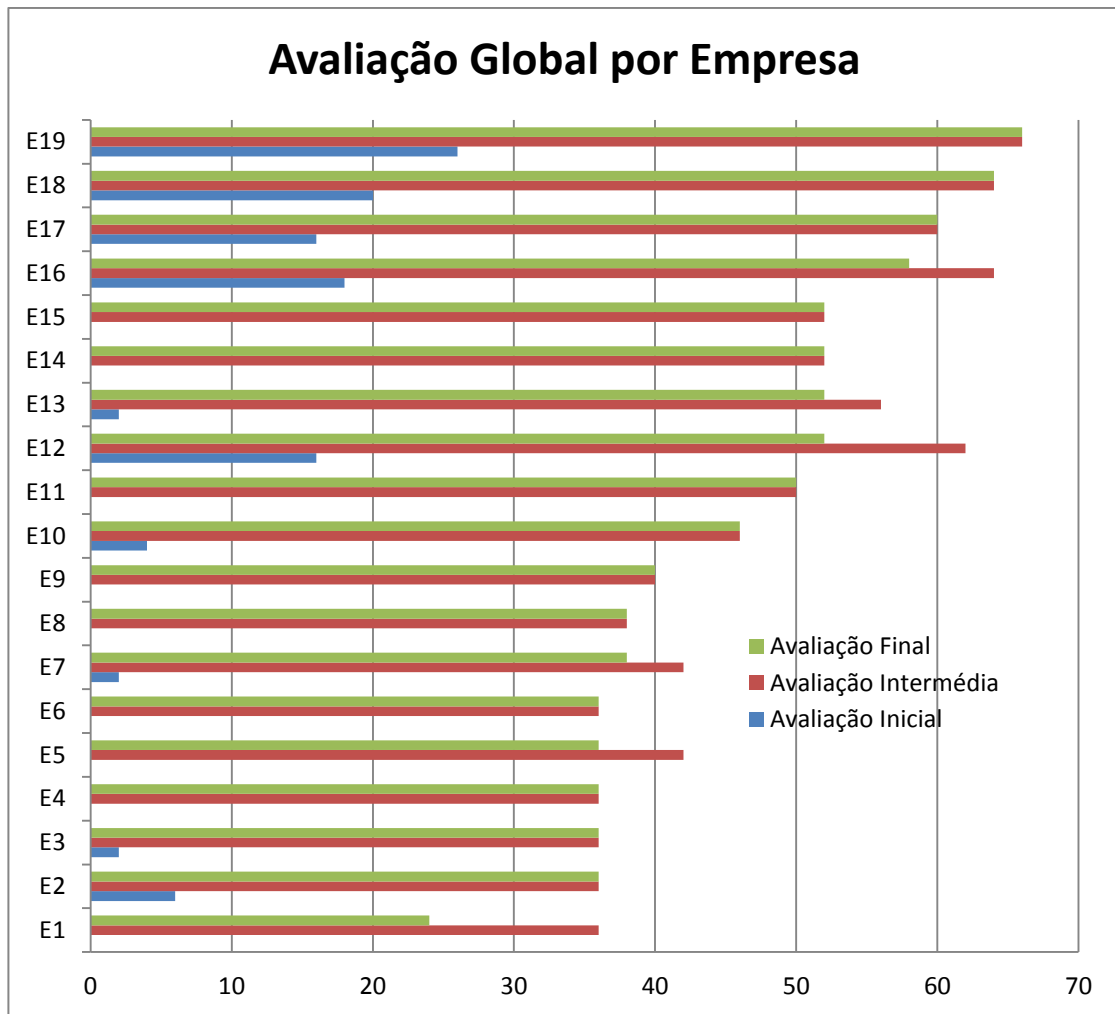


Figura 5.14. Avaliação Global por Empresa

Em relação à segunda origem da diferença de resultados entre a Avaliação Intermédia e a Avaliação Final e para melhor explicação da mesma, devem antes de mais analisar-se as Recomendações implementadas pelas empresas.

O gráfico da Figura 5.15 apresenta o número de Recomendações implementadas em cada empresa, num total de 68 Recomendações implementadas. Foram inicialmente propostas 73 Recomendações no total (cf. Figura 5.9), pelo que não foram implementadas 5 Recomendações.

Uma Empresa não implementou as duas Recomendações propostas por razões de estratégia para o negócio. Em particular, a empresa considerou que outras prioridades no curto prazo eram mais importantes do que a sua participação no projeto, pelo que decidiu não proceder a qualquer ação interna para além do diagnóstico inicial efetuado. Esta decisão constituiu na prática uma desistência temporária, sem que seja possível

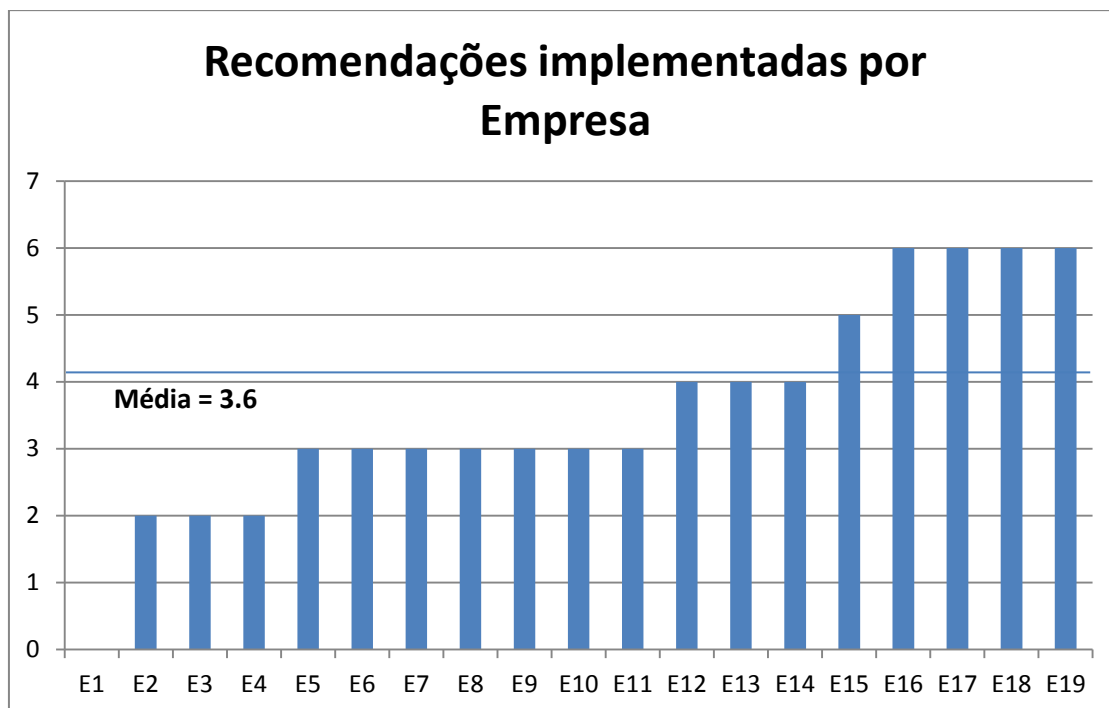


Figura 5.15. Recomendações implementadas por Empresa

prever com certeza se a empresa terá ou não interesse em implementar as Recomendações posteriormente. Consequentemente, a Avaliação Global e-swim ficou-se nos 24 pontos para a Avaliação Final, comparados com os 36 pontos previstos na Avaliação Intermédia, resultantes respetivamente da variação entre 2 e 1 valores no critério “Liderança” da dimensão “Tecnologia” e entre 2 e 0 valores na dimensão “Atividades”. Esta foi a segunda origem das principais diferenças entre a Avaliação Intermédia e a Avaliação Final.

Adicionalmente, uma Empresa não implementou duas das 5 Recomendações propostas por razões táticas, nomeadamente porque no tempo previsto de duração do projeto não seria oportuno dar continuidade ao trabalho desenvolvido. Este facto constituiu na prática uma opção em participar no projeto com menos intensidade que o inicialmente previsto. Consequentemente, a Avaliação Global e-swim ficou-se nos 36 valores para a Avaliação Final, comparados com os 41 valores previstos na Avaliação Intermédia, resultantes respetivamente da variação entre 2 e 1 valores no critério “Liderança” e entre 1 e 0 no critério “Estrutura”, ambos da dimensão “Tecnologia”, e entre 3 e 2 valores na dimensão “Atividades”.

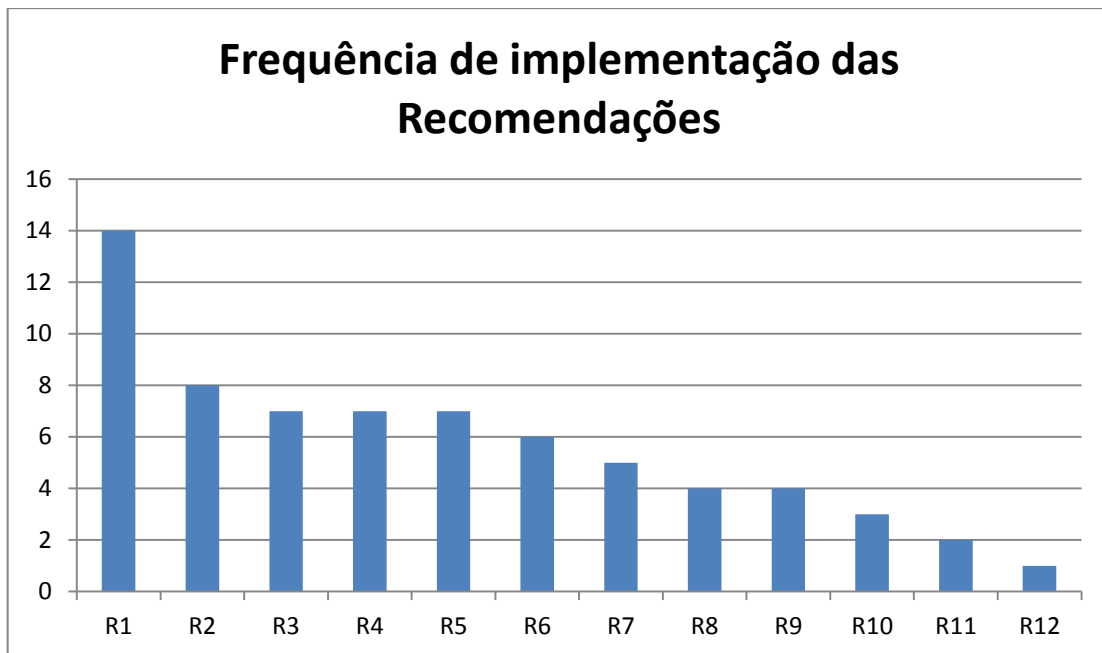


Figura 5.16. Frequência de implementação das Recomendações

Finalmente, uma Empresa implementou 3 das 4 Recomendações, ficando apenas uma por implementar, devido a não ter sido possível dar continuidade a essa tarefa específica, tendo em conta as prioridades de negócio definidas.

A Figura 5.16 apresenta os resultados finais em termos de implementação das Recomendações, que não são substancialmente diferentes daqueles que se propuseram inicialmente para adoção.

Em síntese, nesta fase, podem retirar-se as seguintes conclusões:

- A Avaliação Final foi muito semelhante à Avaliação Intermédia, tendo sido naturalmente influenciada pela intervenção nas empresas segundo o modelo definido
- As Recomendações efetuadas foram quase totalmente implementadas pelas empresas participantes
- A principal diferença entre a previsão de adoção e a realidade da implementação resultou da gestão de prioridades interna de cada empresa, em função dos seus objetivos de negócio

Tendo em conta estas conclusões, procedeu-se a um enquadramento adicional das Recomendações nas dimensões do Modelo e-swim que ajuda a maximizar a sua adequabilidade. Trata-se de um exercício de base empírica que resulta da análise da

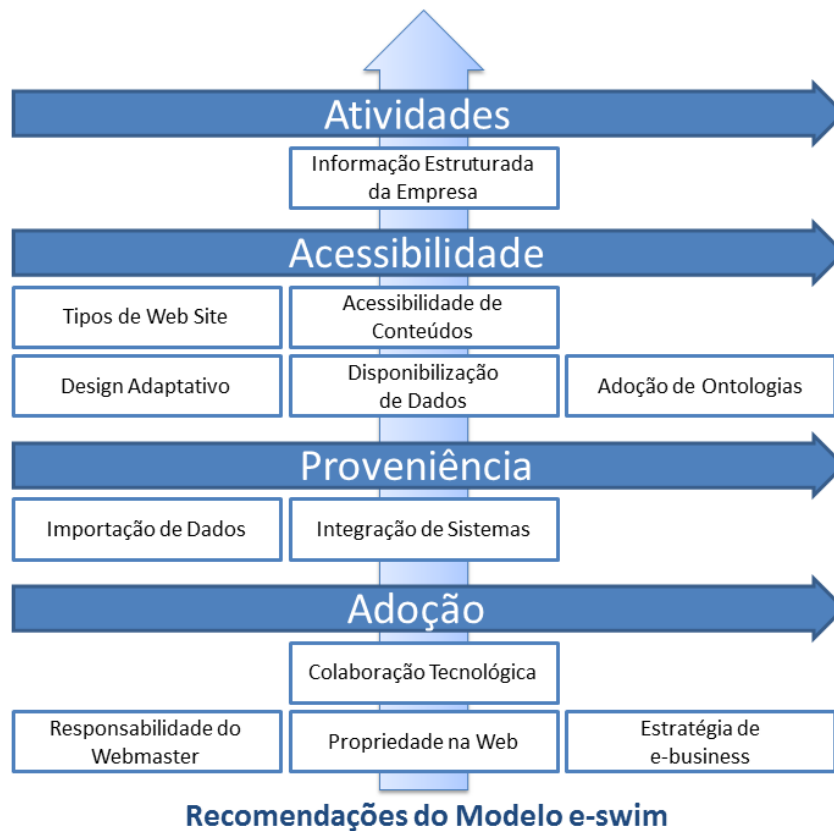


Figura 5.17. Recomendações do modelo e-swim

intervenção efetuada nas empresas mas que considera um pressuposto relevante. A intervenção foi realizada em 19 empresas no mesmo período, partindo de avaliações iniciais distintas e, naturalmente, chegando a avaliações finais também distintas. Presumindo que a reaplicação do modelo em qualquer combinação de empresas e períodos deverá apresentar semelhanças com a evolução analisada neste conjunto de empresas, é possível determinar uma sequência expectável para a adoção das diferentes Recomendações.

A Figura 5.17 apresenta a relação entre as Recomendações elaboradas e as dimensões do Modelo e-swim, fornecendo um guia genérico para a sua adequabilidade a qualquer empresa, que constitui uma referência para a reaplicação generalizada do modelo, a utilizar com a flexibilidade e a personalização necessárias a cada caso particular.

#### 5.6.4 Aplicações e Atividades

A avaliação da dimensão “Atividades” implica a identificação em concreto das atividades dos utilizadores que requerem a utilização da tecnologia. Por outro lado, a resposta à questão central de investigação “*Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?*” implica conhecer ainda quais as aplicações utilizadas pelas empresas participantes neste projeto. Para endereçar estas necessidades recolheu-se e analisou-se a informação apresentada no Anexo 8 e que se comenta de seguida.

Foram identificadas 43 aplicações com utilização da tecnologia Web Semântica. As aplicações do tipo CMS (*Content Management System* ou Sistema de Gestão de Conteúdos) foram as mais frequentes, com 23 ocorrências, 53% do total das aplicações, seguidas da Pesquisa (*Search*) e da EAI (*Enterprise Application Integration* ou Integração de Aplicações) com respetivamente 35% e 12% de representatividade (cf. Figura 5.18). Estes resultados são bastante coerentes com as evidências recolhidas na totalidade do mercado.

Se a maioria das aplicações de Sistemas de Informação utilizadas nas empresas não são baseadas em tecnologia Web, não seria de esperar que este estudo revelasse a utilização frequente da tecnologia em aplicações como, por exemplo, ERP ou CRM, mais associadas aos processos nucleares das organizações. No entanto, é possível que essa

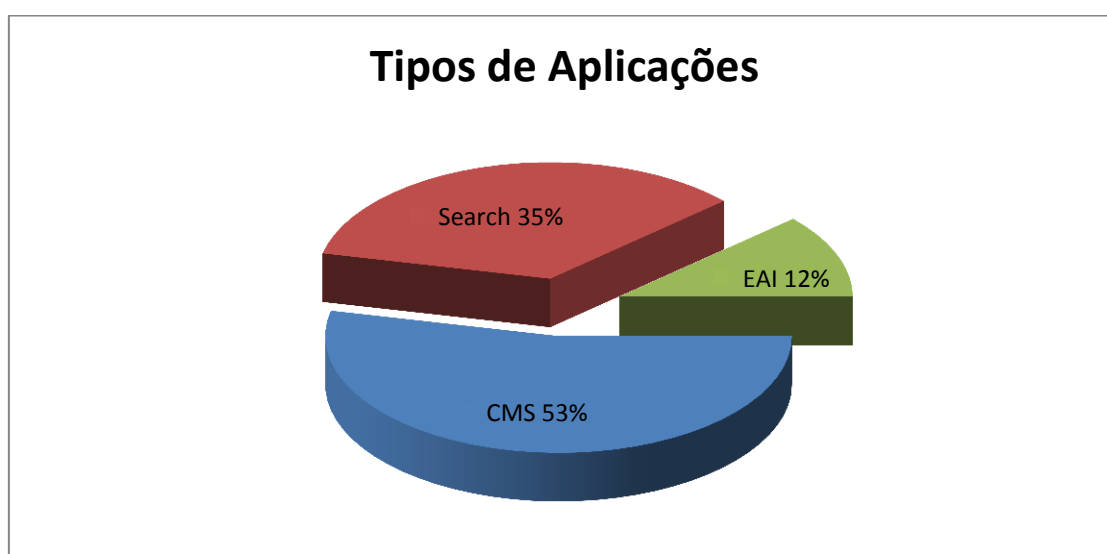


Figura 5.18. Tipos de Aplicações identificados no âmbito do projeto

situação mude, tendo em conta a utilização crescente de aplicações baseadas na *Cloud* que recorrem à tecnologia Web e a adoção de Serviços Web na relação entre Administração Pública e empresas. Se esse caminho implicará a utilização de tecnologia Web Semântica, particularmente em Portugal, será provável mas impossível de prever com exatidão.

Das 43 aplicações com utilização da Web Semântica identificadas, 40 aplicações (93%) são da tipologia Web Site, enquanto apenas 3 (7%) são da tipologia *Web Service*. Não foram registadas ocorrências das tipologias Extranet, Intranet ou *Web App* (cf. Figura 5.19). Este facto é curioso mas também coerente com a tendência de associar a tecnologia à condição de acesso aos dados, ou seja, não se está a observar apenas a Web Semântica, está-se a observar a Web Semântica Aberta, ou, de uma forma mais aplicada, como foi amplamente verificado na literatura e no mercado, não se está apenas a criar *Linked Data*, mas sim *Linked Open Data*.

A outra análise pertinente foca-se nas Atividades desempenhadas pelos utilizadores.

Uma vez que no âmbito da utilização de cada Aplicação foi observada mais que uma Atividade, para essa análise foram consideradas duas séries de valores: a primeira respeitante à percentagem de atividades identificadas num total de 60 atividades observadas; e a segunda respeitante à percentagem de aplicações onde cada atividade foi

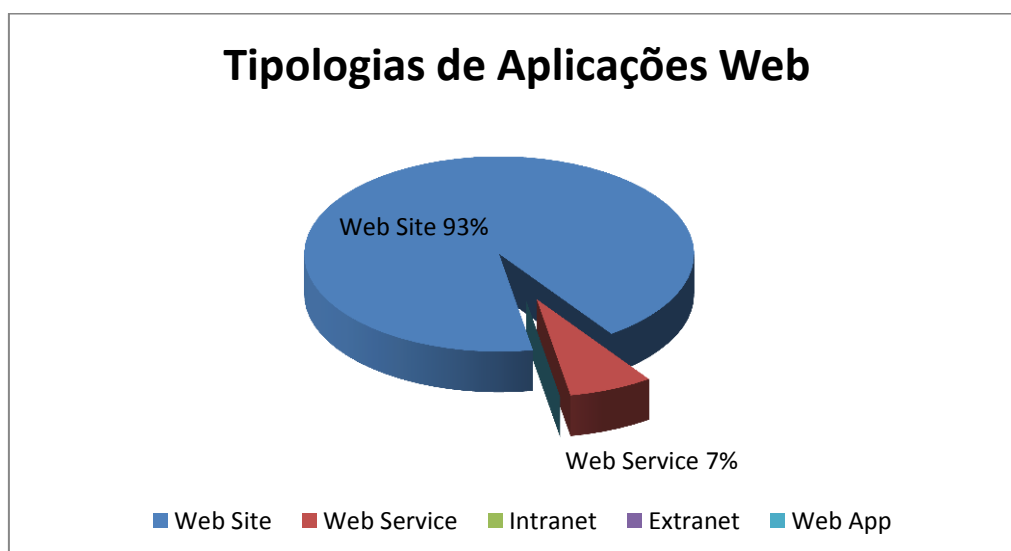


Figura 5.19. Tipologias de Aplicação Web identificadas no âmbito do projeto

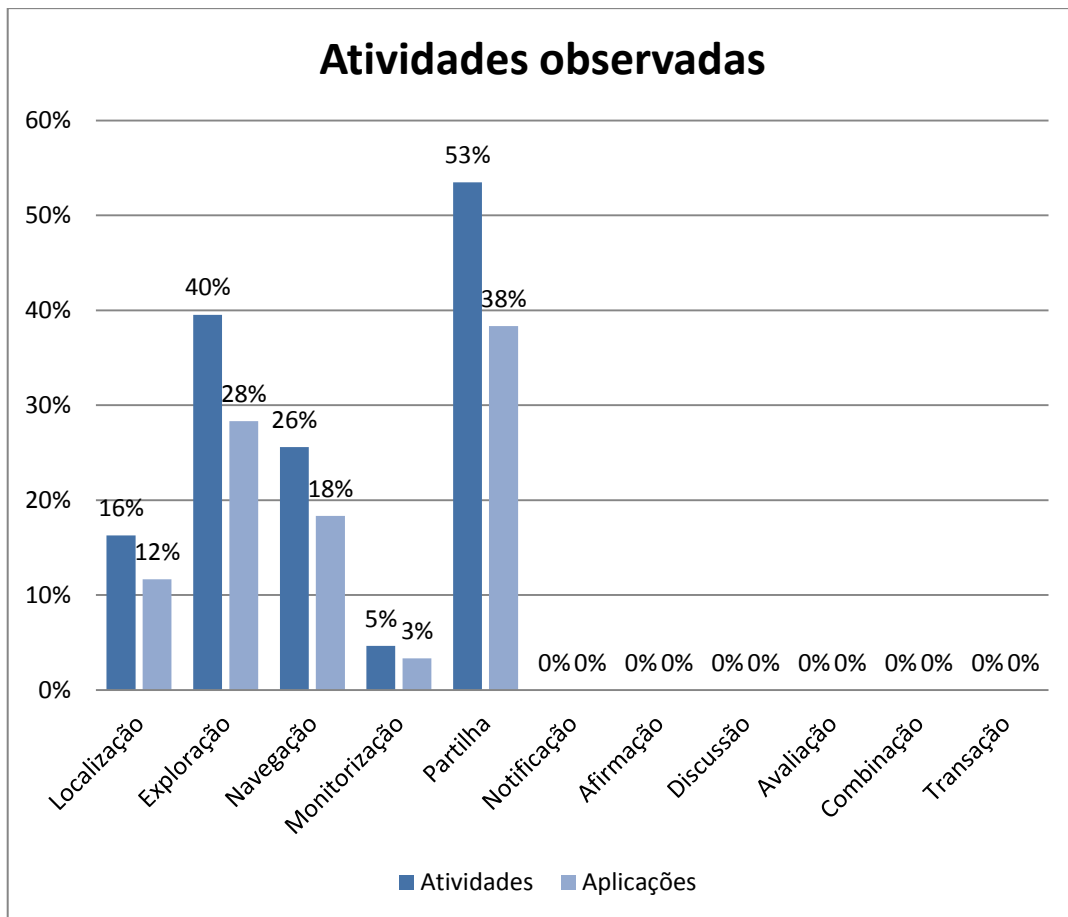


Figura 5.20. Atividades observadas no âmbito do projeto

observada, no total de 43 aplicações já referido. Foram assim observadas 1.4 Atividades por Aplicação, em média.

Conforme se pode verificar pela análise da Figura 5.20, das 11 atividades identificadas na Secção 3.4, foram observadas ocorrências de apenas 5 atividades. A atividade “Partilha” foi a mais frequente, com 38% do total de Atividades e esteve presente em 53% das Aplicações, seguida da atividade “Exploração”, com uma representatividade de 28% e estando presente em 40% das Aplicações.

As Atividades mais frequentes, “Partilha” e “Exploração”, têm curiosamente a particularidade de se complementarem, formando um processo colaborativo. Ainda que não seja evidente tratar-se de uma causa ou de uma consequência, isto é, se são as atividades que influenciam a implementação da tecnologia ou a última que implica a introdução das primeiras, estas atividades confirmam que a colaboração constitui uma alavanca potencial para a implementação da tecnologia Web Semântica.

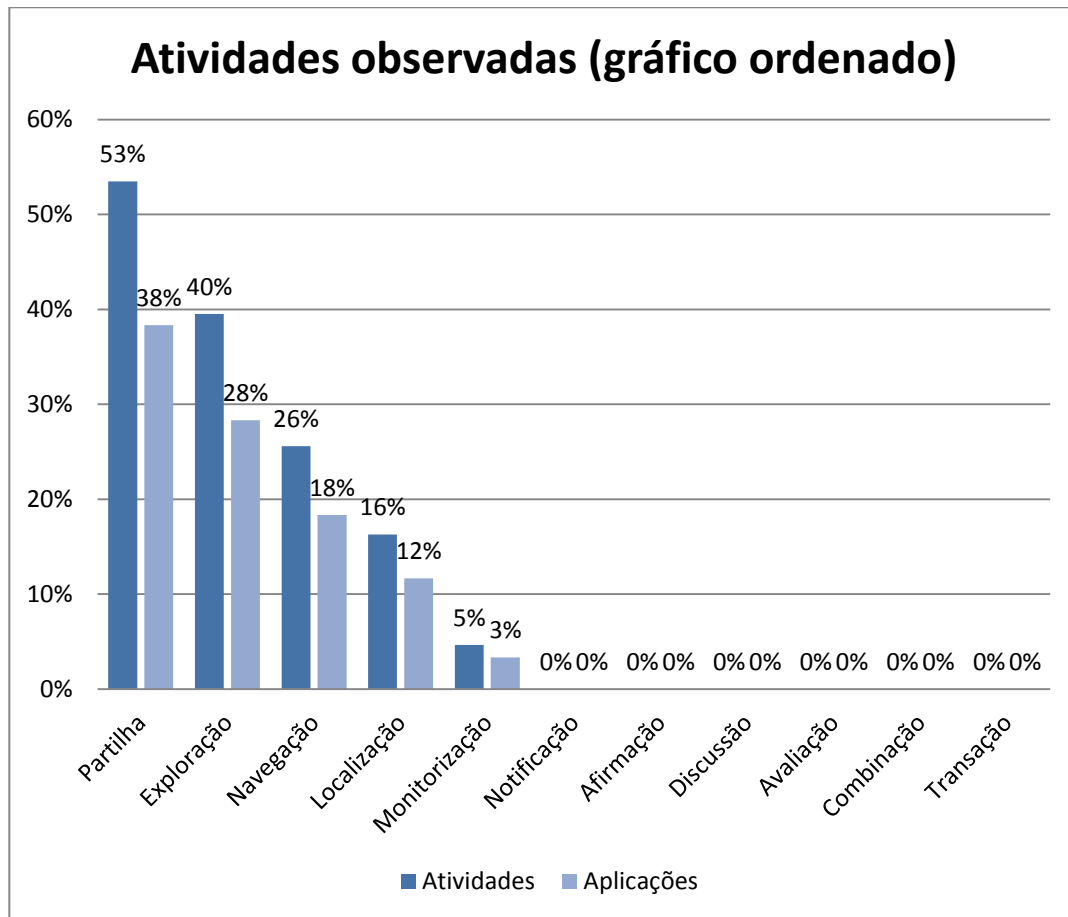


Figura 5.21. Atividades observadas no âmbito do projeto, gráfico ordenado

A transformação da Figura 5.20 na Figura 5.21 suporta uma primeira constatação sobre os resultados obtidos.

Admitindo que as Atividades identificadas surgiram naturalmente com a implementação da tecnologia nas empresas, como pôde verificar-se pela evolução entre a Avaliação Inicial e a Avaliação Final, as atividades mais frequentes deveriam ser as que são mais facilmente ou mais rapidamente implementadas e, por isso, a ordenação das Atividades segundo a sua frequência poderá evidenciar essa indicação.

Adicionalmente, analisando a descrição de cada Atividade bem como o seu grau de conceptualização e diversificação, conforme apresentado na Figura 4.5, pode perceber-se que o seu aparecimento nas organizações pode ter também uma sequência natural, definindo como que uma prioridade para a sua introdução. A utilização mais prolongada do Modelo e-swim poderá introduzir factos adicionais que comprovem esta suposição, mas a experiência empírica e as evidências recolhidas sugerem que as atividades mais diversas sejam as primeiras a surgir e portanto mais frequentes a ser

observadas em fases precoces da implementação, independentemente do seu grau de conceptualização. Ou posto de outra forma, as tarefas que mais rapidamente poderão tirar partido da tecnologia Web Semântica serão as de maior diversificação, independentemente do seu grau de conceptualização, pois se as tarefas mais simples podem ser naturalmente suportadas na tecnologia, independentemente de já o serem com outras tecnologias, também as tarefas mais complexas poderão beneficiar da tecnologia como suporte à atuação do utilizador.

Esta suposição é traduzida na Figura 5.22. O espaço da aplicação da Web Semântica aparece assim mais vasto, incluindo todas as tarefas diversas, mesmo aquelas que não são suficientemente claras para serem associadas a uma atividade específica, que poderíamos designar por tarefas diversificadas. Esse espaço é simultaneamente progressivo, pois à medida que as atividades vão sendo introduzidas ou identificadas, verificar-se-á que dirão respeito a tarefas mais repetitivas que não serão nem tão simples que não possam ser satisfeitas por tecnologias também elas mais simples, nem tão complexas que possam ser tipificadas ou sistematizadas de forma rentável.

## Web Semântica: Tarefas e Atividades



Figura 5.22. Caracterização de Tarefas e Atividades segundo a sua Conceptualização e Diversificação: uma nova proposta

A intervenção nas empresas e o segundo ciclo de Investigação – Ação terminam com a produção de um Relatório Final com as principais conclusões do projeto. O relatório permite que cada empresa analise a sua evolução comparativa com outras empresas e as possibilidades de evolução futura. O Relatório Final foi disponibilizado às empresas através do Web Site do Projeto Raio X.

### 5.6.5 Fatores de Sucesso

Face à apresentação do trabalho efetuado, foram registados os seguintes fatores de sucesso.

#### **Modelo desenvolvido e partilhado com os pares**

O Modelo e-swim foi desenvolvido com base na revisão de literatura efetuada e proposto como ferramenta de apoio à implementação da tecnologia Web Semântica nas empresas. A sua disseminação através da comunidade científica permitiu antecipar as

suas fragilidades e definir uma proposta de investigação que respondesse às necessidades identificadas.

### **Validação do Modelo com recurso a Investigação-Ação**

A validação do Modelo e-swim realizada num conjunto de empresas com base numa metodologia de Investigação-Ação produziu bons resultados, quer do ponto de vista prático para as empresas participantes, quer do ponto de vista teórico, na forma de dados recolhidos que permitem quantificar o impacto da adoção da tecnologia na amostra e perspetivar a generalização da aplicação do modelo a outras empresas.

### **Identificação de aplicações concretas**

A identificação de aplicações concretas a implementar, em alternativa a planos genéricos e tecnológicos, facilita a adoção da tecnologia em ambiente empresarial. O trabalho conseguiu identificar essas aplicações e relacioná-las com as atividades dos utilizadores, contribuindo para a ordenação de ambas segundo a sua maior facilidade ou rapidez de implementação.

#### **5.6.6 [Barreiras identificadas](#)**

A realização do trabalho permitiu identificar barreiras à sua evolução que são relevantes no contexto de extração de conclusões pertinentes.

### **Envolvimento das empresas**

O envolvimento constante das empresas num projeto de média duração como o descrito neste trabalho é muitas vezes difícil de obter e exige uma dedicação acima de qualquer previsão. As prioridades dos negócios mudam a uma velocidade por vezes incompatível com os requisitos e objetivos de médio prazo. E isso é verdade não apenas no que diz respeito ao trabalho desenvolvido no âmbito deste projeto bem como no âmbito mais lato do investimento em tecnologia nas empresas, que não pode ser maximizado apenas com recurso a ações de curto prazo.

### **Diversidade setorial e estratégica**

A diversidade dos negócios e as infinitas abordagens estratégicas instantâneas que cada negócio adota constituem um ambiente adverso à sistematização. A possibilidade de elaborar mais recomendações e de envolver mais empresas ficou condicionada pelo tempo disponível. Assim, o resultado do projeto poderá ser mais interessante se puder ser reaplicado em mais cenários, objetivo a que este trabalho se propôs desde o início.

## 6 Conclusão

### 6.1 Síntese do trabalho realizado

A Web Semântica constitui um importante desafio para a Web na atualidade. A transformação progressiva da Web baseada nos documentos para a Web baseada nos dados está a dar origem a uma base de dados distribuída à escala global.

Sendo a implementação da Web Semântica nas Empresas ainda um tema pouco explorado e encontrando-se numa fase precoce de implementação, este trabalho apresentou a relevância do estudo do tema e propôs um modelo para facilitar e guiar a implementação desta tecnologia nas organizações, sistematizando o caminho a seguir para uma adoção mais rápida e intensiva. O modelo permite ainda a avaliação do impacto da introdução da tecnologia nas aplicações utilizadas por essas organizações e nas tarefas desempenhadas pelos utilizadores, através da medição do progresso da adoção da tecnologia.

Procurando endereçar a questão de investigação “Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?”, desenvolveu-se um trabalho no terreno, junto das empresas, em que os resultados obtidos se traduziram simultaneamente num contributo para um melhor desempenho em termos tecnológicos e com impacto no negócio das empresas e na validação do Modelo de Implementação da Web Semântica em ambiente empresarial.

### 6.2 Análise da contribuição científica

O Modelo e-swim (*Enterprise Semantic Web Implementation Model*) proposto neste trabalho baseia-se no inter-relacionamento da Tecnologia e da Organização e inclui 4 dimensões (Adoção, Proveniência, Acessibilidade e Atividades) que permitem a quantificação da evolução da implementação da tecnologia de Web Semântica (cf. Figura 4.1).

A dimensão Adoção está associada ao conhecimento e à disponibilidade da Organização (a empresa em termos de liderança e restantes colaboradores) e do Exterior (o ambiente externo com que a empresa se relaciona), em explorar as oportunidades apresentadas pela Tecnologia. A dimensão Proveniência endereça a identificação de fontes de dados internas e externas e o fluxo da sua utilização. A dimensão Acessibilidade avalia o grau de adequação dos dados aos standards da tecnologia. Finalmente, a dimensão Atividades pretende identificar as atividades dos utilizadores que são suportadas pela tecnologia.

O Modelo e-swim proposto foi validado com recurso a 2 ciclos de Investigação-Ação com intervenção em 19 empresas. A avaliação da implementação da tecnologia nas empresas foi efetuada com recurso à ferramenta de avaliação apresentada na Figura 5.5 e no Anexo 3, que quantifica cada uma das dimensões e permite o cálculo da variável e-swim, que adota o mesmo acrónimo do Modelo de Implementação. A ferramenta revelou-se adequada para o efeito, permitindo produzir avaliações representativas, bem como previsões fiáveis. A Avaliação Inicial para a variável e-swim, obtida com a ferramenta, foi em termos médios para as empresas intervencionadas de 6 pontos percentuais. A Avaliação Final foi em termos médios de 46 pontos percentuais, sendo reveladora de uma evolução considerável da generalidade das empresas durante a realização da intervenção. A Avaliação Intermédia, realizada como a melhor previsão para a Avaliação Final, foi em termos médios de 48 pontos percentuais, valor muito próximo da Avaliação Final verificada posteriormente, o que traduz a capacidade de previsão ou de planeamento fornecida pelo modelo.

O trabalho propõe um conjunto de 12 Recomendações de ação, relacionadas com as dimensões do Modelo de Implementação (cf. Figura 5.17) e que permitem facilitar e agilizar o processo de implementação da tecnologia. Foram implementadas uma média de 3.6 Recomendações por empresa, com metade das Recomendações a serem adotadas por um terço das empresas. A adequabilidade das Recomendações está relacionada com o grau de adoção da tecnologia, devendo estas ser adotadas pela ordem apresentada na Figura 5.17. Essa ordem de adoção deverá ser considerada como orientação para

aplicação futura nas empresas, podendo ser adaptada em função da avaliação efetuada e das observações recolhidas.

Foi desenvolvido um cenário hipotético de utilização da tecnologia, denominado e-CAE e apresentado na Secção 5.6.1, considerado verosímil por todas as empresas intervencionadas. A consequência de implementação desse cenário seria a necessidade de adoção de todas as Recomendações pela generalidade das empresas Portuguesas, o que se traduziria automaticamente na obtenção de uma avaliação e-swim de 58 pontos, valor superior ao obtido em termos médios na avaliação final das 19 empresas.

O Modelo de Implementação e a ferramenta de Avaliação foram validados no terreno com recurso aos dois ciclos de Investigação-Ação interligados, que permitiram identificar ações bem como a sua reaplicação nas empresas. As empresas participantes no Projeto Raio X, designação adotada para ilustrar a capacidade de observar a Web para além das evidências superficiais, acompanharam a evolução do projeto através de um Web Site criado para o efeito e de sessões de acompanhamento presencial e à distância. O desconhecimento generalizado da tecnologia foi ultrapassado com a adesão ao projeto e, como primeira reação, as empresas colocaram o seu Web Site, a aplicação que imediatamente associaram como principal componente da sua presença na Web, como uma prioridade de intervenção. Para além disso, a necessidade de identificar um Webmaster, um responsável por essa presença Web, foi imediatamente reconhecida como uma necessidade.

Foi possível concluir que as tipologias de aplicações que mais podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas são os Web Sites e os *Web Services*. Outras tipologias de aplicações Web, como as Intranets, poderão ainda ser consideradas relevantes, mas não oferecem benefícios imediatos evidentes, principalmente em fases precoces da adoção da tecnologia, isto é, em casos de classificações e-swim mais baixas.

Foi identificada a utilização da tecnologia de Web Semântica em aplicações do tipo CMS (*Content Management System* ou Sistemas de Gestão de Conteúdos), Pesquisa (*Search*) e EAI (*Enterprise Application Integration* ou Integração de Aplicações). Outros tipos de aplicações poderão ser relevantes no futuro, nomeadamente aquelas aplicações

mais próximas do negócio nuclear das empresas, como os ERP; no entanto, na atualidade, ainda se verifica uma separação natural entre estas aplicações de Sistemas de Informação Empresariais e as Aplicações Web, que poderá num futuro próximo vir a ser atenuada.

Foi também possível concluir que estas aplicações são mais propícias à implementação da tecnologia quando exibem propósitos colaborativos, nomeadamente quando focadas nas atividades de Partilha de informação da empresa e de Exploração de temas relevantes. Outras atividades relevantes identificadas foram as atividades de Navegação ou pesquisa livre, Localização de recursos e Monitorização de informação.

Com base nestes resultados, considerou-se que o espaço para adoção precoce das tecnologias da Web Semântica inclui tarefas diversas, mesmo aquelas que não são suficientemente claras para serem associadas a uma atividade específica, que podem ser designadas por tarefas diversificadas. Com a adoção progressiva da tecnologia, outras atividades tornam-se mais relevantes, nomeadamente as tarefas mais repetitivas que não são nem tão simples de forma a poderem ser satisfeitas por tecnologias mais simples que a Web Semântica, nem tão complexas que não possam ser tipificadas ou sistematizadas de forma rentável (cf. Figura 5.22).

Retoma-se então a questão central de investigação colocada nesta tese e apresenta-se a resposta sintetizada à mesma.

- **Que aplicações podem contribuir para uma adoção rápida e generalizada da Web Semântica nas Empresas?**

A adoção rápida e generalizada da tecnologia Web Semântica nas empresas pode ser conseguida com aplicações de tipologia Web Site ou *Web Service*, principalmente aplicações baseadas em gestão de conteúdos, pesquisa ou integração, utilizadas em tarefas diversificadas, principalmente em atividades de exploração de temas relevantes e partilha de informação.

Responde-se assim à questão central de investigação proposta nesta tese, fornecendo uma pista relevante para a futura adoção da tecnologia nas organizações.

Com base na partilha efetuada do trabalho realizado com a comunidade científica, o envolvimento das empresas participantes e a disponibilização dos resultados obtidos,

considera-se que o Modelo de Implementação e-swim, a ferramenta de avaliação desenvolvida e as Recomendações elaboradas permitem avaliar o grau de adoção da tecnologia e determinar ações para a sua mais rápida e eficaz implementação.

## 6.3 Trabalho Futuro

### 6.3.1 O futuro do Modelo e-swim

O trabalho desenvolvido permitiu analisar as ações necessárias e identificar um conjunto de Recomendações que generalizam a intervenção e que tiveram uma recetividade elevada por parte das empresas. No entanto, considera-se que este conjunto de Recomendações não esgota todo o espectro de intervenção e poderá ser complementado com Recomendações adicionais.

O défice de conhecimento na tecnologia de Web Semântica por parte das empresas que se verificou era previsível, mas foi acompanhado de várias lacunas nas competências dos colaboradores das empresas para utilizarem a tecnologia, principalmente por falta de conhecimentos de base. Tendo sido verificadas também dificuldades no domínio da língua inglesa e a inexistência de uma formação de qualidade em língua portuguesa, iniciou-se, por parte dos promotores do projeto “Raio X”, a criação de um “Curso de *e-business*” que pudesse transmitir conhecimentos introdutórios da tecnologia e incorporar as conclusões do trabalho realizado. Este curso começou a ser desenvolvido e ministrado em 2014 e deverá terminar no final do ano de 2015.

As empresas demonstraram um interesse elevado nos temas abordados, mas alguma ansiedade pela dificuldade de acompanhamento de alguns tópicos. Tornou-se evidente que “ter um Web Site” é claramente insuficiente e que a gestão da presença Web de uma empresa envolve vários aspetos que apenas algumas empresas conseguirão dominar internamente. A necessidade de considerar um Webmaster com competências diversificadas e que garanta a gestão dessa presença Web na empresa sob diferentes perspetivas foi também verificada, constatando-se que estas competências não são frequentes nos profissionais disponíveis no mercado. Idealmente, as empresas desejam um “produto integrado” que combine todas estas características e lhes seja disponibilizado de forma comercialmente simplificada, uma espécie de “Seguro Web”,

onde o pagamento de um “prémio” fixo, eventualmente com algumas correções variáveis, asseguraria a criação, manutenção e evolução da sua Presença Web, em todos os aspetos.

Será também interessante analisar a evolução da avaliação e-swim em função do tempo, ou seja, para além da avaliação efetuada às 19 empresas e restrita à duração do projeto, qual será a evolução registada nessas empresas ao longo de períodos de tempo de maior dimensão. O alargamento desta avaliação a um número mais elevado de empresas e com características diversas será também de considerar.

É expetável que empresas com avaliações superiores evoluam a velocidades mais lentas, principalmente quando se aproximam do final da escala da implementação da tecnologia. Será igualmente interessante verificar que tipo de tendência em termos de função matemática a avaliação e-swim poderá seguir, e se a sua introdução em organizações com avaliações muito baixas é sistematicamente originadora de variações quantitativamente significativas, como se verificou no trabalho efetuado.

### 6.3.2 Oportunidades de Investigação

Considera-se ainda que temas adjacentes aos abordados nesta tese poderão ser alvo de trabalho futuro, dando continuidade e utilizando os resultados obtidos neste trabalho:

- **Que contribuições podem fornecer as Aplicações Web quanto à utilização da Web Semântica, nomeadamente Web Sites, Extranets, Intranets, *Web Apps* ou *Web Services*?**

As diferentes tipologias de Aplicações Web introduzem aspetos de utilização muito concretos que influenciam a sua contribuição para a adoção da tecnologia Web Semântica pelas Empresas. Esses aspetos estão associados a tipos de Sistemas de Informação e a estádios da adoção da tecnologia que as tornam mais ou menos adequadas.

Os Web Sites e os *Web Services* estão melhor posicionados para serem adotados precocemente na adoção da tecnologia. A introdução das restantes tipologias, Extranet,

Intranet e *Web App* estará associada a estádios de desenvolvimento mais adiantados e a outros fatores a determinar.

- **Que tipos de Sistemas de Informação poderão evoluir consideravelmente com a adoção da Web Semântica?**

Os diferentes tipos de sistemas de informação serão influenciados de forma distinta pela tecnologia em causa. Os resultados deste trabalho evidenciam essa distinção e apontam para a gestão de conteúdos, a pesquisa e a integração como os mais significativos em fases precoces. No entanto, outros tipos de sistemas poderão conhecer desenvolvimentos significativos, com destaque para os sistemas mais ligados aos processos nucleares das empresas. O cenário hipotético e-CAE apresentado mostra como o *e-government* pode influenciar significativamente a utilização da tecnologia em aplicações desse tipo.

- **Que tipo de empresas poderá mais beneficiar com a adoção da Web Semântica?**

Conforme referido na Secção 3.5.1, a diferenciação entre setores, dimensão ou qualquer outra característica das empresas que possa facilitar ou prejudicar a adoção da Web Semântica parece pouco significativa quando comparada com o potencial dessa adoção para o universo das empresas como um todo.

Os casos de estudo identificados apontam para uma adoção da tecnologia em setores da economia onde o conteúdo (Media, Cultura, Direito) ou o conhecimento (Farmacêutico, Educação, Turismo) desempenham um contributo significativo na criação de valor económico.

Não foram registadas evidências durante o trabalho no sentido de recomendar a utilização da tecnologia a um tipo de empresa ou setor específico, pois foi possível encontrar vários casos práticos com diferentes características e aplicações de utilização generalizada que facilitam o acesso à tecnologia a todas as empresas. Em particular, o esforço dos principais motores de pesquisa na utilização do *schema.org* torna-o recomendável para ser utilizado por todas as empresas.

No entanto, um estudo mais compreensivo e detalhado em relação aos aspetos de diferenciação das empresas e respetivo impacto na implementação da tecnologia, assim como dos benefícios resultantes poderá constituir um trabalho adicional relevante.

- **Como pode a Web Semântica contribuir para a Inovação?**

Conforme referido na Secção 3.5.2, o estado atual de reduzida adoção desta tecnologia poderá impedir conclusões sobre a sua adoção em áreas ou atividades específicas das empresas; no entanto, o interesse particular na Inovação poderá facilitar o próprio processo de adoção.

As evidências recolhidas apontam para que as organizações, de uma forma geral, e as empresas, em particular, devam considerar a adoção da tecnologia para promover a inovação nos seus processos, principalmente através da colaboração interna, entre colaboradores, e externa, entre colaboradores de organizações distintas. Foi possível observar, no âmbito da intervenção realizada nas empresas, que a introdução da tecnologia influenciou as metodologias de trabalho e contribuiu para o desenvolvimento individual. A contribuição das políticas e práticas públicas poderá ter uma influência determinante na massificação da utilização da tecnologia, mas a inovação pode ser liderada pelas empresas, com benefícios que poderão transformar o investimento efetuado na tecnologia em vantagens competitivas e melhores resultados.

Assim, a este respeito, um estudo que envolva a correlação entre casos de adoção de utilização da Web Semântica e implementação de práticas inovadoras ou disponibilização de produtos e serviços inovadores nas empresas poderá ser feito para sustentar estas constatações.

- **Que resultados da adoção da Web Semântica nas Empresas podem ser reaplicados na adoção pelos Estados e respetivos organismos?**

A potencial reaplicação dos resultados obtidos na área das fontes de dados governamentais será útil para a investigação nessa direção. A adoção da tecnologia estará dependente da atuação do Estado, como foi proposto com o cenário hipotético e-CAE. Mas a motivação para a adoção será distinta, com o Estado a assumir naturalmente um papel de líder, influenciador ou mesmo de prescritor, por razões do interesse público. As oportunidades associadas ao Estado Português existem, estando a sua concretização

dependente da opção por parte do Estado em tomar um papel mais pedagógico ou controlador.

- **A utilização da Web Semântica nas Empresas contribuirá definitivamente para a sua disseminação?**

Este trabalho contribuiu para uma resposta positiva a esta questão através da análise do seu impacto nas empresas intervencionadas. As gerações anteriores da Web também constituíram uma resposta do mesmo sentido em relação à geração Web 3.0. No entanto, somente o tempo e uma adoção mais alargada da tecnologia poderão confirmar esta suspeita. Um caso de estudo relativo a uma implementação específica de âmbito alargado poderá ser um forte contributo para a resposta. Atualmente, o schema.org é um forte candidato para a realização desse estudo.

- **Quais as ferramentas necessárias para a medição adequada da utilização da Web em ambiente empresarial?**

A sobreposição existente das diversas gerações da Web é uma evidência, pois uma presença Web empresarial deve continuar a ser uma fonte de informação direcionada ao mercado e tirar partido da interação colaborativa e social com os intervenientes externos.

Assim, a extensão da aplicação do Modelo às diferentes gerações Web é uma possibilidade de evolução. A avaliação e-swim poderia tornar-se a avaliação da geração Web 3.0 e serem criadas duas novas avaliações para as gerações anteriores. Relativamente às dimensões do Modelo de Implementação que suportariam essas avaliações, parece adequado que as dimensões organizacionais, Adoção e Atividades, possam ser reaplicadas e que as dimensões tecnológicas, Proveniência e Acessibilidade, sejam revistas para cada geração.



## Bibliografia

- Agência para a Modernização Administrativa. (2015). dados.gov.pt. Obtido 30 de Março de 2015, de <http://www.dados.gov.pt>
- Ahmed, Z., & Gerhard, D. (2010). Role of Ontology in Semantic Web Development. *Knowledge Management*. Obtido de <http://arxiv.org/abs/1008.1723>
- Allemang, D. (2010). Semantic Web and the Linked Data Enterprise. Em D. Wood (Ed.), *Linking Enterprise Data* (pp. 3–23). Obtido de <http://linkeddatadeveloper.com/Projects/Linking-Enterprise-Data/Manuscript/led-allemang.html>
- Anantaram, C. (2007). Case Study: Use of Semantic Web Technologies in Natural language interface to Business Applications. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/Tata/>
- Anderson, C. (2007). *A Cauda Longa*. Lisboa, Portugal: Actual Editora.
- Anderson, C. (2008). Annual «cutting the Long Tail down to size» roundup. Obtido 30 de Março de 2015, de [http://www.longtail.com/the\\_long\\_tail/2008/12/annual-cutting.html](http://www.longtail.com/the_long_tail/2008/12/annual-cutting.html)
- Antunes, G., Bakhshandeh, M., Mayer, R., Borbinha, J., & Caetano, A. (2014). Using Ontologies for Enterprise Architecture Integration and Analysis. *Complex Systems Informatics and Modeling Quarterly*, (1), 1–23.
- Aroyo, L., & Houben, G. (2010). User modeling and adaptive Semantic Web. *Semantic Web*, 1, 105–110. doi:10.3233/SW-2010-0006
- Barabasi, A. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509–512. doi:10.1126/science.286.5439.509
- Baskerville, R. L., & Wood-Harper, A. T. (1996). A critical perspective on action research as a method for information systems research. *Journal of Information Technology*, 11(3), 235–246.
- Benjamins, V. R., Radoff, M., Davis, M., Greaves, M., Lockwood, R., & Contreras, J. (2011). Semantic Technology Adoption: A Business Perspective. Em J. Domingue, D. Fensel, & J. A. Hendler (Eds.), *Handbook of Semantic Web Technologies* (pp. 623–657). Berlin, Heidelberg: Springer.
- Berners-Lee, T. (2006a). Linked Data, Design Issues. Obtido 30 de Março de 2015, de <http://www.w3.org/DesignIssues/LinkedData.html>

- Berners-Lee, T. (2006b). Web future - Building beneficial weblike things. Obtido 30 de Março de 2015, de [http://www.w3.org/2006/Talks/0314-ox-tbl/#\(1\)](http://www.w3.org/2006/Talks/0314-ox-tbl/#(1))
- Berners-Lee, T. (2010). Abstractions in Web architecture - Design Issues. W3C. Obtido 30 de Março de 2015, de <http://www.w3.org/DesignIssues/Abstractions.html>
- Berners-Lee, T., Hendler, J., Hall, W., Shadbolt, N., & Weitzner, D. J. (2006). Computer science. Creating a science of the Web. *Science*, 313(5788), 769–771. Obtido de <http://www.ncbi.nlm.nih.gov/pubmed/16902115>
- Berners-Lee, T., Weitzner, D. J., Hall, W., O’Hara, K., Hendler, J. a., & Shadbolt, N. (2006). A Framework for Web Science. *Foundations and Trends® in Web Science*, 1(1), 1–130. doi:10.1561/1800000001
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, 5(3), 1. doi:10.4018/jswis.2009081901
- Bizer, C., Heath, T., Idehen, K., & Berners-Lee, T. (2008). Linked data on the web (LDOW2008). Em *WWW2008 Workshop on Linked Data on the Web* (Vol. 2008, pp. 1265–1266). Beijing: ACM New York, NY, USA. doi:10.1145/1367497.1367760
- Black, J. (2000). Tiny Tuvalu Profits From Web Name. *The New York Times*. New York. Obtido de <http://www.nytimes.com/2000/09/04/business/tiny-tuvalu-profits-from-web-name.html>
- Brickley, D., & Miller, L. (2015). FOAF Vocabulary Specification. Obtido 30 de Março de 2015, de <http://www.foaf-project.org/>
- Brin, S., & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. Em *Proceedings of the 7th International World Wide Web Conference* (Vol. 30, pp. 107–117). Brisbane. doi:10.1016/S0169-7552(98)00110-X
- Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, 36(2), 3–10.
- Cailliau, R. (1995). A short history of the Web. Obtido 30 de Março de 2015, de [http://www.netvalley.com/archives/mirrors/robert\\_cailliau\\_speech.htm](http://www.netvalley.com/archives/mirrors/robert_cailliau_speech.htm)
- Cardoso, J., Miller, J., Su, J., & Pollock, J. (2008). Academic and Industrial Research : Do their Approaches Differ in Adding Semantics to Web Services? *Lecture Notes in Computer Science*, 3387/2005, 14–21. doi:10.1007/978-3-540-30581-1\_2
- Carlson, S. C. (2015). Graph Theory. *Encyclopedia Britannica*. Obtido 30 de Março de 2015, de <http://www.britannica.com/EBchecked/topic/242012/graph-theory>

- Clarke, C., & Greig, F. (2009). Case Study: A Linked Open Data Resource List Management Tool for Undergraduate Students. Obtido 30 de Março de 2015, de (<http://www.w3.org/2001/sw/sweo/public/UseCases/Talis/>)
- Clarkin, L., & Holmes, J. (2007). Enterprise Mashups. *The Architecture Journal*, (October). Obtido de <https://msdn.microsoft.com/en-us/library/bb906060.aspx>
- Cole, R., Puroo, S., Rossi, M., & Sein, M. K. (2005). Being Proactive : Where Action Research meets Design Research. Em *ICIS 2005 Proceedings* (pp. 1–21). Las Vegas.
- Connolly, D. (2000). A Little History of the World Wide Web. W3C. Obtido 30 de Março de 2015, de <http://www.w3.org/History.html>
- Corcho, O., & García-Castro, R. (2010). Five challenges for the Semantic Sensor Web. *Semantic Web*, 1, 121–125. doi:10.3233/SW-2010-0005
- Cuel, R., Delteil, A., Louis, V., & Rizzi, C. (2004). *Knowledge Web Technology Roadmap - The Technology Roadmap of the Semantic Web. Knowledge Web Network of Excellence*. Obtido de <http://knowledgeweb.semanticweb.org/o2i/index875b.html?page=D141.php>
- Data.gov.uk. (2015). Datasets. Obtido 30 de Março de 2015, de <http://data.gov.uk/data/search>
- Dataversity. (2009). Deploying Semantic Technology Within an Enterprise: A Case Study of the UCB Group Project. Obtido 30 de Março de 2015, de [http://semanticweb.com/deploying-semantic-technology-within-an-enterprise-a-case-study-of-the-ucb-group-project\\_b10693](http://semanticweb.com/deploying-semantic-technology-within-an-enterprise-a-case-study-of-the-ucb-group-project_b10693)
- Dau, F. (2010). Semantic Technologies for Enterprises. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6828, 1–18.
- Davis, M. (2008). *Semantic Wave 2008 Report: Industry Roadmap to Web 3 . 0 & Multibillion Dollar Market Opportunities*. Obtido de <http://project10x.com/semantic-wave-report/>
- Davison, R. M., Martinsons, M. G., & Kock, N. (2004). Principles of canonical action research. *Information Systems Journal*, 14(1), 65–86. doi:10.1111/j.1365-2575.2004.00162.x
- Dick, B. (2000). Action learning, action research and process management: theory, practice, praxis. Em O. Zuber-Skerritt (Ed.), . Brisbane, Australia: Griffith University. Obtido de <http://www.scu.edu.au/schools/gcm/ar/arp/ppar.html>

- Dimitrov, M. (2014). Crossing the Chasm with Semantic Technologies. Em *WASABI 2014 - Workshop on Semantic Web Enterprise Adoption and Best Practice*. Portoroz, Slovenia. Obtido de [2014.wasabi-ws.org/keynote/](http://2014.wasabi-ws.org/keynote/)
- dns.pt. (2015). Estatísticas 2012. Obtido 30 de Março de 2015, de <http://www.dns.pt/pt/estatisticas/?tipo=0&ordem=4&graph=0&ano=2012&subm=Filtrar>
- Ecorys, & European Commission. (2012). *EU SMEs in 2012 : at the crossroads*. Obtido de [http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/performance-review/files/supporting-documents/2012/annual-report\\_en.pdf](http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/performance-review/files/supporting-documents/2012/annual-report_en.pdf)
- European Commission. (2009). *Internet of Things — An action plan for Europe*. Europe. Obtido de [http://ec.europa.eu/information\\_society/policy/rfid/documents/commiot2009.pdf](http://ec.europa.eu/information_society/policy/rfid/documents/commiot2009.pdf)
- European Commission. (2015). Future Internet (Digital Agenda for Europe). Obtido de <https://ec.europa.eu/digital-agenda/en/future-internet>
- Eurostat. (2012). Annual enterprise statistics by size class for special aggregates of activities. Obtido 12 de Dezembro de 2014, de [http://appsso.eurostat.ec.europa.eu/nui/show.do?query=BOOKMARK\\_DS-297817\\_QID\\_352336BB\\_UID\\_-3F171EB0&layout=SIZE\\_EMP,L,X,0;GEO,L,Y,0;NACE\\_R2,L,Z,0;INDIC\\_SB,L,Z,1;TIME,C,Z,2;INDICATORS,C,Z,3;&zSelection=DS-297817TIME,2012;DS-297817NACE\\_R2,B-N\\_S95\\_X\\_K;DS-297](http://appsso.eurostat.ec.europa.eu/nui/show.do?query=BOOKMARK_DS-297817_QID_352336BB_UID_-3F171EB0&layout=SIZE_EMP,L,X,0;GEO,L,Y,0;NACE_R2,L,Z,0;INDIC_SB,L,Z,1;TIME,C,Z,2;INDICATORS,C,Z,3;&zSelection=DS-297817TIME,2012;DS-297817NACE_R2,B-N_S95_X_K;DS-297)
- Eurostat. (2014). Information society statistics - enterprises. Obtido 30 de Março de 2015, de [http://ec.europa.eu/eurostat/statistics-explained/index.php/Information\\_society\\_statistics\\_-\\_enterprises](http://ec.europa.eu/eurostat/statistics-explained/index.php/Information_society_statistics_-_enterprises)
- Feigenbaum, L. (2009). SPARQL By Example. Obtido 30 de Março de 2015, de <http://www.w3.org/2009/Talks/0615-qbe/>
- Fensel, D., Facca, F. M., Simperl, E., & Toma, I. (2011). *Semantic Web Services*. Berlin, Germany: Springer.
- Fernández, M. J., & Campos, A. (2008). Case Study: CRUZAR — An application of semantic matchmaking for eTourism in the city of Zaragoza. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/Zaragoza-2/>
- Ferreira, R. (2012). *Tech Days: empresas, inovação, tecnologia*. Porto, Portugal: CreateSpace.
- Frigg, R., & Hartmann, S. (2012). Models in Science. *The Stanford Encyclopedia of Philosophy (Fall 2012 Edition)*. Obtido 20 de Fevereiro de 2015, de <http://plato.stanford.edu/entries/models-science/>

- Fuchs, C., Hofkirchner, W., Schafranek, M., Raffl, C., Sandoval, M., & Bichler, R. (2010). Theoretical Foundations of the Web: Cognition, Communication, and Co-Operation. Towards an Understanding of Web 1.0, 2.0, 3.0. *Future Internet*, 2(1), 41–59. doi:10.3390/fi2010041
- Giannakouris, K., & Smihily, M. (2010). *ICT usage in enterprises 2010*. Obtido de <http://ec.europa.eu/eurostat/product?lang=en&mode=view&code=KS-QA-10-049>
- Giannakouris, K., & Smihily, M. (2011). *ICT usage in enterprises 2011*. Obtido de [http://epp.eurostat.ec.europa.eu/portal/page/portal/information\\_society/data/main\\_tables](http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/data/main_tables)
- Google. (2015a). Corporate Contacts. Obtido 30 de Março de 2015, de <https://developers.google.com/structured-data/customize/contact-points>
- Google. (2015b). Customizing Your Knowledge Graph. Obtido 30 de Março de 2015, de <https://developers.google.com/structured-data/customize/overview>
- Google. (2015c). Google Maps Embed API. Obtido 30 de Março de 2015, de <https://developers.google.com/maps/documentation/embed/>
- Google. (2015d). Search ranking. Obtido 30 de Março de 2015, de <https://support.google.com/webmasters/answer/34432?hl=en>
- Hall, W., & O'Hara, K. (2008). Web Science. *Association of Learning Technologies Newsletter*. Obtido de <http://eprints.ecs.soton.ac.uk/15682/>
- Heath, T. (2010). A taskonomy for the Semantic Web. *Semantic Web Journal*, 1, 75–81. doi:10.3233/SW-2010-0003
- Heath, T., & Bizer, C. (2011). Linked Data: Evolving the Web into a Global Data Space. Em J. Hendler & F. Van Harmelen (Eds.), *Proceedings of the 17th international conference on World Wide Web* (Vol. 1, pp. 1–136). Beijing, China: Morgan & Claypool. doi:10.1145/1367497.1367760
- Herman, I. (2012). Semantic Web Adoption and Applications. *Networks*, 17(7), 981–982. Obtido de [http://www.jucs.org/jucs\\_17\\_4/web20\\_applications\\_and\\_mechanisms/abstract.html](http://www.jucs.org/jucs_17_4/web20_applications_and_mechanisms/abstract.html)
- Hotho, A., & Stumme, G. (2010). Towards the ubiquitous Web. *Semantic Web*, 1, 117–119. doi:10.3233/SW-2010-0024
- Hyvönen, E. (2010). Preventing ontology interoperability problems instead of solving them. *Semantic Web*, 1, 33–37. doi:10.3233/SW-2010-0014

- IANA. (2015). Introducing IANA. Obtido 30 de Março de 2015, de <http://www.iana.org/about>
- ICANN. (2015a). Program Statistics. Obtido 30 de Março de 2015, de <http://newgtlds.icann.org/en/program-status/statistics>
- ICANN. (2015b). Welcome to ICANN! Obtido 30 de Março de 2015, de <https://www.icann.org/resources/pages/welcome-2012-02-25-en>
- International Telecommunication Union. (2014). *The World in 2014 - ICT Facts and Figures*. International Telecommunication Union. Obtido de <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf>
- International Telecommunications Union. (2010). *The World in 2010*. Obtido de <http://www.itu.int/ITU-D/ict/>
- Internet Live Statistics. (2015). Google Search Statistics. Obtido 30 de Março de 2015, de <http://www.internetlivestats.com/google-search-statistics/>
- Internet Live Stats. (2015). Internet Live Stats. Obtido 30 de Março de 2015, de <http://www.internetlivestats.com/>
- InterPlanetary Networking Special Interest Group. (2015). IPN: A communications system. Obtido 30 de Março de 2015, de <http://ipnsig.org/introducing-delay-disruption-tolerant-networking-dtn/ipn-a-communications-system/>
- ITU. (2008). *Requirements related to technical performance for IMT-Advanced radio interface(s)*. Obtido de <http://www.itu.int/pub/R-REP-M.2134-2008/en>
- Janowicz, K., & Hitzler, P. (2010a). Creating knowledge out of interlinked data. *Semantic Web*, 1, 97–104. doi:10.3233/SW-2010-0019
- Janowicz, K., & Hitzler, P. (2010b). Towards the ubiquitous Web. *Semantic Web*, 1, 117–119. doi:10.3233/SW-2010-0024
- Jung, H., Kim, P., Lee, S., Lee, M.-K., Seo, D. M., Sung, W.-K., ... Kim, S. (2010). Case Study: iLaw—Intelligent Legislation Support System. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/iLaw/>
- Kang, S. S., Yang, J. Y., Lee, S. G. K., Gong, K. H., Myung, J. S., & Park, S. C. (2008). An Enterprise Strategy for Semantic Technology Adoption. Em *The 5th International Conference on Information Technology and Applications*. Cairns. Obtido de [http://ids.snu.ac.kr/w/images/1/1c/an\\_enterprise\\_strategy\\_for\\_semantic\\_technology\\_adoption.pdf](http://ids.snu.ac.kr/w/images/1/1c/an_enterprise_strategy_for_semantic_technology_adoption.pdf)
- Koumenides, C. L., Salvadores, M., Alani, H., & Shadbolt, N. R. (2010). Global Integration of Public Sector Information. Em *Proceedings of the WebSci10: Extending the*

*Frontiers of Society On-Line*. Raleigh, NC, USA. Obtido de <http://journal.webscience.org/303/>

Kuhn, W. (2010). Modeling vs encoding for the Semantic Web. *Semantic Web Journal*, 1, 11–15. doi:10.3233/SW-2010-0012

Laudon, K., & Laudon, J. (2015). *Management Information Systems: Managing the Digital Firm* (14th ed.). London, UK: Pearson.

Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mende, P. N., ... Bizer, C. (2012). DBpedia – A Large-scale , Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web*, 1, 1–5. doi:10.3233/SW-140134

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... Wolff, S. (1997). A Brief History of the Internet. Obtido 30 de Março de 2015, de <http://arxiv.org/html/cs/9901011>

Lopes, F. C., Morais, M. P., & Carvalho, A. J. (2005). *Desenvolvimento de Sistemas de Informação*. Lisboa, Portugal: FCA.

MacManus, R. (2010). How Best Buy is Using The Semantic Web. *ReadWrite*. Obtido 30 de Março de 2015, de [http://readwrite.com/2010/06/30/how\\_best\\_buy\\_is\\_using\\_the\\_semantic\\_web](http://readwrite.com/2010/06/30/how_best_buy_is_using_the_semantic_web)

Marakas, G., & O'Brien, J. (2008). *Management Information Systems*. New York, USA: McGraw-Hill/Irwin.

Mettler, A. (2012). Europe Needs a Real Growth Agenda. <http://innovation.blogactiv.eu/>. Obtido 30 de Março de 2015, de <http://innovation.blogactiv.eu/2012/05/10/europe-needs-a-real-growth-agenda/>

Morris, S. (2012). Case Study: Using the Semantic Web to Enhance the Teaching of Dance. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/LJMU/>

NASA. (2015). Disruption Tolerant Networking for Space Operations (DTN) - 02.25.15. Obtido 30 de Março de 2015, de [http://www.nasa.gov/mission\\_pages/station/research/experiments/730.html](http://www.nasa.gov/mission_pages/station/research/experiments/730.html)

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 33(1), 5–34. doi:10.1016/S0024-6301(99)00115-6

O'Brien, R. (1998). An Overview of the Methodological Approach of Action Research. Obtido 30 de Março de 2015, de <http://www.web.net/~robrien/papers/arfinal.html>

- O'Hara, K., Contractor, N. S., Hall, W., Hendler, J. A., & Shadbolt, N. (2013). Web Science: Understanding the Emergence of Macro-Level Features on the World Wide Web. Em *Foundations and Trends® in Web Science* (Vol. 4, pp. 103–267). doi:10.1561/18000000017
- Oliveira, T., & Martins, M. (2011). Literature Review of Information Technology Adoption Models at Firm Level. *The Electronic Journal Information Systems Evaluation*, 14(1), 110–121. Obtido de <http://www.ejise.com/issue/download.html?idArticle=705>
- OpenCorporates. (2012). *The Closed World of Company Data An examination of how open company data is*. Obtido de <http://blog.opencorporates.com/2012/04/16/how-open-is-company-data-in-open-government-partnership-countries/>
- Planeta Virtual. (2015). Projeto Raio X. Obtido 30 de Março de 2015, de <http://www.planetavirtual.pt/projectoraiox.aspx>
- Pollock, J., & Oracle. (2008). *A Semantic Web Business Case*. W3C. Obtido de <http://www.w3.org/2001/sw/sweo/public/BusinessCase/BusinessCase.pdf>
- Riel, M. (2010). Understanding Action Research. *Center For Collaborative Action Research. Pepperdine University*. Obtido 30 de Março de 2015, de <http://cadres.pepperdine.edu/ccar/define.html>
- Rogers, E. M. (1995). *Diffusion of innovations*. (M. B. Salwen & D. W. Stacks, Eds.) *An integrated approach to communication theory and research* (Vol. 65). Free Press. doi:10.1525/aa.1963.65.5.02a00230
- Salokhe, G., Sini, M., & Keizer, J. (2007). Case Study: The Semantic Web for the Agricultural Domain, *Semantic Navigation of Food, Nutrition and Agriculture Journal*. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/FAO/>
- Schulze, H., & Mochalski, K. (2009). *Internet Study 2008 / 2009. ipoque*. Obtido de [http://www.ipoque.com/resources/internet-studies/internet-study-2008\\_2009](http://www.ipoque.com/resources/internet-studies/internet-study-2008_2009)
- Segaran, T., Evans, C., & Taylor, J. (2009). *Programming the Semantic Web*. (M. Treseler, Ed.). Sebastopol, CA, USA: O'Reilly. Obtido de <http://www.amazon.com/Programming-Semantic-Web-Toby-Segaran/dp/0596153813>
- Sellen, A. J., Murphy, R., & Shaw, K. L. (2002). How Knowledge Workers Use the Web. Em *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 227–234). Minneapolis.
- Shadbolt, N. (2008). Research Roadmap - Web Science Trust. *Web Science Trust*. Obtido 16 de Fevereiro de 2013, de <http://webscience.org/research/roadmap.html>

- Shadbolt, N. (2014). The Semantic Web in an Age of Open Data. Em *13th International Semantic Web Conference, Proceedings*. Trentino, Italy. Obtido de <http://iswc2014.semanticweb.org/proceedings>
- Shadbolt, N., & Berners-Lee, T. (2008). Web Science Emerges. *Scientific American*, 299(4), 76–81. Obtido de <http://eprints.ecs.soton.ac.uk/17143/>
- Shannon, V. (2006). A «more revolutionary» Web. *The New York Times*. Obtido de [http://www.nytimes.com/2006/05/23/technology/23iht-web.html?\\_r=1&scp=2&sq=Tim Berners-Lee&st=cse](http://www.nytimes.com/2006/05/23/technology/23iht-web.html?_r=1&scp=2&sq=Tim Berners-Lee&st=cse)
- Shroff, G. (2013). *The Intelligent Web: Search, Smart Algorithms, and Big Data*. Oxford, UK: Oxford University Press.
- Spivack, N. (2010). The Evolution of the Web: Past, Present, Future | Nova Spivack - Minding the Planet. Obtido 30 de Março de 2015, de <http://www.novaspivack.com/uncategorized/the-evolution-of-the-web-past-present-future>
- Statista. (2011). Number of blogs worldwide from 2006 to 2011 (in millions). Obtido 30 de Março de 2015, de <http://www.statista.com/statistics/278527/number-of-blogs-worldwide/>
- Statista. (2015). Leading social networks worldwide as of March 2015, ranked by number of active users (in millions). Obtido 30 de Março de 2015, de <http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- Syldatke, T., Chen, W., Angele, J., Nierlich, A., & Ullrich, M. (2008). Use Case: How Ontologies and Rules Help to Advance Automobile Development. Obtido 30 de Março de 2015, de <http://www.w3.org/2001/sw/sweo/public/UseCases/Audi/>
- Tapscott, D., & Williams, A. D. (2007). *Wikinomics: A Nova Economia das Multidões Inteligentes*. Lisboa, Portugal: Quidnovi.
- The Wikimedia Foundation. (2015). Wikidata. Obtido 30 de Março de 2015, de [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)
- Tornatzky, L. G., & Fleischer, M. (1990). *The Processes of Technological Innovation*. (A. K. Chakrabarti, Ed.) *Lexington Books* (Vol. 273). Michigan: Lexington Books.
- Turner, R. (2013). The Philosophy of Computer Science. *The Stanford Encyclopedia of Philosophy (Winter 2014 Edition)*. Obtido 20 de Fevereiro de 2015, de <http://plato.stanford.edu/entries/computer-science/#Imp>

- Vafopoulos, M. (2011). Web Science Subject Categorization - Full Listing - Web Science Trust. *The Web Science Trust*. Obtido 12 de Fevereiro de 2015, de <http://oldsite.webscience.org/2010/wssc.html>
- Verizon. (2015). *State of the Market The Internet of Things 2015*. Obtido de <http://www.verizonenterprise.com/state-of-the-market-internet-of-things/>
- W3C. (2004). Resource Description Framework (RDF): Concepts and Abstract Syntax. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>
- W3C. (2008a). Cool URIs for the Semantic Web. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/cooluris/>
- W3C. (2008b). Extensible Markup Language (XML) 1.0 (Fifth Edition). Obtido 30 de Março de 2015, de <http://www.w3.org/TR/2008/REC-xml-20081126/>
- W3C. (2008c). SPARQL Query Language for RDF. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/rdf-sparql-query/>
- W3C. (2010). Description of W3C Technology Stack Illustration. Obtido 30 de Março de 2015, de <http://www.w3.org/Consortium/techstack-desc.html>
- W3C. (2011). Notation3 (N3): A readable RDF syntax. Obtido 30 de Março de 2015, de <http://www.w3.org/TeamSubmission/n3/>
- W3C. (2012). OWL 2 Web Ontology Language Document Overview. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/owl2-overview/>
- W3C. (2013). HTML Microdata. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/microdata/>
- W3C. (2014a). HTML5 A vocabulary and associated APIs for HTML and XHTML. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/html5/introduction.html>
- W3C. (2014b). JSON-LD 1.0. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/jsonld/>
- W3C. (2014c). RDF 1.1 N-Triples. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/n-triples/>
- W3C. (2014d). RDF 1.1 XML Syntax. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/rdf-syntax-grammar/>
- W3C. (2015a). CSS Basic User Interface Module Level 3. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/css3-ui/>

- W3C. (2015b). RDF 1.1 Turtle. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/turtle/>
- W3C. (2015c). RDFa Lite 1.1 - Second Edition. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/rdfa-lite/>
- W3C. (2015d). RDFa Primer. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/xhtml-rdfa-primer/>
- W3C. (2015e). Semantic Web. Obtido 30 de Março de 2015, de <http://www.w3.org/standards/semanticweb>
- W3C. (2015f). Web of Things at W3C. Obtido 30 de Março de 2015, de <http://www.w3.org/WoT/>
- W3C, & IETF. (2001). URIs, URLs, and URNs: Clarifications and Recommendations 1.0. Obtido 30 de Março de 2015, de <http://www.w3.org/TR/uri-clarification>
- Wikipedia. (2015). Web 2.0. Obtido 30 de Março de 2015, de [http://en.wikipedia.org/wiki/Web\\_2.0](http://en.wikipedia.org/wiki/Web_2.0)
- Zaino, J. (2014). Schema.Org: The Fire's Been Lit. Obtido 30 de Março de 2015, de <http://www.dataversity.net/schema-org-fires-lit/>



# Anexos

Anexo 1	Investigadores
Anexo 2	Web Science Subject Categorization (PT)
Anexo 3	Protocolo
Anexo 4	Lista de Empresas
Anexo 5	Ferramenta de Avaliação
Anexo 6	Recomendações
Anexo 7	Webmaster: o papel e o valor de um especialista na Presença Web de uma Empresa
Anexo 8	Aplicações e Atividades



## Anexo 1: Investigadores











## Anexo 2: Web Science Subject Categorization (PT)











## Anexo 3: Protocolo







## Anexo 4: Empresas







## Anexo 5: Ferramenta de Avaliação







## Anexo 6: Recomendações



# Acessibilidade de Conteúdos

## Objetivo

Disponibilizar conteúdos Web compatíveis com as orientações de acessibilidade.

## Resumo

A Web é uma rede de utilização livre suportada em tecnologias abertas. No entanto, as recomendações de utilização e em particular as de acessibilidade dos conteúdos e dos dados existem e devem ser respeitadas e adotadas.

## Desenvolvimento

Foi na Suíça, no final de 1990, que o inglês Tim Berners-Lee criou as primeiras aplicações precursoras dos atuais *browsers* e servidores Web. A comunidade científica adotou rapidamente a invenção e foi seguida de perto por muitos. Mas foi talvez com a importante declaração do CERN, em 30 de Abril de 1993, que a Web começou a ganhar o protagonismo que mantém até hoje: enquanto proprietário da invenção, o CERN considerou que a Tecnologia Web deveria ser livre e gratuita para todos. A gestão da atribuição de endereços Internet é da responsabilidade da *Internet Corporation for Assigned Names and Numbers* (ICANN). A gestão das tecnologias Web é da responsabilidade do *World Wide Web Consortium* (W3C).

Os fundamentos básicos da *World Wide Web* referem-se à teia de documentos e respetivas hiperligações, construídos em HTML (*Hyper Text Markup Language*) e publicados em servidores. Esta linguagem é assim a base de toda a Web. Em 2014, o W3C aprovou a versão 5 do HTML, apesar da mesma já estar em plena utilização há muito tempo, pois fabricantes de *browsers* e desenvolvedores de Web Sites já a tinham adotado sem hesitação.

Mas para além da linguagem propriamente dita, a forma como a mesma é utilizada pode conduzir a resultados distintos. Assim, o W3C desenvolveu um conjunto de

orientações designado por *Web Content Accessibility Guidelines (WCAG)*, primariamente destinadas a tornar o conteúdo da Web acessível a indivíduos com necessidades especiais. No entanto, os princípios adotados tornaram evidente que estas orientações são interessantes para todos os indivíduos, apenas um com mais intensidade que outros. Esses princípios são:

- percepção: tornar o conteúdo perceptível fornecendo alternativas e complementaridades, como texto alternativo a imagens ou possibilidade de utilizar leitora com voz sintetizada
- operacionalidade: facilitar a utilização, proporcionando tempo suficiente, continuidade no conteúdo e navegação intuitiva
- compreensão: tornar o conteúdo legível e previsível, ajudando os utilizadores a prevenir e corrigir erros
- robustez: maximizar a compatibilidade com as ferramentas atuais e futuras

A armadilha da aparência pode ofuscar esta realidade. A preocupação estética excessiva, o desconhecimento da tecnologia fundamental da Web, o desprezo pela diversidade de plataformas de acesso à Web, o excesso de vaidade dos autores ou proprietários e, de uma forma geral, a exploração irresponsável do déficit de conhecimento da Web como meio de comunicação fundamental nos tempos atuais, são alguns dos fatores que podem numa análise superficial relevar para segundo plano a acessibilidade.

Com a introdução do RDF (*Resource Description Framework*), uma estrutura que permite que os dados possam ser identificados, publicados, reutilizados e interligados com recurso aos mesmos endereços que se utilizam para identificar a localização de um documento, o W3C proporcionou a base da concretização da Web Semântica ou Web dos dados. Mais uma vez, a acessibilidade dos dados pressupõe que os mesmos possam ser interpretados de forma mais correta e completa, quer por indivíduos, quer por máquinas (i.e. software). Assim, os níveis de acessibilidade dos dados apresentados de seguida ganham especial importância:

- publicado: apenas disponibilizando dados na Web
- formatado: usando formatos proprietários

- aberto: usando formatos abertos
- semântico: através de standards semânticos
- ligado: com dados hiperligados

Em conclusão, a maior acessibilidade de um Web Site deverá proporcionar maior compreensão e melhor utilização do seu conteúdo, contribuindo de forma decisiva para os objetivos a que se propõe.

#### Ações recomendadas

- Avaliar a acessibilidade do(s) Web Site(s) da empresa
  - O Web Site da empresa satisfaz os requisitos do nível A das WCAG?
  - Os principais objetivos do Web Site da empresa são facilmente satisfeitos pelos utilizadores? - entre outros:
    - Os contactos da empresa estão visíveis ou de fácil acesso?
    - A estrutura do Web Site é clara para o visitante?
    - O Web Site apresenta-se completo e funcional na plataforma ou dispositivo em que é acedido?
- Gerir metadados
  - Os metadados do Web Site são utilizados de acordo com as orientações para Web Masters dos principais motores de pesquisa
  - Os conteúdos apresentados são classificados de acordo com os standards, onde possível, e maximizando a sua acessibilidade, sempre.



# Adoção de Ontologias

## Objetivo

Adotar Ontologias relevantes para o sector de atuação da empresa.

## Resumo

As particularidades de cada sector de atividade dão origem a conhecimento específico que pode ser relevante para o nível de competitividade das empresas. As ontologias podem desempenhar um papel fundamental na forma de sistematizar esse conhecimento.

## Desenvolvimento

A Web Semântica, ou Web dos Dados, coloca no horizonte a possibilidade de construir um repositório global de conhecimento. Os desafios e as barreiras a essa construção são diversos, mas como em todas as construções, os alicerces fornecem a solidez necessária. Os alicerces fundamentais da Web Semântica são as Ontologias, vocabulários são bem definidos, estruturados e interrelacionados que definem e esclarecem o significado das coisas.

Entre os muitos desafios que se apresentam à construção desse repositório global, a possibilidade de fundir todo o conhecimento segundo um único conjunto de definições parece utópica. No entanto, se esse conhecimento for restrito a um conjunto de áreas de conhecimento bem definidas, a utopia dá lugar à ação e com alguma facilidade emergem ontologias específicas que fornecem de imediato a possibilidade de mapear algum conhecimento, numa abordagem "*from bottom to top*" inspirada na experiência colaborativa que a Web 2.0 promoveu.

Parece assim interessante que qualquer interveniente num dado sector económico procure as definições que o mercado constrói para as adotar em função dos seus objetivos de negócio. Alguns exemplos dessa adoção incluem:

- desde logo a própria definição de sectores de atividade económica é orientada pela definição das Nações Unidas e adotada em Portugal na lista de Códigos de Atividades Económicas (CAE)
- a OCDE criou uma definição padrão para as transações comerciais e respetiva fiscalidade que foi adotada em muitos países, em Portugal sob o acrónimo SAF-T PT

Outros exemplos adotam já a orientação semântica e disponibilizam ontologias específicas, como por exemplo:

- o *Automotive Ontology Working Group* para o sector automóvel
- o *Financial Industry Business Ontology* para o sector financeiro

Nalguns casos, essas ontologias podem ainda ser uma ação pioneira da própria empresa, que promova a sistematização de definições internas e contribua para a interoperabilidade dos negócios e dos sistemas de informação que os suportam.

#### Ações recomendadas

- Identificar as áreas temáticas onde a organização se enquadra e quais as oportunidades de definição de significados relevantes para o seu negócio nuclear e áreas de suporte
- Identificar ontologias sectoriais relacionadas com a atividade da organização

# Colaboração Tecnológica

## Objetivo

Aderir a esforços colaborativos de âmbito tecnológico.

## Resumo

O investimento em tecnologia de uma organização não tem de ser realizado em exclusivo pela mesma. Os vários esforços colaborativos disponíveis a todas as organizações permitem evoluir mais rapidamente, com menor custo e melhores resultados.

## Desenvolvimento

O investimento em tecnologia é considerado um fator de promoção da inovação e da competitividade das organizações. As oportunidades de investimento são diversas e em todas é possível encontrar formas de colaborar com outras organizações e indivíduos. Estas oportunidades não são mais que momentos específicos em que determinadas circunstâncias se reúnem e formam um quadro favorável a um acontecimento em concreto.

### 1. Eliminar tarefas sem valor acrescentado

As empresas tendem a sofrer de burocracia aguda, inventando tarefas que visam satisfazer necessidades intermédias, muitas vezes desligadas do valor acrescentado que a empresa como um todo produz. Eliminar essas tarefas não afeta o valor produzido pela empresa e elimina todo o custo a elas associado.

### 2. Reduzir tempo ou competências de execução das tarefas

Obviamente, reduzir o tempo de execução de uma tarefa reduz o seu custo. Mas por outro lado, menos óbvio, a redução das competências necessárias à sua execução também reduzirá os custos, através da combinação do esforço de vários intervenientes que no total desempenhem a tarefa com o mesmo resultado, mas à custa da intervenção racional de recursos com custos horários distintos.

### 3. Eliminar erros

Grande parte da insatisfação dos clientes e dos colaboradores resulta de erros que poderiam ser eliminados. Se a estas consequências dificilmente quantificáveis juntarmos o custo direto de reparação desses erros e a perda de capacidade produtiva daí resultante, podemos encontrar frequentemente oportunidades de valor notável.

### 4. Recolher todos os dados

Aproveitar todos os momentos para recolher dados sobre o processo de valor acrescentado da empresa é um dos grandes objetivos do desenvolvimento de uma relação contínua com os clientes. A possibilidade de combinar esses dados e os organizar de forma a produzir informação relevante é uma das maiores fontes de competitividade das empresas.

### 5. Antecipar ocorrências

A capacidade de antecipar determinadas ocorrências pode evitar situações desagradáveis ou maximizar o efeito de momentos positivos. A extrapolação de informação de negócio é uma prática comum nesse sentido.

### 6. Produzir simulações

Em circunstâncias onde a informação disponível é insuficiente ou mesmo quando se pretende avaliar tendências futuras, o recurso à simulação pode ajudar a identificar elementos de informação adicional.

### 7. Promover a autonomia

A autonomia dos indivíduos pode proporcionar fatores de sucesso do negócio e só é possível se possuírem em cada momento a informação de negócio relevante para esse efeito.

### 8. Facilitar a colaboração

Se o trabalho de uma equipa é mais eficaz que o da soma dos seus elementos isolados, então o maior suporte à colaboração eficaz dos indivíduos, presencial ou remota, síncrona ou assíncrona, será um elemento promotor do seu desempenho.

A colaboração pode ser incentivada no interior da organização, começando no interior das equipas ou organizações e nos seus interfaces, e continuar até ao exterior, abrangendo o interface e o interior de outras organizações. Em qualquer dos casos, as organizações têm pleno controlo sobre o nível de colaboração que pretendem atingir e poderão apenas estar limitadas por receios de que essa colaboração não seja justa, proporcional ou recíproca.

#### Ações recomendadas

- Identificar e formalizar os grupos na organização onde a colaboração tecnológica pode ser desenvolvida
- Identificar e formalizar os interfaces da organização com o exterior onde a colaboração tecnológica pode ser desenvolvida



# Design Adaptativo

## Objetivo

Proporcionar uma experiência de utilizador compatível com os diferentes dispositivos e resoluções.

## Resumo

A evolução das tecnologias de écrans planos e de comunicações móveis criaram um cenário de utilização frequente da Web com recurso a múltiplos dispositivos, com vantagens e condicionantes. O design adaptativo tem um papel fundamental na disponibilização de dados e informação para proporcionar uma experiência de utilizador adequada.

## Desenvolvimento

A geração Web chegou e com ela é necessário alterar hábitos enraizados. Um deles, com origem na indústria gráfica, é o de correlacionar a dimensão física de uma imagem com o número de píxel. É verdade que 300 ppp (pontos por polegada, dpi) são suficientes para proporcionar uma boa experiência visual, mas isso depende de serem disponibilizados num relógio digital observado a 30 cm ou num écran LED de 42" observado a 2 metros de distância. Outro desses hábitos, com origem nas origens do computador pessoal e na televisão tradicional, é o de otimizar um Web Site para uma resolução/dimensão fixa. A expressão "otimizado para Internet Explorer a 640\*400px" e as suas variantes eram parcialmente aceitáveis num momento onde a divergência de plataformas e dispositivos ainda não era um facto. Atualmente, estes são os factos:

- nenhuma resolução representa 30% do mercado (o Top 3 é constituído por 1366\*768 com 27%, 1920\*1080 com 10% e 1024\*768 com 9%)
- o Windows continua a dominar enquanto sistema operativo no conjunto de todas as plataformas, com 56%, logo seguido de Android com 18% e iOS+MacOSx com 18%

- os *Desktops* continuam a dominar com 64%, logo seguidos dos *Smartphones* com 30% e dos *Tablets* com 7%
- o Chrome é o *browser* mais utilizado, com 40%, seguido do Internet Explorer com 14% e do Firefox com 12%

(dados de Outubro de 2014 de StatCounter.org, disponíveis na forma de gráfico e para descarga em formato .csv, i.e. nível 3 de Acessibilidade)

Verifica-se que a divergência de plataformas ainda não é total, mas já é suficiente e segue uma tendência que permite concluir que nenhum padrão pode ser considerado, ou por outras palavras, a única expressão aconselhável a utilizar é "otimizado para todas as plataformas".

Uma das opções seguida por muitas empresas é a de criar um Web Site Mobile em alternativa ao Web Site "normal". Esta possibilidade levanta dois problemas fundamentais:

- é necessário acomodar todas as possíveis variantes num dos dois grupos (por exemplo qual a versão mais adequada para um Tablet com 10"?)
- desenvolver e manter dois Web Sites distintos pode ter custos muito próximos do dobro de apenas um Web Site

E assim surgem o *Adaptive Design* (Design Adaptativo) ou o *Responsive Design* (Design Responsivo), abordagens destinadas a proporcionar uma experiência de visualização ótima em qualquer plataforma: leitura fácil, navegação compatível, redimensionamento e deslocamento mínimos, numa gama alargada de dispositivos. Esta é a resposta da indústria de desenvolvimento de Web Sites ao desafio colocado pelos fabricantes de hardware.

Em termos técnicos, ambos se baseiam na deteção da plataforma do utilizador, sendo que o *Adaptive Design* privilegia o fornecimento de versões das páginas com componentes orientados para a experiência do utilizador (p. ex. um menu que se transforma numa caixa fechada na versão mobile) enquanto o *Responsive Design* utiliza conteúdo flexível para tirar melhor partido dos diferentes ecrãs (p. ex. páginas com artigos dispostos em duas colunas que se organizam em uma coluna apenas na versão mobile). Qualquer das soluções permite que o mesmo Web Site se adapte ao dispositivo

em que está a ser visualizado, proporcionando maior flexibilidade para quem desenvolve e maior acessibilidade para quem utiliza. As tecnologias que possibilitam esta capacidade são o HTML 5 e a CSS 3, standards W3C, logo com a máxima acessibilidade.

Com esta abordagem é possível:

- desenvolver um único Web Site para todas as plataformas, evitando a necessidade de criar versões específicas, por exemplo para Mobile, reduzindo os custos de desenvolvimento
- estruturar os conteúdos de forma a que sejam mais acessíveis a todos os públicos, aumentando a eficácia de um Web Site

No entanto, como consequência, também é necessário:

- implementar uma tecnologia com resultados dinâmicos, mais difíceis de prever
- definir uma estrutura de conteúdos mais robusta e coerente

Ações recomendadas

- Avaliar o design adaptativo do(s) Web Site(s) da empresa
  - Qual é o grau de adequação do(s) Site(s) às diferentes plataformas?
  - Qual é o investimento necessário à manutenção do(s) Web Site(s), no presente e no futuro?
- Avaliar a estrutura de dados e conteúdos disponibilizada pela empresa
  - As estruturas de dados relevantes estão identificadas e são disponibilizadas?
  - Os conteúdos são apresentados de forma estruturada e acessível?



# Disponibilização de Dados

## Objetivo

Disponibilizar dados estruturados e metadados que possam ser automaticamente interpretados.

## Resumo

A forma como os dados são disponibilizados tem um impacto direto no potencial da sua reutilização, nomeadamente tornando possível uma interpretação sistematizada que conduza à sua reutilização sucessiva.

## Desenvolvimento

A Web é um meio de comunicação global onde a restrição de acesso é muito próxima de nula. De uma forma geral, pode dizer-se que tudo o que for publicado na Web é de facto publicado, ou seja, tornado público de forma permanente e irreversível. Este simples facto pode até ser relativizado, mas a simples constatação que a visualização de uma página Web que esteja disponível num dado momento a um dado utilizador lhe permite captar o conteúdo do écran e partilhá-lo numa rede social, torna claro que a Web é de facto pública.

Tratando-se entretanto de dados, informação ou conteúdos que uma empresa pretende tornar públicos, sejam eles mais objetivos, como por exemplo uma ficha técnica de um produto, ou menos objetivos, como por exemplo o título de uma campanha de publicidade, a Web é um meio de transmissão global, eficaz, rápido e barato.

A eficácia da disponibilização de dados na Web pode ser aumentada de várias formas. Para simplificar a compreensão do tema, apresentam-se de forma agrupada e sintética, ordenada pela frequência com que habitualmente são abordados:

- disponibilização de metadados descritivos, que caracterizam as páginas Web e o seu conteúdo, direcionados aos motores de pesquisa, uma parte

substancial do tema *Search Engine Optimization* (SEO, otimização para motores e pesquisa) - (como por exemplo a utilização de uma *meta tag* de HTML para indicar o tema de uma página, facilitar a sua indexação aos motores de pesquisa e aumentar o destaque com que surge a quem a pretende encontrar)

- disponibilização de dados estruturados, que possibilitam a sua reutilização, com mais ou menos intervenção humana, satisfazendo os objetivos subjacentes à sua publicação (como por exemplo disponibilizar uma gama de produtos na forma de uma tabela com os dados técnicos relevantes para quem tenha interesse potencial em utilizá-los)
- disponibilização de metadados estruturais, que caracterizam os dados disponibilizados de forma a poderem ser interpretados e reutilizados de forma automática (como por exemplo a caracterização semântica com recurso a RDFa dos contactos de uma empresa, como a sua identidade, localização e telefone)

Em resumo, os objetivos definidos que levam à publicação de determinados dados podem ser mais eficazmente satisfeitos se for considerada a acessibilidade com que são disponibilizados.

#### Ações recomendadas

- Avaliar a orientação dos metadados do(s) Web Site(s) da empresa em termos de SEO
- Avaliar em que medida a disponibilização (controlada) ou publicação (livre) de dados estruturados contribui para os objetivos definidos para o(s) Web Site(s) da empresa
- Implementar a disponibilização de metadados estruturais no(s) Web Site(s) da empresa

# Estratégia de e-business

## Objetivo

Definir e atualizar uma estratégia de e-business para a empresa.

## Resumo

As Tecnologias de Informação e a Internet são responsáveis por uma transformação radical dos negócios nas últimas décadas. Uma estratégia de e-business ou negócio eletrónico identifica como a empresa pretende tirar partido dessa transformação.

## Desenvolvimento

Uma empresa deve considerar o papel preponderante das Tecnologias de Informação na Economia atual. As Tecnologias de Informação suportam diversas atividades que promovem a eficiência e a inovação nas Organizações. Mais ainda, a utilização intensiva da Internet é visível em todos os aspetos da sociedade. Consequentemente, a interação digital nos negócios assume um peso relevante em todos os sectores. Para além do aumento da eficiência no interior das organizações, a extensão dos processos ao seu exterior, envolvendo os diversos intervenientes (Clientes, Fornecedores, Parceiros,...), possibilita novos cenários de otimização dos processos e ainda a criação de novos modelos de negócio.

O e-business é o ambiente global de negócio fortemente suportado em Tecnologias de Informação, onde os processos de valor acrescentado ultrapassam as fronteiras da empresa e se estendem a vários parceiros, possibilitando a criação de modelos de negócio, produtos e serviços inovadores, e explorando novas formas de gestão e geração de lucro. e-business é um termo abrangente, por vezes também referido como Negócio Eletrónico, Economia Digital ou Nova Economia, sem que se encontrem diferenças substanciais entre eles. e-business é um termo abrangente que inclui os mais frequentemente referidos e-commerce, ou comércio eletrónico, que se refere apenas às transações comerciais, e-government, que se refere aos aspetos de relacionamento com

o Estado, bem como os restantes processos inerentes a um negócio, quer no seu exterior, quer no seu interior.

Uma Estratégia de e-business identifica as opções de uma Organização na utilização das Tecnologias de Informação em geral e da Internet em particular para desenvolvimento do seu negócio, fornecendo linhas de orientação para as funções operacionais. Podemos considerar dois cenários limite que ilustram, a título de exemplo, como uma Estratégia de e-business é aplicável a qualquer empresa:

- uma grande empresa seguradora disponibiliza uma plataforma acessível na Web onde todos os seus parceiros podem identificar os processos abertos, adicionar informação sobre os mesmos, indicar quando estão terminados, conhecer os seus resultados, relacioná-los com os seus processos internos e integrar com a sua contabilidade os valores de compensação a receber
- uma pequena empresa de restauração utiliza uma dezena de Web Sites para gerir o seu negócio, incluindo a disponibilização de campanhas e promoções, a gestão de anúncios em outdoors, a gestão de comentários e reclamações, a encomenda de produtos de limpeza e escritório e o envio de newsletters aos seus clientes habituais

Uma estratégia pode estar bem definida e divulgada numa organização, bem como pode estar apenas dispersa sob várias formas ou mesmo implícita na sua atuação. A sua clareza e disponibilidade permitem que todos os colaboradores e parceiros possam identificar oportunidades de desenvolvimento.

#### Ações recomendadas

- Identificar a estratégia de e-business da Empresa e avaliar o grau de conhecimento de toda a organização sobre a mesma
- Identificar as grandes opções da Empresa para o futuro e integrá-las com a evolução da estratégia definida.

# Importação de Dados

## Objetivo

Adotar fontes de dados externas relevantes para o negócio da empresa.

## Resumo

A utilização da informação disponível e relevante para os negócios é uma necessidade das empresas que se pretendem manter competitivas. Essa informação é progressivamente mais vasta e encontra-se muitas vezes fora das fronteiras da empresa, em repositórios públicos ou associativos, pelo que a sua adoção exige uma atitude ativa.

## Desenvolvimento

A competitividade das empresas é influenciada por diversos fatores, entre os quais a utilização da Informação como um ativo essencial para a gestão. Por um lado a atividade regular de qualquer empresa gera cada vez mais informação, proveniente dos diversos registos, dando origem a uma avalanche de dados a necessitar análise. Por outro lado, o ambiente onde essa atividade se desenrola é também ele uma fonte de informação crescente, seja por razões regulatórias, de que é exemplo a recente obrigatoriedade de todas as empresas comunicarem todas as suas linhas de faturação ao Estado, seja por razões de mercado, de que é exemplo a crescente utilização das redes sociais.

A disponibilidade de informação relevante proveniente do exterior é um facto incontornável, conhecido de qualquer utilizador experiente da Web. Eventualmente, poderão ser as técnicas de pesquisa e as ferramentas de análise que podem ser escassas e comprometer os resultados. Para isso, as organizações deverão ter um papel ativo na seleção de fontes de informação adequadas e na definição de ferramentas para a sua utilização. A ilusão que as "competências na ótica do utilizador" resolvem todos os problemas de uma organização tem provocado enormes problemas de ineficiência e a expectativa que cada utilizador consegue ultrapassar as suas limitações sairá ainda mais

frustrada. Se a formação dos utilizadores é importante, a orientação das tarefas desempenhadas em função da criação de valor e da satisfação dos resultados é essencial.

Com a expansão da abordagem Open Data, onde informação fundamental sem qualquer restrição de direitos, como dados sócio demográficos, informação geográfica, entre muitas outras, são disponibilizadas de forma livre e estruturada, as oportunidades de reutilização da informação nos negócios multiplicam-se e possibilitam novas formas de inovação. Esta disponibilidade não reduziu o interesse por fontes de dados restritas ou comerciais.

#### Ações recomendadas

- Identificar e utilizar fontes de dados externas na atividade regular da empresa
  - Existem necessidades de informação regulares na empresa que possam ser satisfeitas pelo exterior?
  - Quais os dados que a empresa pode utilizar regularmente para apoiar o seu negócio?

# Informação Estruturada da Empresa

## Objetivo

Identificar e classificar a empresa e os seus produtos e serviços com ontologias.

## Resumo

A utilização de ontologias para classificação da informação de uma empresa e dos seus produtos facilita a sua interpretação por sistemas externos, com particular destaque para os indexadores da Web que alimentam os motores de pesquisa.

## Desenvolvimento

Tal como as línguas e os vocabulários em que se baseiam permitem a comunicação inteligível entre indivíduos, também a Informação Empresarial depende de vocabulários comuns, definidos inequivocamente. O que é um Cliente? As respostas variam de empresa para empresa. Nalgumas empresas, Cliente é aquele que comprou pelo menos uma vez ou que comprou nos últimos doze meses. Noutras empresas Cliente é aquele que recebeu pelo menos uma Proposta, independentemente de ter comprado. Noutras ainda, Cliente é qualquer empresa ou indivíduo que reúna as características para poder vir a comprar produtos da empresa.

Quando os vocabulários são bem definidos, estruturados e interrelacionados, apelidam-se de Ontologias. A palavra ontologia tem origem na Filosofia e significa literalmente "Ciência do Ser" mas no contexto da Ciência dos Computadores adquire o sinónimo de significado. As ontologias representam um papel fundamental na representação do conhecimento e em particular na construção da Web Semântica ou Web dos Dados.

A estruturação da informação empresarial essencial, como são as características da própria empresa e as dos seus produtos são uma condição essencial para a sua organização interna mas também uma fonte de promoção num mundo fortemente conectado.

A iniciativa schema.org dos principais motores de pesquisa, nomeadamente Google, Bing, Yahoo e Yandex, destina-se a implementar uma ontologia básica para utilização generalizada em Web Sites. Esta ontologia pretende ser simultaneamente genérica e generalizada, com o objetivo de ser utilizada de forma maciça, o que é simultaneamente coerente com os objetivos dos motores de pesquisa (indexar todas as páginas Web) e das empresas (maximizar a divulgação dos seus produtos e serviços).

#### Ações recomendadas

- Avaliar o schema.org do(s) Web Site(s) da empresa
  - O Web Site da empresa possui metadados baseados em schema.org?
  - Esses metadados cumprem os requisitos e podem ser reutilizados de forma consistente?
- Avaliar a utilização de uma ontologia compatível com schema.org nos Sistemas de Informação da empresa
  - Os metadados utilizados no Web Site podem ser integrados com outros Sistemas de Informação?
  - Os metadados utilizados no Web Site podem ser reutilizados em diferentes objetivos?

# Integração de Sistemas

## Objetivo

Integrar fontes de dados e sistemas de informação com o objetivo de disponibilizar informação de negócio.

## Resumo

A integração de dados é uma das prioridades em muitas empresas, com o objetivo de reduzir a duplicação de tarefas e disponibilizar informação relevante a toda a organização. Com a introdução da Internet não só as redes alargadas ganharam protagonismo sobre as redes locais como também possibilitaram novos cenários de aplicação.

## Desenvolvimento

A proliferação das ferramentas de produtividade pessoal, com destaque para as folhas de cálculo, proporcionou ao indivíduo a facilidade em criar soluções de gestão de informação com elevada disponibilidade. No entanto, estas ferramentas muito orientadas para o indivíduo apresentam diversas limitações no suporte à fiabilidade dos dados, à robustez dos processos e à colaboração entre indivíduos.

Os sistemas de informação baseados em aplicações de utilização generalizada na empresa, com particular destaque para os ERP, podem abordar exaustivamente os mais variados temas e processos mas apresentam frequentemente limitações na adaptabilidade ao cenário individual de cada empresa, na flexibilidade de adaptação às alterações de processos e na interoperabilidade com outros sistemas.

A massificação do acesso à Internet e a utilização ubíqua da Web levaram ao aparecimento de imensos cenários onde "o" sistema de informação empresarial é constituído por muitos subsistemas que necessitam partilhar entre si dados e informação com cada vez mais frequência. Consequentemente, as necessidades de integração de dados são crescentes e dependentes da uniformidade das definições dos diferentes

subsistemas. Reciprocamente, a maior uniformidade nas definições promove e facilita a integração dos diferentes subsistemas.

#### Ações recomendadas

- Definir as prioridades de integração de sistemas da empresa em função das necessidades estratégicas do negócio.

# Propriedade na Web

## Objetivo

Identificar o âmbito da propriedade da empresa em matéria de Web e Propriedade Intelectual.

## Resumo

Os Ativos Intangíveis de uma empresa incluem as suas Marcas, os seus Domínios e também as suas Aplicações de Software, com destaque para aquelas das quais é proprietária integral, mas também o seu conhecimento e a forma como este é distribuído na Web.

## Desenvolvimento

A gestão dos ativos de uma empresa assumiu um grau de complexidade adicional com a Web atual. Há poucos anos, uma empresa podia confiar exclusivamente nos seus recursos internos para controlar quase na totalidade a sua exposição ao exterior. Como exemplo, a escolha de aparecer nas Páginas Amarelas com maior destaque que na lista telefónica simples envolvia uma opção clara de investir no meio publicitário, normalmente confiada a um especialista em Marketing ou Design. A primeira geração Web proporcionou inicialmente uma possibilidade semelhante, uma vez que confrontou as empresas com a necessidade de criar uma presença com controlo integral, normalmente na forma de uma apresentação institucional, mais raramente com possibilidade de permitir a interação direta com os visitantes para além de um simples formulário de contacto. Mas rapidamente se assistiu à evolução da Web para um ambiente onde o controlo passou a ser distribuído e colocado na mão de muitos intervenientes. Inicialmente porque as empresas puderam começar a escolher estar presentes em plataformas de terceiros, como por exemplo em diretórios de empresas. Posteriormente porque algumas plataformas assumiram o direito de representar as empresas de alguma forma sem a sua permissão.

Atualmente, a presença Web de uma empresa pode ser dividida pelo menos nos seguintes tipos de presença:

- Web Sites: as diversas aplicações que a empresa disponibiliza ao público em geral ou a um grupo específico, como por exemplo clientes, para as mais diversas funcionalidades, desde colocação de encomendas ou e-learning, com um grau de controlo de elevado a total
- Diretórios de Empresas e Produtos: aglomeradores de informação com fins específicos, que podem incluir simples diretórios informativos e "*marketplaces*", com grau de controlo de médio a elevado (mais alto naqueles com opção de entrada -"*opt-in*"- que naqueles com opção de saída -"*opt-out*"-) ou de baixo a médio nos casos sem opção de entrada ou saída ("*no-opt*")
- Serviços Informativos: com objetivos variados que podem ir desde a localização geográfica à informação empresarial especializada, exibindo informação da empresa a públicos menos ou mais restritos, com um grau de controlo muito variável, de baixo a elevado
- Motores de Pesquisa: serviços de aquisição e análise de informação, normalmente derivada da restante presença Web, com disponibilização de resultados de pesquisa, com um grau de controlo de baixo a médio

Alguns exemplos comuns de componentes da presença Web de uma pequena ou média empresa portuguesa sem grande especificidade incluem:

Componente	Conteúdo	Controlo	Opção
Web Site	informação institucional e generalista	Elevado	opt-in
Extranet	informação do cliente, produtos, encomendas	Elevado	opt-in
Facebook	novidades e eventos	Médio	opt-in
Google My Business (Google+, Google Places)	informação empresarial, novidades e eventos	Médio	opt-in
Google Maps	informação empresarial	Médio	opt-out
LinkedIn	informação empresarial, recrutamento	Médio	opt-in
Páginas Amarelas na Internet	informação empresarial	Médio	opt-out
Portal da Justiça	informação administrativa	Baixo	no-opt

#### Ações recomendadas

- Identificar as diversas componentes da presença Web da organização e classificá-las quanto à sua importância para os resultados pretendidos.



# Responsabilidade do Webmaster

## Objetivo

Identificar e definir responsabilidades do Gestor de Informação ou Webmaster na organização da empresa.

## Resumo

O Webmaster deve assegurar a coordenação global da presença Web de uma Empresa, garantindo que todas as necessidades estão a ser satisfeitas, seja através das suas competências, seja através das competências dos membros da sua equipa interna e dos fornecedores de serviços. As empresas devem identificar quem desempenha essa função.

## Desenvolvimento

A Informação é um dos ativos mais importantes numa organização moderna.

A Gestão de Sistemas de Informação pode ser dividida em quatro fases:

- Planeamento - a escolha das opções mais importantes para a evolução da empresa a cada momento ("o quê")
- Desenvolvimento - a seleção, adaptação ou criação de sistemas adequados ("como")
- Implementação - a aquisição, instalação, configuração e teste dos sistemas ("quando")
- Exploração - a extensão do período de vida de funcionamento em níveis eficientes ("até quando")

O Gestor de Sistemas de Informação será então o principal responsável por todo o ciclo de vida dos Sistemas de Informação da empresa e das Tecnologias de Informação que os suportam, assegurando que o Ativo Informação é um ingrediente para o sucesso da empresa.

A Web desempenha atualmente um papel muito relevante no contexto dos Sistemas de Informação, pois tornou-se numa plataforma de contacto privilegiada entre os colaboradores das empresas e os seus interlocutores. No entanto, se por um lado as Aplicações Web são uma parte dos Sistemas de Informação, o inverso já não é verdade uma vez que muitos sistemas de informação atuais não são "*web-based*", i.e. não foram desenvolvidos segundo uma arquitetura Web.

A evolução da Web não é apenas tecnológica e envolve diversas áreas de conhecimento ligadas à Ciência da Web. A omnipresença da Web na sociedade fez convergir áreas naturalmente mais dispersas, como a Sociologia e o Direito, a áreas mais nucleares, como a Matemática e a Ciência dos Computadores, criando uma exigência elevada aos profissionais Web. Esses profissionais poderão abranger mais ou menos áreas de especialização ou mesmo cobrir todas as áreas numa perspetiva generalista. Essa opção ou necessidade dependerá naturalmente das exigências e expectativas colocadas por organizações de maior ou menor dimensão, com mais ou menos recursos internos, recorrendo ou não à contratação de colaboradores e empresas externas.

As tarefas esperadas para um Webmaster incluem:

- Coordenação
  - Liderança
  - Coordenação de responsabilidades de equipa
  - Integração com a direção e departamentos da empresa
  - Gestão de Projetos
  - Elaboração de estratégia, planos e relatórios de atividades
- Gestão da Infraestrutura
  - Configuração de Servidores
  - Gestão de Alojamentos (housing, hosting, storage,...)
  - Controlo da disponibilidade (availability, uptime,...)
  - Avaliação de desempenho (speed, traffic,...)
  - Segurança
  - Gestão de acessos individuais
- Gestão da Propriedade
  - Registo e renovação de domínios

- Análise de Propriedade Intelectual (p. ex. registo de marcas)
- Design
  - Logotipos
  - Imagem corporativa
  - Imagem de marca
  - Web Design (páginas)
  - Compatibilidade (utilização em diferentes plataformas por diferentes públicos)
  - Acessibilidade (utilização por públicos com limitações)
- Gestão de Conteúdos
  - Gestão da Informação
  - Produção de conteúdos
    - Apresentação Corporativa
    - Descrição de Produtos/Serviços
  - Gestão de inserções em diretórios e portais
- Desenvolvimento de Aplicações
  - Edição de HTML/CSS
  - Edição de Javascript
  - Correção de erros
  - Gestão e programação de Bases de Dados
  - *Search Engine Optimization*
  - Testes
- Gestão de Aplicações
  - Criação e gestão de contas
    - Portais e Extranets
    - correio eletrónico
    - comércio eletrónico
  - Gestão de componentes e aplicações integradas
- Gestão de Desempenho
  - Avaliação de Tráfego
  - Avaliação de Negócio

- Marketing Digital
  - Elaboração e gestão de campanhas publicitárias
  - Gestão de Promoções
  - Gestão de comunidades

Percebe-se assim que a expectativa sobre as funções de um Webmaster é muito diversa e levará naturalmente a que a função seja desempenhada por um conjunto de indivíduos com diversas especialidades, eventualmente liderados por um indivíduo com capacidades mais generalistas nas diversas áreas de conhecimento essenciais à Web da atualidade.

A dimensão de uma equipa multidisciplinar e o investimento necessário para a criar e manter são uma condicionante a muitas empresas em que a dimensão do seu negócio não o justifica. Nesses casos, a opção recai no recurso a uma equipa externa, disponível no tempo parcial necessário para satisfazer os requisitos da empresa. Em alternativa, um recurso interno menos versátil poderá satisfazer alguns requisitos, ficando ao critério da gestão analisar como poderão satisfazer os restantes que se verifiquem ser importantes para a empresa.

Em conclusão, o papel do Webmaster deverá ser o de assegurar a coordenação global da presença Web de uma Empresa, garantindo que todas as necessidades estão a ser satisfeitas, seja através das suas competências, seja através das competências dos membros da sua equipa interna e dos fornecedores de serviços.

#### Ações recomendadas

- Identificar o Webmaster na organização
  - A empresa possui um Webmaster?
  - O Webmaster pode desempenhar a totalidade das funções relevantes para empresa?
  - Existem fornecedores que assegurem as restantes funções necessárias?
- Alinhar a atividade do Webmaster com a estratégia da empresa
  - O Webmaster tem acesso à gestão de topo da empresa?

- A gestão de topo reconhece o papel do Webmaster na garantia de sucesso na sua estratégia?



# Tipos de Web Site

## Objetivo

Definir os tipos de Web Site e a sua relação com a estratégia da empresa.

## Resumo

Os Web Sites podem ser classificados quanto à proveniência dos seus utilizadores e ao tipo de aplicações que os suportam. A escolha sobre os diferentes tipos de Web Site está diretamente ligada à estratégia da empresa. Pretende-se que as escolhas efetuadas ao longo do tempo estejam sempre associadas à estratégia da empresa em cada momento, fazendo com que os Web Sites sejam de facto uma peça que suporte a prossecução dos objetivos empresariais.

## Desenvolvimento

Os Web Sites podem ser classificados quanto à proveniência dos seus utilizadores, nomeadamente se eles são internos ou externos à empresa, dando origem assim a Aplicações distintas:

- Intranet, quando os utilizadores são exclusivamente os colaboradores da empresa, com acesso restrito dependendo de autenticação das suas credenciais
- Extranet, quando os utilizadores são externos à empresa mas acedem numa área reservada a informação condicionada à sua autenticação
- Web Site (genericamente, ou Web Site público), quando os utilizadores anónimos acedem a uma área pública, sem restrição nem autenticação

Assim, a escolha de um tipo de Web Site depende da identificação das funcionalidades que se pretende implementar e, frequentemente, uma empresa implementa funcionalidades em diversos Web Sites.

Adicionalmente, os Web Sites podem ser classificados quanto ao tipo de aplicações que os suportam. Essencialmente, os Web Sites são criados com recurso a:

- Sistemas de Gestão de Conteúdos (CMS - *Content Management Systems*), essencialmente para permitirem a edição de conteúdos associada a "*templates*" de design
- Desenvolvimento à Medida, com resultados que podem ser enquadrados em diversos Tipos de Sistemas de Informação, como ERP (*Enterprise Resource Planning*), CRM (*Customer Relationship Management*) ou, tendo em conta a especificidade em relação ao negócio da empresa, ser considerados Sistemas Periciais

De uma forma geral, todas as empresas podem adotar a primeira opção e utilizar um CMS para suportar a criação e edição do seu Web Site. De uma forma sucinta, o CMS deverá permitir criar uma ou mais páginas que apresentem a empresa, identifiquem os seus produtos e serviços e caracterizem a sua oferta da forma mais completa possível, nomeadamente quanto a contactos, localização, horário de funcionamento, entre outros. Para muitas empresas em diversos sectores, por exemplo restauração e comércio a retalho, uma estratégia de e-business pode limitar-se à implementação dessas funcionalidades e não identificar mais necessidades. Drupal, Joomla, SharePoint e WordPress são exemplos de CMS muito utilizados.

No outro extremo, algumas empresas definem Modelos de Negócio de elevada complexidade com recurso a ferramentas digitais que implementam funcionalidades exclusivas para a empresa, recorrendo obviamente ao Desenvolvimento à Medida. Algumas empresas, como por exemplo uma empresa que implemente uma rede social, são criadas com este objetivo estratégico bem definido à partida.

Entre ambos os limites encontram-se as empresas que evoluem no sentido de ultrapassar as restrições de um CMS genérico, muitas vezes limitado à mera publicação de conteúdos relevantes. A maioria dos CMS disponíveis no mercado disponibilizam extensões (*add-ons*) que permitem implementar funcionalidades que ultrapassam a Gestão de Conteúdos por si só, como por exemplo um Sistema de Comércio Eletrónico. No entanto, mais tarde ou mais cedo, os pressupostos assumidos em qualquer sistema de carácter genérico constituem uma limitação à estratégia da empresa e surge o tema do desenvolvimento à medida. Se esse desenvolvimento é mais caro ou menos rentável no curto ou médio prazo, poderá influenciar a decisão de avançar com um CMS.

Um exemplo seria o de um restaurante que pretenda efetuar reservas através da Web sem necessidade de intervenção dos seus colaboradores, funcionalidade que terá quase obrigatoriamente que incluir desenvolvimento à medida. É simples criar um formulário que permita ao visitante de um Web Site indicar a data, hora e número de pessoas e formular um pedido de reserva. Pressupondo que o restaurante possui um sistema de gestão que identifica as mesas disponíveis e a sua capacidade, seria necessário que o Web Site recebesse a resposta quanto à viabilidade da reserva solicitada. Com desenvolvimento à medida, será possível construir esta funcionalidade com duas ou três páginas e um *Web Service*. Com um CMS é provável que o custo de desenvolvimento seja superior.

Em conclusão, qualquer empresa deve evoluir no seguinte sentido, ainda que saltando uma ou mais fases:

- disponibilizar um Web Site informativo com base num CMS genérico
- implementar funcionalidades no mesmo Web Site ou em Web Sites adicionais que satisfaçam as necessidades identificadas
- desenvolver funcionalidades à medida do Modelo de Negócio da empresa

#### Ações recomendadas

- Avaliar os objetivos do(s) Web Site(s) da empresa
  - A empresa possui um Web Site que a identifique, bem como aos seus Produtos ou Serviços, fornecendo pública e facilmente a informação básica necessária?
  - O Web Site da empresa pode ser atualizado em tempo útil, internamente ou com recurso a um fornecedor?
- Avaliar o alinhamento estratégico dos Web Sites da empresa
  - Foi definida uma estratégia onde os Web Sites da empresa desempenhem um papel relevante?
  - Os Web Sites da empresa satisfazem as necessidades identificadas?
  - O tipo de aplicações que suportam os Web Sites são adequados à evolução prevista?



## Anexo 7: Webmaster - o papel e o valor de um especialista na Presença Web de uma Empresa



















## Anexo 8: Aplicações e Atividades





