



**Universidade da Beira Interior**

Departamento de Informática

**Editor para a Web Semântica Integrando Anotações**

**Semânticas, Ontologias e RDF**

Rui Alexandre da Rocha Gonçalves Pereira

Covilhã

2006





**Universidade da Beira Interior**

Departamento de Informática

**Editor para a Web Semântica Integrando Anotações  
Semânticas, Ontologias e RDF**

Rui Alexandre da Rocha Gonçalves Pereira

Dissertação submetida à Universidade da Beira Interior para obtenção do grau de  
Mestre em Engenharia Informática

Covilhã

2006



Dissertação realizada sob a orientação do  
Professor Doutor Mário Marques Freire,  
Professor Associado do Departamento de Informática da  
Universidade da Beira Interior



## Resumo

Uma das principais consequências resultantes do aparecimento da *World Wide Web*, também designada simplesmente por Web, foi o de popularizar a Internet à escala mundial. Tim Berners-Lee, conhecido como o inventor da Web, idealizou-a como sendo constituída por uma rede informática onde a informação seria definida de forma clara e objectiva e poderia ser facilmente utilizada para melhorar a cooperação entre as pessoas e as aplicações informáticas. Mas, ao contrário do previamente idealizado, a actual Web é caracterizada como sendo um meio de comunicação exclusivamente humano. A interpretação de forma simples e universal da informação existente na Web por parte de aplicações informáticas genéricas ainda não é uma realidade.

A Web Semântica surge assim como uma tentativa de materializar todo o potencial da Web idealizada e apresentada por Tim Berners-Lee. Ela é definida como sendo uma evolução da actual Web, na medida em que lhe adiciona um modelo universal para estruturar e representar o significado da informação. Segundo os investigadores da Web Semântica, a utilização conjunta desse modelo e da Web facilitará a manipulação e a interpretação da informação por parte de aplicações informáticas genéricas. É, no entanto, importante referir que, na Web Semântica, processar e relacionar informação não significa um comportamento inteligente com o mesmo significado que lhe é atribuído pelos investigadores da área da Inteligência Artificial, mas apenas a capacidade de interpretação e manipulação da informação por parte de aplicações informáticas genéricas.

Actualmente, vivemos um período de grande entusiasmo devido ao exponencial aparecimento de ferramentas associadas às normas e tecnologias agregadas à arquitectura da Web Semântica. Contudo, a utilização dessas ferramentas não é acessível a todos os actuais utilizadores da Web, uma vez que, obrigam que estes estejam familiarizados com a arquitectura da Web Semântica.

---

Com o intuito de ultrapassar este importante obstáculo, e dado que o sucesso de uma tecnologia depende inevitavelmente da sua aceitação por parte dos seus utilizadores finais, a presente dissertação contribui com uma exposição elucidativa dos conceitos, normas e tecnologias associadas à Web Semântica, com a categorização da maioria das suas principais ferramentas e através do desenvolvimento de uma nova ferramenta, denominada SWedt, a qual permite que utilizadores pouco familiarizados com a Web Semântica possam desenvolver páginas para a Web Semântica.



## **Abstract**

One of main contributes resulting from the World Wide Web, also named Web, was to get the internet worldly known. Tim Berners-Lee, also called the Web inventor, has thought it as constituted by a computing network where information would be clearly and objectively defined, and it could be easily used to ease cooperation among people, people and computing applications, and among computing applications. Nevertheless, unlike what was previous idealized, the nowadays web is characterized as a means of communication exclusively human. The simple and universal interpretation of the existing information in the Web by the generic computing applications is not yet a reality.

Semantic Web comes, therefore, as an attempt to materialize all the Web potential designed and presented by Tim Berners-Lee. It is defined as resulting from the nowadays Web, as it adds a universal structure in order to understand the meaning of the information. According to the Semantic Web researcher, the joint usage of such structure and of the Web will ease the manipulation and interpretation of the information, by generic computing applications. It is, nonetheless, important to refer that an intelligent behaviour in processing and relating information in the Semantic Web does not have the same meaning as that used by researchers in the Artificial Intelligence area, but solely the capacity of interpretation and manipulation of the information by the generic computing applications.

Currently, we are going though a period of great enthusiasm due to the exponential appearance of tools associated to the norms and technologies of the Semantic Web architecture. However, the use of these tools is not accessible to all the current users for it requires a previous knowledge of the Semantic Web architecture. Aiming at overcoming this crucial obstacle, and once the success of a certain technology depends, inevitably, upon its acceptance by the end-users, this thesis contributes a clear approach of concepts, norms and technologies associated to the

---

Semantic Web, with the classification of the majority of its main tools, and by means of developing a brand new tool, called SWedt, which allows non-familiar users to develop Semantic Web pages.

# Agradecimentos

Em primeiro lugar, quero agradecer ao Professor Doutor Mário Marques Freire, meu orientador, pelo seu constante entusiasmo, apoio e disponibilidade.

Quero agradecer à Universidade da Beira Interior pelas óptimas condições de trabalho disponibilizadas.

Quero agradecer aos professores de mestrado pela disponibilidade e apoio proporcionado durante a realização da parte lectiva do mestrado.

Quero agradecer aos colegas do Departamento de Informática pelo constante apoio e motivação.

Quero agradecer à minha família, em especial aos meus pais, pelo carinho, apoio e compreensão.

Finalmente, quero agradecer aos colegas de mestrado e a todos os meus amigos pelo apoio manifestado.

---

# Prefácio

Na primeira edição de Mestrado em Engenharia Informática da Universidade da Beira Interior (UBI), foi apresentado um leque alargado de temas interessantes para a realização de diversas dissertações. De entre os temas expostos, escolhi o tema **Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF**, proposto pelo Professor Doutor Mário Freire. A razão da minha escolha prendeu-se com a actualidade e extrema importância, quer para os engenheiros informáticos, quer para a sociedade em geral, que a implementação da Web Semântica ostenta.

Apesar de a ideia base da Web Semântica se apresentar de forma relativamente simples, a sua materialização é, pelo contrário, extremamente complexa: transformar a Web de forma a que todo o seu conteúdo, o qual pode e deverá continuar a ser inserido por qualquer pessoa, seja facilmente interpretado por pessoas e por aplicações informáticas genéricas. Assim, a Web Semântica, além de ser constituída por um extenso conjunto de tecnologias associadas, possui também um objectivo de ambição imensurável. Além disso, essas tecnologias abrangem um extenso campo de actuação, de forma a que, na minha opinião, cada uma delas constitui, por si só, matéria suficiente para ser abordada noutras dissertações ou teses. No entanto, tenho como forte convicção que a materialização da Web Semântica e das suas envolventes irá, seguramente, influenciar e modificar a nossa própria forma de viver e de interagir com o mundo. Desta forma, ficarei satisfeito se a presente dissertação contribuir para uma melhor compreensão da estrutura da Web Semântica, bem como, constituir uma base de informação a todos os actuais e futuros investigadores e utilizadores da Web Semântica.

---

A presente dissertação pretende apresentar uma descrição mais aprofundada do conceito da Web Semântica, bem como, desenvolver um editor para a criação de páginas web para a Web Semântica. Assim sendo, foi estruturada de acordo com quatro pontos essenciais, distribuídos segundo a seguinte ordem de trabalho:

- Trabalho de pesquisa relacionado com o estado da arte das principais tecnologias associadas à Web Semântica;
- Estudo de avaliação e categorização das actuais ferramentas para a Web Semântica;
- Definição da arquitectura a utilizar para a implementação de um editor para a Web Semântica;
- Desenvolvimento do editor.

Apesar da constante preocupação em realizar uma descrição, o mais exaustiva possível, das tecnologias associadas à Web Semântica, a dimensão do tema delimitou o nível de profundidade na abordagem de algumas delas mas, essa limitação não constituiu prejuízo na descrição geral e objectiva da Web Semântica apresentada nesta dissertação.

Por fim, quero salientar que o trabalho conducente à elaboração desta dissertação foi complementado através de um conjunto de reuniões, efectuadas com o meu orientador, extremamente úteis para a definição de metodologias, selecção de percursos, exposição de críticas construtivas e esclarecimento de dúvidas.

---

## Notação Utilizada

Neste trabalho são utilizados alguns estilos que têm significados próprios, dos quais se destacam os seguintes:

- *Formato Itálico*: Utilizado em palavras não referentes à Língua Portuguesa para as quais não existe uma tradução adequada ou existe a necessidade de apresentar o termo na sua língua de origem com o intuito de tornar mais explícito o seu significado;
- Fonte de letra do tipo **Sans-serif**: Utilizada para apresentar excertos em linguagem de programação e para realçar algumas frases apresentadas em situações específicas.

## Lista de Abreviaturas

Sempre que é apresentada uma abreviatura (sigla ou acrónimo), uma marca registada ou uma tradução que suscite alguma dúvida, esse facto é complementado através de uma nota de rodapé. No entanto, devido à enorme quantidade de abreviaturas utilizadas na literatura da especialidade e de forma a facilitar uma consulta rápida, são aqui expostas, por ordem alfabética, todas as abreviaturas utilizadas nesta dissertação:

<i>ADODB</i>	: <i>ActiveX Data Object Data Base.</i>
<i>AIML</i>	: <i>Astronomical Instrument Markup Language.</i>
<i>AML</i>	: <i>Astronomical markup Language.</i>
<i>ANSI</i>	: <i>American National Standards Institute.</i>
<i>API</i>	: <i>Application Programming Interface.</i>
<i>ASL</i>	: <i>Apache Software License.</i>
<i>ASP</i>	: <i>Active Server Pages.</i>
<i>BD</i>	: Base de Dados.
<i>BSD</i>	: <i>Berkeley Software Distribution.</i>
<i>BSE</i>	: Encefalopatia Espongiforme Bovina.

---

*CERN* : *European Organization for Nuclear Research* (Laboratório Europeu de Física de Partículas).

*CML* : *Chemical Markup Language*.

*CODIP* : *Components for Ontology Processing*.

*CSS* : *Cascading Style Sheet*.

*CVS* : *Concurrent Versions System*.

*DAML* : *DARPA Agent Markup Language*.

*DAML+OIL* : *DARPA Agent Markup Language + Ontology Inference Layer*.

*DARPA* : *Defense Advanced Research Projects Agency*.

*DCML* : *Dental Charting Markup Language*.

*DIG* : *Description Logic Interface*.

*DISA* : *Defense Information Systems Agency*.

*DFKI* : *German Research Center for Artificial Intelligence*.

*DL* : *Description Logic*.

*DLG* : *Directed Labeled Graphs*.

*DSML* : *Directory Services Markup Language*.

*DSSSL* : *Document Style Semantics and Specification Languages*.

*DTD* : *Document Type Definition*.

*DOM* : *Document Object Mode*.

*EJB* : *Enterprise JavaBeans*.

*ERCIM* : *European Research Consortium for Informatics and Mathematics*.

*FNC* : *Federal Networking Council*.

*FOL* : *First-Order Logic*.

*FZI* : *Forschungszentrum Informatik*.

*GEML* : *Gene Expression Markup Language*.

*GIF* : *Graphics Interchange Format*.

*GNU* : *GNU is Not Unix*.

*GPL* : *General Public License*.



---

<i>GSS</i>	: <i>Graph Stylesheets.</i>
<i>GUI</i>	: <i>Graphical User Interface.</i>
<i>HP</i>	: <i>Hewlett Packard Company.</i>
<i>HTML</i>	: <i>Hyper Text Markup Language.</i>
<i>HTTP</i>	: <i>Hypertext Transfer Protocol.</i>
<i>I&amp;D</i>	: <i>Investigação e Desenvolvimento.</i>
<i>IA</i>	: <i>Inteligência Artificial.</i>
<i>IE</i>	: <i>Internet Explorer.</i>
<i>IDE</i>	: <i>Integrated Development Enviroment.</i>
<i>INRIA</i>	: <i>Institut National de Recherche en Informatique et Automatique.</i>
<i>INTAP</i>	: <i>Interoperability Technology Association for Information Processing.</i>
<i>IP</i>	: <i>Internet Protocol.</i>
<i>IR</i>	: <i>Information Retrieval.</i>
<i>IRML</i>	: <i>Investment Research Markup Language.</i>
<i>IST</i>	: <i>Information Science Technologies.</i>
<i>iTQL</i>	: <i>Tucana SQL-like query language.</i>
<i>JDBC</i>	: <i>Java Database Connectivity API.</i>
<i>JDT</i>	: <i>Java Development Tooling.</i>
<i>JSP</i>	: <i>Java Server Pages.</i>
<i>KIF</i>	: <i>Knowledge Interchange Format.</i>
<i>KR</i>	: <i>Knowledge Representation.</i>
<i>KSL</i>	: <i>Knowledge Systems Laboratory.</i>
<i>LGPL</i>	: <i>GNU Lesser General Public License.</i>
<i>LPOO</i>	: <i>Linguagens de Programação Orientada a Objectos.</i>
<i>LSN</i>	: <i>Large Scale Networking.</i>
<i>MathML</i>	: <i>Mathematical Markup Language.</i>
<i>MIF</i>	: <i>Management Information Format.</i>
<i>MIT</i>	: <i>Massachusetts Institute of Technology.</i>

---

<i>MML</i>	: <i>Medical Markup Language.</i>
<i>MPL</i>	: <i>Mozilla Public License.</i>
<i>N3</i>	: <i>Notation3 Language.</i>
<i>N3QL</i>	: <i>Notation3 Query Language.</i>
<i>NCSA</i>	: <i>National Center for Supercomputing Applications.</i>
<i>NG4J</i>	: <i>Named Graphs API for Jena.</i>
<i>NML</i>	: <i>News Markup Language.</i>
<i>NIST</i>	: <i>National Institute of Standards and Technology.</i>
<i>NOKOS</i>	: <i>Nokia Open Source License.</i>
<i>OIL</i>	: <i>Ontology Inference Layer.</i>
<i>OKBC</i>	: <i>Open Knowledge Base Connectivity.</i>
<i>OleDB</i>	: <i>Object Linking and Embedding Data Base.</i>
<i>OMF</i>	: <i>Weather Observation Definition Format.</i>
<i>ORIENT</i>	: <i>Ontology engineRIng EnvironmenT.</i>
<i>OWL</i>	: <i>Web Ontology Language.</i>
<i>OWL-DL</i>	: <i>OWL-Description Logic.</i>
<i>OWL-QL</i>	: <i>OWL-Query Language.</i>
<i>OWL-S</i>	: <i>Web Services Ontology Language.</i>
<i>OWLS-TC</i>	: <i>OWL-S Service Retrieval Test Collection.</i>
<i>PCL</i>	: <i>Printer Control Language.</i>
<i>PDA</i>	: <i>Personal digital assistant.</i>
<i>PDE</i>	: <i>Plug-in Development Environment.</i>
<i>PDF</i>	: <i>Portable Document Format.</i>
<i>PetroXML</i>	: <i>Petroleum XML.</i>
<i>PHP</i>	: originalmente utilizada como abreviatura de <i>Personal Home Page Tools</i> mas actualmente é utilizada como abreviatura de <i>PHP: Hypertext Pre-processor.</i>

---

<i>PL/SQL</i>	: <i>Procedural Language extensions to SQL.</i>
<i>PNG</i>	: <i>Portable Network Graphics.</i>
<i>PNL</i>	: <i>Pseudo Natural Language.</i>
<i>POSIX</i>	: <i>Portable Operation System Interface.</i>
<i>RDF</i>	: <i>Resource Description Framework.</i>
<i>RDFS</i>	: <i>RDF Schema (RDF vocabulary description language).</i>
<i>RDFQL</i>	: <i>RDF Query Language.</i>
<i>RDQL</i>	: <i>RDF Data Query Language.</i>
<i>RIC</i>	: <i>RDF Instance Creator.</i>
<i>ROWL</i>	: <i>Rule Extension of OWL Mobile Commerce.</i>
<i>SAX</i>	: <i>Simple API for XML.</i>
<i>SGML</i>	: <i>Standard Generalized Markup Language.</i>
<i>SHOE</i>	: <i>Simple HTML Ontology Extensions.</i>
<i>SMIL</i>	: <i>Synchronized Multimedia Integration Language.</i>
<i>SO</i>	: <i>Sistema Operativo.</i>
<i>SOAP</i>	: <i>Simple Object Access Protocol.</i>
<i>SOFA</i>	: <i>Simple Ontology Framework API.</i>
<i>SPARQL</i>	: <i>Protocol And RDF Query Language.</i>
<i>SPC</i>	: <i>Simples Programas de Computador ou outro tipo de entidade computacional não-humana.</i>
<i>SPOO</i>	: <i>Sistemas de Programação Orientada a Objectos.</i>
<i>SQL</i>	: <i>Structured Query Language.</i>
<i>SVG</i>	: <i>Scalable Vector Graphics.</i>
<i>SWRL</i>	: <i>Semantic Web Rule Language.</i>
<i>SWT</i>	: <i>Standard Widget Toolkit.</i>
<i>TCP</i>	: <i>Transmission Control Protocol.</i>
<i>ThML</i>	: <i>Theological Markup Language.</i>
<i>TREC</i>	: <i>Text Retrieval Conference.</i>

---

<i>UBI</i>	: Universidade da Beira Interior.
<i>UDDI</i>	: <i>Universal Description Discovery Integration.</i>
<i>UI</i>	: <i>User Interface.</i>
<i>UML</i>	: <i>Unified Modeling Language.</i>
<i>URI</i>	: <i>Uniform Resource Identifier</i> ou <i>Universal Resource Identifier</i>
<i>URL</i>	: <i>Uniform Resource Locator.</i>
<i>URN</i>	: <i>Uniform Resource Name.</i>
<i>VIML</i>	: <i>Virtual Instruments Markup language.</i>
<i>W3C</i>	: <i>World Wide Web Consortium.</i>
<i>W3L</i>	: <i>W3C Software License.</i>
<i>WQL</i>	: <i>WMI Query Language.</i>
<i>WMI</i>	: <i>Windows Management Instrumentation.</i>
<i>WML</i>	: <i>Wireless Markup Language.</i>
<i>WSDL</i>	: <i>Web Services Description Language.</i>
<i>XAML</i>	: <i>Transaction Authority Markup Language.</i>
<i>XBEL</i>	: <i>Bookmark Exchange Language.</i>
<i>XHTML</i>	: <i>EXtensible HyperText Markup Language.</i>
<i>XML</i>	: <i>Extensible Markup Language.</i>
<i>XOL</i>	: <i>XML-based Ontology Exchange Language.</i>
<i>XSD</i>	: <i>XML Schema Definition.</i>
<i>XSL</i>	: <i>eXtensible Stylesheet Language.</i>
<i>YARS</i>	: <i>Yet Another RDF Store.</i>

# Índice

<b>Resumo</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Agradecimentos</b>	<b>v</b>
<b>Prefácio</b>	<b>vii</b>
Notação Utilizada . . . . .	ix
Lista de Abreviaturas . . . . .	ix
<b>Índice</b>	<b>xv</b>
<b>Lista de Figuras</b>	<b>xix</b>
<b>Lista de Tabelas</b>	<b>xxiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Enquadramento da Dissertação . . . . .	1
1.2 Objectivos . . . . .	4
1.3 Principais Contribuições . . . . .	6
1.4 Organização da Dissertação . . . . .	7
<b>2 Limitações da Web e Respectiva Evolução</b>	<b>9</b>
2.1 Introdução . . . . .	9
2.2 A Internet . . . . .	10

2.3	Os Primórdios da Web . . . . .	14
2.4	As Limitações da Web . . . . .	16
2.5	Evolução da Web . . . . .	20
2.6	Resumo . . . . .	21
<b>3</b>	<b>Web Semântica</b>	<b>23</b>
3.1	Introdução . . . . .	23
3.2	Perspectiva Geral sobre a Web Semântica . . . . .	25
3.3	Representação da Informação e do Conhecimento . . . . .	30
3.4	Princípios da Web Semântica . . . . .	39
3.4.1	Primeiro Princípio: Tudo pode ser Identificado Através de um URI . . . . .	40
3.4.2	Segundo Princípio: Os Recursos e as Relações Podem ser Tipificados . . . . .	42
3.4.3	Terceiro Princípio: Tolerância a Quebras de Relações . . . . .	44
3.4.4	Quarto Princípio: Desnecessidade da Existência de uma Confiança Absoluta . . . . .	45
3.4.5	Quinto Princípio: Permitir Evolução . . . . .	47
3.4.6	Sexto Princípio: Implementação Minimalista . . . . .	48
3.5	Arquitetura Estratificada da Web Semântica . . . . .	48
3.5.1	XML . . . . .	51
3.5.2	Arquitetura RDF . . . . .	59
3.5.3	Ontologias . . . . .	67
3.6	A Imagem da Web Semântica . . . . .	78
3.7	Resumo . . . . .	81
<b>4</b>	<b>Ferramentas para a Web Semântica</b>	<b>85</b>
4.1	Introdução . . . . .	85
4.2	Classificação em Categorias das Ferramentas para a Web Semântica . . . . .	86
4.3	Ferramentas associadas a cada Categoria . . . . .	92
4.3.1	Categoria - Anotação . . . . .	92
4.3.2	Categoria - API . . . . .	93

4.3.3	Categoria - Armazenamento . . . . .	96
4.3.4	Categoria - Conversão . . . . .	98
4.3.5	Categoria - Edição . . . . .	99
4.3.6	Categoria - Integração . . . . .	101
4.3.7	Categoria - Inferência . . . . .	104
4.3.8	Categoria - Interpretador Web . . . . .	106
4.3.9	Categoria - Interrogação . . . . .	108
4.3.10	Categoria - Validação . . . . .	110
4.3.11	Categoria - Visualização . . . . .	110
4.4	Resumo . . . . .	112
<b>5</b>	<b>SWedt: Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF</b>	<b>113</b>
5.1	Introdução . . . . .	113
5.2	A Ferramenta SWedt . . . . .	115
5.2.1	Plataforma <i>Eclipse</i> . . . . .	116
5.2.2	Arquitetura da ferramenta SWedt . . . . .	122
5.2.3	APIs utilizadas . . . . .	128
5.3	Implementação da Ferramenta SWedt . . . . .	132
5.3.1	<i>Plug-In Development Environment</i> (PDE) . . . . .	133
5.3.2	Principais Entidades e Acções Implementadas . . . . .	134
	Entidade Documento XML . . . . .	136
	Entidade Texto Simples . . . . .	138
	Entidade Tabela de Associações . . . . .	139
	Entidade Lista de Ontologias . . . . .	140
	Entidade Tabela de Asserções . . . . .	141
5.4	Modo e Funcionamento da Ferramenta SWedt . . . . .	142
5.4.1	Componentes gráficas da ferramenta . . . . .	142

Vista Gestor de Documentos . . . . .	144
Editor Múltiplo . . . . .	145
Vista Recursos . . . . .	149
Vista Asserções . . . . .	150
Vistas Tarefas e Problemas . . . . .	151
5.4.2 Modo de Funcionamento . . . . .	151
5.4.3 Exemplo de Utilização . . . . .	155
Dados e Informação na Web Semântica . . . . .	155
Utilização do editor SWedt . . . . .	169
5.5 A Ferramenta SMORE . . . . .	171
5.6 Resumo . . . . .	174
<b>6 Conclusões e Perspectivas de Trabalho Futuro</b>	<b>177</b>
6.1 Conclusões . . . . .	178
6.2 Sugestões para Trabalho Futuro . . . . .	181
<b>Referências</b>	<b>182</b>
<b>Apêndices:</b>	<b>218</b>
<b>A Sintaxe RDF/XML</b>	<b>219</b>
<b>B OWL</b>	<b>227</b>
<b>C Diagramas de Classes em UML da ferramenta SWedt</b>	<b>235</b>



# Lista de Figuras

2.1	O que, realmente, compreendem os cães (adaptado de [57]). . . . .	18
2.2	Como as ferramentas web interpretam as páginas web. . . . .	19
3.1	Ferramentas web na Web e SPC na Web Semântica (adaptado de [70]). . . . .	27
3.2	Recursos do mundo físico e os seus URIs na Web Semântica. . . . .	41
3.3	Ligações entre objectos na Web. . . . .	42
3.4	Os recursos e as relações podem ser tipificados na Web Semântica. . . . .	43
3.5	A Web na perspectiva de Tim Berners-Lee em 1989. Adaptado de [37]. . . . .	44
3.6	Página a informar a indisponibilidade de se poder aceder a um endereço web. . .	45
3.7	Uma cadeia de confiança para definir direitos de acesso. . . . .	46
3.8	Combinação de nova informação com informação anterior. . . . .	47
3.9	A arquitectura estratificada da Web Semântica. . . . .	49
3.10	Exibição de um documento XML utilizando diferentes regras de apresentação. .	53
3.11	Representação gráfica da asserção: "Rui Pereira é o autor do documento sem- web.pdf". . . . .	61
3.12	Representação gráfica da asserção: "Rui Pereira é o autor do documento sem- web.pdf, o qual tem <b>Web Semântica</b> como título". . . . .	61
3.13	Relação existente entre um vocabulário Esquema RDF e asserções RDF definidas numa página Web Semântica. (Adaptado de [139]). . . . .	65
3.14	Níveis de riqueza semântica associados às Ontologias [146]. . . . .	72

3.15 O modelo conceptual como uma representação comum do conhecimento (Adaptado de [87]). . . . .	74
3.16 Interligação dos vários componentes da Web Semântica (adaptado de [159]). . .	79
4.1 Classificação das ferramentas disponíveis para a Web Semântica. . . . .	87
4.2 Ferramentas agrupadas por linguagem utilizada no seu desenvolvimento. . . . .	88
4.3 Ferramentas agrupadas por licença. . . . .	90
4.4 Ferramentas agrupadas por anos. . . . .	91
5.1 Principais tecnologias associadas à ferramenta SWedt. . . . .	116
5.2 A plataforma <i>Eclipse</i> . . . . .	118
5.3 Principais componentes da ferramenta SWedt. . . . .	123
5.4 Arquitectura da Ferramenta SWedt. . . . .	125
5.5 APIs utilizadas pela ferramenta SWedt. . . . .	129
5.6 Representação RDF do vCard de Rui Pereira. . . . .	129
5.7 Interligação das tecnologias e componentes associados à ferramenta Swedt. . . .	132
5.8 <i>Plug-In Development Environment</i> . . . . .	133
5.9 Entidades e Acções associadas ao desenvolvimento da ferramenta SWedt. . . . .	135
5.10 Ícone SWedt. . . . .	142
5.11 Ferramenta SWedt. . . . .	143
5.12 Vista Gestor de documentos. . . . .	145
5.13 Editor Simples da ferramenta SWedt. . . . .	146
5.14 Interpretador Web da ferramenta SWedt. . . . .	147
5.15 Editor Completo da ferramenta SWedt. . . . .	148
5.16 Vista Recursos da ferramenta SWedt. . . . .	149
5.17 Vista Asserções da ferramenta SWedt. . . . .	150
5.18 Vista Tarefas da ferramenta SWedt. . . . .	151
5.19 Vista Problemas da ferramenta SWedt. . . . .	151
5.20 Interacções existentes entre os componentes gráficos do Editor SWedt. . . . .	152

## LISTA DE FIGURAS

---

5.21	Estrutura XML da frase <i>Port is served with Stilton cheese</i> . . . . .	157
5.22	KSL Wine Agent: <i>Course Type: CHEESE-NUTS-DESSERT</i> . . . . .	167
5.23	<i>Inference Web Explainer</i> . <i>Course Type: CHEESE-NUTS-DESSERT</i> . . . . .	168
5.24	SWedt: <i>Port is served with Stilton cheese</i> . . . . .	169
5.25	A ferramenta SMORE. . . . .	172
A.1	Rui Pereira é o autor do documento <i>semweb.pdf</i> . . . . .	219
A.2	Utilização de "recursos anónimos" em asserções RDF. . . . .	221
A.3	Colecção de Objectos do tipo Bag. . . . .	222
C.1	Diagrama de Classes da ferramenta SWedt. . . . .	236
C.2	Diagrama de Classes do Pacote DOM. . . . .	237
C.3	Diagrama de Classes do Pacote RDF. . . . .	238
C.4	Diagrama de Classes do Pacote RDFS. . . . .	239
C.5	Diagrama de Classes do Pacote OWL. . . . .	240
C.6	Diagrama de Classes do Pacote Editors. . . . .	241
C.7	Diagrama de Classes do Pacote Simple. . . . .	242
C.8	Diagrama de Classes do Pacote Complete. . . . .	243
C.9	Diagrama de Classes do Pacote Rules. . . . .	244
C.10	Diagrama de Classes do Pacote Views. . . . .	245
C.11	Diagrama de Classes do Pacote Resources. . . . .	246
C.12	Diagrama de Classes do Pacote Perspective. . . . .	247



# Lista de Tabelas

4.1	Categoria - Anotação . . . . .	92
4.2	Categoria - API . . . . .	94
4.3	Categoria - Armazenamento . . . . .	96
4.4	Categoria - Conversão . . . . .	98
4.5	Categoria - Edição . . . . .	99
4.6	Categoria - Integração . . . . .	102
4.7	Categoria - Inferência . . . . .	104
4.8	Categoria - Interpretador Web . . . . .	107
4.9	Categoria - Interrogação . . . . .	108
4.10	Categoria - Validação . . . . .	110
4.11	Categoria - Visualização . . . . .	111



# Capítulo 1

## Introdução

"All men dream, but not equally. Those who dream by night in the dusty recesses of their minds, wake in the day to find that it was vanity: but the dreamers of the day are dangerous men, for they may act on their dreams with open eyes, to make them possible."

Thomas Edward Lawrence (of Arabia)

### 1.1 Enquadramento da Dissertação

A Web Semântica, proposta por Tim Berners-Lee *et.al.* [1] e corroborada pelo *World Wide Web Consortium* (W3C)<sup>1</sup>, tem sido alvo de uma intensa actividade de Investigação e Desenvolvimento (I&D). Ela representa uma evolução da actual Web que é caracterizada pela adição de uma estrutura organizacional e semântica à informação existente na Web, de forma a que o significado dessa informação possa ser menos ambíguo e facilmente partilhado, interpretado,

---

<sup>1</sup>O consórcio W3C promove e encoraja a interactividade e o aprofundamento das opiniões sobre questões relacionadas com a Web, tendo nos últimos anos apresentado mais de trinta especificações técnicas para a sua infra-estrutura. A Web Semântica é encorajada pelo W3C através do projecto *W3C Semantic Web Activity* [2], tendo sido, Eric Miller, o pesquisador escolhido por Tim Berners-Lee para chefiar esse projecto.

processado e relacionado através de simples programas de computador ou outro tipo de entidade computacional não humana (SPC)<sup>2</sup> e não, exclusivamente, pelos seres humanos como acontece na actual Web.

A ideia da Web Semântica obteve inspiração a partir de anteriores projectos de Recuperação de Informação<sup>3</sup> (IR) e de Representação do Conhecimento<sup>4</sup> (KR), mas ao contrário deles, apresenta um modelo descentralizado para definir, extrair, partilhar e contextualizar a informação de forma universal e interpretável pelos SPC. Desta forma, abandona a solução centralizadora e tradicional de apresentar apenas um vocabulário, ou um conjunto muito restrito de vocabulários, para a representação de termos ou conceitos de uma forma universal e encorajar a existência e utilização simultânea de milhares de diferentes vocabulários<sup>5</sup>. Assim, num ambiente em constante crescimento e mudança, como o da Web, a imposição consciente ou subliminar de mecanismos de selecção por parte dos utilizadores e dos SPC fortalecerá apenas os vocabulários mais profícuos e fiáveis levando-os inevitavelmente à sua aceitação generalizada e consequente "sobrevivência".

É indubitável que a informação disponibilizada na Web Semântica sobre um determinado recurso<sup>6</sup> poderá ser disponibilizada ou apresentada de diferentes formas<sup>7</sup>. Apesar de ser fácil para uma pessoa comparar diferentes descrições de um mesmo recurso fornecidas por diferentes autores e identificar que se trata do mesmo recurso, para um algoritmo informático essa tarefa é extremamente complicada. A utilização de vários e diferentes vocabulários de uma forma eficiente e útil estará dependente do desenvolvimento de SPC capazes de compreender<sup>8</sup> o signi-

---

<sup>2</sup>A sigla SPC irá ser utilizada ao longo desta dissertação para designar **Simples Programas de Computador ou outro tipo de entidade computacional não-humana**.

<sup>3</sup>Do termo em inglês: *Information Retrieval*.

<sup>4</sup>Do termo em inglês: *Knowledge Representation*.

<sup>5</sup>Alguns contraditórios entre si ou até pérfidos.

<sup>6</sup>Na Web Semântica, documentos existentes na Web, pessoas, livros, locais, objectos do mundo físico e até conceitos abstractos são designados por Recursos.

<sup>7</sup>Cada autor disponibiliza a informação de acordo com a suas preferências, experiências, possibilidades, etc.

<sup>8</sup>Neste contexto, o significado de compreensão não é semanticamente rico como aquele que é utilizado pelos seres humanos mas, mesmo assim é mais rico que a simples associação de meta-dados aos dados como é actualmente utilizado, como por exemplo através da utilização da especificação XML. Desta forma, os SPC além de



## 1.1. ENQUADRAMENTO DA DISSERTAÇÃO

---

ficado e as relações existentes entre os recursos de cada um desses vocabulários. Desta forma, os vocabulários necessitarão, obrigatoriamente, de ser expressos através de modelos que possibilitem a sua interpretação por parte dos SPC. Na Web Semântica, esses modelos universais são fundamentalmente baseados na *Extensible Markup Language* (XML) [3], na arquitectura *Resource Description Framework* (RDF) [4], na linguagem ontológica *Web Ontology Language* (OWL) [5] e em regras de inferência que poderão ser definidas quer a nível dos SPC, que a nível das ontologias<sup>9</sup>. Através da definição de ontologias que representem recursos e as suas relações de forma lógica<sup>10</sup> e interpretável pelos SPC, a autonomização de regras poderá ser facilmente utilizada por parte dos SPC para inferir nova informação.

Durante algum tempo, a Web Semântica foi considerada por muitos investigadores como sendo uma obra de ficção mas, cada vez mais, está a tornar-se uma realidade. Os principais desafios para a sua concretização e implementação consistem na especificação de uma arquitectura universal para a representação da informação e na definição de um conjunto de regras de inferência que permita relacionar e interpretar essa informação, de forma automática e autónoma, por parte dos SPC. Actualmente, ela representa um enorme desafio para os actuais grupos de I&D da Web, compreendendo-se, assim, porque é que a sua alargada divulgação constitui uma das actuais e principais preocupações do consórcio W3C.

---

reagirem como se realmente tivessem compreendido a informação, deverão também ser capazes de trocar entre si a informação resultante do seu processamento interno e complementá-la com informação descritiva, realçando o raciocínio por eles realizado.

<sup>9</sup>Regra geral, uma ontologia representa uma descrição de todos os conceitos, significados e as suas relações dentro de uma área do conhecimento em particular, como por exemplo: a área da Medicina; da Matemática; da Reparação Automóvel; etc. As ontologias serão apresentadas, mais detalhadamente, na secção 3.5 do capítulo 3.

<sup>10</sup>A partir do momento que um sistema é baseado na lógica, ele pode ser utilizado para provar coisas.

### 1.2 Objectivos

Devido ao facto de a Web Semântica ser uma das mais recentes áreas de I&D e de abranger um extenso conjunto de tecnologias<sup>11</sup>, não existe uma ideia clara por parte da generalidade dos utilizadores da Internet e dos profissionais de informática, que desenvolvem as suas actividades na área das tecnologias da Internet, sobre o seu conceito, a sua arquitectura e as suas potencialidades. Mesmo assim, ela não deixa de apresentar-se como sendo um novo e empolgante desafio para os actuais grupos de I&D. Prova disso é a enorme quantidade de novas ferramentas<sup>12</sup> associadas às camadas<sup>13</sup> da sua arquitectura que têm surgido constantemente nos últimos cinco anos [6]. Dois tipos de ferramentas muito comuns desenvolvidas para a Web Semântica são os seguintes:

- Editores para criar conteúdos RDF, tais como: o *RDF Editor* (RDFedt) [7] e o *RDF Instance Creator* (RIC) [8];
- Editores para criar e manipular ontologias, tais como: o *Protégé* [9], o *OILEd* [10] e o *OntoEdit* [11].

Apesar de os exemplos referidos no parágrafo anterior serem adequados quando a acção se concentra na criação de documentos RDF ou de ontologias, a ausência de editores que integrem, de forma simples e única, as diversas funcionalidades de cada uma das camadas da Web Semântica implica que, de momento, ainda seja necessário utilizar várias etapas para a criação de documentos para a Web Semântica. Além disso, por mais promissora que a Web Semântica venha a ser, esta dificilmente se tornará uma realidade enquanto não forem desenvolvidas ferramentas que permitam a sua utilização de forma simples por todos os potenciais utilizadores da Web. O desenvolvimento de ferramentas deste género, que permitam aproximar os actuais

---

<sup>11</sup>Sendo a maioria delas também muito recente.

<sup>12</sup>As principais ferramentas para a Web Semântica serão apresentadas no capítulo 4 dedicado às Ferramentas para a Web Semântica.

<sup>13</sup>A arquitectura da Web Semântica é definida como sendo uma estrutura em camadas (conforme se pode verificar na secção 3.5 do capítulo 3, dedicada à arquitectura estratificada da Web Semântica), constituindo cada uma delas por si só, um desafio de I&D.

## 1.2. OBJECTIVOS

---

utilizadores da Web aos conceitos e tecnologias da Web Semântica, é actualmente quase que inexistente, como se poderá constatar através do estudo e classificação apresentado no capítulo 4, dedicado às actuais ferramentas para a Web Semântica.

O actual e generalizado desconhecimento, quer parcial ou inexacto, do correcto modo de funcionamento da Web Semântica e a inexistência de ferramentas que permitam a fácil integração dos seus princípios por parte de utilizadores não familiarizados com a sua arquitectura, podem identificar-se e agregar-se como sendo um problema de ineficiente envolvimento, divulgação e encorajamento dos actuais utilizadores e investigadores da Web na temática da Web Semântica. Como o sucesso e a materialização da Web Semântica está dependente da adopção da sua arquitectura e princípios, pelo maior número de utilizadores possível, existe a necessidade de uma urgente resolução do problema enunciado. Desta forma, os objectivos da presente dissertação consistem em:

- i) Descrever de forma simples o nível de desenvolvimento actual das principais tecnologias associadas à Web Semântica;
- ii) Apresentar um estudo sobre a classificação da panóplia de tecnologias disponíveis para a Web Semântica;
- iii) Apresentar uma proposta de arquitectura e respectiva implementação, com vista ao desenvolvimento de uma ferramenta que permita integrar, de forma simples, as principais tecnologias ao nível da edição para a Web Semântica, bem como, facilitar a criação de conteúdos, de acordo com os princípios da Web Semântica, por parte de utilizadores não familiarizados com a Web Semântica.

### 1.3 Principais Contribuições

De modo a concretizar os objectivos referidos na secção anterior, foi realizado o trabalho de investigação e de desenvolvimento relatado nesta dissertação e do qual resultaram algumas contribuições científicas. A presente secção descreve, na opinião do autor, essas contribuições.

A primeira contribuição do trabalho conducente à presente dissertação consiste na apresentação, de forma detalhada e compreensiva, da arquitectura da Web Semântica e do respectivo estado da arte, a qual é apresentada no capítulo 3. Esta contribuição foi parcialmente publicada como uma entrada na *Encyclopedia of Multimedia Technology and Networking*, publicada pela Idea Group Reference [12].

A segunda contribuição consiste num estudo de classificação das tecnologias existentes para a Web Semântica, o qual envolveu a análise de mais de oitenta ferramentas. Esta contribuição, apresentada no capítulo 4, foi aceite para publicação como entrada na *Encyclopedia of Information Science and Technology, Second Edition*, a publicar pela Idea Group Reference [6].

A terceira contribuição consiste na proposta da arquitectura e implementação de um editor para a Web Semântica, integrando anotações semânticas, ontologias e RDF, a qual é apresentada no capítulo 5. A proposta da arquitectura para o referido editor foi apresentada na *International Conference on Information Networking* [13], a qual aceitou 141 artigos de um total de 468 artigos submetidos. Este artigo foi posteriormente seleccionado, após novo processo de revisão, para publicação em livro incluído na série *Lecture Notes in Computer Science* da Springer-Verlag [14]. Detalhes sobre a modelação e implementação da arquitectura para o editor proposto foram ainda apresentados na *International Conference on Internet and Web Applications and Services*, a qual teve uma taxa de aceitação de cerca de 37%, tendo sido apresentados 110 artigos [15].

### 1.4 Organização da Dissertação

A presente dissertação encontra-se estruturada em seis capítulos e três apêndices. Em relação às temáticas abordadas, a dissertação divide-se em quatro partes principais:

- Capítulo 2 - Limitações da Web e Respectiva Evolução;
- Capítulo 3 - A Web Semântica;
- Capítulo 4 - As Ferramentas para a Web Semântica;
- Capítulo 5 - A Ferramenta SWedt.

Em seguida, é apresentada uma breve descrição acerca do conteúdo de cada um dos capítulos que compõem a dissertação. O presente capítulo, **Introdução**, apresenta o contexto em que se enquadra o tema, os principais objectivos, a organização da dissertação e as principais contribuições. O segundo capítulo, **Limitações da Web e Respectiva Evolução**, apresenta uma breve descrição sobre a Internet e o seu modo de funcionamento. Esta breve descrição sobre a Internet serve de base para a apresentação da origem histórica, evolução e de algumas limitações da tecnologia Web. Esta introdução relacionada com a Internet e a Web é muito importante na medida em que apresenta os alicerces que servem de base para justificar a necessidade actual de uma nova geração da Web, a Web Semântica. O terceiro capítulo, **Web Semântica**, consiste na apresentação da ideia da Web Semântica, dos seus principais princípios e da sua arquitectura estratificada. É dada especial ênfase às tecnologias da Web Semântica que mais têm evoluído nos últimos tempos: a XML, a arquitectura RDF e as Ontologias. O quarto capítulo, **Ferramentas para a Web Semântica**, apresenta o resultado de um estudo de categorização de oitenta das principais ferramentas desenvolvidas para a Web Semântica. O quinto capítulo, **SWedt: Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF**, é dedicado à apresentação da arquitectura, do modo de funcionamento e das principais vantagens da ferramenta SWedt, a qual foi desenvolvida durante a realização da presente dissertação. As principais conclusões do trabalho realizado, tendo como base os objectivos previamente estabelecidos, encontram-se

## 1.4. ORGANIZAÇÃO DA DISSERTAÇÃO

---

enunciadas no sexto capítulo. Na sequência do trabalho de investigação desenvolvido são apresentadas, também nesse capítulo, sugestões para trabalho futuro.

Em complemento a estes capítulos a dissertação inclui ainda três apêndices. Nesses apêndices serão abordadas de forma mais específica as seguintes temáticas:

- Apêndice A - Sintaxe RDF/XML, apresenta uma breve descrição da sintaxe RDF/XML;
- Apêndice B - OWL, expõe uma breve descrição da linguagem OWL e da sua sintaxe;
- Apêndice C - Diagrama de Classes em UML da ferramenta SWedt, apresenta o conjunto de diagramas de classes em *Unified Modeling Language* (UML) utilizado no desenvolvimento da nova ferramenta.

# Capítulo 2

## Limitações da Web e Respectiva Evolução

"I cannot say whether things will get better if we change;  
what I can say is they must change if they are to get better."

G. C. Lichtenberg

### 2.1 Introdução

A Web Semântica não é fruto de uma ideia espontânea, ela deve a sua origem a um passado histórico, sendo na realidade uma tentativa de aperfeiçoamento de algo que já existe, a Web<sup>14</sup>. Esta última é, por sua vez, normalmente definida como um repositório de informação à escala mundial, sendo também considerada responsável pela banalização da Internet. Por seu lado, a Internet, apesar de actualmente ser um espaço democrático e universal de expressão e acesso à informação, teve as suas raízes no desenvolvimento de projectos de defesa confidenciais realizados pelo Departamento de Defesa dos Estados Unidos da América a partir da década de

---

<sup>14</sup>Como já tinha referido no prefácio, Web é utilizada como uma designação reduzida de *World Wide Web*. Outras duas designações reduzidas também utilizadas são: WWW e W3.

sessenta. Desta forma, antes de se aprofundar o tema da Web Semântica, é indispensável fazer uma pequena abordagem acerca das tecnologias que lhe servem de base dando ênfase à origem, principais limitações e evolução da actual Web.

O presente capítulo é constituído por seis secções. Nesta primeira secção é apresentada uma breve introdução ao tema que vai ser abordado com maior detalhe ao longo deste capítulo, bem como uma pequena descrição da organização temática de todo o capítulo. Na segunda secção, **Internet**, é apresentada uma breve descrição sobre a Internet e um dos seus principais serviços, a Web. Na terceira secção, **Os Primórdios da Web**, é apresentada uma breve descrição acerca da Web e da sua origem. Na quarta secção, **As Limitações da Web**, são abordadas algumas das limitações da actual Web, as quais podem ser consideradas como a justificação para a necessidade de uma nova arquitectura para a Web. Na quinta secção, **Evolução da Web**, é apresentada a Web Semântica como sendo uma possível solução para ultrapassar as actuais limitações da Web. Por último, é apresentado na sexta secção um pequeno resumo acerca do tema abordado neste capítulo.

## 2.2 A Internet

Estamos a assistir a uma profunda transformação da sociedade em que vivemos. Essa transformação está a ser protagonizada pela Revolução da Informação e do Conhecimento, cuja evolução já foi comparada, por Peter Drucker [16] em 1999, à da Revolução Industrial. Segundo este autor, a grande descoberta tecnológica da revolução industrial foi a máquina a vapor. Posteriormente surgiram novos produtos e serviços, tendo-se também observado descobertas tecnológicas distintas da que provocou a mudança inicial, nomeadamente a descoberta dos caminhos de ferro e todos os serviços e alterações de mentalidade que surgiram em consequência dessa descoberta. No caso da Revolução da Informação e do Conhecimento, a grande descoberta tecnológica foi o computador que rapidamente se tornou num objecto de uso quotidiano. À semelhança da Revolução Industrial, observam-se também, na Revolução da Informação e do Conhecimento,



## 2.2. A INTERNET

---

descobertas tecnológicas distintas da que provocou a mudança inicial. A Internet representa uma grande inovação que está actualmente a mudar por completo as formas de encarar a economia e a sociedade em que vivemos [16].

No presente momento, assiste-se, por uma significativa parte da população mundial, a uma banalização da ideia da Internet. Mas, uma tão alargada diversidade de experiências e de culturas existentes no mundo conduz a que também existam várias e diferentes definições sobre a Internet. Por exemplo, uma pequena pesquisa a partir da informação publicada através dos meios de comunicação revela uma panóplia de diferentes definições para a Internet. A seguir são apresentados alguns exemplos dessas definições, as quais se podem encontrar na própria Internet:

- É o protótipo da auto-estrada da informação do futuro [17];
- É um local onde as pessoas podem comunicar com os amigos, professores, família, desconhecidos, etc [17];
- É uma ferramenta de investigação [17];
- É uma plataforma de negócio [17];
- É uma biblioteca universal [18];
- É um sistema do género do *Big Brother*<sup>15</sup> do escritor George Orwell [19];
- É um meio de comunicação não controlado [20];
- É uma das maiores invenções de todos os tempos [21];
- É um ninho para o crime organizado [22].

---

<sup>15</sup>A definição de Sistema *Big Brother* (Irmão mais Velho) surgiu em 1949 no livro de George Orwell intitulado de '1984'. Neste tipo de sistema, os habitantes de um país são permanentemente vigiados por uma única entidade suprema - no caso do livro '1984' essa entidade era o governo.

## 2.2. A INTERNET

---

Como podemos constatar, a maioria das definições encontradas são incompletas ou incorrectas. Segundo o Dicionário de Língua Portuguesa da Porto Editora [23], a Internet é um neologismo que designa a rede mundial de comunicação por computadores, de origem norte-americana, a qual, permite aos seus utilizadores a troca de mensagens e o acesso a grande quantidade de informação.

A 24 de Outubro de 1995, o *Federal Networking Council* (FNC)<sup>16</sup>, aprovou, por unanimidade, uma resolução sobre a definição do termo 'Internet' [24]. Segundo essa resolução o termo 'Internet' representa um Sistema de Informação Global que:

- Encontra-se logicamente interligado através de um Espaço de Endereçamento Global único baseado no protocolo de comunicação denominado *Internet Protocol* (IP) ou numa das suas extensões;
- Permite suportar comunicações através da utilização em conjunto dos protocolos *Transmission Control Protocol* (TCP) e do IP, numa das suas extensões ou através de outros protocolos que sejam compatíveis com o IP;
- Fornece, utiliza ou torna acessíveis, quer publicamente quer privadamente, Serviços de Alto Nível associados às comunicações e respectivas infra-estruturas.

De uma maneira geral, a Internet é considerada como sendo o maior repositório de informação a nível mundial. Ela não é suportada nem controlada por uma única companhia ou organização, nem por um conjunto específico de companhias ou organizações. Além disso, não deve ser definida como uma única rede mundial, mas sim como o resultado da interligação de milhões de redes, pertencentes e utilizadas por milhões de pessoas em centenas de países. Desta forma, representa, em termos muito genéricos, um sistema interligado à escala mundial de computadores e de pessoas, cuja principal vantagem consiste em permitir a troca, de uma forma muito

---

<sup>16</sup>Instituição, de prestígio mundial, relacionada com a Internet. Em Outubro de 1997, algumas das suas funções foram transferidas para o grupo *Large Scale Networking* (LSN) [25].

## 2.2. A INTERNET

---

simples, de dados<sup>17</sup> entre todos os seus utilizadores.

Actualmente, não restam dúvidas que a Internet é um fenómeno mundial. Mas, inversamente ao que a maioria das pessoas possa julgar, a situação actual e objectiva da Internet não foi previamente planeada. A Internet, tal e qual nós hoje a conhecemos, não resulta de facto de um projecto previamente planeado, mas tão somente de uma evolução arbitrária permanentemente alimentada em múltiplas direcções através do trabalho de investigadores e do apoio de grandes empresas. A exemplo disso, temos a Web que além de ser um dos serviços mais conhecidos da Internet, também começou por ser um pequeno e bem direccionado projecto que inesperadamente transbordou do seu leito original<sup>18</sup>.

Segundo Greg Elmer [26], em termos teóricos, a Web pode ser designada como uma representação de um espaço virtual, onde é simulado um meio de interacção mundial, o qual existe como propriedade, mas sem exercício ou estado físico. Na prática e de uma forma muito simples, a Web pode ser considerada como uma interface gráfica à escala mundial que suporta uma vastíssima quantidade de dados que podem ser partilhados, acedidos e actualizados de forma extremamente fácil por milhões de utilizados em todo o mundo. Ela é o resultado do trabalho de vários investigadores ao longo de vários anos, sendo porém atribuído, merecidamente, o maior contributo a Tim Berners-Lee, o qual é apelidado por vários autores [27] de Pai da Web.

Devido, sobretudo, à notável facilidade de transferência de dados através da Internet existe um crescimento diário e planetário do número de utilizadores e de serviços com ela relacionados. Desta forma, a Internet possui uma dimensão<sup>19</sup> de tal maneira prodigiosa que não existe outro tipo de entidade conhecida que lhe possa servir de comparação se considerarmos o seu, extraordinariamente, rápido crescimento durante os últimos quinze anos. É de tal maneira

---

<sup>17</sup>Ficheiros, documentos web, dados multimédia, etc.

<sup>18</sup>Como iremos ver com maior detalhe na secção 2.3 - "Os primórdios da Web".

<sup>19</sup>Segundo dados recolhidos a partir da (*Internet World Stats* [28], em 31 de Dezembro de 2005 existiam 1.018.057.389 utilizadores da Internet, sendo a Ásia o continente com maior número de utilizadores: 364.270.713 utilizadores.

## 2.3. OS PRIMÓRDIOS DA WEB

---

grandiosa e útil que é considerada como a maior fonte de dados e o maior meio de exposição de informação existente. Apesar de ser um sistema completamente descentralizado e parcialmente desorganizado, consegue, graças, por um lado, à enorme quantidade de dados que abarca e, por outro lado, ao recurso a algoritmos de pesquisa de dados, ser o meio mais valioso de disponibilização de dados relativamente às diversas áreas do conhecimento humano. Segundo Alexander Voiskounsky [29], graças à sua universalidade, a Internet influenciou e continua a influenciar mudanças profundas nos comportamentos de todas as sociedades existentes.

A descrição, de forma, pormenorizada acerca do modo de funcionamento da Internet e da Web ultrapassa o âmbito do tema tratado nesta dissertação. Uma descrição, mais geral e precisa, sobre a Internet e a Web pode ser encontrada através da consulta de livros da especialidade, por exemplo em [30, 31, 32, 33].

## 2.3 Os Primórdios da Web

Desde o início da década de sessenta que a ideia de uma rede mundial de dados interligada era alimentada através de diversos projectos de I&D [34, 35, 36]. A ideia começou, verdadeiramente, a ganhar corpo no final da década de oitenta, quando Tim Berners-Lee iniciou um projecto [37] que possuía como objectivo gerir a informação existente no *European Center of Nuclear Research* (CERN) [38]. A ideia principal do projecto era a de ligar entre si os vários documentos existentes no CERN, muitos deles de sistemas heterogéneos, e implementar um sistema que permitisse uma fácil visualização e navegabilidade entre eles. Esse sistema deveria permitir, sempre que fosse necessário, obter mais informação sobre um dado termo ou objecto presente num documento acedendo facilmente a outros documentos ou dados com ele relacionados. Actualmente, um sistema com estas capacidades é designado por sistema de Hipertexto<sup>20</sup> ou Hipermedia<sup>21</sup>. No início da década de 90, Tim Berners-Lee realizou experiên-

---

<sup>20</sup>Sistema constituído por documentos que contêm partes do seu texto interligadas com partes de texto de outros documentos. O termo foi apresentado pela primeira vez por Ted Nelson em 1965.

<sup>21</sup>Sistema que apresenta uma extensão ao sistema Hipertexto, na medida e que permite também interligar objectos multimédia.

### 2.3. OS PRIMÓRDIOS DA WEB

---

cias de forma a implementar o seu modelo de sistema de Hipertexto/Hipermédia. Segundo Dan Connolly [39], foi durante o mês de Agosto de 1991 que foi, oficialmente, apresentada ao mundo a Web idealizada por Tim Berners-Lee. No início, apenas as grandes instituições, como Universidades e Centros de I&D, aderiram à Web. Por volta de 1992 surgiram os primeiros Navegadores Web<sup>22</sup> os quais, promoveram quer o desenvolvimento de páginas web quer a adesão do público à Web. Além da revolução originada pelo aparecimento dos Navegadores Web, a adopção por parte do CERN da filosofia de que a Web deveria ser um espaço aberto, universal e gratuito foi de extrema importância, quer para a sua sobrevivência, quer para o seu rápido crescimento. Em 1993, Tim Berners-Lee, em conjunto com outros investigadores, fundou o consórcio W3C, o qual tem vindo a publicar normas para as tecnologias utilizadas na Web. Na altura, o Navegador Web que obteve maior sucesso, foi o chamado Mosaic [45], desenvolvido por Marc Andreessen<sup>23</sup> ao serviço da *National Center for Supercomputing Applications* (NCSA) [46].

Desde essa altura que a criação e divulgação de novas páginas web cresce, rapidamente, ano após ano. Depois do mundo científico, foi a vez das grandes empresas e organizações se renderem às potencialidades disponibilizadas pela Web, sobretudo no que diz respeito à divulgação de produtos e serviços a nível mundial. Pouco a pouco, a Web começou a impor-se de tal forma que inevitavelmente provocou profundas alterações de estratégia na generalidade das grandes empresas, sendo o caso da empresa Microsoft [48] o exemplo mais conhecido<sup>24</sup>.

Actualmente vivemos numa sociedade de informação onde a utilização da Web é considerada uma parceira cada vez mais indispensável. Ela é o palco de novas tecnologias, de ideias de negócio e até de uma nova Economia Mundial.

---

<sup>22</sup>Termo utilizado na língua Inglesa: *Browsers*. São programas cliente de Navegação para a Web, utilizados para visualizar páginas web e interagir com os seus conteúdos. O Internet Explorer (IE) [40], o Netscape Navigator [41], o Opera [42], o Mozilla Firefox [43] e o Lynx [44] são cinco exemplos de Navegadores Web.

<sup>23</sup>Também conhecido como o fundador da Netscape [47].

<sup>24</sup>Segundo John Naughton [49], a Microsoft sempre se mostrou céptica em relação à Web de modo a desprezá-la por completo e só em meados de 1995 é que assumiu o seu erro. Primeiro começou por se manifestar incrédula devido ao seu desajustamento e atraso relativamente à nova ordem mundial. Depois, desorientada e coberta de um súbito pavor tentou por todos os meios "conquistar" a posição da, na altura, detentora do mercado de Navegadores Web, a Netscape, o que viria a conseguir mais tarde.

### 2.4 As Limitações da Web

Segundo Tim Berners-Lee *et.al.* [1], a Web já ultrapassou a sua primeira barreira, a qual era caracterizada pela disponibilização de grandes quantidades de dados a milhões de pessoas, estando estas distribuídas pela quase totalidade dos países do mundo [28, 50]. Mas, devido à actual imensurável quantidade de dados presente na Web, a procura de informação de forma fácil, eficiente e rápida, tornou-se num problema que urge ultrapassar. O facto de, na maioria das vezes, os utilizadores acrescentarem dados não estruturados à Web leva-a a um excesso de dados não estruturados disponíveis, dificultando a pesquisa e o acesso a esses mesmos dados. Neste sentido ganham evidência três tipos de preocupações:

- Como encontrar na Web, de forma simples e rápida, dados e a informação pretendida?
- Que técnicas de selecção e de triagem se devem implementar para melhorar a selecção de dados/informação úteis?
- Como aceder, de forma eficiente, aos dados/informação disponibilizados na "Web Profunda"<sup>25</sup> [51]?

O aparecimento, cada vez mais frequente, deste tipo de questões apenas realça o facto de que o rápido e parcialmente desorganizado crescimento da Web tem dificultado a localização, acesso, apresentação e manutenção de dados/informação de forma eficaz e para um número, também, crescente de utilizadores. A disponibilização de dados/informação não estruturados, por vezes heterogéneos, e a subjectividade real associada à relevância da própria informação acabam por gerar grandes dificuldades na sua pesquisa. A pesquisa de informação útil tem as suas raízes nas chamadas técnicas de IR popularizadas por Vannevar Bush [52]. Estas técnicas foram evoluindo lentamente ao longo do tempo, utilizando, como base para testes, colecções de documentos relativamente pequenas e muito específicas. No princípio dos anos 90, a situação começou a mudar

---

<sup>25</sup>Também designada por "Web Invisível", representa todos os dados/informação que se encontram na Web mas que não são directamente acessível por parte dos actuais algoritmos de pesquisa de dados na Web. A utilização de dados/informação armazenada em bases de dados especializadas é uma das suas principais formas de manifestação.

## 2.4. AS LIMITAÇÕES DA WEB

---

principalmente devido à promoção, por parte de Donna Harman, de uma conferência anual cognominada *Text Retrieval Conference* (TREC) [53]. As conferências TREC têm como objectivo a apresentação e o teste de vários sistemas de IR com base em grandes colecções de documentos. A pesquisa e a IR na Web tem vindo a ser facilitada, através da utilização de dois tipos específicos de ferramentas web robustas e práticas, os Motores de Busca<sup>26</sup> e os Directórios Web<sup>27</sup>. Contudo, apesar dos grandes avanços tecnológicos que se têm vindo a registar nesta área, as actuais técnicas de pesquisa e de IR são ainda muito limitadas. Se fizermos uma analogia entre a Web<sup>28</sup> e uma base de dados (BD) de elevada dimensão, verificamos que a estrutura da Web é muito diferente da estrutura base presente nas BD ditas clássicas, nas quais, os elementos são organizados, guardados e indexados de maneira a facilitar a rápida recuperação da informação. Assim, devido à falta de estrutura, organização e indexação da informação presente na Web, os actuais Motores de Busca, na sua grande maioria, baseiam as suas pesquisas nas páginas web através da utilização de uma palavra específica ou em simples conjuntos de palavras a procurar. Como as palavras são desprovidas de conteúdo, estas ferramentas web retornam uma panóplia de endereços<sup>29</sup> de páginas web onde foi encontrada a palavra, ou conjunto de palavras, pesquisada. Infelizmente, este tipo de resultados nem sempre é útil para os humanos e é considerado fraco em termos de utilidade para os SPC, pois as palavras podem ter muitos significados<sup>30</sup> como podem também possuir diferentes contextos associados<sup>31</sup>. Além disso, são considerados factores de grande insatisfação<sup>32</sup> a demora na obtenção de resultados das pesqui-

---

<sup>26</sup>No âmbito da Web, um Motor de Busca é uma ferramenta utilizada para facilitar a procura e a localização de dados na Web. Por exemplo, quando se pretende aceder a um dado específico, mas não se possui o seu endereço de localização na web, utiliza-se um Motor de Busca com o intuito de se obter o *Uniform Resource Locator* (URL) do dado pretendido. O Google [54], o Altavista [55], e o SAPO [56] são três exemplos de Motores de Busca de dados para a Web.

<sup>27</sup>Directórios Web é uma estrutura definida através de uma estrutura hierárquica de categorias, a qual é utilizada para classificar documentos web. Os endereços electrónicos de todos os documentos web que sejam classificados numa mesma categoria são guardados num mesma pasta virtual.

<sup>28</sup>Ou seja, o conjunto de todas as suas páginas web.

<sup>29</sup>O endereço de uma página web específica uma localização na Web, da mesma forma que o endereço referente à nossa caixa do correio específica uma localização no planeta terra.

<sup>30</sup>Por exemplo, a palavra *Ferreiro* pode estar associada a diferentes significados: nome de uma profissão; nome próprio; nome de uma localidade, etc.

<sup>31</sup>Por exemplo, ao pesquisar pelo nome do autor de um livro tanto recebemos as páginas web que se referem aos livros ou documentos por ele publicado, como documentos de outrem que se referem ao autor em questão.

<sup>32</sup>Insatisfação sentida por parte dos actuais utilizadores da Web.

## 2.4. AS LIMITAÇÕES DA WEB

---

sas e a necessidade de realizar manualmente trabalho de filtragem e de selecção para se atingir a informação realmente útil. Contudo, a principal limitação da Web não reside no facto de as ferramentas web de busca e pesquisa de informação serem pouco complexas e eficientes. Pelo contrário, apesar de as ferramentas web serem complexas, eficientes e estarem constantemente a serem melhoradas, a estruturação e definição da informação presente nas páginas web não se tem alterado muito desde o aparecimento da Web<sup>33</sup>. Neste cenário, estas ferramentas web necessitam de utilizar algoritmos cada vez mais complexos que poderiam ser simplificados e utilizados com melhores resultados, se as actuais páginas web evoluíssem para páginas com informação mais organizada e catalogada de forma universal. Deste ponto de vista, a principal limitação da Web baseia-se na inexistência de uma estruturação e conseqüente catalogação da informação presente nos seus documentos.

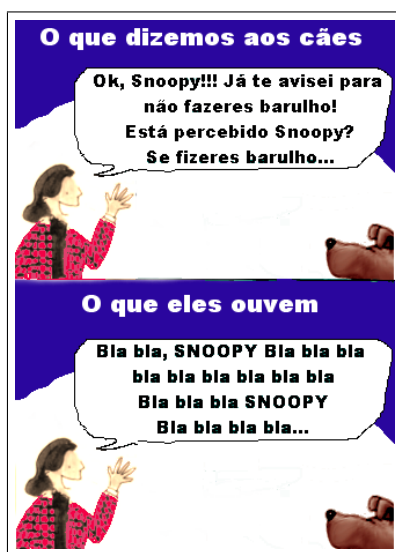


Figura 2.1: O que, realmente, compreendem os cães (adaptado de [57]).

Outra limitação muito importante, prende-se com o facto de a Web ser caracterizada como um meio de comunicação exclusivamente orientado para humanos. Apenas é utilizada a linguagem humana, ou seja, a informação é concebida tendo como consumidor os humanos e não os SPC. Nestas circunstâncias, apenas os humanos conseguem perceber ou manipular o real significado

---

<sup>33</sup>Com algumas excepções relacionadas com a recente utilização de regras estruturais e de meta-dados, como por exemplo, a utilização da especificação XML.



## 2.4. AS LIMITAÇÕES DA WEB

---

da informação presente na Web. Podemos fazer uma analogia entre o modo como os animais, por exemplo os cães, compreendem as nossas ordens e como as actuais ferramentas web interpretam as páginas na Web. A figura 2.1 ilustra como os cães realmente compreendem os humanos e a figura 2.2 exemplifica como realmente as actuais ferramentas web interpretam as páginas web.

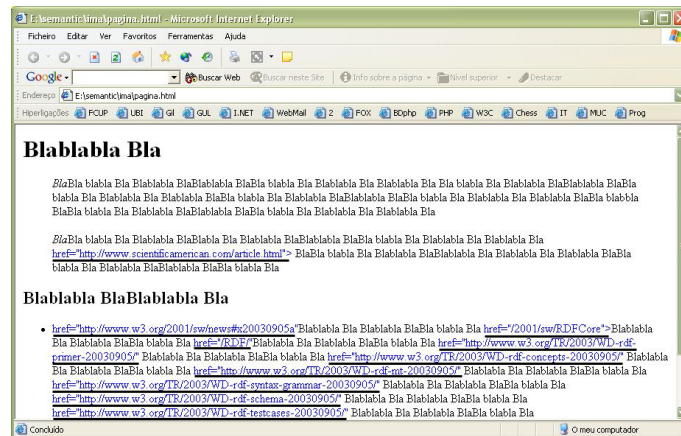


Figura 2.2: Como as ferramentas web interpretam as páginas web.

De acordo com Tim Berners-Lee *et al.* [1] e Ana Moura [58], os principais factores que contribuem para as actuais limitações da Web podem, resumidamente, ser agrupadas nos seguintes pontos:

- Constante aumento de publicações de documentos não estruturados na Web;
- Inexistência de uma classificação e integração universal da informação presente nos documentos publicados na Web de forma a permitir uma eficiente interpretação e manipulação por parte de SPC; e
- Reduzida autonomia por parte das actuais ferramentas web.

A identificação e o diagnóstico das actuais limitações da Web contribui, por um lado, para um endurecimento das críticas dos actuais utilizadores da Web, mas por outro lado, para motivar e estimular as actividades de I&D na procura de uma nova abordagem para a arquitectura da Web. Essa nova abordagem, deveria definir e implementar um modelo normalizado para

representar de forma fácil e universal os recursos disponibilizados na Web. Além disso, a interpretação e manipulação dos recursos definidos nesse modelo deveria ser facilmente conseguida quer pelos humanos quer de forma automática e autónoma pelos SPC.

## 2.5 Evolução da Web

A evolução da Web é aqui abordada, apenas, em termos do seu futuro próximo, ou seja, como a tentativa de ultrapassar as actuais limitações num curto prazo de tempo. Apesar de existirem visões sobre o futuro da Web a longo prazo [59], nesta dissertação não são apresentadas especulações de longo prazo sobre o futuro da Web por existir um extremamente elevado grau de incerteza na previsão da sua evolução a longo prazo.

Segundo Ana Moura [58], encontramos-nos no início de uma nova geração da Web. Ela é conhecida como a terceira geração da Web e é cognominada por Tim Berners-Lee de Web Semântica. A primeira geração da Web foi caracterizada pela criação manual de páginas web através da utilização da *Hypertext markup language*(HTML)<sup>34</sup> [60] e a segunda geração foi caracterizada pela criação automática<sup>35</sup> de páginas web. Estas duas gerações caracterizaram-se pelo facto de a informação disponibilizada na Web se encontrar definida com o único propósito de vir a ser interpretada pelos humanos. Esta última característica é o factor de ruptura entre as duas primeiras gerações e a terceira geração da Web. A Web Semântica, ou seja a terceira geração da Web, caracteriza-se por disponibilizar a informação de forma a poder ser facilmente interpretada e manipulada pelas pessoas, mas também e principalmente, pelos SPC.

Segundo Tim Berners-Lee *et.al.* [1], enquanto que a actual Web banalizou o intercâmbio e a interligação de documentos entre diferentes sistemas computacionais, a Web Semântica pretende

---

<sup>34</sup>É uma linguagem de marcação utilizada na criação de páginas web. A sua principal vantagem reside no facto de permite criar, de forma muito simples, ligações entre diversas páginas web.

<sup>35</sup>Através da utilização de aplicações servidoras e de linguagens de programação incorporadas no código HTML das páginas web, como por exemplo as linguagens *Active Server Pages* (ASP) [61], *PHP: Hypertext Preprocessor* (PHP) - originalmente conhecida como *Personal Home Page Tools* [62], *Java Server Pages* (JSP) [63], etc.

ir mais além, permitindo que os sistemas computacionais e outros tipos de dispositivos que se encontrem conectados partilhem, processem e interpretem a informação que se encontrar dentro dos próprios documentos. Esta extensão da Web permitirá a troca, manipulação e interpretação autónoma e eficiente, por parte dos SPC, da informação existente na Web. Os SPCs da Web Semântica serão capazes de reunir informação de documentos dispersos, processá-la e deduzir nova informação. Em síntese, a ideia da Web Semântica é a de aproveitar ao máximo os amplos recursos disponíveis na Web e transformar, os milhões de conceitos independentes que existem hoje, numa única e extensa rede de informação auto-descrita e ao mesmo tempo simplificar, automatizando, as interações através dessa mesma rede. Tim Berners-Lee, fundador e director do W3C, e a sua equipa de visionários, são de opinião de que a Web Semântica terá um impacto positivo para todos os seus utilizadores, quer para os utilizadores individuais com um simples computador pessoal, quer para as grandes empresas, organizações e instituições científicas [64].

## 2.6 Resumo

São poucos os fenómenos ocorridos na história da humanidade que tenham alterado significativamente o modo de vida de uma tão grande quantidade de pessoas num curtíssimo intervalo de tempo como o que se verificou com a Internet. Actualmente, ela influencia, a vários níveis, o comportamento de grande parte das sociedades existentes. A sua importância estratégica a nível mundial tornou-se inegável, vindo, cada vez mais, a ser considerada indispensável para o sucesso de algumas das actuais actividades humanas. Ela é considerada como sendo o maior repositório de informação a nível mundial. Disponibiliza diversas formas de acesso, actualização e distribuição de dados, sendo as mais utilizadas o correio electrónico e a Web.

A Web, por seu lado, representa um espaço virtual de troca e disponibilização de informação, o qual é proporcionado ao utilizador através de um ambiente simples e, normalmente, de forma gráfica. A sua história é curta em idade mas longa em persuasão. Só graças ao constante entusiasmo, principalmente de Tim Berners-Lee, é que a Web é hoje uma realidade. A incre-

dibilidade na Web era de tal maneira abrangente, que até mesmo a maior empresa de *software* a nível mundial, a Microsoft, por pouco não sofreu um rombo irremediável simplesmente por a ter menosprezado. Apesar disso, o crescimento rápido e desorganizado da Web tem dificultado a localização, acesso, apresentação e manutenção da informação. Mesmo utilizando poderosas ferramentas web nem sempre se consegue aceder, de forma fácil, à informação que procuramos. E quando isso acontece, é resultante de um esforço manual e individual suportado pelos utilizadores na triagem e selecção dos resultados devolvidos por essas ferramentas. Assim, a demora na obtenção de resultados úteis é um importante factor de insatisfação sentido pelos actuais utilizadores da Web.

A parcial desorganização da informação na actual Web e a frustração crescente dos utilizadores, resultante da baixa eficiência, em termos de informação considerada útil, das respostas retornadas pelas ferramentas web, estão na base de uma crescente necessidade em promover uma revolução na actual Web. Tim Berners-Lee, uma vez mais, aparece como o fundador da próxima geração da Web, à qual deu o nome de Web Semântica.

A Web Semântica caracteriza-se por disponibilizar informação que pode ser facilmente interpretada e manipulada quer por pessoas quer por SPC. A ênfase da Web Semântica é dirigida principalmente para a interpretação e manipulação das páginas web por parte dos SPC, com o objectivo de se construir uma Web mais automatizada e útil. No fundo, a Web Semântica pretende, apenas, expandir as actuais potencialidades da Web.

# Capítulo 3

## Web Semântica

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

Tim Berners-Lee, James Hendler, Ora Lassila,

"The Semantic Web", Scientific American, May 2001 [1]

### 3.1 Introdução

Ultimamente, observa-se uma intensa actividade na investigação e no desenvolvimento de ferramentas eficientes para a recuperação de informação e para a adição de informação semântica às páginas web. Por vezes, a Web Semântica é definida como sendo mais uma dessas técnicas, o que na realidade não é verdadeiro, uma vez que, ela é muito mais abrangente que isso. Apesar de, inspirada em modelos de IR e de KR, ela apresenta, por um lado, uma arquitectura robusta e descentralizada, e por outro, um modo de interacção completamente novo para a actual Web, fazendo com que seja defendida pelos seus mentores como a essência da terceira geração da Web. Além disso, ao contrário da maioria das actuais ferramentas para a Web, nas quais a importân-

### 3.1. INTRODUÇÃO

---

cia da interpretação da informação está do lado dos algoritmos utilizados por aplicações clientes, na Web Semântica a importância da interpretação da informação estará maioritariamente dependente da forma como a informação é especificada nas páginas web. Desta forma, não é de todo obrigatório que os SPC necessitem de ser extremamente complexos, será apenas suficiente que os conteúdos das páginas web possam ser facilmente processados de forma lógica pelos SPC.

A temática da Web Semântica pode ser considerada como simples mas, extensa e complexa de implementar. A sua objectiva implementação está directamente dependente da sua adopção por parte dos actuais utilizadores da Web. Mas, a ineficiente e inexacta divulgação da sua ideologia tem contribuído para um parco envolvimento dos actuais utilizadores e investigadores da Web na sua temática. Desta forma, torna-se urgente fomentar acções que permitam alterar a actual situação. O presente capítulo pretende ser uma valiosa contribuição nesse sentido, pois apresenta de uma forma muito simples e íntegra a temática da Web Semântica, contribuindo, assim, para a sua eficiente divulgação e rápida implementação. O conteúdo deste capítulo é parcialmente baseado no artigo [12].

O presente capítulo é constituído por sete secções. Nesta primeira secção é apresentada uma breve introdução à definição de Web Semântica que, por sua vez, será aprofundada durante o capítulo. É, também, apresentada uma descrição da estrutura temática de todo o capítulo. Na segunda secção, **Perspectiva Geral sobre a Web Semântica**, é apresentada uma descrição geral sobre a Web Semântica. Na terceira secção, **Representação da Informação e do Conhecimento**, é apresentado o modelo da Web Semântica como sendo mais abrangente que os actuais modelos de representação da informação e do conhecimento. Na quarta secção, **Princípios da Web Semântica**, são apresentados seis princípios considerados imprescindíveis para a implementação de forma consistente da Web Semântica. Na quinta secção, **Arquitectura Estratificada da Web Semântica**, é apresentado o modelo de camadas que é actualmente defendido como sendo a arquitectura estratificada da Web Semântica. É dada especial ênfase às normas e tecnologias associadas às três camadas que mais têm evoluído nos últimos tempos: a XML, a sintaxe/modelo RDF

e as ontologias para a Web Semântica. Na sexta secção, *A Imagem da Web Semântica*, é apresentada a Web Semântica como sendo o resultado da combinação de um conjunto alargado de diferentes componentes mutuamente dependentes. Por último, é apresentado na sétima secção um pequeno resumo acerca do tema abordado neste capítulo.

### 3.2 Perspectiva Geral sobre a Web Semântica

Segundo Tim Berners-Lee *et.al.* [1], a Web é definida como sendo uma simples e excelente plataforma mundial de interligação e transferência de documentos entre sistemas computacionais heterogéneos. Mas, apesar disso, é extremamente ineficiente quando se pretende a interligação ou a transferência, de uma forma simples e autónoma, dos conceitos definidos dentro dos próprios documentos. Outra das suas limitações reside no facto de a informação existente na Web ser desenvolvida e preparada tendo em conta, exclusivamente, a sua compreensão pelos humanos. Desta forma, podemos resumir que a fonte de grande parte das limitações da Web reside nos seus próprios documentos, mais especificamente na inexistência de uma estrutura universal e de uma definição, também universal, relativamente ao tipo de informação neles contida. Assim, a forma como actualmente os dados/informação são definidos e estruturados na Web pode ser, na melhor das hipóteses, considerada como caótica. Conforme referido no Capítulo 2, foi a partir da consciencialização destas limitações que surgiu a necessidade de desenvolver uma nova arquitectura para a Web. Esta nova arquitectura permitirá tornar a Web num meio de colaboração por excelência quer entre as pessoas quer entre os SPC. Através da utilização desta arquitectura, os SPC, facilmente, conseguirão analisar a informação existente na Web, o seu conteúdo, as respectivas ligações e estabelecer relações com outros SPC e/ou com os humanos. Segundo Tim Berners-Lee [65] e Sandro Hawke [66], a materialização desta arquitectura está dependente da utilização de uma semântica que atribua um significado bem definido à informação disponibilizada na actual Web. Neste contexto, a informação semântica significa informação que pode ser interpretada e processada pelos SPC. Desta forma, a utilização de informação semântica firmará uma forte colaboração entre as pessoas e os SPC, pois

### 3.2. PERSPECTIVA GERAL SOBRE A WEB SEMÂNTICA

---

permitirá a interpretação, manipulação e a troca de informação entre ambos. Assim, uma das condições chave para o sucesso desta nova arquitectura da Web é a de que a colaboração venha a ser de carácter universal, ou seja, independente de raças, culturas, línguas e comunidades existentes. Esta nova visão<sup>36</sup>, do funcionamento da Web, foi apelidada de Web Semântica por Tim Berners-Lee e pelo consórcio W3C.

Segundo Tim Berners-Lee *et.al.* [1] e Marja-Riitta Koivunen e Eric Miller [67], o principal objectivo da Web Semântica é o de acoplar à actual Web um modelo universal de partilha e de definição da informação, a qual poderá ser facilmente definida e relacionada de maneira a ser, também facilmente, interpretada e processada por SPC. Estas novas funcionalidades não se limitarão à simples formatação da informação a ser apresentada aos humanos<sup>37</sup>, mas sobretudo à disponibilização de mecanismos que facilitarão a integração e a reutilização de forma autónoma da informação por parte de outras e possivelmente diferentes ferramentas web. A partir do momento em que os SPC consigam interpretar o conteúdo de qualquer documento, podem automaticamente filtrar a sua informação, e assim, executar facilmente tarefas que são actualmente consideradas extremamente complexas<sup>38</sup>. Este processo será progressivo, começando pelo desenvolvimento de SPC que ainda necessitarão de, prévias, avaliações humanas até ao desenvolvimento de SPC que atingirão elevados níveis de autonomismo e automatismo.

Segundo Tim Berners-Lee [68], é importante clarificar que a Web Semântica não é o mesmo que a Inteligência Artificial (IA), ou seja, o objectivo não é fazer com que os computadores compreendam a linguagem humana. Processar e relacionar informação na Web Semântica não significa um procedimento inteligente no mesmo sentido que lhe é atribuído pelos investigadores da área da IA. O objectivo que se pretende atingir com a Web Semântica é o de especificar um modelo universal para a definição da informação e um conjunto de regras de inferência que

---

<sup>36</sup>Também designada por revolução, ou simplesmente por uma extensão da actual Web

<sup>37</sup>Como acontece na maioria das ferramentas web existentes.

<sup>38</sup>Os actuais Motor de Busca não apresentam respostas minimamente aceitáveis para questões complexas, como por exemplo: *Tell me what wines I should buy to serve with each course of the following menu. And, by the way, I don't like Sauternes*[69].



### 3.2. PERSPECTIVA GERAL SOBRE A WEB SEMÂNTICA

permita que os SPC possam facilmente manipular essa informação da mesma forma como se verdadeiramente a compreendessem.

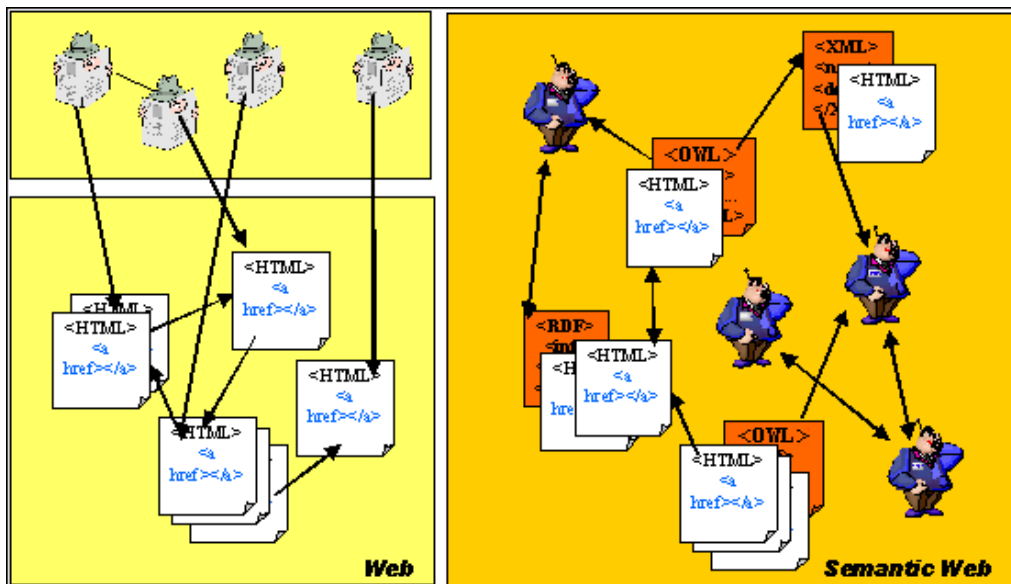


Figura 3.1: Ferramentas web na Web e SPC na Web Semântica (adaptado de [70]).

Ainda, segundo Tim Berners-Lee *et.al.* [71, 72], a Web Semântica não se resume apenas ao somatório de todas as páginas web mais as suas ligações<sup>39</sup>, mas, e principalmente, à definição de uma camada<sup>40</sup> constituída pelas relações existentes entre recursos<sup>41</sup> e na definição universal de cada recurso através da utilização de ontologias. O termo ligação é utilizado para definir uma conexão entre dois documentos, como acontece na Web. O termo relação, utilizado na Web Semântica, também define uma conexão entre dois documentos ou recursos, mas, neste caso, esta conexão pode, e deverá, ser tipificada.

A figura 3.1 ilustra, de uma forma muito simplista, o modelo de interacção existente entre as entidades constituintes da Web e da Web Semântica. Ela encontra-se dividida em duas partes, esquerda e direita. A parte da esquerda da figura está por sua vez dividida em duas secções,

<sup>39</sup>Como é o caso da actual Web.

<sup>40</sup>Baseada num modelo universal de representação de recursos, o qual será apresentado na secção 3.5 dedicada à apresentação da arquitectura estratificada da Web Semântica.

<sup>41</sup>Na Web Semântica podemos considerar como recurso: os documentos web, os documentos físicos, as pessoas, as cidades, isto é, tudo aquilo que nos rodeia.

### 3.2. PERSPECTIVA GERAL SOBRE A WEB SEMÂNTICA

---

superior e inferior. A secção superior engloba as ferramentas web<sup>42</sup> e a secção inferior engloba as páginas web. A parte da direita da figura engloba os SPC e as páginas web da Web Semântica. A interacção entre as diversas entidades na Web e na Web Semântica são representadas por setas, as quais podem ser unidireccionais ou bidireccionais.

Na parte esquerda da figura 3.1, a comunicação existente entre as ferramentas web e as páginas web é caracterizada como sendo maioritariamente unidireccional, ou seja, as ferramentas web acedem às páginas web. Como a informação contida nas páginas web é heterogénea e não está definida segundo uma estrutura normalizada, as ferramentas web têm que colmatar essas deficiências para serem capazes de obter resultados considerados úteis. Mas a diversidade de dados e de tipos de apresentação desses dados é de tal modo colossal que as ferramentas web necessitam de um esforço hercúleo para conseguirem extrair, de forma eficiente, informação considerada útil a partir das páginas web. Além disso, as diversas ferramentas web baseiam-se em diferentes, e por vezes opostos, paradigmas o que leva a que não consigam interagir e conjugar esforços entre si. Este tipo de sistema<sup>43</sup> em que as páginas web não evoluem e pouco contribuem para a definição da informação que comportam, leva a que as ferramentas web necessitem de ser cada vez mais complexas mas em contrapartida apresentem melhorias pouco significativas em termos de resultados úteis.

Na parte direita da figura 3.1, a comunicação existente entre os SPC e as páginas da Web Semântica é caracterizada como sendo bidireccional. Podemos verificar que não existe uma separação rígida entre os SPC e as páginas da Web Semântica. Nesta solução, as páginas da Web Semântica contribuem positivamente para a estruturação e definição da informação que contêm. Deste modo, como a informação se encontra auto-especificada<sup>44</sup>, os SPC já não necessitam de serem tão complexos como no caso das ferramentas web apresentadas no parágrafo

---

<sup>42</sup>Tais como Navegadores Web, Motores de Busca, Directórios Web, etc.

<sup>43</sup>A actual Web.

<sup>44</sup>Através da utilização da especificação XML, do modelo RDF e de linguagem ontológica OWL. Este grupo de tecnologias será apresentado, em maior pormenor, na secção 3.5 dedicada à arquitectura estratificada de tecnologias da Web Semântica.

### 3.2. PERSPECTIVA GERAL SOBRE A WEB SEMÂNTICA

---

anterior, pois não necessitam de manipular e interpretar a informação existente no seu estado bruto. Enquanto que na actual Web, as ferramentas web necessitam de ser muito complexas para interpretar uma grande diversidade de dados não-estruturados, na Web Semântica os SPC serão muito mais simples e muito mais eficientes<sup>45</sup>. A obtenção de melhores resultados por parte dos SPC prender-se-á com o menor esforço necessário para interpretar os recursos da Web Semântica. Contudo, deve ser enfatizado que, a Web Semântica não é uma nova linguagem para escrever e descrever a informação presente nas páginas na Web. Se assim fosse cairíamos no problema da limitação das linguagens. O que a Web Semântica apresenta é apenas um modelo para estruturar, auto-descrever e inter-relacionar a informação de uma forma universal.

Os mentores do projecto inicial da Web Semântica destacam a universalidade da actual Web como uma das suas propriedades vitais e almejam que a Web Semântica também venha a ser o mais descentralizada possível. Segundo Ana Moura [58], embora seja muito promissora, a implementação da Web Semântica é de enorme complexidade. Os principais desafios para a sua implementação são os seguintes:

- Estabelecer normalizações que definam de uma forma inteligente, simples e universal a informação presente nas páginas web, facilitando assim uma posterior pesquisa e interpretação;
- Desenvolver SPC que consigam, de forma autónoma/automática, interpretar e partilhar informação obtida a partir de diversas e diferentes origens;
- Desenvolver SPC que consigam, de forma autónoma/automática, gerar informação nova e útil a partir da informação recebida.

Quando estes desafios forem alcançados as pesquisas realizadas na Web serão mais precisas e profundas, e a veracidade e a validade da informação encontrada será mais facilmente testada.

---

<sup>45</sup>Pois apenas terão que suportar parte do trabalho de classificação e estruturação da informação, e isto, a um nível universal, porque as páginas web da Web Semântica contribuirão para a simplificação de todo o processo.

## 3.3 Representação da Informação e do Conhecimento

Um dos principais objectivos da Web Semântica é o de permitir que os SPC tenham acesso a dados estruturados, os consigam interpretar, e os possam utilizar de uma maneira eficiente e fiável. A utilização de regras de inferência, especificadas através de asserções<sup>46</sup> RDF e da utilização de linguagens ontológicas<sup>47</sup>, tornará possível a representação da informação de forma explícita. Só assim, a informação poderá ser facilmente manipulada pelos SPC, mas, para isso, é necessário definir um modelo simples e universal que permita representar a informação e por sua vez o Conhecimento<sup>48</sup> existente na Web actual. Há já algum tempo que se têm vindo a desenvolver diversos modelos de KR, sendo dois exemplos o *Knowledge Interchange Format* (KIF) [73] e o *Cycorp* (Cyc) [74]. Estes modelos de KR, actualmente classificados como tradicionais, são centralizados o que obriga a que os seus utilizadores apenas possam usar definições de conceitos pré-definidos. Por exemplo, a ontologia Cyc sugere que todos os seus utilizadores concordem com os mesmos termos para as palavras inglesas comuns, o que não será de todo praticável na Web Semântica. Segundo Ana Moura [58], o controlo centralizado tem muitas desvantagens. Além de ser muito rígido, possui um número limitado de tipos de interrogações que podem ser realizadas de modo a que o sistema possa responder correctamente. Além disso, cada sistema de KR ajusta as suas próprias regras e idiossincrasias para fazer a inferência da sua informação. Nestas situações, mesmo quando é possível transferir a informação existente de um sistema para outro sistema, geralmente as regras associadas à informação do primeiro sistema, porque se encontram definidas em formatos específicos para a aplicação, raramente podem ser aproveitadas no segundo sistema. De acordo com Tim Berners-Lee *et.al.* [1], Ana Moura [58] e Tim Berners-Lee [65], o modelo pretendido para a Web Semântica é diferente dos anteriores sistemas, porque é mais abrangente na medida em que permite o relacionamento universal entre conceitos de áreas de conhecimento diferentes. Desta forma, ao contrário dos sistemas ditos tradicionais, onde um conceito possui uma única localização num sistema de conhecimento, na

---

<sup>46</sup>Uma asserção é uma frase declarativa empregue para afirmar ou negar algo.

<sup>47</sup>Como por exemplo a OWL, que será apresentada em maior pormenor na subsecção 3.5.3 - Ontologias; e no apêndice B - OWL.

<sup>48</sup>A Representação do Conhecimento é uma das área de investigação da actual IA.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

Web Semântica os conceitos serão definidos de forma independente por toda a Web. Neste tipo de modelo, qualquer pessoa poderá, de uma forma fácil, contribuir com a especificação da sua própria informação. Para isso, os investigadores da Web Semântica terão de desenvolver regras, o mais expressivas possíveis, que permitam que a Web Semântica venha a ter uma distribuição o mais abrangente possível, aceitando, assim, a possibilidade de paradoxos e a existência de perguntas sem resposta, como uma condição indispensável para se poder atingir uma maior versatilidade. Para isso, pretendem adoptar, uma filosofia similar à que foi adoptada pela Web: disponibilização dispersa de dados<sup>49</sup> e a utilização de ferramentas que permitam em tempo real uma pesquisa da informação disponibilizada.

Nos modelos ditos tradicionais, cada campo do conhecimento possui a definição dos seus conceitos centralizada, se não em termos filosóficos, pelo menos nas suas implementações práticas, o que os impede de se disseminarem globalmente. Segundo Tim Berners-Lee [65], a área de KR tem a reputação de ser muito interessante mas limitada a pequenos e bem definidos projectos, ficando muito aquém de atingir resultados práticos quando a área de acção é mais alargada. Isto era exactamente o que se passava com o Hipertexto e o Hipermedia antes do aparecimento da Web, uma vez que eram apenas considerados úteis se a sua utilização se limitasse a aplicações pequenas e especializadas. Pelo contrário, actualmente o Hipertexto e o Hipermedia são omnipresentes na maioria das tecnologias web. Da mesma forma, a Web Semântica será o que obteremos se globalizarmos a KR. Para que isso seja possível e para que possamos extrair todo o potencial da Web Semântica é necessário remover definitivamente os conceitos centralizados e de confiança absoluta.

A informação, na Web Semântica, não será descrita utilizando uma linguagem natural mas mantida numa estrutura lógica a partir da qual será facilmente interpretada e manipulada pelas pessoas e pelos SPC. Essa estrutura descreverá a informação através da utilização e da

---

<sup>49</sup>Mas neste caso estruturados e associados a significados possíveis de serem facilmente interpretados pelos SPC.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

combinação de asserções descritivas entre dois ou mais recursos. É claro que um nome, por exemplo Rui Pereira, não é muito adequado para identificar uma pessoa em todo o mundo. Por isso, na Web Semântica, os recursos e as relações entre recursos serão identificados de maneira única através de um identificador unívoco chamado *Uniform Resource Identifier* (URI)<sup>50</sup> [75]. Desta forma, a gigantesca quantidade de informação que possa ser expressa através de uma asserção descritiva, como por exemplo: Carla é a mãe da Joana, Joana vive na Maia, Peça n.º.123 custa 12 euros, O presidente da Republica Portuguesa é o Prof. Doutor Aníbal Cavado Silva, poderá ser representada de uma forma única e precisa na Web Semântica.

Segundo Sandro Hawke [66], a forma mais agradável para descrever estas asserções descritivas, de uma forma única e interpretável pelos SPC, é através da utilização da arquitectura RDF, a qual é uma Recomendação W3C<sup>51</sup> de 22 de Fevereiro de 1999 [76], sendo a sua última versão designada por arquitectura RDF Revista que, por sua vez, é uma Recomendação W3C de 10 de Fevereiro de 2004 [4] mas, as suas raízes são muito mais antigas. A arquitectura RDF é composta pelo Modelo RDF e pela Sintaxe RDF<sup>52</sup>. O Modelo RDF não é mais que um modelo muito simples que permite definir as relações existentes entre vários recursos através da utilização de propriedades. A Sintaxe RDF, por sua vez, é constituída por um conjunto de instruções<sup>53</sup> definidas em XML que podem ser utilizadas para descrever sintacticamente o Modelo RDF. Uma das vantagens da utilização de asserções no formato RDF prende-se com o facto de serem facilmente convertidas para outras linguagens, alargando desta forma o leque de informação compreensível quer pelas pessoas quer pelos SPC. Desta forma, apesar de na

---

<sup>50</sup>Também conhecido como *Universal Resource Identifier*. É um conjunto de caracteres utilizado na Web Semântica para identifica univocamente um recurso. Um recurso na Web Semântica é qualquer entidade que possa ser referenciada. O *Uniform Resource Locator* (URL) é o sub-tipo mais conhecido do URI, o qual é utilizado como um apontador para localizar informação na Internet. Outro exemplo de um sub-tipo de URI é o *Uniform Resource Name* (URN), o qual é utilizado para definir localizações de informação na Web de forma persistente.

<sup>51</sup>Uma Recomendação W3C é interpretada, quer pela industria, quer pela comunidade web, como sendo um sinónimo de normalização para a Web. Cada Recomendação W3C não é mais do que uma especificação estável desenvolvida por um Grupo de Trabalho W3C (*W3C Working Group*) e revista pelos membros do W3C. Este tipo de recomendações promove a interoperabilidade de tecnologias web a partir de consensos obtidos entre a industria e a academia que se encontram representados em Grupos de Trabalho do W3C.

<sup>52</sup>O Modelo e a Sintaxe RDF serão apresentados em maior pormenor na secção 3.5 deste capítulo.

<sup>53</sup>Instruções que serão apresentadas em maior pormenor no apêndice A - Sintaxe RDF/XML.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

futura Web Semântica muita da informação continuar a ser expressa através da utilização de linguagens naturais, a sua grande maioria será representada através de asserções RDF<sup>54</sup>.

As asserções RDF podem ser escritas através da XML. A XML é simplesmente um conjunto de regras e convenções para a escrita de informação estruturada em simples ficheiros de texto. É considerada como sendo o suporte mais utilizado na partilha e troca de informação entre diferentes organizações em todo o mundo [77]. Além disso, é uma Recomendação W3C de 10 de Fevereiro de 1998 [78], sendo a sua terceira e última versão uma Recomendação W3C de 4 de Fevereiro de 2004 [3]. As suas principais vantagens são:

- Independência: É independente, quer das aplicações quer dos sistemas operativos; Simplifica a criação de documentos e de informação independente das aplicações;
- Estruturação: Minimiza o problema da falta de estrutura na informação presente nos actuais documentos web; Todos os documentos XML organizam a informação que possuem numa estrutura em forma de árvore; Este tipo de estrutura possibilita uma fácil manipulação e validação do seu conteúdo;
- Extensibilidade: Não são definidos elementos de marcação à priori; Cada utilizador possui a liberdade de criar os seus próprios elementos;
- Desenvolvimento de aplicações mais flexíveis: Os documentos XML são facilmente perceptíveis pelas pessoas e também facilmente processáveis e interpretáveis pelos SPC;
- Várias formas de visualização: Separa a informação da sua possível forma de formatação, desta forma um mesmo documento em XML pode ser facilmente utilizado em centenas de formatações diferente<sup>55</sup>;
- Suporte: Existem centenas de linguagens baseadas em XML<sup>56</sup>.

---

<sup>54</sup>Ou através de um modelo que lhe seja equivalente e desta forma facilmente interpretável pelos SPC.

<sup>55</sup>Por exemplo, nos formatos: *EXtensible HyperText Markup Language* (XHTML) [79], *Portable Document Format* (PDF) [80], *PostScript* [81], *Scalable Vector Graphics* (SVG) [82], *Management Information Format* (MIF) [83], *Printer Control Language* (PCL) [84], etc.

<sup>56</sup>Algumas delas serão apresentadas a título de exemplo na subsecção 3.5.1 - XML.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

Segundo Renato Iannella e Andrew Waugh [85], os documentos XML são estruturados através da utilização de anotações<sup>57</sup> genéricas<sup>58</sup>. Essas anotações são geralmente utilizadas para associar outros dados aos dados que são englobados pela anotação. Este tipo especial de dados, que se encontram associados a outros dados como o objectivo de os classificar, denominam-se meta-dados<sup>59</sup>. A utilização de meta-dados tem um impacto positivo na pesquisa e descoberta de informação na Web. A sua utilização tem variadas aplicações que vão desde um resumo do significado da informação até ao controlo do acesso<sup>60</sup> à própria informação. Eles podem ser utilizados para descrever qualquer tipo de recurso para a Web Semântica: documentos web, imagens de vídeo, áudio, etc. Basicamente, os meta-dados permitem descrever o tipo de atributos de um recurso, por exemplo, título, autor, etc. Cada atributo possui um ou mais valores. De acordo com Renato Iannella e Andrew Waugh [85], o modelo básico utilizado para a definição de meta-dados é conhecido como: tipo de atributo e valor. Além disso, os meta-dados para poderem ser utilizados com eficácia em grande escala tem que ser normalizados, isto é, é necessário definir um conjunto de atributos que podem e devem ser utilizados para descrever os recursos. Em geral essa normalização define [85]:

- O tipo de dados que pode ser utilizada na descrição;
- Quais são os atributos obrigatórios e quais são os opcionais;
- O que significa, precisamente, cada atributo;
- Regras sobre o formato e a utilização de valores.

Têm vindo a ser propostos diversos tipos de vocabulários<sup>61</sup> de meta-dados para a classificação e estruturação das páginas web, sendo o mais conhecido o "Dublin Core Metadata Initiative"[86],

---

<sup>57</sup>Uma anotação, também conhecida pelo termo inglês *tag*, é composta por uma marcação de início (e.g. <img>) e uma marcação de fim (</img>), sendo o nome que aparece dentro dos sinais < e > o nome da anotação, que neste caso é *img* e representa a inserção de uma imagem na notação da linguagem HTML. O nome das anotações é geralmente utilizado como meta-dados.

<sup>58</sup>Na Especificação XML, o nome a dar à anotação não está sujeito a nenhum tipo de limitações, podendo ser um qualquer conjunto de caracteres.

<sup>59</sup>O termo utilizado na língua inglesa: *metadata*.

<sup>60</sup>Como acontece por parte dos pais que pretendem limitar o acesso a alguns tipos de informação na Web aos seus filhos menores de idade.

<sup>61</sup>Normalizados ou em vias de normalização.



### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

também conhecido por apenas *Dublin Core*. O *Dublin Core* define quinze elementos de meta-dados que representam os tipos de atributos possíveis de utilizar na classificação de um documento:

Título	Colaboradores	Fonte
Autor ou criador	Data	Língua
Assunto e palavras chaves	Tipo de recurso	Relação
Descrição	Formato	Domínio
Editor	Identificador do recurso	Direitos de autor

Na descrição de um documento, os elementos *Dublin Core* podem ser repetidos e são todos de utilização opcional. Os meta-dados, deste vocabulário, são geralmente classificados como descritivos, porque são exteriores ao significado dos recursos e encontram-se mais relacionados com a forma como os recursos foram criados.

De acordo com Renato Iannella e Andrew Waugh [85], os meta-dados não são uma coisa nova. Os bibliotecários catalogam livros e jornais desde há centenas de anos. E, de facto, a catalogação realizada pelos bibliotecários é, efectivamente, meta-dados, pois pode ser posteriormente utilizada para procurar livros e jornais a partir de um determinado assunto.

Segundo Tim Berners-Lee *et.al.* [1], a XML, actualmente largamente utilizada, tem a grande potencialidade de permitir aos utilizadores especificarem uma estrutura arbitrária para os seus dados durante a criação das suas páginas web, mas, em contrapartida, essa especificação não atribui nenhum significado aos dados. Por outro lado, a arquitectura RDF, através da utilização da sintaxe XML, especifica um modelo para a representação das relações existentes entre os diferentes recursos, embora, tal como a XML, também não atribua nenhum significado aos recursos nem especifica como é que eles possam ser interpretados.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

Infelizmente, não existe um esquema/vocabulário de meta-dados que satisfaça as diferentes comunidades existentes, e por isso a ênfase das actividades de I&D actuais em meta-dados é focada na procura de uma forma de definir as relações existentes entre diferentes esquemas/vocabulários de meta-dados. É aqui que entram as ontologias. Segundo Renato Iannella e Andrew Waugh [85], uma ontologia estabelece uma terminologia acerca das relações existentes entre os membros de uma determinada comunidade ou área do conhecimento. Mas, apesar de uma ontologia descrever o significado dos recursos e das suas relações, ela não define como os recursos e as relações deverão ser utilizados.

Segundo John Sowa [87], a normalização das notações XML e RDF foi o primeiro passo para a construção de um modelo uniforme de partilha de informação. Apesar disso, é ainda um resultado insuficiente pois não disponibiliza formas de comparação, relacionamento ou tradução de vocabulários. Segundo Simon Phipps [88], a XML fornece o meio para expressar a nossa compreensão acerca do significado da informação, mas não é suficiente para discernir entre diferentes realidades e conceitos quando pretendemos utilizá-lo no processo de troca de informação. Desta forma, mais importante do que normalizar vocabulários é necessário desenvolver métodos que permitam definir e traduzir vocabulários. Para se obter uma semântica eficiente e um pragmatismo sadio é necessário desenvolver métodos que relacionem os termos existentes nos vocabulários com os recursos a que se referem e com os SPC, de forma a que possam ser utilizados de forma útil.

De acordo com Tim Berners-Lee *et.al.* [1] e Ana Moura [58], a Web Semântica pode ser considerada como a composição de um grande número de pequenos componentes ontológicos que apontam entre si. Dessa forma, empresas, instituições e grupos de interesse específico procurarão ter os seus recursos web ligados a conteúdos ontológicos, já que serão disponibilizadas ferramentas com o objectivo de partilhar e combinar essa informação entre as ferramentas web. Assim, a Web Semântica além de espelhar uma nova gestão no domínio da informação, também poderá contribuir para a evolução do próprio conhecimento. Esses conteúdos ontológicos

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

poderão ser definidos através de ontologias, as quais, são entidades de alto nível que permitem aos SPC:

- Facilmente relacionar e manipular informação proveniente de diferentes fontes de meta-dados<sup>62</sup>;
- Compreender as relações existentes entre os recursos e os meta-dados;
- Suportar, de forma simples, a evolução da informação e dos esquemas/vocabulários de meta-dados;
- Criar uma interligação mundial entre os recursos e os meta-dados<sup>63</sup>.

A utilização de regras de inferência nas ontologias é de extrema importância, porque apesar de os SPC, na realidade, não compreenderem a informação representada pelas ontologias, poderão facilmente e eficientemente manipular essa informação de uma forma útil para os seus utilizadores. De acordo com Tim Berners-Lee *et.al.* [1], os SPC da Web Semântica irão possibilitar a avaliação da informação através de inferências lógicas obtidas a partir do processamento de regras e de conhecimento especificado em ontologias. Além disso, poderão ser configurados para se comportarem com diferentes graus de cepticismo ou optimismo relativamente a toda a informação que não seja assinada digitalmente e em que a sua fonte não tenha sido verificada em termos de autenticidade.

De acordo com Paul Ford [89], Tim Berners-Lee *et.al.* [71, 72], a Web Semântica não é mais do que a descrição de recursos e das suas relações de uma forma universal, através da utilização de um modelo lógico relacionado com as tecnologias XML, RDF e as ontologias, o qual será facilmente interpretado pelos SPC. Afirmções lógicas do tipo *Se A é irmão de B, então B é irmão de A* serão facilmente analisadas, combinadas, processadas e interpretadas pelos SPC na Web Semântica. Um processo típico envolverá a criação de uma cadeia de valor onde sub-conjuntos

---

<sup>62</sup>Quer sejam heterogéneas ou distribuídas.

<sup>63</sup>Criar um espaço simples e único de interacção mundial, i.e. a Web Semântica.

### 3.3. REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

---

de informações serão trocados, sempre que necessário entre SPC, permitindo desta forma que cada um desses SPC possa adicionar valor na construção de um resultado final que satisfaça um pedido complexo realizado por um utilizador. Na criação autónoma de complexas cadeias de valor, alguns SPC explorarão tecnologias baseadas em diversas outras tecnologias e áreas de investigação, como por exemplo a IA<sup>64</sup>.

É interessante imaginar uma rede virtual de SPC interagindo sobre a Web Semântica, onde cada um deles pudesse cooperar facilmente com os outros<sup>65</sup>, formando assim uma valiosa cadeia de valor mundial que elevará a capacidade da comunicação e da acção humana a um novo zénite na história da humanidade. Nessa situação, a autenticidade e a confiança das fontes de informação adquirirá um novo significado quando os SPC automaticamente e autonomamente, através da manipulação e do raciocínio sobre a informação, conseguirem chegar a conclusões que afectem a acção humana de forma positiva. Pois, de acordo com o resultado obtido relativo à realização de um teste de autenticidade a uma assinatura digital, um SPC poderá alterar o grau de certeza que associa ao resultado do seu raciocínio: certeza total; certeza parcial; ou mesmo ignorar a informação, se a fonte de informação não for considerada fidedigna. Contudo, o verdadeiro poder da Web Semântica só será perceptível quando realmente apresentar um valor real para os utilizadores finais. Desta forma, é necessário que sejam criados SPC que facilitem e apoiem os utilizadores a entrar no mundo da Web Semântica. Segundo Tim Berners-Lee *et.al.* [1], as principais vantagens<sup>66</sup> obtidas pela implementação do modelo de representação da informação e do conhecimento da Web Semântica podem ser agrupar nos seguintes pontos:

- A informação nas páginas web poderá ser descrita de forma facilmente interpretável e universal. Para isso, o modelo actual da Web Semântica baseia-se na XML, no Modelo e na Sintaxe RDF e na ontologia OWL;

---

<sup>64</sup>A IA representa uma colaboração muito importante porque ela fornece as fundações e a estrutura básica que torna praticável a combinação de algumas das tecnologias utilizadas pela Web Semântica.

<sup>65</sup>Cada SPC poderá publicitar as suas funcionalidades ou procurar um SPC que ofereça determinadas funcionalidades.

<sup>66</sup>As quais foram referidas ao longo desta secção.

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

---

- A informação poderá ser facilmente associada a conceitos universais e precisos. Existirá uma maior facilidade de relacionar a informação existente na Web a conceitos que se encontram descritos através de estruturas universais e regras precisas;
- As páginas web elaboradas segundo o modelo da Web Semântica permitirão aos SPC retirar maior conhecimento sinérgico da informação disponibilizada na Web. Desta forma, os SPC poderão de forma simples, automática e autónoma realizar tarefas sofisticadas a pedido dos utilizadores;
- A Procura de informação será mais eficiente. Os SPC conseguirão interpretar e manipular a informação, não só de uma forma isolada, mas também através de uma forma conjugando várias páginas web;
- A Web Semântica tornar-se-á menos impessoal, mais útil e de maior confiança. Qualquer utilizador poderá especificar as suas preferências pessoais e os seus níveis de confiança relativamente à informação disponibilizada.

## 3.4 Princípios da Web Semântica

A Web Semântica é uma iniciativa do consórcio W3C [2], que apresenta como principal objectivo estender a actual Web de forma a que ela possibilite a implementação de serviços automáticos e autónomos de interpretação da informação, uma organização descentralizada e mecanismos de verificação e de validação da confiança associada à informação. Este consórcio apresenta seis princípios que considera imprescindíveis para a sustentação de uma forma consistente da ideologia Web Semântica:

- Primeiro Princípio: Tudo pode ser identificado através de um URI;
- Segundo Princípio: Os recursos, e as relações, podem ser tipificados;
- Terceiro Princípio: Tolerância a quebras de relações;
- Quarto Princípio: Desnecessidade da existência de uma confiança absoluta;

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

---

- Quinto Princípio: Permitir evolução;
- Sexto Princípio: Implementação minimalista.

Estes princípios serão descritos em detalhe nas seguintes sub-secções.

#### 3.4.1 Primeiro Princípio: Tudo pode ser Identificado Através de um URI

Segundo Aaron Swartz [90], todos os documentos existentes na Web bem como as pessoas, os livros, os locais, os objectos do mundo físico e até conceitos abstractos<sup>67</sup>, são designados na Web Semântica por recursos. Como, sempre que queremos falar de ou referirmo-nos a um recurso temos que o identificar primeiro, na Web Semântica, um recurso deverá ser sempre referenciado através um identificador único. A este identificador único de um recurso dá-se o nome de URI, o qual não é mais que um conjunto de caracteres que identifica univocamente um recurso. Na Web Semântica, as relações existentes entre recursos são, também, consideradas como recursos e desta forma devem também ser referenciadas através de URIs.

A definição e atribuição de URIs é flexível e descentralizada: qualquer entidade<sup>68</sup> pode definir os seus próprios URIs para identificar de forma única os seus recursos, ou simplesmente utilizar URIs definidos por outras entidades. Mas, nem sempre é necessário atribuir um URI para cada recurso, podemos definir recursos indirectamente, ou seja, a partir de outros recursos que estejam identificados por um URI. Por exemplo, na frase seguinte podemos nos referir a uma pessoa cujo nome é Rui Pereira utilizando, apenas, o URI do seu endereço do correio electrónico:

A pessoa cujo endereço de correio electrónico é `rpereira@di.ubi.pt`  
chama-se Rui Pereira.

Desta forma, é possível especificar características acerca de uma pessoa sem haver a necessidade de lhe atribuir vários URIs. Outro exemplo, é a possibilidade de identificar uma instituição de

---

<sup>67</sup>Como as ideias em geral.

<sup>68</sup>Que pode ser uma pessoa ou uma organização.

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

---

ensino, como por exemplo a UBI, atribuindo um URI igual ao endereço da sua página na web, como se pode ver na figura 3.2.



Figura 3.2: Recursos do mundo físico e os seus URIs na Web Semântica.

Cada vez mais, a descrição dos recursos e das suas relações é realizada através da utilização de simples vocabulários normalizados de termos<sup>69</sup> ou através de vocabulários e estruturas complexas que normalmente se englobam nas chamadas ontologias<sup>70</sup>.

Ainda segundo Aaron Swartz [90] a criação descentralizada de URIs, apesar de possuir inúmeras vantagens, levanta também alguns problemas. Como qualquer entidade pode definir os seus próprios URIs, existe uma grande probabilidade de aparecerem diferentes URIs a identificar exactamente o mesmo recurso ou de um mesmo URI poder vir a ser definido para identificar diferentes recursos. Para minorar este problema de possibilidade de existência de ambiguidades na atribuição e na utilização de URIs, podem e devem ser utilizados Espaços de Nomes<sup>71</sup>. Os Espaços de Nomes não são mais do que a simples atribuição e associação de um URI a um vocabulário ou a uma ontologia. Desta forma, um Espaço de Nomes é composto por duas partes separadas pelo símbolo ":" como por exemplo:

`nsprefix:localname`

---

<sup>69</sup>O vocabulário mais conhecido e já referido anteriormente é o *Dublin Core* [86], o qual descreve o tipo de elementos que compõem um documento.

<sup>70</sup>Que serão apresentas na secção 3.5 - "Arquitectura Estratificada de tecnologias da Web Semântica".

<sup>71</sup>Termo utilizado em língua Inglesa: *Namespaces*.

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

---

Neste exemplo, `ns` prefix representa o URI do Espaço de Nomes e `localname` representa os termos utilizados nesse vocabulário ou ontologia.

#### 3.4.2 Segundo Princípio: Os Recursos e as Relações Podem ser Tipificados

Segundo Marja-Riitta Koivunen e Eric Miller [67], os `objectos web`<sup>72</sup> são criados tendo como única preocupação a sua interpretação por parte dos humanos. Desta forma, não costumam conter informação que especifique a sua finalidade ou o tipo de relações<sup>73</sup> existente com outros `objectos`.

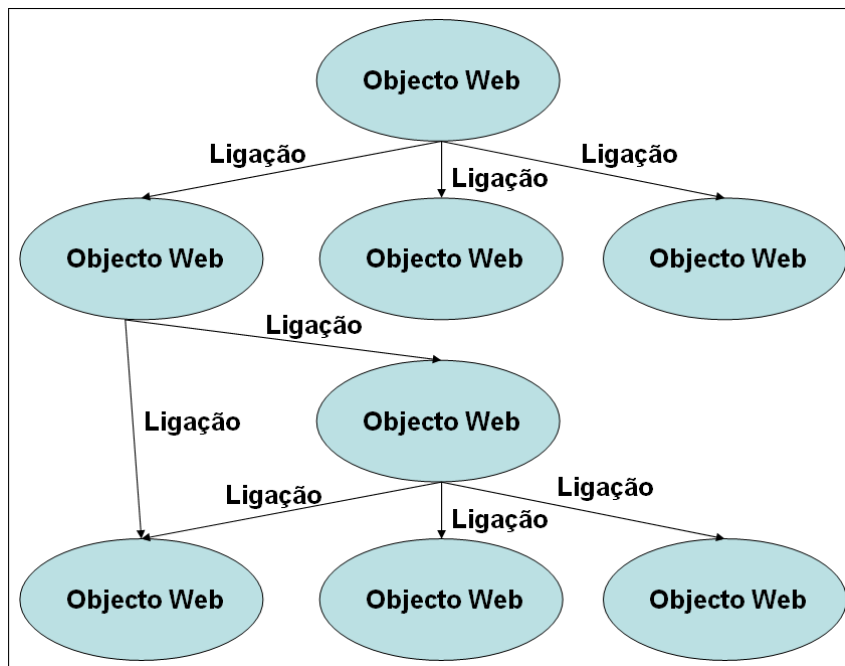


Figura 3.3: Ligações entre objectos na Web.

Conforme se pode observar na figura 3.3, a Web é constituída por `objectos web` e por `ligações`. Não há dúvida que para a grande maioria dos humanos é relativamente simples identificar o significado de um `objecto` e também o tipo de relações que podem existir com outros `objec-`

---

<sup>72</sup>Maioritariamente páginas web.

<sup>73</sup>Relações e não apenas simples ligações.



### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

tos<sup>74</sup>. Mas, para as actuais ferramentas web, o problema é mais complexo pois não possuem uma descrição do tipo de relações associadas às ligações.

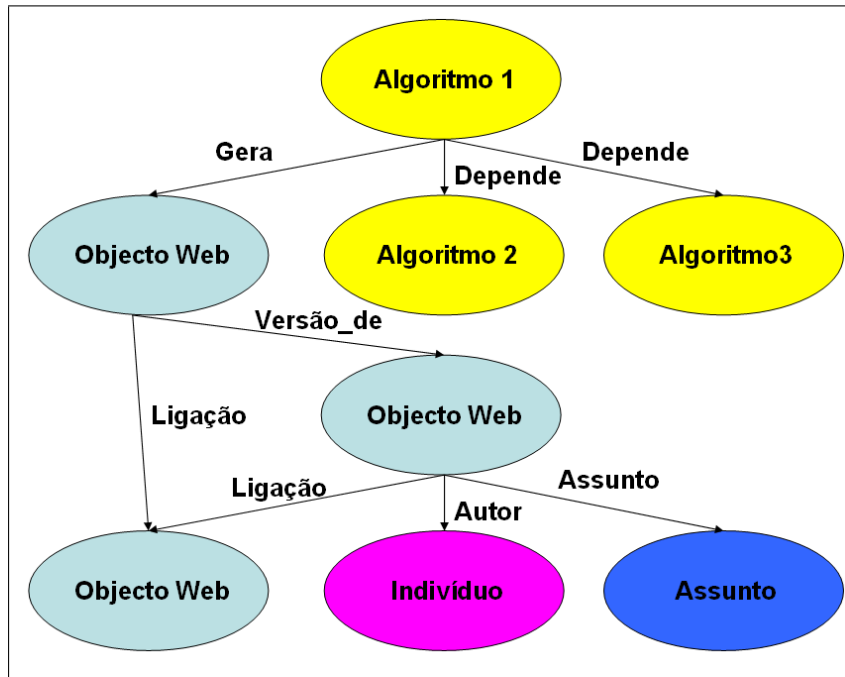


Figura 3.4: Os recursos e as relações podem ser tipificados na Web Semântica.

A Web Semântica, por seu turno, será constituída por recursos e relações. Neste caso, os recursos e as relações podem ser tipificados de maneira a definir conceitos que possibilitem uma melhor compreensão por parte dos SPC sobre o tipo de relações existentes entre os diferentes recursos web. A figura 3.4 apresenta um conjunto de recursos e de relações tipificadas. Algumas relações informam que um recurso foi escrito por um recurso que descreve uma indivíduo em particular, ou que é uma versão de um outro recurso ou que é do tipo algoritmo/*software*, etc. *Depende*; *Versão\_de* e *Autor*, são apenas, alguns exemplos de meta-dados que podem ser associados às relações existentes entre os recursos na Web Semântica.

Conforme se pode observar na figura 3.5 a proposta inicial para a Web, apresentada por Tim Berners-Lee em 1989, já especificava diferentes tipos de recursos e de relações. Como podemos

<sup>74</sup>Por exemplo, a relação existente entre o *Curriculum Vitae* de uma pessoa e a sua página web pessoal.

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

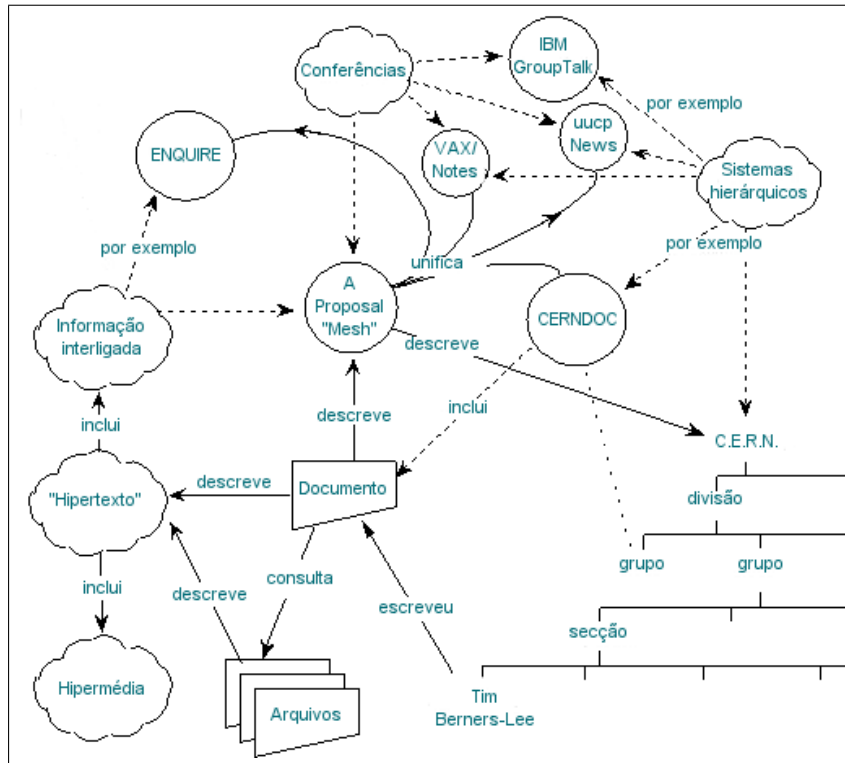


Figura 3.5: A Web na perspectiva de Tim Berners-Lee em 1989. Adaptado de [37].

facilmente constatar, essa proposta encontra-se mais próxima da actual ideia apresentada como sendo a Web Semântica do que da própria Web.

#### 3.4.3 Terceiro Princípio: Tolerância a Quebras de Relações

A Web menospreza o controlo da integridade das suas ligações em troca de uma maior versatilidade. Desta forma, os criadores de páginas web, podem facilmente criar ligações com outros documentos sem terem que se preocupar com as ligações que existem para o seu próprio documento. Como não existe uma maneira de actualizar automaticamente as ligações quando os documentos são mudados de localização ou simplesmente são eliminados, é universalmente aceite a apresentação automática de uma página a informar que a ligação para esse recurso já não existe<sup>75</sup>, conforme exemplificado na figura 3.6. Este facto é considerado, por partes dos utilizadores da Web, como um pequeno preço a pagar como contribuição para a simplificação da actual Web.

<sup>75</sup>Antigamente designado como erro '404 - File Not Found'.

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

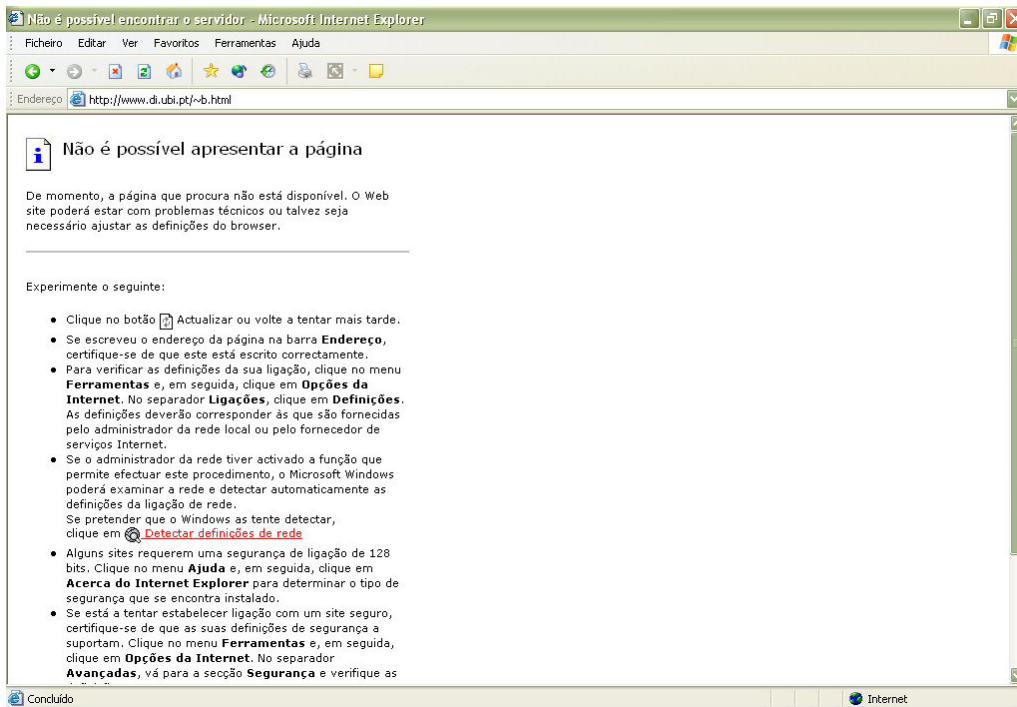


Figura 3.6: Página a informar a indisponibilidade de se poder aceder a um endereço web.

Como a Web Semântica também pretende ser versátil, os SPC deverão ser capazes de compreender, tolerarem e contornarem o problema da quebra de relações através da implementação de soluções dinâmicas e não se limitarem, apenas, à apresentação de uma página a informar a indisponibilidade de se poder aceder a um endereço web em concreto. Desta forma, continuará a ser possível e desejável, tal como acontece na actual Web, que qualquer entidade possa facilmente declarar tudo acerca de tudo.

#### 3.4.4 Quarto Princípio: Desnecessidade da Existência de uma Confiança Absoluta

Analisando a última frase da secção anterior, podemos afirmar que nem toda a informação que se encontra na Web é de confiança, e que o mesmo vai acontecer com a informação contida na Web Semântica. Por isso, uma avaliação fácil e eficiente da fidelidade da informação é um dos objectivos da Web Semântica. Essa avaliação será realizada individualmente por cada SPC, ou seja, a ideia é não existir um controlo geral e centralizado, mas sim um controlo que é efectuado

### 3.4. PRINCÍPIOS DA WEB SEMÂNTICA

localmente por cada SPC. Os SPC avaliarão o nível de fidedignidade da informação através da análise do contexto de asserções do tipo: quem diz o quê?; quando?; que credenciais são utilizadas para sustentar essas asserções?; etc.

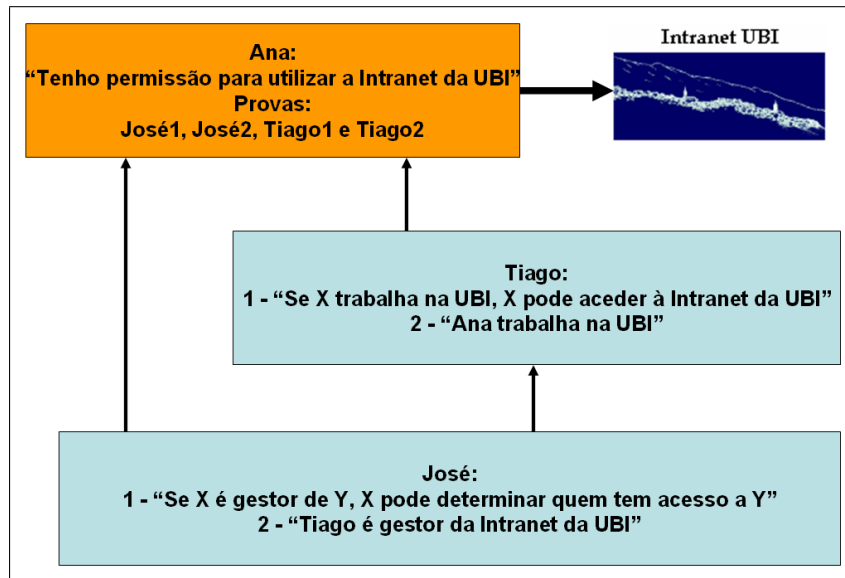


Figura 3.7: Uma cadeia de confiança para definir direitos de acesso.

Por exemplo, na figura 3.7, temos uma pessoa chamada Ana que trabalha na UBI e que deseja aceder à página da Intranet<sup>76</sup> da Universidade. Para isso, ela necessita de provar que tem o direito de poder aceder a essa Intranet. Nesse sentido, ela faz referência a quatro declarações feitas por outros dois funcionários da Universidade, o Tiago e o José, que por sua vez têm o direito de fornecer essas declarações, por causa do seu estatuto, um como gestor e o outro como o administrador da Intranet da UBI. A aplicação da Intranet aceita as provas sabendo que pode confiar nas asserções do José, porque é o administrador, e nas asserções do Tiago, porque uma existe uma asserção do José que delega ao Tiago a responsabilidade de definir a lista de funcionários com direitos de acesso às páginas da Intranet da Universidade. É este modelo descentralizado de confiança que a Web Semântica pretende implementar.

<sup>76</sup>Rede privada que funciona com a mesma arquitectura e as mesmas ferramentas da Internet.

#### 3.4.5 Quinto Princípio: Permitir Evolução

A arquitectura que a Web Semântica usará para a definição e a estruturação da informação será suficientemente maleável para permitir a evolução dos conceitos e das suas descrições, isto é, utilizará convenções descritivas que se expandirão da mesma maneira que o conhecimento humano se expande. Desta forma, poderá ser facilmente adicionada nova informação sem haver necessidade de alterar ou apagar a informação anterior.

Sendo a Web Semântica uma entidade global, ela irá frequentemente disponibilizar diferentes definições para o mesmo conceito, ou variações da definição de um conceito ao longo do tempo. Nestas situações terá, necessariamente, de fornecer mecanismos e SPC que eliminem essas ambiguidades e que permitam uma eficiente cooperação entre diferentes comunidades ou entre diferentes ontologias. Por outro lado, também as próprias ontologias poderão ser capazes de evoluir.

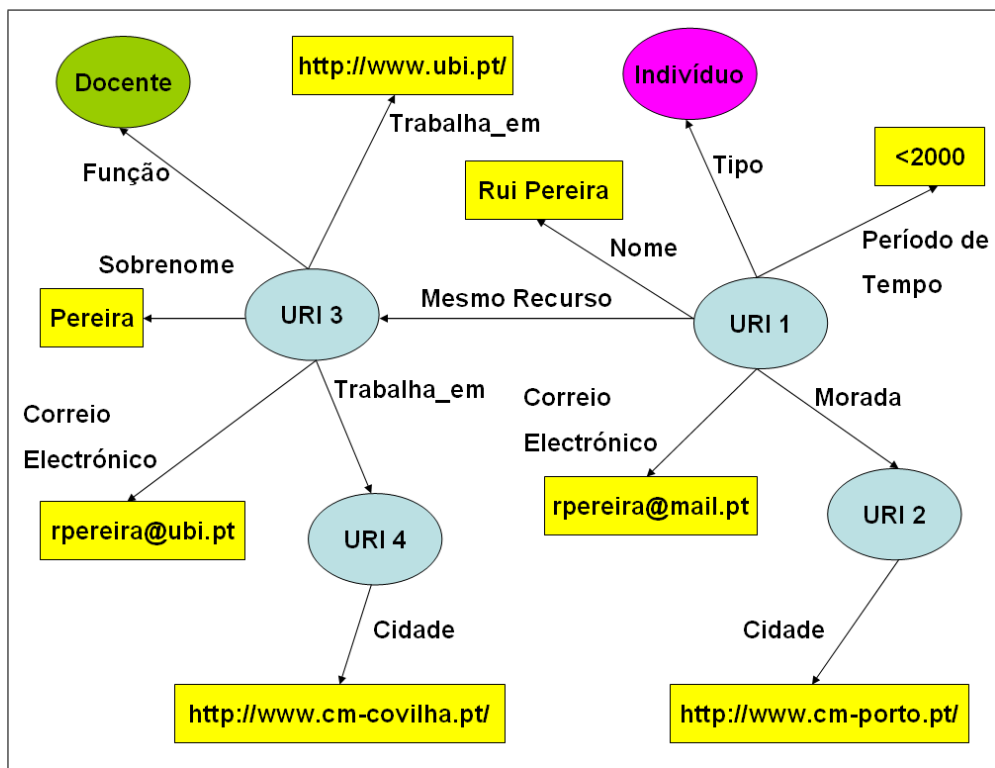


Figura 3.8: Combinação de nova informação com informação anterior.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

Na figura 3.8 é apresentada a informação de uma pessoa chamada Rui Pereira. Na parte direita da figura é apresentada informação acerca do Rui Pereira que já se encontra definida na Web Semântica. Na parte esquerda da figura é apresentada nova informação e também parte da informação anterior mas redefinida estruturalmente. A informação anterior pode ser combinada com a informação actual de várias maneiras. Podemos adicionar uma nova propriedade, por exemplo *Mesmo Recurso*, que ligue os dois tipos de informação entre si. Podemos também definir uma transformação a partir do vocabulário actual para o vocabulário usado para a informação antiga como sendo uma relação equivalente. É também possível adicionar uma nova propriedade que defina o intervalo de tempo em que a informação poderá ser considerada válida.

#### 3.4.6 Sexto Princípio: Implementação Minimalista

A Web Semântica só se tornará uma realidade se, além de poderosa e útil, for também simples de utilizar. Desta forma, os investigadores da Web Semântica deverão procurar soluções simples para problemas simples e soluções acessíveis para problemas complexos, ou seja, não deverão complicar mais do que o estritamente necessário.

## 3.5 Arquitectura Estratificada da Web Semântica

Segundo Tim Berners-Lee *et.al.* [1], "uma propriedade essencial da Web é a sua universalidade". Para que a actual Web possa permitir interações semânticas de carácter universal por parte dos SPC, é necessário definir uma arquitectura que tenha por base o conceito de universalidade<sup>77</sup>. Mas, devido à grande heterogeneidade de assuntos e de abordagens existentes na actual Web, é muito difícil desenvolver uma estrutura que especifique, de forma simples, todo o conteúdo das páginas web de uma forma universal. Assim, é necessário conjugar esforços e estabelecer acordos e compromissos para ser estabelecida uma estrutura descentralizada, global e acessível. Esta estrutura, que não é mais do que uma extensão da actual Web, é designada por Web Semântica.

---

<sup>77</sup>Qualquer pessoa pode criar ligações para qualquer coisa.



Figura 3.9: A arquitectura estratificada da Web Semântica.

Na proposta de desenvolvimento da Web Semântica apresentada por Tim Berners-Lee e pelo consórcio W3C, figura 3.9, é sugerida uma arquitectura estratificada englobando diversas especificações. Apesar de, as diversas especificações constituintes da arquitectura da Web Semântica, como por exemplo o URI [75], a XML [3], o Modelo e a Sintaxe RDF [4], o Esquema RDF<sup>78</sup> [91], as ontologias, etc., serem apresentadas em várias camadas, não significa que sejam entidades isoladas, pelo contrário estão de tal maneira interligadas que por vezes é difícil analisá-las separadamente. Estas especificações são apresentadas numa arquitectura estratificada porque, de uma maneira geral, as especificações das camadas inferiores definem as bases para as especificações das camadas superiores. Desta forma, à medida que percorremos as camadas da pilha em ordem ascendente as diversas especificações vão representando a informação de uma forma, cada vez mais, expressiva e rica em significado. A maioria das especificações encontram-se em fase de investigação ou de normalização. Sendo assim, de uma forma geral podem ser

<sup>78</sup> *RDFS - RDF Schema - RDF vocabulary description language.*

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

classificadas em três grupos:

- Especificações normalizadas - camadas inferiores da pilha: URI/Unicode<sup>79</sup> e XML/Espaço de Nomes;
- Especificações recém-normalizadas - camadas centrais da pilha: Modelo RDF & Sintaxe RDF, Esquema RDF e a linguagem ontológica OWL [5];
- Especificações experimentais - camadas superiores da pilha: Regras, Estrutura Lógica, Prova e Confiança<sup>80</sup>.

As camadas inferiores da pilha preservam a utilização de um conjunto de caracteres internacionais e possibilitam um meio de identificar, integrar e de partilhar os recursos na Web Semântica. As camadas centrais da pilha, permitem a definição de vocabulários, ontologias, tipos de recursos, relações e asserções RDF com base nos URIs dos recursos. Actualmente, existe uma grande investigação relacionada com a criação autónoma de asserções RDF e de ontologias. As camadas superiores da pilha, possibilitam a escrita, execução e avaliação de regras lógicas e o estabelecimento de forma distribuída de níveis de confiança para a informação. A definição de regras para deduzir inferências, escolher caminhos de acção e responder a interrogações complexas, constitui, também, uma das actuais actividades de I&D de maior ênfase na área da Web Semântica. Uma mistura de decisões matemáticas e de engenharia complica esta tarefa, porque a lógica utilizada deverá ser bastante poderosa para poder descrever, de forma perceptível para os SPC, as propriedades complexas dos recursos [1].

Na arquitectura estratificada da Web Semântica, existe uma multi-camada normalmente denominada por Assinatura<sup>81</sup>/Encriptação<sup>82</sup>, a qual é utilizada para codificar e detectar possíveis

---

<sup>79</sup>A Web é puramente internacional, mas, no entanto, a maior parte da informação que podemos encontrar está especificada através do conjunto de caracteres definidos na norma ASCII. O formato Unicode, ao contrário do ASCII, oferece um suporte para quase todas as especificações de caracteres internacionais.

<sup>80</sup>Termos utilizados na língua Inglesa: *Rules; Logic framework; Proof; Trust*, respectivamente.

<sup>81</sup>ou também referida como *XML Signatures* [92].

<sup>82</sup>ou também referida como *XML Encryption* [93].



### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

alterações sofridas pela informação. A assinatura digital e a encriptação deverão ser implementadas em sintaxe XML, tornando-se desta forma extremamente versáteis e poderosas.

A partir do momento em que as camadas da estrutura da Web Semântica sejam facilmente acedidas, interpretadas e manipuladas pelos SPC, eles, por sua vez, serão capazes de compreender o significado e as relações existentes entre a informação que se encontra definida segundo a arquitectura da Web Semântica. Além disso, serão capazes de trocar entre si informação recolhida, resultados de processamento e informação relacionada com o tipo de raciocínio por eles realizado. A autenticidade e confiança das fontes adquirirá, assim, um novo significado, pois os SPC ao raciocinarem sobre a informação poderão chegar a conclusões que afectem a acção humana [12].

Nas próximas sub-secções serão apresentadas, com maior pormenor, três componentes essenciais para o actual desenvolvimento da ideologia da Web Semântica:

- XML;
- Arquitectura RDF;
- Ontologias.

#### 3.5.1 XML

A crescente influência da Web na vida quotidiana das pessoas e a possibilidade de qualquer entidade<sup>83</sup> conseguir, por seu intermédio, interactuar facilmente com várias outras entidades, banalizaram de tal forma a sua utilização que é indiscutivelmente considerada, a nível mundial, como um importante meio de comunicação. Os principais factores que contribuíram, ao longo dos últimos anos, para a sua massificação prendem-se com a simplicidade de utilização, a disponibilização instantânea de informação a uma escala mundial e a utilização de modelos

---

<sup>83</sup>Individual ou colectiva.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

simples e internacionalmente aceites como a linguagem HTML, a linguagem *JavaScript*<sup>84</sup> e a utilização de formatos de imagem simples como o *Graphics Interchange Format* (GIF) [95] e o *Portable Network Graphics* (PNG) [96]. No entanto, estes modelos, internacionalmente aceites, não apresentam uma estrutura simples e normalizada para a representação da informação de uma forma universal. Surge assim, a necessidade de definir um modelo simples, versátil e universal para a representação da informação na Web. É neste contexto que é introduzida a XML, a qual é definida como sendo a resposta à actual necessidade de definir dados de forma simples e estruturada.

Segundo Ana Moura [58], a XML é uma tecnologia, considerada recente, que é utilizada com o intuito de fornecer uma interactividade sintáctica para os dados definidos na Web. Ela separa os dados da respectiva formatação, o que não acontece com a linguagem HTML. Além disso, supera as limitações da linguagem HTML, permitindo a criação de anotações definidas pelo próprio utilizador e proporcionando, desta forma, uma maior versatilidade na descrição dos recursos em termos de meta-dados. Mas, a XML apenas define uma estrutura para a representação dos dados. O modo como esses dados poderão vir a serem apresentados está para além da sua competência. A partir de um documento XML, qualquer pessoa ou qualquer SPC dotado de um Interpretador<sup>85</sup> XML pode definir a maneira como deseja apresentar os seus dados. Essa especificação pode ser realizada através de várias técnicas diferentes, sendo a mais utilizada a criação de folhas de estilo<sup>86</sup>. As técnicas de criação de folhas de estilo mais utilizadas são a especificação *Cascading Style Sheet* (CSS) [97] e a *eXtensible Stylesheet Language* (XSL) [98].

Desta forma, a formatação de um documento XML é tratada separadamente de sua estrutura,

---

<sup>84</sup>É uma linguagem de programação do tipo interpretada, desenvolvida pela Netscape. Uma linguagem interpretada, ao contrário das linguagens compiladas, apenas é interpretada no momento de execução. A linguagem *JavaScript* pode ser inserida em páginas HTML e é interpretável pela generalidade dos Navegadores Web actuais, o que a torna muito utilizada na Web [94]. Pode também ser utilizada em conjunto com as linguagens ASP [61], PHP [62] e JSP [63].

<sup>85</sup>Na língua inglesa denomina-se *Parser*. É um tipo de ferramenta informática que processa documentos de texto em função dos seus elementos constituintes, geralmente palavras. É utilizado para a verificação da grafia e de regras associadas aos elementos além de normalmente servir de base para muitas ferramentas especializadas na análise de documentos de texto.

<sup>86</sup>Especificações ou linguagens que permitem definir estilos visuais (cores, tipos de fontes, etc.) geralmente utilizados para a formatação da informação.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

resolvendo assim um dos principais problemas da linguagem HTML. Além disso evita a necessidade de recorrer a sistema de marcações complexas como é o caso da *Standard Generalized Markup Language* (SGML)<sup>87</sup>.

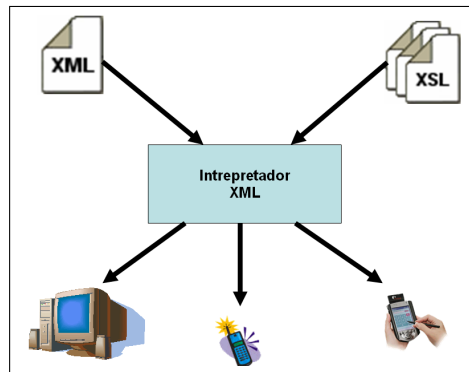


Figura 3.10: Exibição de um documento XML utilizando diferentes regras de apresentação.

A linguagem XSL é conhecida como sendo uma linguagem sofisticada de folha de estilo, a qual permite adicionar estilos visuais (cores, tipos de fontes, etc.) aos documentos XML. Ao contrário da especificação CSS, é definida em XML o que a torna muito versátil e poderosa. Existe desde 1999, mas apenas adquiriu o carácter de Recomendação W3C a 15 de Outubro de 2001 [100]. Ela define um conjunto de regras de formatação e de apresentação que podem ser facilmente associadas à informação estruturada existente nos documentos XML. A utilização de dois documentos com propósitos diferentes, um apenas com a informação estruturada (documento XML) e um outro apenas com as regras de formatação associadas aos elementos do primeiro (documento XSL), permite a criação de forma muito simples de vários tipos de apresentação diferentes e independentes para a mesma informação. Por exemplo, a figura 3.10 apresenta um único documento XML, mas, a informação nele contida pode ser facilmente disponibilizada, através da utilização de um Interpretador XML e de diferentes ficheiros de estilos XSL, em diferentes dispositivos (computador pessoal, telemóvel, *Personal digital assistant* (PDA), etc.).

Qualquer documento XML pode ser representado como sendo uma simples estrutura definida

---

<sup>87</sup>Surgiu como uma meta-linguagem de marcação, a partir da qual poderiam ser criadas outras linguagens de marcação. As principais desvantagens da SGML são a de ser considerada extensa e complexa [99].

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

em forma de árvore, onde cada nó representa um elemento XML do documento XML. A XML é, cada vez mais, utilizada como base para a definição e extensão de centenas de linguagens de marcação em Código Aberto<sup>88</sup>, das quais são apresentadas as seguintes como exemplo:

- AIML - *Artificial Intelligence Markup Language* [103];
- AML - *Astronomical markup Language* [104];
- CML - *Chemical Markup Language* [105];
- DCML - *Dental Charting Markup Language* [106];
- DSML - *Directory Services Markup Language* [107];
- GEML - *Gene Expression Markup Language* [108];
- IRML - *Investment Research Markup Language* [109];
- MathML - *Mathematical Markup Language* [110];
- MML - *Medical Markup Language* [111];
- NML - *News Markup Language* [112];
- OMF - *Weather Observation Definition Format* [113];
- PetroXML - *Petroleum XML* [114];
- RDF - *Resource Description Framework*;

---

<sup>88</sup>Do termo inglês *Open-source*. A definição do Código Aberto está associada a um tipo de licença de *software* denominada por *Copyleft* ou, também, por *General Public License* (GPL) [101], a qual é substancialmente diferente da licença de *software* mundialmente conhecida como *Copyright*. A licença GPL determina que todo o *software* por ela definido é de carácter gratuito, isto é, não existe qualquer encargo financeiro associado à sua utilização ou distribuição por parte de qualquer entidade. Além disso o código fonte do *software* deverá ser obrigatoriamente fornecido. Qualquer entidade poderá modificar esse código fonte se assim o desejar, mas ficará obrigada a registar essa modificação através de uma licença GPL. Desta forma, qualquer *software* licenciado como GPL será sempre aberto, pois esta licença não permite a possibilidade de criação de *software* proprietário baseado em *software* GPL. Esta licença foi idealizada e concebida em 1984 por Richard Stallman para reger todo o software criado pelo projecto GNU (*GNU is not Unix*). Um tipo de *software* que é regido por esta licença é o, mundialmente conhecido e utilizado, sistema operativo Linux[102].

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

- SMIL - *Synchronized Multimedia Integration Language* [115];
- SVG - *Scalable Vector Graphics*;
- ThML - *Theological Markup Language* [116];
- UDDI - *Universal Description Discovery Integration* [117];
- VIML - *Virtual Instruments Markup language* [118];
- WML - *Wireless Markup Language* [119];
- WSDL - *Web Services Description Language* [120];
- XAML - *Transaction Authority Markup Language* [121];
- XBEL - *Bookmark Exchange Language* [122].

A XML, além de ser utilizada para estruturar o conteúdo das páginas web, é também utilizada como uma norma para a transferência de informação de forma independente por parte da maioria das novas tecnologia para a Web, como por exemplo os Serviços Web<sup>89</sup>.

Os elementos pertencentes a um documento XML chamam-se elementos XML e são individualmente limitados por uma anotação. Os elementos XML podem possuir várias características que são designadas por atributos XML. A XML baseia-se num conjunto de regras simples para a estruturação da informação [124]:

- Só pode existir um elemento XML raiz. Só existe um elemento XML para o qual a sua anotação não se encontra dentro de qualquer outro elemento XML;
- Os elementos XML não raiz necessitam de estar definidos dentro de um elemento XML;
- Todos os elementos XML devem estar hierarquicamente aninhados. A última anotação a ser aberta deve ser a primeira a ser fechada;

---

<sup>89</sup>Termo em língua Inglesa: *Web Services* [123]

- Os atributos XML devem ser únicos.

Quando um documento XML cumpre as regras anteriormente enunciadas é denominado por documento XML Bem Formado. Os documentos XML precisam de ser Bem Formados para que os SPC os possam manipular de forma correcta. Um documento XML também pode ser considerado Válido, para isso necessita de ser um documento XML Bem Formado e respeitar um conjunto de restrições acerca da formatação dos seus elementos constituintes. Esse conjunto de restrições é geralmente armazenado num documento de restrições XML que poderá ser associado ao documento XML. Os dois tipos de documentos de restrições XML mais utilizados são o *Document Type Definition* (DTD) [125] e o Esquema XML (*XML Schema*) [126]. A utilização destes documentos de restrições XML, permite definir uma estrutura robusta para a informação existente nos documentos XML e deste modo contribuir para a simplificação quer da validação de documentos XML, quer da criação de ferramentas baseadas em XML. De acordo com Stefan Decker *et.al.* [127], é importante referir que o documento DTD e o documento Esquema XML apenas especificam convenções sintácticas, pelo que qualquer tentativa de planeamento semântico está fora do domínio da XML. Um documento de restrições XML possui as seguintes características [124]:

- Especifica quais as anotações disponíveis para serem utilizadas como meta-dados associados aos elementos XML;
- Especifica quais os atributos XML disponíveis e a que elementos XML estão associados;
- Define a estrutura do conteúdo de cada elemento XML. Isto é, que sub-elementos possui, por que ordem aparecem, onde pode aparecer texto simples, etc.

Cada vez mais, o consórcio W3C está a substituir a utilização do documento DTD pela utilização do documento Esquema XML. Ainda, segundo Stefan Decker *et.al.* [127], o Esquema XML

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

permite, ao contrário do DTD, definir tipos de dados e modelos de conteúdos complexos<sup>90</sup>. Mas acima de tudo, a sua maior vantagem face ao DTD é a de ser um documento XML Bem Formado.

As ferramentas que interpretam, validam e navegam na estrutura em árvore dos documentos XML são designadas por Interpretadores XML. As duas *Application Programming Interface* (API) mais utilizadas como Interpretadores XML são a *Simple API for XML* (SAX) [128] e a *Document Object Mode* (DOM) [129]. Segundo o consórcio W3C, a API SAX caracteriza-se por ser pouco exigente em termos de necessidade de processamento e por ser simples de implementar. Ela define um conjunto de *interfaces*<sup>91</sup> em Java muito úteis para o acesso e a interpretação de documentos XML. A sua aplicabilidade estende-se por um conjunto extremamente variado de aplicações. O seu modo de funcionamento caracteriza-se pela travessia completa da estrutura, em forma de árvore, de um documento XML sempre que é realizado um pedido de acesso a um dos elementos XML. Por cada novo pedido, é percorrida novamente toda a estrutura do documento XML. Por seu lado, a API DOM permite carregar em memória, de uma forma estruturada, toda a informação que se encontra definida num documento XML. Os documentos XML são assim, representados na memória principal do computador através de uma árvore de objectos onde cada objecto está associado a um componente na estrutura do documento XML. Desta forma, a API DOM permite acessos rápidos e bastante facilidade na manipulação da informação existente em documentos XML. Tanto a API SAX com a API DOM são largamente recomendadas e utilizadas. As principais vantagens da API SAX podem-se resumir a [130]:

- Possibilidade de analisar documentos XML de qualquer tamanho. Como não necessita de armazenar na memória principal do computador a estrutura do documento XML, é muito eficiente quando o tamanho físico dos documentos XML é elevado. Desta forma, pode-se afirmar que não se encontra limitada à quantidade de memória RAM disponibilizada pelo hardware;

---

<sup>90</sup>Tipos de dados como por exemplo: 'string', 'integer', 'real', 'time', 'date', etc. E a utilização de um modelo do tipo Orientado aos Objectos (Objectos, Herança, etc.).

<sup>91</sup>Um *interface* Java não é mais do que uma colecção de definições de métodos e constantes, desenvolvidos para serem utilizados a partir de várias linguagens de programação.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

- Possibilidade de trabalhar apenas com parte de um documento XML;
- É extremamente rápida e eficiente quando utilizada para pequenos números de acessos a documentos XML.

Por sua vez, as principais vantagens da API DOM podem-se resumir a [130]:

- Representação lógica em memória, em forma de uma árvore, da estrutura de um documento XML. Assim, é facilmente identificada a sua estrutura de uma forma global;
- Suportar o acesso aleatório aos elementos de um documento XML;
- Facilidade em efectuar pesquisas complexas em documentos XML;
- Possibilidade de alterar o conteúdo ou a estrutura lógica de um documento XML. Através da adição, remoção e alteração do valor dos elementos;
- Facilidade de interacção com BD relacionais e hierárquicas.

Actualmente existem centenas de ferramentas desenvolvidas com o objectivo de manipular documentos XML. Geralmente são classificadas como sendo pertencentes a um dos seguintes grupos [131]:

- Ferramentas para a criação, modificação e documentação de folhas de estilo;
- Ferramentas para a criação e modificação de documentos XML certificados;
- Ferramentas com suporte de gestão de documentos, como BD de documentos e Motores de Busca;
- Ferramentas para a criação, modificação e composição de documentos XML de uma forma interactiva;
- Ferramentas para a distribuição e apresentação de forma electrónica de documentos XML;
- Interpretadores XML.



### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

Por fim, é necessário realçar que a XML apenas organiza estruturalmente os dados existentes num documento. O estabelecimento de uma interpretação comum para a informação presente nos documentos não é uma competência da XML. Ela apenas apresenta um modelo simples e coerente para a estruturação de dados, o qual é fielmente utilizado por muitas outras tecnologias, como por exemplo, a arquitectura RDF e as linguagens ontológicas definidas de acordo com os princípios da Web Semântica.

#### 3.5.2 Arquitectura RDF

Através da utilização da XML podemos associar meta-dados aos dados presente em documentos de texto e multimédia. A um nível mais elevado em termos de riqueza de significado, podemos utilizar um modelo que relacione esses recursos<sup>92</sup> entre si. Esse modelo, será extremamente útil se conseguir demonstrar que é simples de compreender, implementar, manipular e alterar, quer pelas pessoas quer pelos SPC. Além disso, a possibilidade da sua utilização de forma distribuída será também considerada como uma característica muito valiosa.

Existe uma enorme variedade de arquitecturas definidas com o objectivo de representar as relações existentes entre diferentes recursos. Essas arquitecturas apresentam como prioridades, promover uma melhor descrição semântica da informação e incrementar uma eficiente procura de informação na Web. Uma das arquitecturas que mais provas tem apresentado nesse sentido é a arquitectura *Resource Description Framework* (RDF), a qual é, conforme foi anteriormente referido, uma Recomendação W3C.

A arquitectura RDF permite descrever os recursos, as suas propriedades e as relações existentes entre esses recursos, através da utilização de um modelo, denominado Modelo RDF, e de uma sintaxe, denominada Sintaxe RDF. Uma das principais vantagens da utilização da arquitectura RDF prende-se com o facto de a informação especificada através do Modelo RDF e da Sintaxe

---

<sup>92</sup>O termo recurso é utilizado quando se refere a um dado ou meta-dado que esteja representado através de um URI.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

RDF ser facilmente interpretável quer pelos humanos, quer pelos SPC.

O Modelo RDF utiliza como primitiva a representação de uma simples asserção composta por três elementos, a que se dá o nome de asserção RDF ou de tripla RDF. Uma asserção RDF define uma relação tripla existente entre um recurso (Sujeito), o seu valor (Objecto) e uma propriedade (Predicado), a qual é utilizada para caracterizar a relação existente entre o recurso e o seu valor. O elemento valor de uma asserção RDF é geralmente representado por um valor simples e sem conotação sintáctica, mas também pode ser representado por um recurso. Uma propriedade, por sua vez, é sempre representada por um recurso. Na representação de uma asserção RDF utilizam-se os URI de cada um dos seus elementos. A seguir estão representados, de uma forma abstracta, dois exemplos da utilização de uma asserção RDF:

\* Asserção RDF entre 2 recursos e um valor:

```
<URI do Recurso>  
    <URI da propriedade>  
        <Valor>
```

\* Asserção RDF entre 3 recursos:

```
<URI do Recurso>  
    <URI da propriedade>  
        <URI do recurso>
```

Na representação de asserções RDF, são, cada vez mais, utilizadas ferramentas gráficas baseadas em *Directed Labeled Graphs* (DLG) [76]. Os elementos do DLG são representados por nós e arcos, onde:

- Um nó, representado por uma elipse, identifica um recurso RDF;
- Um nó, representado por um rectângulo, identifica um valor;
- Um arco, representado por uma seta, identifica uma propriedade. A direcção da seta é importante, pois o arco parte sempre do sujeito e aponta para o objecto da asserção;

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

- Um nó, representado por uma elipse vazia, identifica um recurso anónimo;

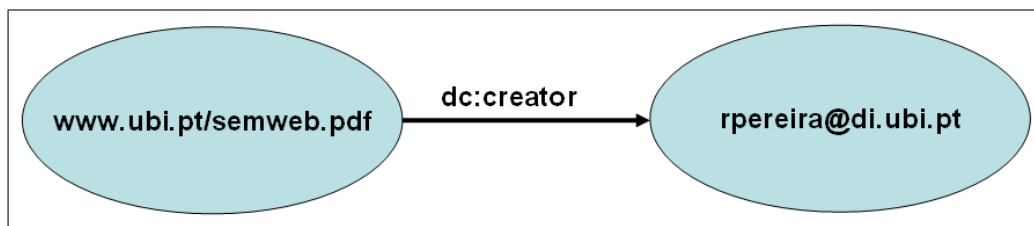


Figura 3.11: Representação gráfica da asserção: "Rui Pereira é o autor do documento semweb.pdf".

A figura 3.11 apresenta uma asserção RDF criada através de uma ferramenta gráfica chamada IsaViz [132]. A asserção RDF da figura, associa o URI da pessoa Rui Pereira ao URI de um documento do tipo PDF. A associação é estabelecida através de uma propriedade chamada *creator* que se encontra definida no vocabulário *Dublin Core*.

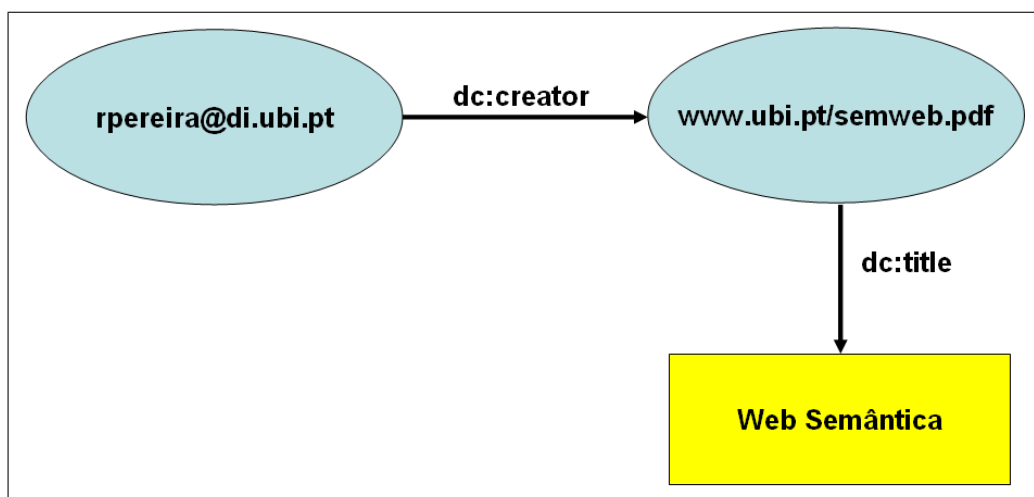


Figura 3.12: Representação gráfica da asserção: "Rui Pereira é o autor do documento semweb.pdf, o qual tem Web Semântica como título".

A Arquitectura RDF permite que o conjunto dos três elementos de uma asserção RDF possa ser referenciado como um único recurso RDF. Desta forma, as asserções RDF podem ser facilmente encadeadas, criando assim correntes de informação, conforme se pode observar na figura 3.12.

A sintaxe RDF é constituída por um vocabulário muito simples, o qual, tem por objectivo descrever textualmente o Modelo RDF. Quando essa descrição é feita em XML, como é o caso do exemplo apresentado a seguir, designa-se por sintaxe RDF/XML:

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="rpereira@di.ubi.pt">
    <dc:publisher
      rdf:resource="www.di.ubi.pt/semweb.pdf"/>
  </rdf:Description>
</rdf:RDF>
```

Este exemplo não é mais do que a sintaxe RDF/XML da asserção RDF apresentada graficamente na figura 3.11. A primeira linha do exemplo informa aos SPC que o código está escrito em XML. Na terceira e quarta linha são declarados, através da utilização dos seus Espaços de Nomes, os vocabulários<sup>93</sup> utilizados. O elemento RDF *Description* através do seu atributo *about* identifica o primeiro recurso a ser descrito<sup>94</sup>. O elemento *publisher* do *Dublin Core* está a ser utilizado como uma propriedade<sup>95</sup>, a qual associa o primeiro recurso descrito com o recurso<sup>96</sup> descrito pelo elemento RDF *resource* e que se encontra agregado à propriedade. Desta forma, esta sintaxe RDF/XML descreve textualmente a asserção RDF apresentada na figura 3.11, que tem como significado:

Rui Pereira é o editor do documento *semweb.pdf*.

Um documento RDF possui uma ou mais *Descrições RDF*, ou seja, asserções RDF acerca de um ou mais recursos da Web Semântica. A análise de documentos RDF é auxiliada através da utilização de Interpretadores RDF. Os tipos de paradigmas mais utilizados pelos Interpretadores RDF podem ser classificados como [133]:

---

<sup>93</sup>Neste caso o prefixo *rdf* está referenciado como as definições do modelo básico do RDF, e o prefixo *dc* está referenciado com o esquema de definição dos elementos do vocabulário *Dublin Core*.

<sup>94</sup>Sujeito da asserção.

<sup>95</sup>Predicado da asserção.

<sup>96</sup>Objecto da asserção.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

- Centrados no Modelo RDF<sup>97</sup>: A unidade básica é o modelo de RDF completo; Apenas permite abrir, gravar e apagar modelos de RDF;
- Centrados nas Asserções RDF<sup>98</sup>: A unidade básica são as Asserções RDF; A informação RDF é manipulada como um conjunto de triplas RDF;
- Centrados nos Recursos RDF<sup>99</sup>: A unidade básica são os Recursos RDF; O modelo é manipulado como sendo constituído por Recursos RDF que podem possuir Propriedades RDF;
- Centrados nas Ontologias<sup>100</sup>: Possui suporte para o tipo de objectos que são esperados existir numa Ontologia:
  - Classes: Numa hierarquia de classes;
  - Propriedades: Numa hierarquia de propriedades;
  - Indivíduos: Entidades ímpares.

Actualmente, a API mais conhecida e utilizada para a manipulação da arquitectura RDF, é designada por *Jena - A Semantic Web Framework* [134] e foi desenvolvida pelo *HP Labs Semantic Web Programme* [135]. Ela é definida como sendo uma estrutura que facilita a criação de aplicações para a Web Semântica, sendo a sua principal característica a manipulação de modelos RDF em sintaxe RDF/XML, Notation3 (N3) [136] e N-Triples [137].

De acordo com Ana Moura [58], Tim Berners-Lee [68] e Dan Brickley *et.al.* [138], apesar de a arquitectura RDF especificar um modelo simples para a descrição de relações existentes entre diferentes recursos através da utilização de propriedades e valores, não fornece um mecanismo

---

<sup>97</sup> *Model-centric.*

<sup>98</sup> *Statement-centric.*

<sup>99</sup> *Resource-centric.*

<sup>100</sup> *Ontology-centric.*

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

universal para a classificação das relações existentes entre os diversos recursos. Para implementar esse mecanismo universal foi desenvolvido o Esquema RDF [91], o qual é uma linguagem constituída por uma colecção de recursos RDF pré-definidos que podem ser utilizados na descrição de vocabulários. Esses vocabulários podem ser usados para definir classes, propriedades e restrições que podem, por sua vez, ser utilizados para descrever outras classes, propriedades e restrições. O sistema de representação de classes, propriedades e restrições é semelhante ao utilizado pelas Linguagens de Programação Orientada a Objectos (LPOO). No entanto, enquanto que nos sistemas de LPOO as classes são definidas de acordo com as propriedades que possuem, no Esquema RDF as propriedades é que são definidas de acordo com as classes RDF que as utilizam. Desta forma, esta abordagem centrada nas propriedades facilita a adição de novas definições de propriedades sem haver a necessidade de redefinir a descrição original das classes.

É apresentado na figura 3.13 um diagrama que ilustra uma possível interligação entre um vocabulário Esquema RDF, na parte superior da imagem, e um conjunto de asserções RDF, na parte inferior da imagem, definidas, por exemplo, numa página Web Semântica.

O exemplo representado na figura 3.13 ilustra a forma como o RDF pode ser utilizado para descrever conceitos reais como pessoas e documentos, as classes a que pertencem, como `eg:Person` e `eg:Document`, bem como as propriedades que são utilizadas para relacionar membros dessas classes, como `eg:author`. Através da utilização do Esquema RDF é possível descrever relações entre as propriedades RDF e classes de recursos definidas em RDF. Neste exemplo em concreto, o Esquema RDF está a ser utilizado para associar documentos a pessoas através da propriedade `eg:author`. Além disso, associa a classe documentos à classe trabalho e a classe pessoas à classe agentes.

O vocabulário do Esquema RDF [91] é definido através da utilização do modelo RDF na especificação de entidades como classes, propriedades e restrições. De seguida é apresentada uma

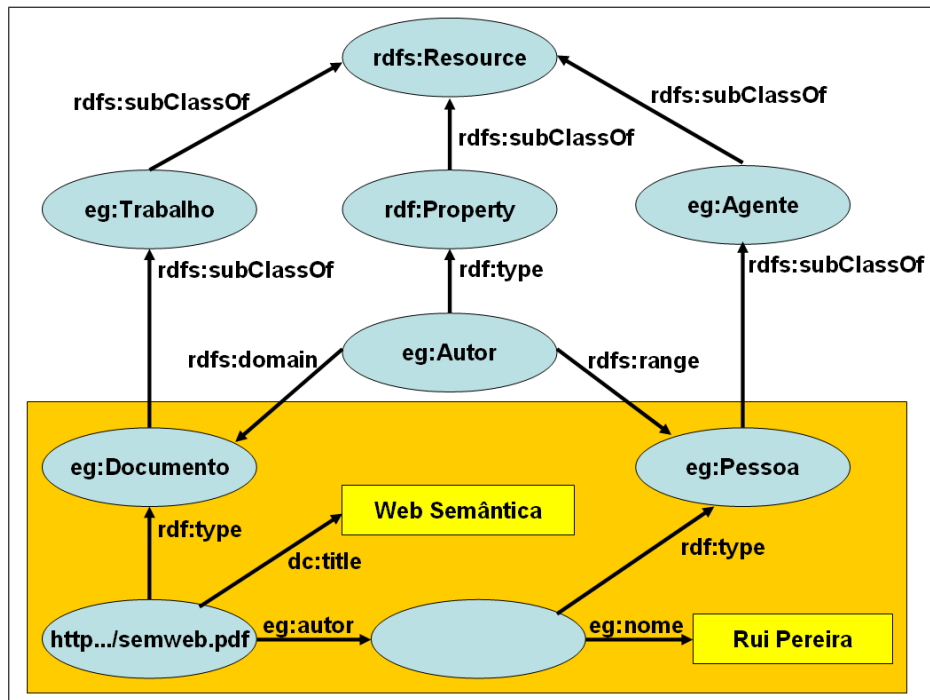


Figura 3.13: Relação existente entre um vocabulário Esquema RDF e asserções RDF definidas numa página Web Semântica. (Adaptado de [139]).

descrição sucinta das principais classes e propriedades pré-definidas no Esquema RDF:

- Classe `rdfs:Class` - Corresponde ao conceito genérico de tipo ou categoria de um recurso. Idêntico ao conceito de classe utilizado nas LPOO;
- Classe `rdf:Property` - Representa os recursos que são propriedades RDF;
- Classe `rdfs:Resource` - Representa uma qualquer entidade que possa ser descrita através do RDF;
- Classe `rdf:Statement` - Representa uma qualquer Tripla RDF;
- Classe `rdfs:Literal` - Representa um qualquer valor simples como um valor inteiro ou uma palavra;
- Propriedade `rdf:type` - Indica que um recurso é membro de uma classe;
- Propriedade `rdfs:subClassOf` - Indica uma relação hierárquica entre duas classes. É uma propriedade transitiva;

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

- Propriedade `rdfs:subPropertyOf` - Indica uma relação hierárquica entre duas propriedades. É uma propriedade transitiva;
- Propriedade `rdfs:domain` - Um domínio de uma propriedade limita os indivíduos aos quais a propriedade pode ser aplicada. Se uma propriedade relacionar um indivíduo a um outro indivíduo, e a propriedade tiver uma classe como um dos seus domínios, então o indivíduo deve pertencer a essa classe;
- Propriedade `rdfs:range` - Uma propriedade pode limitar, através do termo `rdfs:range`, o intervalo de indivíduos que a propriedade pode ter como seu valor. Se uma propriedade relacionar um indivíduo com um outro indivíduo e a propriedade tiver uma classe específica no seu intervalo de indivíduos, então o outro indivíduo deve obrigatoriamente pertencer a essa classe específica;
- Propriedade `rdfs:label` - Fornece uma descrição textual do nome do recurso;
- Propriedade `rdfs:comment` - Fornece uma descrição textual do recurso.

Através da utilização do Esquema RDF, os recursos RDF podem ser definidos como instâncias de uma ou mais classes utilizando para isso a propriedade `type`. Por sua vez, a propriedade `subClassOf` permite que um utilizador especifique uma organização hierárquica de classes<sup>101</sup>. A propriedade `subPropertyOf` funciona da mesma forma que a propriedade `subClassOf` mas direcionada para as propriedades.

De acordo com Stefan Decker *et.al.* [127] e Dan Brickley *et.al.* [138], é importante sublinhar que a arquitetura RDF apenas define um modelo normalizado para a representação de relações entre recursos. Tal como a XML, a arquitetura RDF também não providencia mecanismos para definir quais os conceitos e as propriedades que podem ou devem ser utilizados. A utilização de URIs garante uma definição única dos conceitos. No entanto, o mesmo conceito pode ser

---

<sup>101</sup>Por exemplo, uma classe denominada "Gato" poderia ser considerada como uma subclasse de uma classe denominada "Mamífero" que por sua vez poderia ser considerada uma subclasse da classe de "Animal". Isto significaria que qualquer recurso do tipo `rdf:type Gato` também seria do tipo `rdf:type Animal`.



### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

expresso de formas diferentes e em línguas diferentes. Por exemplo, duas BD podem armazenar a mesma informação mas utilizando terminologias diferentes. Para que a informação possa ser processada e relacionada de forma eficiente e útil, é necessário que exista uma definição única e universal para cada um dos seus conceitos e relações. A solução proposta para a Web Semântica é a utilização de Ontologias, principalmente aquelas que são desenvolvidas em conformidade com a XML e a arquitectura RDF.

O desenvolvimento de uma linguagem universal para a representação de todo o conhecimento existente na Web, é o sonho de muitos dos actuais investigadores da Web. Mas, por uma quantidade de razões pragmáticas esse sonho é actualmente inexecutável. Pelo contrário, a realidade passa, cada vez mais, pela definição e utilização de uma multiplicidade de representações de meta-dados. Uma das vantagens da arquitectura RDF é o de permite representar os dados de forma a serem facilmente partilhados pela generalidade das linguagens de meta-dados existentes. Além disso, a linguagem do Esquema RDF é suficientemente prática e poderosa para permitir a definição de novas, e semanticamente mais ricas, linguagens de meta-dados tendo como base o modelo RDF.

#### 3.5.3 Ontologias

A procura de uma estrutura universal que permita definir e representar univocamente recursos, é uma preocupação que se pode encontrar na generalidade dos campos da intervenção humana<sup>102</sup>. Um dos resultados desta procura resume-se à definição e criação de vocabulários específicos, normalmente utilizados no âmbito de uma área de conhecimento em particular. Mas este tipo de resultados, apesar de bastante úteis, não satisfaz a crescente necessidade de combinar recursos que se encontram definidos em diversas áreas do conhecimento. Cada vez mais, a partilha de informação é considerada como um requisito indispensável para a subsistência saudável de qualquer área de intervenção humana.

---

<sup>102</sup>Por exemplo, os filósofos procuram categorias universais para classificar todos os recursos existentes, os lexicógrafos procuram terminologias universais para definirem tudo o que pode ser linguisticamente expresso, os bibliotecários procuram formas de catalogar tudo o que se encontra publicado, etc.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

Um dos objectivos da Web Semântica é o de estender o âmbito local e específico das actuais representações de recursos para um âmbito universal. Para isso, baseia-se na utilização de uma estrutura que possibilita, de forma extremamente simples, definir e representar qualquer tipo de recursos, independentemente de se encontrarem definidos em sistemas de informação heterogéneos ou distribuídos. Essa estrutura pode ser definida como uma rede mundial de ontologias desenvolvidas de acordo com os princípios da Web Semântica. Essa rede mundial de ontologias Web Semânticas permitirá classificar, anotar, definir, integrar e utilizar, de forma universal e interpretável pelos SPC, a maioria da informação que se encontra actualmente armazenada na Web, a qual é, como já vimos, considerada como o maior repositório de informação humana.

Uma Ontologia pode ser definida como sendo o estudo das categorias de todos os recursos que podem existir numa área do conhecimento em particular. Segundo John Sowa [87], "o resultado desse estudo, chamado Ontologia, é um catálogo do tipo de recursos que existem num domínio de interesse, designado por D, a partir da perspectiva de uma pessoa que utiliza uma linguagem, designada por L, com o objectivo de se expressar acerca de D". Ainda segundo John Sowa, "os tipos de recursos definidos numa ontologia representam os predicados, os sentidos das palavras ou os conceitos e tipos de relações existentes na linguagem L quando esta é utilizada na discussão de tópicos no domínio D".

De acordo com Ana Moura [58] e T. Gruber [140], o termo ontologia teve origem na Filosofia, onde é definido como "uma especificação explícita de um conceito", isto é, tem como pressuposto a existência de algo. O uso desta definição gerou alguma controvérsia na comunidade científica da IA, porque, em sistemas inteligentes, o que existe é aquilo que pode ser representado através de um formalismo. Quando uma área do conhecimento é representada através de um formalismo declarativo, o conjunto de objectos que pode ser representado é denominado de universo de discurso. Assim, uma ontologia, segundo a óptica da IA, deverá ser considerada como a descrição explícita e precisa de conceitos e relações que existem num domínio em particular. As próprias estruturas primitivas de representação utilizadas para a descrição de conceitos e

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

relações podem ser consideradas como uma ontologia<sup>103</sup>.

Segundo John Sowa [87], o conceito de ontologia já vem do filósofo grego Aristóteles<sup>104</sup>, o qual além de especificar um método para classificar categorias através das suas propriedades e da utilização de regras de herança<sup>105</sup>, também desenvolveu um método para raciocinar sobre as ideias e os conceitos, o qual é conhecido como a Lógica Formal [141]. A sua invenção dos Silogismos<sup>106</sup> como padrões formais para a representação de regras de inferência, constitui a base para um conjunto alargado de ferramentas, tais como diversas aplicações desenvolvidas para a área da IA, os Sistemas de Programação Orientada a Objectos (SPOO), os dicionários, as ontologias, etc.

Uma ontologia pode classificar-se como Simples, se apenas definir um vocabulário específico para a representação dos recursos associados a uma área do conhecimento. Por outro lado, uma ontologia pode classificar-se como Complexa se além de especificar um determinado vocabulário, também definir um conjunto de axiomas<sup>107</sup> lógicos de forma a atribuir uma semântica ao significado pretendido pelos termos desse vocabulário. De seguida é apresentado um conjunto de características associadas à definição de uma ontologia [143]:

- Controlo - Utilizar um vocabulário controlado;
- Clareza - Descrever de um modo objectivo e total o significado dos seus termos, classes e relações, de forma a não serem criadas interpretações ambíguas;
- Coerência - Possuir relacionamentos hierárquicos rigorosos entre classes e subclasses. Uma asserção inferida dos axiomas de uma ontologia não deve contradizer uma definição da própria ontologia;

---

<sup>103</sup>Um exemplo de uma linguagem definida como uma ontologia é o Esquema RDF, o qual é definido em termos de si próprio.

<sup>104</sup>Mundialmente conhecido (384 a.C. - 322 a.C).

<sup>105</sup>A classificação de uma nova categoria é realizada através do agrupamento das propriedades existentes entre ela e a raiz de uma árvore de categorias e de propriedades das categorias.

<sup>106</sup>Silogismo é definido como sendo "um argumento ou raciocínio formado por três proposições: a maior, a menor (premissas) e a conclusão, sendo esta última deduzida da maior, por intermédio da menor" [142].

<sup>107</sup>Conjunto de proposições que não carecem de demonstração.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

- Expansibilidade - Permitir a definição de novos termos com base no vocabulário existente, sem que tal implique a necessidade de alterar as suas definições originais;
- Codificação simples - Normalizar ao nível do conhecimento, sem envolver codificações particulares ao nível dos símbolos;
- Modelo de Classes - Especificar propriedades em forma de classes;
- Entidades Individuais - Incluir entidades individuais;
- Restrições - Especificar restrições de valores em forma de classes;
- Relações semanticamente ricas - Especificar classes **disjuntas**, distinção entre relações tais como **inverso** e **parte/totalidade**;
- Modelo Lógico - Especificar relações lógicas de forma arbitrária entre os termos.

As primeiras cinco características (Controlo, Clareza, Coerência, Expansibilidade e Codificação simples) são consideradas propriedades típicas e imperativas de uma ontologia. A sexta, sétima e oitava propriedades apresentadas são consideradas propriedades típicas mas não imperativas. As restantes propriedades são consideradas desejáveis, mas não típicas nem imperativas.

Uma ontologia, tipicamente, contém uma hierarquia dos recursos existentes numa área de conhecimento e descreve as propriedades associadas a cada um dos seus recursos através de um mecanismo de atributo-valor. As relações existentes entre os recursos são descritas através de axiomas lógicos adicionais. Os valores individuais, como "Janeiro", são atribuídos a um ou mais recursos, como "Mês", de maneira a associar entre si tipos de propriedades. A utilização de uma Lógica matemática, como o Cálculo de Predicados<sup>108</sup>, a qual não impõe nenhuma restrição associada aos recursos ou ao modo como os recursos devem ser utilizados, em conjugação com

---

<sup>108</sup>em 1879, Gottlob Frege desenvolveu a *Begriffsschrift* - uma notação em árvore para a Lógica. Em 1883, Charles Sanders Peirce desenvolveu uma notação algébrica para o Cálculo de Predicados (*Predicate Calculus*). Actualmente, esta última notação é a mais utilizada, mas especificada com um conjunto diferente de símbolos, os quais foram propostos por Giuseppe Peano em 1889.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

um vocabulário, possibilita a definição de linguagens que podem, facilmente, expressar relações semânticas entre os diversos recursos de uma ou várias áreas de conhecimento.

Ainda segundo John Sowa [87], existem dois tipos de linguagens universais que podem ser utilizadas para descrever os recursos:

- Linguagem Natural - Permite descrever qualquer aspecto da experiência humana. Por exemplo, o Português, Inglês, Chinês, etc;
- Linguagem Lógica<sup>109</sup> - Tudo o que possa ser expressado de forma clara e precisa em Linguagem Natural, ou possa ser implementado num computador através de uma linguagem de programação, pode ser convertido em Linguagem Lógica.

Apesar de tudo o que possa ser expressado de forma clara e precisa, possa ser convertido em Linguagem Lógica, nem sempre essa conversão é facilmente conseguida. Desta forma, a tentativa de relacionar diferentes sistemas de Lógica é muitas das vezes extremamente complicada [87]:

- Apesar de a partilha de informação entre diferentes variações de implementações de Lógica ser normalmente possível, por vezes, poderá existir perda ou alteração da informação;
- Um recurso pode ser definido de maneira diferente em diferentes Ontologias; Outras vezes, pode acontecer ser utilizada uma mesma definição para recursos diferentes;
- A ordem de implementação, ou processamento, da informação e das regras pode gerar, sobretudo em diferentes sistemas, diferentes resultados.

---

<sup>109</sup>Gottlob Frege e Charles Sanders Peirce são conhecidos por terem desenvolvido a Lógica de Primeira-Ordem (FOL - *First-Order Logic*). Apesar de ter sido definida há um século continua a ser a base de todas as novas teorias lógicas apresentadas pelos lógicos e pelos matemáticos. Toda a matemática existente, qualquer computador digital que tenha sido alguma vez construído, a semântica da Lógica Fuzzy [144], a Lógica Modal, as Redes Neurais [145], a Lógica de Elevada-Ordem (*Higher-Order Logic*) e incluindo ela própria podem ser definidas em FOL. Em 1935, Alfred Tarski estabeleceu a semântica actual da FOL.

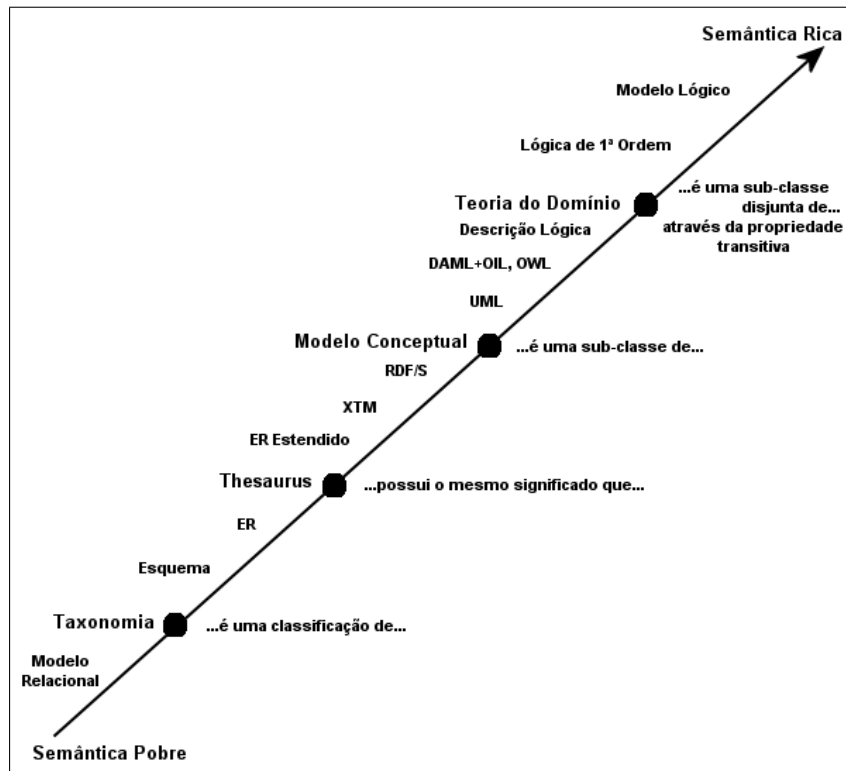


Figura 3.14: Níveis de riqueza semântica associados às Ontologias [146].

As ontologias são geralmente classificadas de acordo com o grau de riqueza semântica que representam, conforme se pode ver na figura 3.14. Elas podem ser agrupadas em três grupos de acordo com o seu grau de complexidade e nível de representação de riqueza semântica: simples, semi-complexas e complexas. Na sua forma simples, representam o conhecimento através de uma estrutura hierárquica, geralmente referida como uma taxonomia. Na sua forma semi-complexa, representam o conhecimento através de thesaurus e de modelos conceptuais e finalmente, na sua forma complexa, representam o conhecimento através de Teorias Lógicas [146].

Uma taxonomia, o tipo mais básico de uma ontologia, é definida através de um modelo hierárquico que representa recursos e as suas relações, através da definição de classes de recursos e das relações existentes entre elas. Desta forma, é possível expressar um grande número de relações entre recursos atribuindo propriedades às classes e permitindo que as sub-classes herdem tais propriedades. Os recursos posicionados perto do topo de uma taxonomia representam conceitos

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

gerais, à medida que a percorremos de forma descendente, os recursos vão representando conceitos cada vez mais específicos. As taxonomias são muito utilizadas no estabelecimento de uma semântica simples para a definição de um espaço de informação ou de uma área de conhecimento.

O termo thesaurus foi apresentado pela primeira vez por Peter Mark Roget através da publicação da *Roget's Thesaurus of English Words and Phrases* em 1852 [147]. Um Thesaurus é uma Taxonomia enriquecida com relações semânticas simples entre os seus termos<sup>110</sup>. De uma maneira geral, é definido como um vocabulário estruturado onde as relações de equivalência, homografia<sup>111</sup>, hierarquia e associação entre os termos são apresentadas de uma forma clara e facilmente identificável a partir de indicadores de relacionamento normalizados. Os principais objectivos de um Thesaurus consistem no armazenamento consistente da informação e a sua fácil recuperação [148]. O Thesaurus mais conhecido e utilizado é o *WordNet Thesaurus* [149]. Ele foi desenvolvido, tendo em consideração as teorias psico-linguísticas sobre a memória lexical humana e pode ser utilizado localmente ou consultado interactivamente através da Web [146].

Segundo D. Tschritzis e A. Klug [150], desde a década de 70 que existe a preocupação em criar uma base normalizada para a representação do conhecimento de uma forma universal. Em 1977, a *American National Standards Institute* (ANSI) propôs que todo o conhecimento relevante associado a um domínio específico deveria ser representado através de um modelo conceptual. Desta forma, um modelo conceptual é descrito como sendo um modelo que representa e descreve os conceitos, eventos, atributos e relações existentes, dentro de uma área do conhecimento específica.

A figura 3.15, apresenta um sistema integrado baseado num modelo conceptual. Cada círculo representa uma área de intervenção especializada (BD, Aplicações e interações com o utilizador), mas todas foram desenvolvidas a partir de uma representação comum do conhecimento (modelo

---

<sup>110</sup>palavras ou frases.

<sup>111</sup>dependência recíproca entre dois termos.

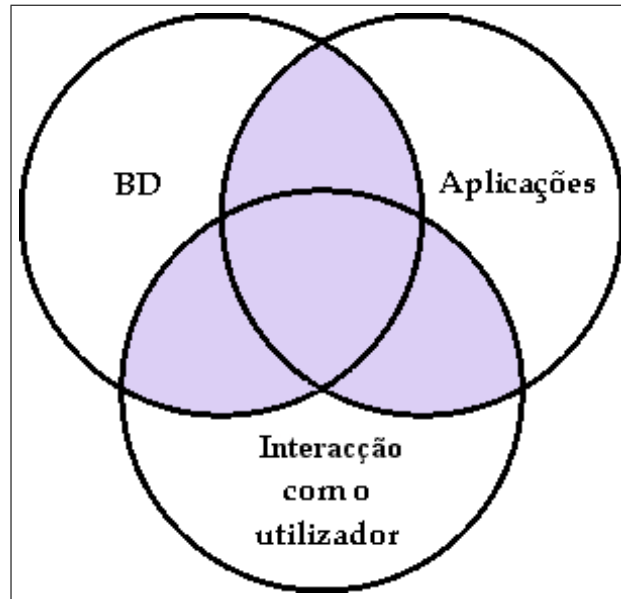


Figura 3.15: O modelo conceptual como uma representação comum do conhecimento (Adaptado de [87]).

conceptual - representado pela zona central sombreada). A este tipo de modelo encontram-se associadas regras simples definidas em termos de axiomas e restrições.

As ontologias complexas na forma de Teorias Lógicas são, normalmente, constituídas por axiomas e regras de inferência. As regras de inferência permitem a geração de conclusões válidas a partir de asserções. Os axiomas e as regras de inferência são utilizados para provar teoremas que envolvam a representação de uma área do conhecimento. O conjunto de todos os axiomas, regras de inferência e os teoremas, constituem a Teoria Lógica.

Segundo Michael Daconta *et.al.* [146], todas as ontologias possuem como base estrutural uma taxonomia. Desta forma, podemos dizer que as taxonomias fornecem a estrutura básica para a representação de um espaço de informação. Mas, uma taxonomia, por si só, não possui suficiente informação semântica para concretizar o objectivo da Web Semântica. Ao contrário destas, as ontologias na forma de Teoria Lógica possuem uma semântica rica e são facilmente interpretáveis pelos SPC, o que as coloca no actual estado da arte para a representação de informação e conhecimento na Web Semântica.



### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

Além da sua complexidade, uma ontologia pode ser classificada como [87]:

- Informal - Quando é especificada por um catálogo de tipos indefinidos ou definidos apenas numa linguagem natural;
- Formal - Quando é especificada por um colecção de nomes de conceitos e de tipos de relações organizados por uma ordem particular através da relação tipo/sub-tipo. Por sua vez, as ontologias formais ainda podem ser divididas em mais dois tipos, de acordo com a forma como os seus sub-tipos são distinguidos dos seus super-tipos:
  - Axiomática - Os sub-tipos são distinguidos através da utilização de axiomas e de definições baseadas numa linguagem formal, como a Lógica ou outra linguagem informática que possa ser convertida para Lógica;
  - Protótipa - Os sub-tipos são distinguidos através da comparação com um membro típico para cada sub-tipo.

Ainda segundo John Sowa [87], a maioria das grandes ontologias utilizam uma mistura de diferentes métodos de definição, sendo os axiomas e as definições formais mais utilizadas na matemática, física e engenharia e os protótipos na biologia e nos recursos utilizados no dia-a-dia.

Qualquer ontologia pode ser facilmente estendida, refinada e utilizada por outras ontologias. Desta forma, as ontologias permitem construir modelos com representações expressivas para a generalidade dos objectos e conceitos do mundo real. Mas, à medida que se vai alargando o domínio de interacção, também se vai aumentando a complexidade do sistema de partilha de informação. A procura, definição e utilização de incontáveis normalizações para a troca de informação, ao contrário de contribuir para a simplificação do problema, tornou-se uma das componentes do problema - cada área de conhecimento possui as suas próprias normalizações, terminologias e convenções. Desta forma, existe a necessidade de criar formas normalizadas

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

que possibilitem uma correcta partilha de informação entre as actuais normalizações específicas. O problema de associar conceitos e relações entre várias ontologias, é idêntico ao de alinhar palavras de vocabulários pertencentes a línguas naturais diferentes. Sendo assim, duas ontologias podem referir-se a uma mesma especificação de um objecto ou conceito, através de uma interligação entre as suas definições, muito embora possam apresentar vocabulários distintos de representação, tal como por exemplo, o uso de dois idiomas distintos para descrever uma mesma ontologia.

Segundo John Sowa [87], é relativamente frequente surgirem desalinhamentos entre diferentes ontologias que especificam conceitos comuns. Esses desalinhamentos são geralmente resultantes de diferenças culturais, geográficas, linguísticas, técnicas de implementação e, por vezes, até de diferenças aleatórias. Estes desalinhamentos representam inconvenientes mas não criam, obrigatoriamente, inconsistência entre ontologias desde que sejam tratados através da adição de conceitos que especifiquem essas distinções. A acção de adicionar novos elementos a uma ontologia é designado por Refinamento.

Actualmente, existem vários tipos de linguagens utilizadas para a criação de ontologias sendo as apresentadas a seguir aquelas que foram desenvolvidas no âmbito da camada Ontologia da arquitectura estratificada de tecnologias da Web Semântica. Com excepção das linguagens *Simple HTML Ontology Extensions* (SHOE)<sup>112</sup>[151] e *XML-based Ontology Exchange Language* (XOL)<sup>113</sup> [152], a principal característica que estas linguagens apresentam é a capacidade de representarem a informação no formato RDF/XML<sup>114</sup>:

- SHOE - É uma extensão para a linguagem HTML e para a XML, a qual fornece um meio de incorporar informação semântica processável pelos SPC nos documentos web. O seu principal objectivo é o de definir a informação existentes nos documentos web de forma a

---

<sup>112</sup>A linguagem SHOE é referida aqui por ter sido uma das primeiras linguagens ontológicas a ser utilizada no âmbito da ideia da Web Semântica por Sean Luke e Jeff Heflin

<sup>113</sup>Referida pelas mesmas razões da linguagem SHOE.

<sup>114</sup>Formato consagrado pelo consórcio W3C para a especificação e troca de informações na Web Semântica.

### 3.5. ARQUITECTURA ESTRATIFICADA DA WEB SEMÂNTICA

---

melhorar os resultados das ferramentas de pesquisa de informação na Web;

- XOL - É uma linguagem de especificação de ontologias, especificada em DTD e XML. Utiliza um modelo semântico, denominado OKBC [153], que se baseia em estruturas de IA normalmente conhecidas por Quadros<sup>115</sup>;
- *Ontology Inference Layer* (OIL) [154] - É uma linguagem para a especificação de Ontologias que reúne as seguintes características:
  - Fornece primitivas de modelação em estruturas IA denominadas Quadros;
  - Compatível com a XML e o Esquema RDF;
  - Possui uma semântica bem definida, simples e clara, baseada em descrições lógicas;
  - Apresenta um suporte para deduções automáticas.
- *DARPA Agent Markup Language* (DAML) [155] - A linguagem DAML é uma iniciativa da agência *Defense Advanced Research Projects Agency* (DARPA) que também foi desenvolvida como uma extensão da XML e do Esquema RDF;
- *DARPA Agent Markup Language+Ontology Inference Layer* (DAML+OIL) [156] - É uma linguagem que surgiu da combinação da linguagem DAML com a linguagem OIL. A combinação DAML+OIL sofre muita influência da linguagem OIL original, embora não utilize o seu conceito de Quadros;
- *Web Ontology Language* (OWL) - A linguagem OWL é a mais recente e é considerada pelo consórcio W3C como a linguagem ontológica de maior relevância no actual contexto de desenvolvimento de ferramentas para a Web Semântica. É uma Recomendação W3C de 10 de Fevereiro de 2004 [5].

Uma das características da ferramenta desenvolvida durante esta dissertação, e que será descrita no capítulo 5, é a de aceder, de forma automática e transparente, a conceitos definidos em Ontologias OWL.

---

<sup>115</sup>Do termo em inglês: *frames*.

A Web Semântica só poderá ser materializada através de um inter-relacionamento automático entre diversas ontologias. Elas poderão ser desenvolvidas de forma totalmente independente e, regra geral, serão específicas para domínios de actuação específicos. A ferramenta, actualmente, mais utilizada para a edição de ontologias é a *Protégé-2000*. Esta e outras ferramentas para a edição de ontologias serão apresentadas, de forma breve, no capítulo 4 que é dedicado às ferramentas para a Web Semântica.

## 3.6 A Imagem da Web Semântica

A arquitectura RDF é considerada como o modelo central da arquitectura da Web Semântica. Ela é geralmente referida como a Grande Teoria Unificadora da Web [157] porque permite a criação de uma rede de informação que se comporta como uma espinha dorsal onde diversas e diferentes estruturas de informação se podem conectar. Ela é a base de suporte da Web Semântica da mesma forma que a linguagem HTML foi a base de suporte da Web original. Ela funciona como uma estrutura comum de suporte para a descrição de recursos e das suas relações, que pode ser facilmente utilizada na interligação de informação descrita no formato XML. De uma forma simples, podemos dizer que a Arquitectura RDF está para a os SPC como as páginas web desenvolvidas em linguagem HTML estão para a espécie humana. Como qualquer pessoa consegue interpretar facilmente a informação que se encontra inserida numa página desenvolvida em HTML, também qualquer SPC conseguirá interpretar a informação que encontre descrita em RDF.

A imagem da figura 3.16 apresenta a visão da Web Semântica como sendo o resultado da combinação de um conjunto alargado de diferente componentes mutuamente dependentes. A utilização de cada um destes componentes é extremamente útil, mas a sua utilização encadeada com os outros componentes criará um efeito sinérgico poderosíssimo, o qual é identificado pelos investigadores da Web Semântica como sendo um Efeito em Rede [158] que corresponderá à materialização da Web Semântica. Assim, a Web Semântica só será uma realidade quando

### 3.6. A IMAGEM DA WEB SEMÂNTICA

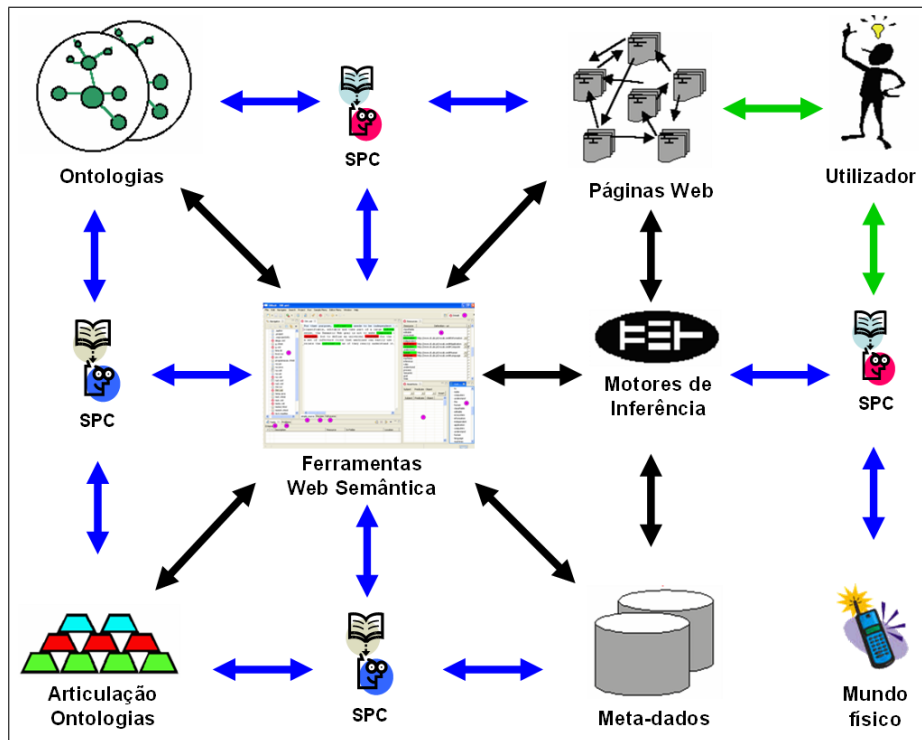


Figura 3.16: Interligação dos vários componentes da Web Semântica (adaptado de [159]).

conseguir a integração dinâmica e autónoma entre milhares/milhões/biliões desses diferentes componentes, como sejam as ferramentas Web Semânticas, as ontologias, os SPCs, as páginas Web, etc. Desta forma, é importante referir que uma única página Web Semântica não possui qualquer valor por si só, sendo que a conjugação e integração dinâmica de várias páginas Web Semânticas, Ontologias e SPCs é que constituirá a verdadeira potencialidade da Web Semântica.

Na zona central da figura 3.16, estão representadas as actuais ferramentas para Web Semântica. Essas ferramentas incluem a criação e gestão de Ontologias, as anotações de páginas Web Semânticas, etc. A sua importância actual é extremamente elevada, como se pode verificar pela sua posição central da imagem, mas será suplantada pelo aparecimento de mais e melhores SPCs num futuro próximo. As ferramentas para a Web Semântica apresentam relações, representadas por setas direccionais, e por vezes servem de ponte às diversas tecnologias associadas à Web Semântica. Na figura apenas aparecem as principais tecnologias relacionadas com a Web Semântica:

### 3.6. A IMAGEM DA WEB SEMÂNTICA

---

- As Ontologias - Representadas na parte superior esquerda da figura;
- As articulações entre ontologias - Representadas na parte inferior esquerda da figura;
- As páginas Web - Representadas na parte superior da figura;
- Os Motores de Inferência<sup>116</sup> - Representadas na parte central da figura;
- Os Meta-dados - Representadas na parte inferior da figura;
- Os SPCs - Espalhados por toda a figura.

Através das inúmeras posições que os SPCs ocupam na figura, podemos constatar que eles actuam em qualquer uma das fases do processo encadeado da Web Semântica. Segundo os investigadores da Web Semântica, eles serão capazes, através da utilização de procedimentos muito simples, de:

- Interpretar e raciocinar acerca do significado da informação, bem como, das relações existentes entre a informação disponibilizada na Web Semântica;
- Trocar entre si informação resultante do processamento interno da informação adquirida através da Web Semântica;
- Trocar entre si informação acerca do tipo de raciocínio por cada um efectuado durante a fase de processamento da informação adquirida; e
- Definir níveis de confiança para a informação e entre si.

Outra característica importante da Web Semântica e representada na parte inferior direita da figura 3.16 através da imagem de um telemóvel, prende-se com o facto de ela transbordar do mundo virtual e abraçar o mundo físico [1]. Nesta nova perspectiva, a Web Semântica irá ser

---

<sup>116</sup>Um Motor de Inferência é definido como um sistema que apresenta a habilidade em raciocinar, interrogar e expressar relações lógicas.

utilizada para descrever dispositivos como telemóveis, televisores, rádios, frigoríficos, automóveis, etc. Estes dispositivos, através da utilização da Web Semântica, poderão anunciar as suas funcionalidades aos SPC através da utilização de uma representação universal da informação, despontando um novo mundo de potencialidades. Como objectivo final da Web Semântica temos os utilizadores da Web, os quais se encontram representados na parte superior direita da figura 3.16.

## 3.7 Resumo

A Web possui uma quantidade imensurável de informação que foi, e ainda é, disponibilizada por diferentes organizações, comunidades e indivíduos. Os utilizadores da Web conseguem aceder a essa informação através da utilização de endereços URL, fazendo pesquisas, ou seguindo as ligações existentes entre diversas páginas web. A sua simplicidade de utilização constitui o factor que mais contribui para a sua popularidade, ao ponto de já se tornar difícil imaginar a nossa vida sem ela. Mas a sua simplicidade também tem um preço, a utilização de ferramentas de pesquisa de informação têm demonstrado resultados ligeiramente ineficientes na obtenção de informação útil por partes dos seus utilizadores. Isto é, a maioria das ferramentas de pesquisa disponibilizam enormes quantidades de informação que normalmente são completamente irrelevante dentro do tema pesquisado.

A Web foi conseqüentemente sendo idealizada como um espaço mundial de armazenamento e de distribuição de informação. Os seus principais objectivos eram os de facilitar a troca de informação e o de melhorar a cooperação existente entre os humanos, entre os humanos e os SPC e entre os SPC. Mas actualmente, um dos maiores obstáculos à concretização desta visão prende-se com o facto de a maioria da informação existente na Web ser disponibilizada com a única preocupação de ser utilizada e compreendida pelos humanos.

Pelo contrário, a Web Semântica pretende desenvolver modelos e tecnologias que permitam

### 3.7. RESUMO

---

definir univocamente o significado da informação, facilitar a interpretação da informação por parte dos SPC e adicionar níveis de confiança associados à informação. Desta forma, a Web Semântica possibilitará um enriquecimento do processo de descoberta, integração e de navegação através da informação, além de servir de suporte ao desenvolvimento de ferramentas web relacionadas com a manipulação autónoma de informação.

A Web Semântica ao permitir estruturar a informação, a definição de heurísticas<sup>117</sup>, o acesso autónomo à informação por parte dos SPC e a associação de informação descritiva e de carácter universal a qualquer tipo de recurso, tal como pessoa, conceito, relação, etc, e não apenas limitar-se à definição de URLs associados a cada um documentos web, como acontece na actual Web, possibilitará que:

- A informação existente nas páginas web venha a ser descrita de forma facilmente interpretável e interligada com conceitos universais;
- Os SPC consigam retirar maior conhecimento sinérgico da informação disponibilizada em páginas web elaboradas segundo o modelo da Web Semântica;
- A actual Web se torne menos impessoal, mais útil e de maior confiança;
- A pesquisa de informação venha a ser mais eficiente.

Apesar de a investigação actual em torno da Web Semântica ser de carácter mundial, ela começou por ser uma iniciativa de Tim Berners Lee e do consórcio W3C, tendo este último definido seis princípios que são actualmente considerados como imprescindíveis para sustentar o desenvolvimento consistente e descentralizado da Web Semântica:

- 1.º princípio: Tudo pode ser identificado através de um URI;

---

<sup>117</sup>Regras do senso comum que são obtidas após um olhar atento sobre um determinado problema ou tipo de problema. Geralmente levam à obtenção de bons resultados, embora não garantam que se obtenha o melhor resultado ou sequer um resultado. Ajudam a reduzir a complexidade na resolução de problemas e são particularmente úteis na resolução de problemas sujeitos à explosão combinatória. Na área científica da AI são tratadas como conhecimento. Em certos casos podem ser vistas como regras do senso comum que se aplicam e que não se discutem, as quais se convencionou designar por *Rules of Thumb*.



- 2.º princípio: Os recursos, e as relações, podem ser tipificados;
- 3.º princípio: Tolerância a quebras de relações;
- 4.º princípio: Desnecessidade da existência de uma confiança absoluta;
- 5.º princípio: Permitir evolução;
- 6.º princípio: Implementação minimalista.

Devido à enorme heterogeneidade de modelos e de abordagens possíveis de associar à Web Semântica, é muito difícil especificar uma estrutura universal e única para a representação do significado da informação. Neste momento assiste-se a um esforço, por parte dos investigadores da Web Semântica, no sentido de definirem e estabelecerem uma arquitectura estratificada para a Web Semântica. Assim, a actual arquitectura apresenta algumas camadas compostas por especificações que ainda se encontram em investigação ou em discussão. Na base da arquitectura, encontra-se a camada de **Unicode e URI**, a qual preserva a utilização de conjuntos de caracteres internacionais e possibilita um meio para identificar os recursos na Web Semântica. Um degrau acima na arquitectura, encontram-se as camadas de **XML + Espaços de Nomes e Esquema XML**, as quais asseguram a normalização sintáctica da representação da informação. Continuando a subir na arquitectura, encontramos a camada **RDF + Esquema RDF**, a qual assegura a normalização estrutural na definição de asserções baseadas nos URIs dos recursos. As asserções RDF são utilizadas para definir os recursos da Web Semântica e as suas relações. Por cima desta camada encontra-se a camada de **Ontologias**, a qual permite a especificação de vocabulários para áreas de conhecimento específicas, definindo relações complexas e universais entre os diferentes recursos da Web Semântica. As camadas superiores: **Lógica, Prova e Confiança**, encontram-se ainda em fase de investigação. A camada Lógica possibilitará a escrita de regras semânticas enquanto a camada de Prova permitirá executar essas regras e avaliar em conjunto com a camada de Confiança quando se deverá confiar na informação fornecida, isto é, atestam a validade da informação. Em paralelo com as camadas de RDF, Ontologia, Lógica e Prova

### **3.7. RESUMO**

---

encontra-se a camada de Assinatura/Encriptação, a qual é utilizada na detecção de alterações sofridas ao longo do tempo pela informação na Web Semântica. São estas as camadas que estão neste momento a ser desenvolvidas e normalizadas quer pelos grupos de trabalho existentes no consórcio W3C como, também, por grande parte da comunidade científica mundial.

# Capítulo 4

## Ferramentas para a Web Semântica

"The pessimist complains about the wind,  
the optimist expects it to change,  
the realist adjusts the sails."

William Arthur Ward

### 4.1 Introdução

Um dos objectivos da presente dissertação está relacionado com o desenvolvimento de um editor para a Web Semântica. Esse editor deveria ter como principais características o facto de ser simples de utilizar pelos actuais utilizadores da Web e de integrar a maioria das especificações associadas à arquitectura da Web Semântica. Numa primeira fase, antes de iniciar o desenvolvimento dessa ferramenta, foi realizado um trabalho de pesquisa na Web com o intuito de verificar se existia de facto a necessidade, tendo em conta o desenvolvimento actual da Web Semântica, de implementar uma ferramenta com estas características. Esse trabalho, além de confirmar a necessidade de uma ferramenta com as características referidas no anterior parágrafo, também, revelou que existe um número elevado de novas ferramentas para a Web Semântica e que esse número cresce semanalmente. Desta forma, tornou-se evidente que seria necessário fazer um estudo sobre as principais ferramentas desenvolvidas para a Web Semântica. Assim, este capí-

## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

tulo apresenta o resultado de um trabalho de classificação de oitenta das principais ferramentas desenvolvidas para a Web Semântica [6].

O presente capítulo é constituído por três secções. Nesta primeira secção é apresentada uma breve introdução ao tema que vai ser abordado neste capítulo. Além disso apresenta, também, uma pequena descrição da estrutura adoptada neste capítulo. Na segunda secção, **Ferramentas para a Web Semântica**, é apresentado o resultado de um estudo de categorização de oitenta das principais ferramentas desenvolvidas para a Web Semântica. Por último, é apresentado na terceira secção um pequeno resumo acerca do tema abordado neste capítulo.

### 4.2 Classificação em Categorias das Ferramentas para a Web Semântica

Se até à pouco tempo existiam dúvidas por parte da comunidade científica sobre a credibilidade e a fiabilidade da arquitectura proposta para a Web Semântica por Tim Berners-Lee e pelo W3C, a qual fora apontada como sendo a mais provável para a futura Web, essas dúvidas foram sendo progressivamente dissipadas devido ao imparável aparecimento e a uma escala mundial, de uma enorme quantidade de ferramentas para a Web Semântica. Esse galopante aparecimento de novas ferramentas para a Web Semântica, bem como a diversidade de áreas e de formas de actuação que apresentam, resulta numa exigência crescente em termos de conhecimentos das actuais abordagens utilizadas no desenvolvimento de ferramentas por parte dos investigadores e criadores de ferramentas para a Web Semântica. O presente capítulo pretende contribuir para a clarificação dos tipos e características de ferramentas disponíveis, sendo apresentado um estudo de categorização e de análise das características de oitenta ferramentas desenvolvidas para a Web Semântica [6]. Essas ferramentas foram classificadas e catalogadas nas seguintes onze categorias:

- Categoria Anotação: 6 ferramentas;

## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

- Categoria API: 11 ferramentas;
- Categoria Armazenamento: 9 ferramentas;
- Categoria Conversão: 6 ferramentas;
- Categoria Edição: 8 ferramentas;
- Categoria Integração: 8 ferramentas;
- Categoria Inferência: 12 ferramentas;
- Categoria Interpretador Web: 7 ferramentas;
- Categoria Interrogação: 7 ferramentas;
- Categoria Validação: 3 ferramentas;
- Categoria Visualização: 3 ferramentas.

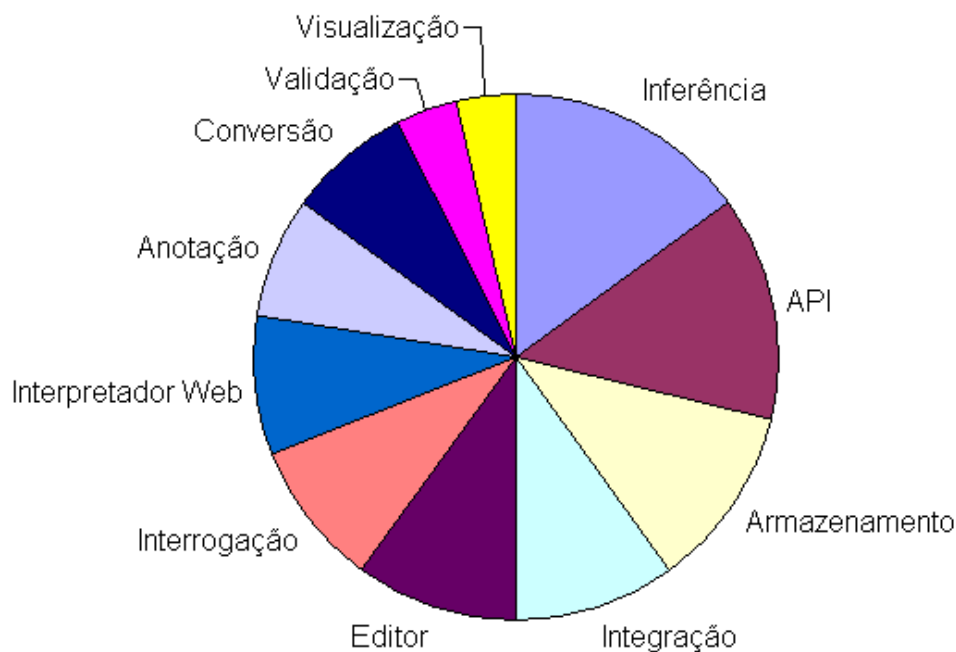


Figura 4.1: Classificação das ferramentas disponíveis para a Web Semântica.

## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

A figura 4.1 apresenta a distribuição das oitenta ferramentas, utilizadas no estudo, pelas onze categorias apresentadas anteriormente. Para facilitar a realização de uma análise mais profunda sobre as principais características apresentadas por cada uma destas ferramentas são também apresentadas três novos agrupamentos das mesmas ferramentas de acordo com as seguintes características específicas:

- Linguagens de programação utilizadas no desenvolvimento da ferramenta;
- Tipo de licença de utilização associada à ferramenta;
- Ano de distribuição da última versão da ferramenta.

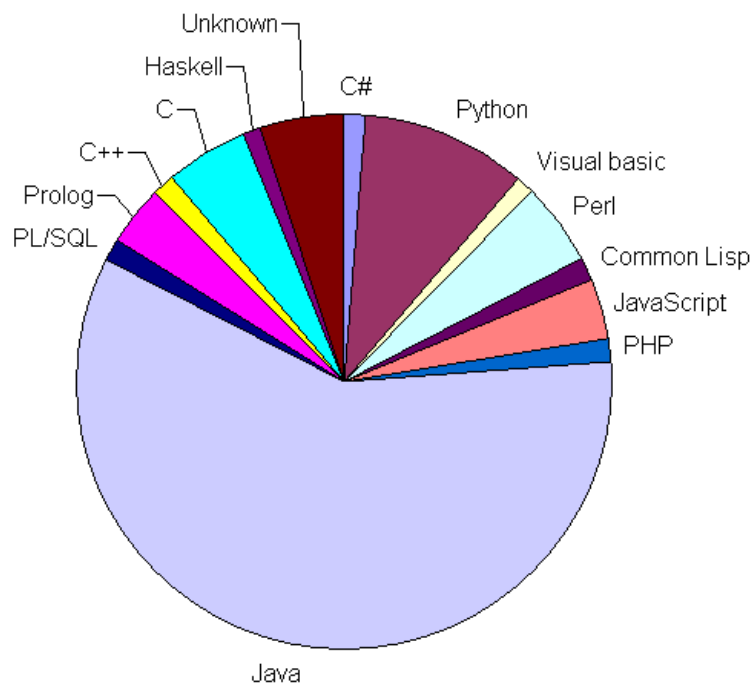


Figura 4.2: Ferramentas agrupadas por linguagem utilizada no seu desenvolvimento.

A figura 4.2 apresenta a distribuição das oitenta ferramentas por treze linguagens de programação utilizadas no seu desenvolvimento:

- *Java* [160]: 59% das ferramentas. Uma possível justificação para o facto de a linguagem *Java* ser mais utilizada com uma percentagem próxima do 60% deve-se ao facto

## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

de as ferramentas desenvolvidas em linguagem *Java* serem independentes dos Sistemas Operativos;

- *Python*<sup>118</sup>: 10% das ferramentas;
- C: 5% das ferramentas;
- *Perl*: 5% das ferramentas;
- *Prolog*: 4% das ferramentas;
- *JavaScript*: 4% das ferramentas;
- C++: 1% das ferramentas;
- C#: 1% das ferramentas;
- *Common Lisp*: 1% das ferramentas;
- *Visual Basic*: 1% das ferramentas;
- *Haskell*<sup>119</sup>: 1% das ferramentas;
- PHP: 1% das ferramentas;
- *Procedural Language extensions to SQL* (PL/SQL) [163]: 1% das ferramentas.

Na figura 4.3 podemos observar a distribuição das oitenta ferramentas pelos seguintes nove tipos de licença de utilização:

- *Berkeley Software Distribution* (BSD) [164]: 25% das ferramentas;
- *GNU Lesser General Public License* (LGPL) [165]: 19% das ferramentas;
- *GNU General Public License* (GPL) [101]: 18% das ferramentas;

---

<sup>118</sup>É uma linguagem de programação interpretável, interactiva e orientada aos objectos. É geralmente comparada com as linguagens *Perl* e *Java* [161].

<sup>119</sup>É uma linguagem de programação funcional [162].

## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

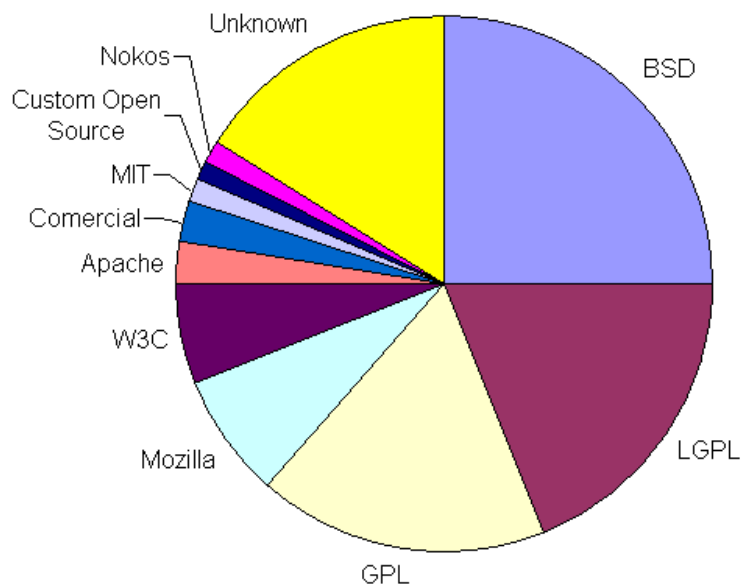


Figura 4.3: Ferramentas agrupadas por licença.

- *Mozilla Public License* (MPL) [166]: 8% das ferramentas;
- *W3C Software License* (W3L) [167]: 6% das ferramentas;
- *Apache Software License* (ASL) [168]: 3% das ferramentas;
- Proprietário/Comercial: 3% das ferramentas;
- *Massachusetts Institute of Technology License* (MIT License) [169]: 1% das ferramentas;
- *Nokia Open Source License* (NOKOS) [170]: 1% das ferramentas.

A figura 4.4 apresenta a distribuição das ferramentas por ano de versão. Através da análise dos dados dos anos podemos verificar que a taxa de aparecimento de novas ferramentas está a crescer muito rapidamente:

- 2005: 30% das ferramentas. De notar que apenas é contabilizado o tempo desde Janeiro até Maio de 2005;
- 2004: 51% das ferramentas;



## 4.2. CLASSIFICAÇÃO EM CATEGORIAS DAS FERRAMENTAS PARA A WEB SEMÂNTICA

---

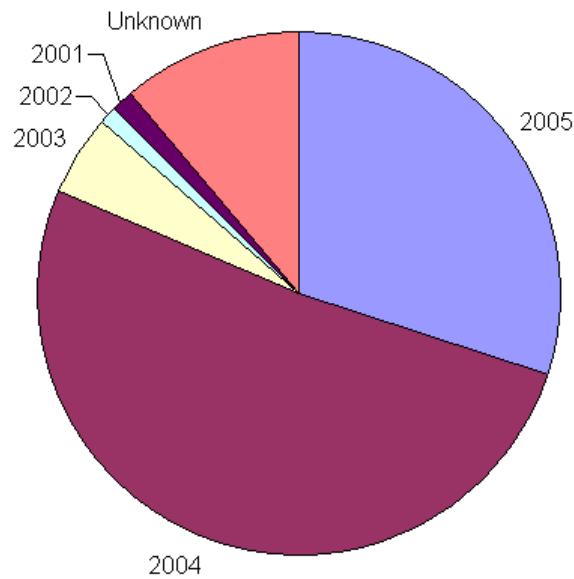


Figura 4.4: Ferramentas agrupadas por anos.

- 2003: 5% das ferramentas;
- 2002: 1% das ferramentas;
- 2001: 1% das ferramentas.

As próximas onze secções descrevem cada uma das categorias atrás referidas, sendo apresentada uma breve caracterização das ferramentas para cada tipo das categorias, através de tabelas. Cada tabela é constituída pelo seguinte conjunto de características associadas às diversas ferramentas:

- Autor: Indivíduo ou entidade que desenvolveu a ferramenta;
- Versão: Número da última versão e a data da sua publicação;
- Licença: Tipo de licença utilizada na sua distribuição;
- Linguagem: Linguagem de programação utilizada no seu desenvolvimento;
- Formatos de entrada (*INPUT*): Tipos de ficheiros de dados que a ferramenta conhece;

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

- Formatos de saída (*OUTPUT*): Tipos de ficheiros de dados que a ferramenta produz através da sua utilização.

## 4.3 Ferramentas associadas a cada Categoria

### 4.3.1 Categoria - Anotação

A categoria Anotação inclui seis ferramentas. Na tabela 4.1 podemos verificar o nome de cada uma dessas ferramentas e, também, as suas principais características.

Tabela 4.1: Categoria - Anotação

Anotação						
	Autor	Versão	Licença	Linguagem	<i>INPUT</i>	<i>OUTPUT</i>
AeroSWARM	UBOT Team	-	-	-	HTML	OWL
OntoMat Annotizer	Handschuh, S., Braunv, M., Buerkle, C., Kühn, K., Meyer, L. e Krekeler, T.	0.8.2 (Feb-5)	LGPL	Java	N3 OWL RDF	N3 OWL RDF
PhotoStuff	Mindswap	2.11 (Mar05)	MPL	-	-	-
RIC	Michael Grove (Mindswap)	3.0 Alpha	-	Java	-	-
SemanticWord	Teknowledge Corporation	1.0 (Aug-04)	-	Visual Basic	OWL	OWL
Swangler	Clay Fink, James Mayfield	1.0.1 (Apr-05)	GPL	Java	OWL RDF	OWL

Com a excepção da ferramenta *SemanticWord* [171, 172], que foi desenvolvida para o Sistema Operativo (SO) *Windows NT/2000*, todas as outras ferramentas são independentes do SO. Em termos de interacção com o utilizador, as ferramentas *SemanticWord*, *OntoMat-Annotizer* [173], *PhotoStuff* [174, 175] e a *RDF Instance Creator* (RIC) [8, 176] disponibilizam uma interacção com o utilizador de forma gráfica (GUI)<sup>120</sup>, a ferramenta *Swangler* [177] disponibiliza uma interacção através de uma API e a ferramenta *AeroText Semantic Web Automated Relation Markup* (*AeroSWARM*) [178, 179] disponibiliza uma interacção através da Web e do protocolo *Simple Object Access Protocol* (SOAP) [180].

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

---

<sup>120</sup> *Graphical User Interface*.

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

- *AeroSWARM*: É um Serviço Web que tem como principal função analisar e extrair informação genérica a partir do conteúdo de páginas web. Por cada entidade de informação encontrada, tal como pessoas, locais, organizações, datas, nacionalidade, etc, a ferramenta gera marcações em formato OWL;
- *OntoMat-Annotizer*: É uma ferramenta que possibilita a anotação de página web em OWL;
- *PhotoStuff*: É uma ferramenta de anotação de imagens. Ela fornece ao utilizador a possibilidade de anotar regiões de images de acordo com conceitos definidos em ontologias;
- *RIC*: É uma ferramenta que permite facilitar o processo de anotação de informação. Ela possibilita a geração de triplas RDF válidas através da utilização de simples formulários. O seu principal objectivo é libertar o utilizador da necessidade de conhecer e dominar as triplas RDF;
- *Semantic Word*: É uma ferramenta que permite efectuar associações entre marcações OWL e regiões específicas em documentos do editor de texto *Microsoft Word* [181];
- *Swangler*: É uma ferramenta que permite criar documentos anotados a partir de documentos baseados em RDF. Num modelo mais abrangente de categorização, esta ferramenta poderá, também, ser incluída na categoria API.

#### 4.3.2 Categoria - API

A categoria API inclui onze ferramentas. Na tabela 4.2 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Enquanto que a ferramenta *OWL Semantic Search Services* [182] foi implementada para funcionar no SO *Linux* [183] as ferramentas *Components for Ontology Processing (CODIP)* [184], *HAWK* [185], *Kazuki* [186], *Named Graphs API for Jena (NG4J)* [187], *OWL API* [188] e

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

*Simple Ontology Framework API* (SOFA) [189] são independentes do SO. Em termos de interação com o utilizador, a ferramenta *OWL Semantic Search Services* disponibiliza interações através da Web e através de uma API e as ferramentas CODIP e *Kazuki* disponibilizam uma interações através de uma API e através da linha de comandos. Como paradigmas utilizados pelos Interpretadores RDF na manipulação de modelos RDF, as ferramentas CARA [190] e *Sparta* [191] utilizam o paradigma Centrado nos Recursos RDF, as ferramentas *Python Rdf Api* (*Pyrple*) [192] e NG4J utilizam o paradigma Centrado nas Asserções RDF e a ferramenta SOFA utiliza os paradigmas Centrado nos Recursos RDF e Centrado nas Ontologias. Como modelo de armazenamento, as ferramentas CARA e *Pyrple* utilizam simplesmente a memória, a ferramenta NG4J utiliza memória e a BD *MySQL* [193] e a ferramenta SOFA utiliza a memória e BDs compatíveis com a *Java Database Connectivity API* (JDBC) [194].

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

Tabela 4.2: Categoria - API

API						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
CARA	Stefan Kokkelink	0.001 (Mar-01)	GPL	PERL	RDF/XML N-Triples	RDF/XML N-Triples
CODIP	DARPA	0.9.0 (Dec-04)	BSD	Java	OWL UML XMI	OWL UML XMI
HAWK	SWAT Lab, Lehigh University	1.1 (Dec-04)	GPL	Java PL/SQL	OWL RDF	OWL RDF
Kazuki	Self, T., Lerner, J. e Rager D.	1.2 (Jun-02)	BSD	Java	OWL	-
NG4J	Chris Bizer Richard Cyganiak Rowland Watkins	0.4 (Fev-05)	BSD	Java	RDF/XML N-Triples N3 TriX TriG	RDF/XML N-Triples N3 TriX TriG
OWL API	Bechhofer, S., Volz, R., Kalyanpur, A., Crowther, P., Horan, B., Turi, D., e Lord, P.	1.4.2 (Mar-05)	LGPL	Java	OWL	OWL
OWL Semantic Search Services	Bhanu Vasireddy John Li	0.1 (Dec-04)	-	Java JavaScript Prolog	-	-
OWL-S API	Paolucci, M., Srinivasan, N. Softagents Group	0.1 beta (Dec-04)	LGPL	Java	OWL	OWL
Pyrple	Sean B. Palmer	Jun-04	-	Python	RDF/XML N3 N-Triples	RDF/XML N3 N-Triples
SOFA	Alishevskikh, A. Mihalik, I. e Ganesh, S.	0.3 (Mar-05)	LGPL	Java	OWL RDF	OWL RDF
Sparta	Mark Nottingham	0.6	-	Python	RDF/XML N-Triples	RDF/XML N-Triples

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

- CARA: É uma API desenvolvida com base no modelo RDF. Permite manipular grafos RDF e armazená-los de forma persistente. Possui um Interpretador RDF;
- CODIP: É uma API que fornece uma colecção de módulos em linguagem *Java* para o processamento, criação e gestão de ontologias. Esta ferramenta, num modelo mais abrangente de categorização, pode também ser incluída na categoria Conversão;
- HAWK: É uma ferramenta que fornece APIs para interpretação, edição, manipulação e armazenamento de ontologias OWL. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Interrogação e Armazenamento;
- *Kazuki*: É uma ferramenta que permite gerar uma API *Java* para manipular instâncias de informação no formato OWL directamente a partir de um conjunto de ontologias OWL;
- NG4J: É uma API de extensão para a ferramenta *Jena*. A sua função é permitir a interpretação, a manipulação e a serialização de conjuntos de grafos etiquetados<sup>121</sup>;
- OWL API: É uma API que permite interagir com estruturas definidas em ontologias OWL;
- *Web Services Ontology Language (OWL-S) API* [196]: É uma API que permite manipular informação que se encontre no formato OWL-S;
- *OWL Semantic Search Services*: É uma API que permite procurar e indexar informação existente na Web e que se encontre especificada no formato DAML/OWL. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Interrogação e Armazenamento;
- *Pyrp*: É uma API que permite interpretar informação disponibilizada nos formatos RDF/XML, N3 e N-Triples;
- SOFA: É uma API que permite representar uma ontologia através de um modelo de objectos abstracto e independente da linguagem de especificação utilizada na ontologia;

---

<sup>121</sup> *Named Graphs* [195].

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

- *Sparta*: É uma API que permite associar nodos RDF a objectos da linguagem *Python* e arcos RDF a atributos desses objectos.

#### 4.3.3 Categoria - Armazenamento

A categoria Armazenamento inclui nove ferramentas. Na tabela 4.3 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.3: Categoria - Armazenamento

Armazenamento						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
3store	Riddoch, A., Gibbins, N. Harris, S., Beckett, D. Dawes, P.	2.2.18 (Nov-04)	GPL	C	-	-
An Entry Sub-ontology of OWL-Time	Feng Pan	Dec-04	-	-	OWL	OWL
ParkaSW	Mindswap	1.1b (Apr-04)	MIT	C	-	-
PySesame	Tom Hoffman	0.1 (Mar-04)	LGPL	Python	-	-
Time Zone Resource in OWL	Feng Pan	Dec-04	LGPL	Python	OWL	OWL
RDFStore	Alberto Reggiori	0.50 (Aug-04)	BSD	C	RDF/XML N-Triples XML	RDF/XML N-Triples XML
Sesame	Aduna NLnet Foundation	1.1 (Nov-04)	LGPL	Java	RDF/XML N-Triples N3 Turtle	RDF/XML N-Triples N3 Turtle
Kowari	Tucana Technologies	1.1.0 (Dez-04)	Mozilla	Java	RDF/XML N-Triples	RDF/XML N-Triples
YARS	Andreas Harth Stefan Decker Hannes Gassert	Jan-05	BSD	Java	RDF/XML N-Triples N3	RDF/XML N-Triples N3

A ferramenta *PySesame* [197] é independente do SO e disponibiliza uma API para interacção com o utilizador. As ferramentas *An Entry Sub-ontology of OWL-Time* [198] e *Time Zone Resource in OWL* [199] disponibilizam uma interacção com o utilizador através do *Hypertext Transfer Protocol* (HTTP) [200]. Como paradigmas utilizados pelos Interpretadores RDF na manipulação de modelos RDF, as ferramentas *3Store* [201], *RDFStore* [202] e *Yet Another RDF Store* (YARS) [203] utilizam o paradigma Centrado no Modelo RDF, a ferramenta *Kowari* [204] utiliza os paradigmas Centrado nos Recursos RDF, Centrado nas Asserções RDF e Centrado nas Ontologias e, por fim, a ferramenta *Sesame* [205] utiliza os paradigmas Centrado no Recursos RDF e Centrado nas Asserções RDF. Como modelo de armazenamento, a ferramenta *3Store* utiliza a BD *MySQL*, a ferramenta *RDFStore* utiliza memória, ficheiros e a BD *Berkeley*

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

[206] e a ferramenta *Sesame* utiliza a memória, ficheiros e as BDs *MySQL*, *PostgreSQL* [207], *Oracle* [208] e *Microsoft SQL Server* [209]. Como linguagem de interrogação, a ferramenta *3Store* utiliza as linguagens *RDF Data Query Language* (RDQL) [210] e OKBC [153], a ferramenta *Kowari* utiliza a linguagem *Tucana SQL-like query language* (iTQL) [211], a ferramenta *RDFStore* utiliza as linguagens RDQL e *Protocol and RDF Query Language* (SPARQL) [212], a ferramenta *Sesame* utiliza as linguagens RDFS e *OWL-Lite* [213] e a ferramenta YARS utiliza a linguagem *Notation3 Query Language* (N3QL) [214]. Como suporte do Motor de Inferência, as ferramentas *Kowari* e *Sesame* utilizam as linguagens RDFS e *OWL-Lite* e a ferramenta *RDFStore* utiliza apenas a linguagem RDFS.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de funcionamento de cada uma destas ferramentas:

- *3Store*: É uma ferramenta que permite armazenar triplas RDF;
- *Kowari*: Disponibiliza uma BD implementada propositadamente para o armazenamento de meta-dados. É extremamente escalável e possibilita transacções seguras;
- *An Entry Sub-ontology of OWL-Time*: Disponibiliza uma ontologia para descrever temporalmente o conteúdo de páginas Web e propriedades de Serviços Web. Esta sub-ontologia de *OWL-Time* é uma ontologia mais simples que a *OWL-Time* completa, mas que fornece a maioria dos conceitos básicos e relações que a maioria das aplicações necessitam;
- *ParkaSW* [215]: Disponibiliza uma BD para regras de inferência;
- *PySesame*: É uma ferramenta desenvolvida em *Python* que serve de invólucro para a ferramenta *Sesame*;
- *RDFStore*: É uma ferramenta que permite armazenar informação do tipo RDF;
- *Sesame*: Disponibiliza uma BD RDF que suporta inferências e questões para o Esquema RDF. *Sesame* foi desenvolvida tendo como principal objectivo fornecer a máxima flexibilidade no desenvolvimento de ferramentas RDF e ferramentas de Esquemas RDF;

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

- *Time zone resource in OWL*: Disponibiliza uma ontologia que define o vocabulário acerca de regiões e zonas horárias;
- YARS: É uma ferramenta que permite armazenar informação do tipo RDF.

#### 4.3.4 Categoria - Conversão

A categoria Conversão inclui seis ferramentas. Na tabela 4.4 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.4: Categoria - Conversão

Conversão						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
java2owl-s	Naveen Srinivasan Softagents Group	Jun-04	-	-	Java	OWL-S
owl-s2uddi	Naveen Srinivasan Softagents Group	Jun-04	-	-	OWL-S	UDDI
owl2dig	Lei Zhang Jian Zhou	0.1 (Jun-04)	GPL LGPL	Java	OWL	DIG
wSDL2owl-s	ATLAS Group	1.0.1 (Jan-05)	-	Java	XML XML Schema	OWL RDF
Owl Converter	Mindswap	1.2 (Dec-03)	-	Perl	DAML+OIL	OWL
SWeHG	Semantic Computing Research Group	Out-03	-	Perl Prolog	RDF	HTML

A ferramenta *Owl2dig* [216] é independente do SO e utiliza para interacção com o utilizador uma API e a linha de comando. A ferramenta *WSDL2owl-s* [217] utiliza como interacção uma API, o protocolo SOAP e a Web.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *java2owl-s* [218]: É uma ferramenta que permite gerar descrições no formato OWL-S [219] a partir de código fonte disponibilizado em linguagem de programação *Java*;
- *owl2dig*: É uma ferramenta que permite converter informação no formato OWL para o formato *Description Logic Interface* (DIG) [220], um formato XML;
- *Owl Converter* [221]: É uma ferramenta que permite converter Ontologias do tipo DAML+OIL em Ontologias do tipo OWL;



### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

- *owl-s2uddi* [222]: É uma ferramenta que permite converter informação no formato OWL-S para o formato UDDI;
- *SWeHG* [223]: A arquitectura RDF é utilizada para descrever conteúdos de forma a que os SPC os consigam interpretar facilmente. O objectivo desta ferramenta é um pouco o reverso da medalha: como cada vez mais os ficheiros RDF são muito complexos, a ideia é transformar esses ficheiros em páginas HTML e de forma a serem facilmente compreendidas pelos seres humanos;
- *wSDL2owl-s*: É uma ferramenta que permite converter informação no formato *Web Services Description Language* (WSDL) para o formato OWL-S.

#### 4.3.5 Categoria - Edição

A categoria Edição inclui oito ferramentas. Na tabela 4.5 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.5: Categoria - Edição

Edição						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
ORIENT	APEX Lab	0.1.1 (May-04)	-	Java	OWL RDF XMI	OWL RDF XMI
OWL Filetype Plugin for VIM	Jeremy Lerner Troy Self	(Jan-04)	BSD	-	OWL	OWL
OWL Mode for Emacs	BBN Technologies	(Jan-04)	BSD	-	-	OWL
Protégé	Stanford Medical Informatics	3.0 (Fev-05)	Mozilla	Java	-	-
SMORE	Mindswap	5.0 (Aug-05)	LGPL	Java	-	-
SWeDE	BBN Technologies	1.0.2 (Mar-05)	BSD	Java	-	OWL
SWedt	UBI	0.1 (Jun-05)	BSD	Java	XML	XML
SWOOP	Mindswap	2.2 (Mar-05)	LGPL	Java	OWL RDF/XML N3 Turtle	OWL RDF/XML N3 Turtle

Enquanto que a ferramenta *OWL Mode for Emacs* [224] foi desenvolvida para os SO *Windows* e *Portable Operating System Interface* (POSIX) [225], as ferramentas *Ontology engineRING ENvironment* (ORIENT) [226], *Protégé* [9], SMORE [227], SWeDE [228], Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF (SWedt) [13, 14, 15] e SWOOP

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

[229] são independentes do SO. Em termos de interacção com o utilizador, as ferramentas SMORE, SWeDE e SWedt disponibilizam uma interacção através de uma GUI. A ferramenta ORIENT disponibiliza interacções através de uma GUI e uma API e a ferramenta *OWL Mode for Emacs* disponibiliza interacções através de uma GUI e da linha de comandos.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- **ORIENT**: É uma ferramenta desenvolvida para a plataforma *Eclipse* [230] que permite a criação e a manipulação de ontologias. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Armazenamento e Visualização;
- ***OWL Filetype Plugin for VIM*** [231, 232]: É uma ferramenta que fornece um conjunto de funções que podem ser utilizadas para extrair informação a partir de ontologias OWL;
- ***OWL Mode for Emacs***: É um módulo para a manipulação da sintaxe OWL para o editor *Emacs* [233];
- ***Protégé***: É uma ferramenta que permite criar e manipular ontologias. É implementada em *Java* e pode, de uma forma geral, ser definida de três formas:
  - Como uma ferramenta que possibilita a construção e a manipulação de ontologias;
  - Como uma plataforma que pode se estendida ou incorporada em outros sistemas de KR;
  - Como uma biblioteca que pode ser utilizada por outros sistemas de KR.
- **SMORE**: É uma ferramenta que permite a criação e utilização de asserções RDF associadas a ontologias. Os utilizadores necessitam de possuir conhecimentos mínimos sobre a especificação RDF e de fornecer a fonte das ontologias ao sistema;
- **SWeDE**: É uma ferramenta que apresenta um ambiente de desenvolvimento integrado e extensível para ferramentas Web Semânticas. Foi desenvolvida para a plataforma *Eclipse*

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

e inclui um editor de ontologias no formato OWL, um verificador de sintaxe OWL, um conversor<sup>122</sup> de OWL para *Java* e de um visualizador gráfico<sup>123</sup> de ontologias. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Validação e Visualização;

- SWedt: É uma ferramenta desenvolvida para a plataforma *Eclipse* que facilita o processo de criação de páginas web de acordo com os princípios da Web Semântica. Possui como principal objectivo, despertar o interesse dos actuais utilizadores da Web pela Web Semântica. As principais características e o seu modo de funcionamento serão apresentadas, mais detalhadamente, no capítulo 5;
- SWOOP: É uma ferramenta, baseada em hipermédia, que permite a edição de ontologias. Ao contrário da maioria das actuais ferramentas de edição de ontologias, as quais são um pouco complexas de manusear pelos utilizadores da Web, apresenta uma simples e consistente interface para a criação e manipulação de ontologias OWL por parte de todos os utilizadores da Web. A sua principal característica prende-se com o facto de permitir uma navegação simples, do tipo hipermédia, através das entidades definidas nas ontologias OWL.

#### 4.3.6 Categoria - Integração

A categoria Integração inclui oito ferramentas. Na tabela 4.6 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

A ferramenta *RDF Gateway* [235] foi implementada para o SO *Windows*. Como paradigmas utilizados pelos Interpretadores RDF na manipulação de modelos RDF, a ferramenta *Wilbur* [236] utiliza os paradigmas Centrado nos Recursos RDF e Centrado nas Asserções RDF, a ferramenta *4Suite* [237] utiliza o paradigma Centrado nas Asserções RDF, a ferramenta *RDF*

---

<sup>122</sup>API Kazuki.

<sup>123</sup>API DumpOnt [234].

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

*Gateway* utiliza os paradigmas Centrado no Modelo RDF, Centrado nos Recursos RDF e Centrado nas Asserções RDF e as ferramentas *Jena* [134], *KAON* [238], *PerlRDF* [239], *RAP* [240] e *Redland* [241] utilizam os paradigmas Centrado nas Asserções RDF, Centrado nos Recursos RDF e Centrado nas Ontologias. Como modelo de armazenamento, a ferramenta *Wilbur* utiliza simplesmente a memória, as ferramentas *4Suite* e *PerlRDF* utilizam memória, ficheiros e a BD *PostgreSQL*, a ferramenta *Jena* utiliza memória, ficheiros e as BDs *Berkeley*, *MySQL* e *SQLite* [242], a ferramenta *KAON* utiliza memória e qualquer BD compatível com *SQL2* [243], a ferramenta *RAP* utiliza memória e qualquer BD compatível com *ActiveX Data Objects Data Base (ADOdb)* [244], a ferramenta *RDF Gateway* utiliza memória, ficheiros e qualquer BD compatível com *Object Linking and Embedding Data Base (OleDB)* [245] e, por fim, a ferramenta *Redland* utiliza memória, ficheiros e as BDs *MySQL*, *PostgreSQL*, *Oracle*, *Berkeley*, *Microsoft SQLServer* e *DB2* [246]. Como linguagem de interrogação, a ferramenta *4Suite* utiliza a linguagem *Versa* [247], as ferramentas *Jena* e *Redland* utilizam as linguagens *RDQL* e *SPARQL*, a ferramenta *KAON* utiliza a linguagem *KAON*, a ferramenta *PerlRDF* utiliza uma linguagem Centrado nos Recursos RDF, a ferramenta *RAP* utiliza a linguagem *RDQL*, a

Tabela 4.6: Categoria - Integração

Integração						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
4Suite	Fourthought, Inc	1.0a4 (Nov-04)	Apache	Python	RDF/XML	RDF/XML
Jena	HP Labs Semantic Web Research	2.2 (Jan-05)	BSD	Java	RDF/XML N-Triples N3 Turtle	RDF/XML N-Triples N3 Turtle
KAON	FZI WIM e AIFB LS3	1.2.7 (Nov-04)	LGPL	Java	RDF/XML	RDF/XML
PerlRDF	Ginger Alliance	0.31 (Mar-2)	GPL Mozilla	PERL	RDF/XML N3	RDF/XML N3
RAP	Westphal, D., Bizer, C. Oldakowski, R., Gauß, T. Dawes, P., Grimnes, G. Köstlbacher, A., Auer, S. Smith, L., Lopez, L. Catanzani, R., Willy, S.	0.91 (Dec-04)	LGPL	PHP	RDF/XML N-Triples N3 GRDDL	RDF/XML N-Triples N3 GRDDL
RDF Gateway	Intellidimension	2.2.2 (Jan-05)	Commercial	-	RDF/XML N-Triples N3 Turtle	RDF/XML N-Triples N3 Turtle
Redland RDF	Dave Beckett	1.0.0 (Jan-05)	GPL LGPL Apache	C	RDF/XML N-Triples Turtle RSS	RDF/XML N-Triples Turtle RSS
Wilbur	Nokia Research Center	(Nov-04)	NOKOS 1	Common Lisp	RDF/XML	RDF/XML

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

ferramenta *RDF Gateway* utiliza a linguagem *RDF Data Query Language* (RDFQL) [248] e a ferramenta *Wilbur* utiliza a linguagem *Windows Management Instrumentation Query Language* (WQL) [249]. Como suporte ao Motor de Inferência, a ferramenta KAON utiliza as linguagens RDFS e *OWL-Lite*, a ferramenta RAP utiliza a linguagem RDFS e algumas regras da linguagem OWL, a ferramenta *RDF Gateway* utiliza as linguagens RDFS e OWL, a ferramenta *Redland* utiliza as linguagens RDFS, *OWL-Lite* e DIG e a ferramenta *Wilbur* utiliza a linguagem RDFS.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *4Suite*: É uma ferramenta que permite aos seus utilizadores retirarem vantagens das tecnologias XML e RDF no desenvolvimento de aplicações para a Web;
- *Jena*: É uma ferramenta que permite servir de apoio à construção de aplicações Web Semânticas. Ela fornece um ambiente de programação para a utilização de RDF, Esquema RDF, OWL e um Motor de Inferência baseado em regras. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída na categoria API;
- KAON: É uma ferramenta que apresenta uma estrutura que permite criar e manipular ontologias;
- *PerlRDF*: É uma ferramenta que permite interpretar, armazenar e interrogar informação no formato RDF;
- RAP: É uma ferramenta que permite procurar, interpretar, manipular, serializar e apresentar modelos RDF;
- *RDF Gateway*: É uma ferramenta que permite o desenvolvimento e a distribuição de aplicações Web Semânticas;
- *Redland*: É uma ferramenta que permite o desenvolvimento e a manipulação de modelos RDF;

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

- *Wilbur*: É uma ferramenta que permite o desenvolvimento de aplicações que utilizem XML, RDF e DAML+OIL.

#### 4.3.7 Categoria - Inferência

A categoria Inferência é composta por doze ferramentas. Na tabela 4.7 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.7: Categoria - Inferência

Inferência						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
Euler	Jos De Roo	(Fev-05)	W3C	Java	N3	N3
EulerMoz	Doeblin, A., Roo, J., Alos, O. Hernandez, M., Sanchez, J.	(Jan-05)	Mozilla W3C	JavaScript	N3	N3
F-OWL	Youyong Zou	-	BSD	Java Prolog	OWL RDF SWRL	OWL RDF
FaCT++	Tsarkov, D. e Horrocks, I.	0.99	GPL	C++	-	-
Hoolet	-	-	LGPL	Java	OWL SWRL	-
Instance Store	Turi, D., Bechhofer, S. Li, L., Lord, P. e Roberts, D.	1.4.1 (Jul04)	LGPL	Java	OWL	-
Metalog	Marchiori, M., Epifani, A. Trevisan, S. e Saarela, J.	2.1 (Out-03)	W3C	Python	RDF/XML N-Triples	RDF/XML N-Triples
OWLJessKB	Joe Kopena	(Jan-05)	GPL	Java	RDF/XML	RDF/XML
Pellet	Mindswap	1.1.0 (Dec-04)	W3C	Java	OWL RDF/XML	OWL RDF/XML
Pychinko	Mindswap	0.1 (Jan-05)	-	Python	N3	N3
RdfLib	Ian Davis James Carlyle	0.21 (Jan-05)	MIT	C#	RDF/XML N-Triples	RDF/XML N-Triples
Swish	Graham Klyne	0.2.1 (Fev-04)	GPL	Haskell	N3	N3

A ferramenta *FaCT++* [250] foi implementada para o SO *Windows* e para o SO *Linux*, as ferramentas *F-OWL* [251] e *Hoolet* [252] para o S.O. *Linux* e as ferramentas *EulerMoz* [253] e *Instance Store* [254] são independentes do SO. Em termos de interação com o utilizador, a ferramenta *F-OWL* disponibiliza uma GUI, uma API e a linha de comandos, a ferramenta *Instance Store* disponibiliza apenas GUI e uma API. Como paradigmas utilizados pelos Interpretadores RDF na manipulação de modelos RDF, as ferramentas *Euler* [255], *EulerMoz* e *Swish* [256] utilizam o paradigma Centrado no Modelo RDF, a ferramenta *Metalog* [257] utiliza os paradigmas Centrado nos Recursos RDF e Centrado nas Asserções RDF, a ferramenta *OWLJessKB* [258] utiliza os paradigmas Centrado nas Asserções RDF e Centrado nas Ontologias, a ferramenta *Pychinko* [259] utiliza os paradigmas Centrado no Modelo RDF e Centrado nas Asserções RDF

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

e a ferramenta *RdfLib* [260] utiliza os paradigmas Centrado nos Recursos RDF e Centrado nas Ontologias. Como modelo de armazenamento, as ferramentas *Euler*, *EulerMoz*, *OWLJessKB*, *Pychinko* e *Swish* utilizam simplesmente a memória, a ferramenta *Metalog* utiliza memória e ficheiros e a ferramenta *RdfLib* utiliza a memória e a BD *MySQL*. Como linguagem de interrogação, a ferramenta *Euler* utiliza a linguagem N3QL, a ferramenta *Metalog* utiliza a linguagem *Pseudo Natural Language* (PNL) [261], a ferramenta *OWLJessKB* utiliza a linguagem *Jess* [262] e a ferramenta *Pycmhinko* utiliza a linguagem N3. Como suporte ao Motor de Inferência, as ferramentas *Euler* e *EulerMoz* utilizam as linguagens RDFS e OWL, a ferramenta *Metalog* utiliza a linguagem *Prolog*, a ferramenta *OWLJessKB* utiliza as linguagens RDFS e *OWL-Lite*, a ferramenta *Pychinko* utiliza as linguagens N3, RDFS, OWL e *XML Schema Definition* (XSD) [263] e as ferramentas *RdfLib* e *Swish* utilizam a linguagem RDFS.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de funcionamento de cada uma destas ferramentas:

- *Euler*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato RDFS e OWL;
- *EulerMoz*: É uma ferramenta baseada na ferramenta Euler e que também disponibiliza um Motor de Inferência para a informação descrita no formato RDFS e OWL;
- F-OWL: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato OWL;
- *FaCT++*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato *OWL-Lite*;
- *Hoolet*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato OWL-DL [213];
- *Instance Store*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato OWL-DL. Num modelo mais abrangente de categorização,

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

esta ferramenta pode também ser incluída na categoria Armazenamento;

- *Metalog*: É uma ferramenta que apresenta um poderoso Motor de Inferência e uma interação simples como o utilizador através da utilização de uma linguagem de programação muito prática, denominada PNL;
- *OWLJessKB*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato OWL;
- *Pellet* [264]: É uma ferramenta que disponibiliza um Motor de inferência para a informação descrita no formato OWL-DL [213];
- *Pychinko*: É uma ferramenta que disponibiliza um Motor de Inferência para informação a descrita nos formatos RDFS, N3 e XSD;
- *RdfLib*: É uma ferramenta que disponibiliza um conjunto de serviços RDF e um Motor de Inferência para a informação descrita no formato RDFS;
- *Swish*: É uma ferramenta que disponibiliza um Motor de Inferência para a informação descrita no formato RDFS.

#### 4.3.8 Categoria - Interpretador Web

A categoria Interpretador Web inclui sete ferramentas. Na tabela 4.8 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Enquanto a ferramenta *FlinkCommands* [265] foi desenvolvida para o SO *Linux*, a ferramenta *OWL-S IDE* [266] foi desenvolvida para o SO *Linux* e para o SO *Windows*, a ferramenta *Visual Variable-Depth Info Display* [267] foi desenvolvida para o S.O. *Windows NT/2000* e as ferramentas *DumpOnt* [234], *HyperDAML* [268], *Object Viewer* [269] e *OWL-P* [270] são independentes do SO. Em termos de interacção com o utilizador, a ferramenta *Visual Variable-Depth Info Display* disponibiliza uma interacção Web, as ferramentas *Object Viewer*, *OWL-P*, *OWL-S*



### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

IDE e *FlinkCommands* disponibilizam uma interacção do tipo GUI e as ferramentas *DumpOnt* e *HyperDAML* disponibilizam uma interacção Web, uma API e a linha de comandos.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *DumpOnt*: É uma ferramenta que apresenta graficamente as classes e propriedades existentes em ontologias OWL;
- *FlinkCommands*: É uma ferramenta que permite formalizar e apresentar semanticamente comandos em *Linux* através de uma ontologia OWL;
- *HyperDAML*: É uma ferramenta que permite criar uma representação HTML a partir de conteúdo OWL;
- *Object Viewer*: É uma ferramenta que permite visualizar grafos RDF através da Web;
- OWL-P: É uma ferramenta que define uma ontologia OWL que pode ser utilizada como um protocolo para ontologias OWL. Esta ferramenta, num modelo mais abrangente de categorização, pode também ser incluída na categoria Conversão;
- *Visual Variable-Depth Info Display*: É uma ferramenta que permite apresentar estrutu-

Tabela 4.8: Categoria - Interpretador Web

Interpretador Web						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
DumpOnt	Moore, D., Kolas, D. Lerner, J., Dean, M. Blace, R. e Self, T.	1.2 (Feb-04)	BSD	Java	OWL RDF	HTML
FlinkCommands	Barreau, G., Rocha, R. Machado, M., Gama, C. Abdalla, D., Gagnon, M. Oliveira, K., Anquetil, N.	-	GPL	Java	OWL	OWL
HyperDAML	Lerner, J. Dean, M. e Self, T.	Jan-04	BSD	Java	OWL RDF	HTML
Object Viewer	Jeremy Lerner Troy Self	Jul-04	BSD	Java	OWL RDF	-
OWL-P	Mallya, A. e Desai, N.	1 (Nov-04)	-	Java	OWL	OWL
OWL-S IDE	Naveen Srinivasan Softagents Group	0.1 (Nov-04)	-	Java	OWL XML	OWL
Visual Variable-Depth Info Display	Baoshi Yan	1.1 (Set-04)	Mozilla	Java JavaScript XMI	OWL RDF	HTML

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

ras RDF e OWL. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Interrogação e Visualização.

#### 4.3.9 Categoria - Interrogação

A categoria Interrogação é composta por sete ferramentas. Na tabela 4.9 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.9: Categoria - Interrogação

Interrogação						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
Cwm	W3C	(Jan-05)	W3C	Python	N3 OWL RDF/XML	N3 OWL RDF/XML
Lehigh University Benchmark	SWAT Lab	1.1 (Jun-04)	GPL	Java	OWL	OWL
OWL-QL	Robert Mccool	04-12-07 (Dec-04)	BSD	Java	-	-
OWLS-TC	Klusch, M. Fries, B. e Khalid, M.	1.0 (Apr-05)	GPL	Java	-	-
ROWL	Norman Sadeh	(Jan-05)	-	Java	OWL RDF	-
Semantic Discovery Service	Dan Mandell Sheila McIlraith	0.5 (Jun-04)	GPL	Java	OWL XML XML Schema	-
TAP	Feigenbaum, E., Fikes, R. Guha, R., McGuinness, D. McIlraith, S., McCool, R. Miller, E., Brickley, D. Sundarajan, A. e Joly, K.	0.75 (Apr-03)	BSD	C	RDF/XML	RDF/XML

Enquanto que a ferramenta *Semantic Discovery Service* [271] foi implementada para os SOs *MacOS* [272], *Windows NT/2000* e *Linux* as ferramentas *Closed World Machine(Cwm)* [273], *Lehigh University Benchmark* [274] e *OWL-S Service Retrieval Test Collection (OWLS-TC)* [275] são independentes do SO. Em termos de interação com o utilizador, a ferramenta *Semantic Discovery Service* disponibiliza um acesso através do protocolo SOAP e através da linha de comandos e as ferramentas *Cwm* e *Lehigh University Benchmark* disponibilizam um acesso através de uma API e da linha de comandos. Como paradigmas utilizados pelos Interpretadores RDF na manipulação de modelos RDF, a ferramenta *Cwm* utiliza o paradigma Centrado no Modelo RDF e a ferramenta *TAP* [276] utiliza os paradigmas Centrado nas Asserções RDF e Centrado nos Recursos RDF. Como linguagem de interrogação, a ferramenta *Cwm* utiliza a

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

linguagem N3 e a ferramenta TAP utiliza a linguagem/protocolo GetData [277].

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *Cwm*: É uma ferramenta que permite interrogar, verificar e filtrar informação descrita no formato RDF/XML e N3. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída na categoria Inferência;
- *Lehigh University Benchmark*: É uma ferramenta que permite facilitar a avaliação de modelos de armazenamento e de interrogação de informação para a Web Semântica;
- *OWL Query Language (OWL-QL)* [278]: É uma ferramenta/linguagem que utiliza uma linguagem e um protocolo formal de interrogação, especificados em OWL, que pode ser utilizado na comunicação entre agentes;
- *OWLS-TC*: É uma ferramenta que permite avaliar a performance de algoritmos de serviços de OWL-S;
- *Rule Extension of OWL Mobile Commerce Lab (ROWL)* [279]: É uma ferramenta/linguagem que apresenta uma extensão às regras utilizadas na inferência de informação em OWL. Num modelo mais abrangente de categorização, esta ferramenta pode também ser incluída nas categorias Inferência e Conversão;
- *Semantic Discovery Service*: É uma ferramenta que permite procurar e interrogar serviços Web;
- *TAP*: Fornece um conjunto de mecanismos simples que podem ser utilizados para publicar informação na Web ou fornecê-la a SPC através de uma linguagem de interrogação simples.

## 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

### 4.3.10 Categoria - Validação

A categoria Validação inclui três ferramentas. Na tabela 4.10 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

Tabela 4.10: Categoria - Validação

Validação						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
ConsVISor	Versatile Information Systems	-	-	Java	OWL RDF	HTML OWL
OWL Validator	Rager, D., Lerner, J. Self, T.	(Jul-04)	BSD	Java	OWL	Validation Report XML
SWRL Vaidator	Troy Self David Kolas	Nov-04	BSD	Java	SWRL	Validation Report

Todas as ferramentas apresentadas nesta categoria são independentes do SO.

Em termos de interacção com o utilizador, as ferramentas *OWL Validator* [280] e *SWRL Validator* [281] disponibilizam um acesso Web, uma API e a linha de comandos.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *ConsVISor* [282]: É uma ferramenta que permite verificar a consistência de documentos OWL;
- *OWL Validator*: É uma ferramenta que permite verificar a consistência de documentos OWL;
- *SWRL Validator*: É uma ferramenta que permite verificar a consistência de anotações em *Semantic Web Rule Language* (SWRL) [283].

### 4.3.11 Categoria - Visualização

A categoria Visualização inclui três ferramentas. Na tabela 4.11 são apresentados os nomes e as principais características associadas a cada uma dessas ferramentas.

### 4.3. FERRAMENTAS ASSOCIADAS A CADA CATEGORIA

---

A ferramenta *VisioOWL* [284] foi implementada para o SO *Windows*.

Como referência rápida, é apresentada a seguir uma breve descrição acerca do modo de actuação de cada uma destas ferramentas:

- *IsaViz* [132]: É uma ferramenta que permite a edição e a manipulação de modelos RDF através de uma forma gráfica. As suas principais características são:
  - Possui um interface gráfico que possibilita uma fácil navegação visual através de modelos RDF;
  - Permite a criação e edição de triplas RDF através da utilização de elipses, rectângulos e arcos;
  - Reconhece os formatos: RDF/XML, Notation 3 e N-Triple;
  - Permite exportar ficheiros nos formatos: RDF/XML, Notation 3, N-Triple, SVG e PNG;
  - Permite criar grafos RDF através da utilização de Folhas de Estilo Gráfico, denominadas *Graph Stylesheets* (GSS) [285].
- *SVG-OWL Viewer* [286]: É uma ferramenta que permite apresentar ontologias de forma gráfica, através da utilização do formato SVG. A principal ideia é a de tentar perceber a estrutura e os relacionamentos inerentes das ontologias, através da sua análise gráfica;
- *VisioOWL*: É uma ferramenta da *Microsoft* que permite criar representações gráficas de ontologias OWL.

Tabela 4.11: Categoria - Visualização

Visualização						
	Autor	Versão	Licença	Linguagem	INPUT	OUTPUT
IsaViz	Emmanuel Pietriga	2.1 (Out-04)	-	Java	RDF/XML N3 N-Triple	RDF/XML N3 N-Triple SVG PNG
SVG-OWL Viewer	Aditya Kalyanpur	3.0	BSD	Java	OWL	SVG
VisioOWL	John Flynn	Jun-04	BSD	-	OWL RDF	Graph

## 4.4 Resumo

O aparecimento de grande quantidade de ferramentas para a Web Semântica nos últimos cinco anos, apenas reflecte a importância que ela tem vindo a consolidar na comunidade científica mundial. Neste capítulo foi apresentado um estudo de catalogação de oitenta ferramentas para a Web Semântica. Além da catalogação, foram também apresentadas as principais características de cada ferramenta.

Não restam dúvidas de que estamos a viver um momento que se apresenta como crucial, quer para a avaliação, quer para a implementação de toda a arquitectura da Web Semântica. Sendo assim, a catalogação das mais importantes ferramentas para a Web Semântica contribuirá com mais uma base de apoio para o sucesso da materialização da própria Web Semântica.

A análise dos resultados deste estudo demonstra que a necessidade de interacção de informação entre diferentes ferramentas está a crescer enormemente, e que na Web Semântica essa necessidade será colossal. Desta forma, esta necessidade de interacção de informação entre diferentes ferramentas deverá ser considerada como a base para qualquer nova ferramenta que venha a ser desenvolvida para a Web Semântica.

Outra conclusão, que se pode retirar a partir dos resultados obtidos do estudo desenvolvido, resume-se na constatação da inexistência de ferramentas que utilizem de forma integrada as diversas tecnologias associadas à Web Semântica e, simultaneamente, permitam a criação de páginas web de acordo com os princípios da Web Semântica por parte de utilizadores não familiarizados com a Web Semântica. Se a Web Semântica não conquistar o interesse dos actuais utilizadores da Web, possivelmente poderá vir a ser desvalorizada e esquecida como muitas outras tecnologias que, apesar de serem poderosas e úteis não se conseguiram impor com sucesso e acabaram apenas como referências associadas à história.

# Capítulo 5

## SWedt: Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF

"Genius is one percent inspiration and  
ninety-nine percent perspiration."

"I have not failed.  
I've just found 10,000 ways that don't work."

Thomas Alva Edison

### 5.1 Introdução

Com vimos através da análise do estudo de categorização das actuais ferramentas para a Web Semântica apresentado no capítulo 4 - Ferramentas para a Web Semântica -, durante os últimos anos o aparecimento de novas ferramentas para a Web Semântica está a apresentar um crescimento cada vez mais elevado. No entanto, a maioria dessas ferramentas possui um campo de

## 5.1. INTRODUÇÃO

---

actuação geralmente limitado a apenas uma das camadas da estrutura da Web Semântica, como seja por exemplo editores para RDF ou para OWL. De facto, o desenvolvimento de ferramentas que, por um lado apresentem como campo de actuação a integração simultânea de várias camadas da arquitectura da Web Semântica e, por outro lado, que sejam fáceis de manusear e testar por qualquer utilizador da Web independentemente dos seus conhecimentos acerca da arquitectura da Web Semântica, ainda não é uma realidade. Desta forma, um dos objectivos da actual dissertação consistiu na definição de uma arquitectura e no desenvolvimento de um editor que colmatasse as limitações apresentadas pela maioria das ferramentas para a Web Semântica no estudo anteriormente referido. A ferramenta desenvolvida denomina-se "Editor para a Web Semântica Integrando Anotações Semânticas, Ontologias e RDF", ou simplesmente "Editor para a Web Semântica"(SWedt), e a sua arquitectura e implementação serão apresentadas ao longo deste capítulo. O conteúdo deste capítulo é parcialmente baseado nos artigos [13, 14, 15].

O presente capítulo é constituído por seis secções. Nesta primeira secção é apresentada uma breve introdução à temática que vai ser abordada neste capítulo, além disso apresenta, também, uma pequena descrição da estrutura adoptada neste capítulo. Na segunda secção, **A ferramenta SWedt**, é apresentada uma descrição da arquitectura da ferramenta SWedt, a qual foi desenvolvida durante a realização desta dissertação. Na terceira secção, **Implementação da ferramenta SWedt**, é apresentada a implementação e as principais API associadas à ferramenta SWedt. Na quarta secção, **Exemplo de Utilização**, é realizada uma demonstração prática da utilização da ferramenta SWedt. Na quinta secção, **A Ferramenta SMORE**, é apresentada a ferramenta para a Web Semântica que mais se assemelha com a ferramenta desenvolvida nesta dissertação. Por último, é apresentado na sexta secção um pequeno resumo acerca do tema abordado neste capítulo.



### 5.2 A Ferramenta SWedt

A ideia de criar a ferramenta SWedt numa altura em que a arquitectura, o modo de implementação e de funcionamento da Web Semântica ainda se encontra em franco desenvolvimento, prendeu-se com a premente necessidade de criar uma ponte entre os investigadores da Web Semântica e os actuais programadores/utilizadores da Web<sup>124</sup>. Desta forma, um dos principais objectivos do desenvolvimento da ferramenta SWedt é o de despertar o interesse dos utilizadores da Web pela Web Semântica.

A ferramenta desenvolvida, durante esta dissertação, abrange e integra um conjunto extremamente diversificado de tecnologias, estando, por sua vez, muitas delas ainda em desenvolvimento. Da mesma forma, no desenvolvimento da ferramenta SWedt foram utilizadas várias ferramentas/tecnologias que permitiram de certa forma acelerar o processo de desenvolvimento. Duas das principais vantagens em utilizar ferramentas/tecnologias previamente disponibilizadas com licenças abertas de utilização prende-se com a desnecessidade de, passo a expressão, "reinventar constantemente a roda" e com a robustez em termos de funcionamento que geralmente apresentam<sup>125</sup>. De entre todas as ferramentas/tecnologias utilizadas, a Plataforma *Eclipse*<sup>126</sup> [230] e a API *Jena* [134] foram as que mais contribuíram para o desenvolvimento da ferramenta SWedt.

A figura 5.1 apresenta as principais tecnologias associadas à ferramenta desenvolvida nesta dissertação. Ela é composta por quatro blocos coloridos. O bloco localizado mais à esquerda e de cor de laranja representa a plataforma *Eclipse*. O bloco central de cor roxa representa a ferramenta SWedt. O bloco localizado mais à direita e de cor azul representa a API *Jena*.

---

<sup>124</sup>Existem, na história, centenas de exemplo de tecnologias promissoras que foram ultrapassadas por tecnologias inferiores apenas porque não conseguiram conquistar os utilizadores, como por exemplo as cassetes de vídeo Beta (ou Betamax, da Sony [287]), os processadores de texto AmiPro (Lotus Development [288]) e Word Perfect (Satellite Software International [289]), entre outros.

<sup>125</sup>Geralmente estas ferramentas/tecnologias possuem um grupo alargado de pessoas que contribuem activamente para a sua estabilidade e robustez.

<sup>126</sup>Plataforma suportada por um consórcio de grandes empresas. Começou por ser um projecto proprietário que foi iniciado em Abril de 1999 pela OTI e pela IBM, mas em Novembro de 2001 foi disponibilizado como um projecto em código aberto. Actualmente encontra-se na sua terceira edição.

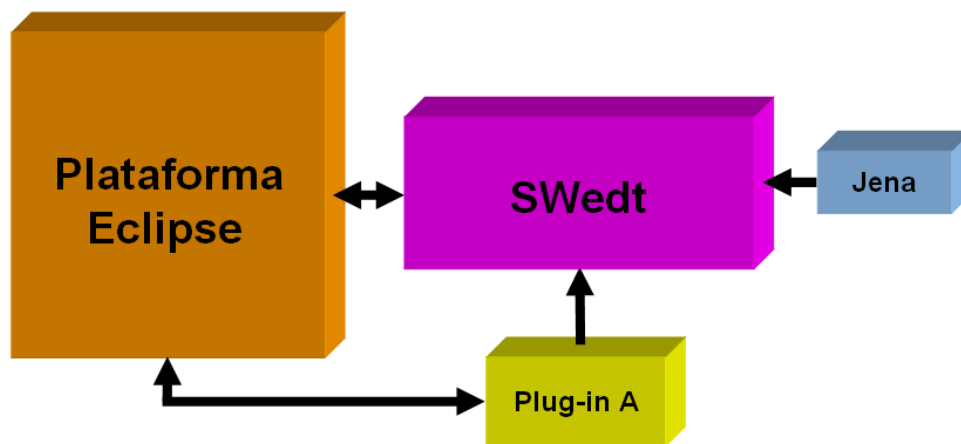


Figura 5.1: Principais tecnologias associadas à ferramenta SWedt.

Todas estas tecnologias serão apresentadas em maior pormenor nas próximas secções. O bloco de cor amarela que se encontra localizado na parte inferior da imagem representa uma ou mais tecnologias/ferramentas que poderão vir a ser facilmente associados à plataforma *Eclipse* ou à ferramenta SWedt. A figura é apresentada para realçar o facto de a ferramenta SWedt ter sido desenvolvida de forma a ser integrada na plataforma *Eclipse* e que se apoia principalmente na API *Jena*.

### 5.2.1 Plataforma *Eclipse*

Cada vez mais, a necessidade de relacionar informação resultante de diferentes ferramentas é crucial. Na Web Semântica essa necessidade será colossal. A maioria das actuais ferramentas, além de apresentarem diferentes formas de interações de utilização, necessitam constantemente de trocar (importar e exportar) informação entre elas. Estes factores contribuem para uma ineficiente integração da informação resultante dessas ferramentas, os quais devem ser levados em consideração por partes de todos os programadores que queiram fidelizar os seus utilizadores.

Uma outra perspectiva, diferente da apresentada no anterior parágrafo, passa pela utilização de uma plataforma aberta que utilize uma interacção universal com os seus utilizadores e que possibilite uma integração fácil, automática e simbiótica de diferentes ferramentas. Uma pla-

## 5.2. A FERRAMENTA SWEDT

---

taforma que permita aos utilizadores escolher a ferramenta ou conjunto de ferramentas que melhor se adapte aos seus interesses e, por outro lado, possibilite o desenvolvimento de novas ferramentas ou a sua extensão por parte dos programadores. Uma plataforma aberta, de integração e expansão de ferramentas, seria extremamente valiosa, quer para os utilizadores, quer para os programadores.

Um exemplo deste tipo de plataforma é a plataforma *Eclipse*, a qual de acordo com o *Eclipse White Paper* [230] foi implementada em código aberto, é compatível com a maioria dos SO<sup>127</sup>, pode ser facilmente estendida de forma a permitir a utilização de qualquer tipo de linguagem<sup>128</sup> e possui como principal objectivo a integração de diferentes ferramentas através da utilização de um IDE agradável e universal. Nesta plataforma as ferramentas são desenvolvidas em forma de *Plug-ins*. Os *Plug-ins* são pacotes de código estruturado que podem ser acoplados a uma aplicação/plataforma<sup>129</sup> em particular, contribuindo, desta forma, com a adição de novas funcionalidades. Além disso, os *Plug-ins* podem definir Pontos de Extensão, os quais são localizações bem definidas no código de um *Plug-in* que podem ser estendidas por outros *Plug-ins*. A plataforma *Eclipse* mantém um registo de todos os *Plug-ins* instalados e das suas funcionalidades. A procura, carregamento e execução dinâmica de *Plug-ins* é uma das tarefas da plataforma *Eclipse*. Mas, o mais extraordinário e fascinante é o facto de a funcionalidade associada a cada *Plug-in* estar intimamente relacionada com a imaginação do programador.

A arquitectura de *Plug-in* do *Eclipse* é definida por:

- *Plug-ins*: Conjunto de funcionalidades que podem ser integradas no *Eclipse*;
- Ponto da Extensão: Entidades que permitem a interligação entre *Plug-ins*;
- Extensão: Um *Plug-in* que se interliga a um *Plug-in* existente.

---

<sup>127</sup> *Microsoft Windows, Linux, etc.*

<sup>128</sup> Exemplo de linguagens que podem ser utilizadas: HTML, *Java*, C, JSP, *Enterprise JavaBeans* (EJB), XML, etc.

<sup>129</sup> Neste caso a própria plataforma *Eclipse*.

## 5.2. A FERRAMENTA SWEDT

---

Por sua vez, cada *Plug-in*:

- Expande um ou mais Pontos de Extensão;
- Opcionalmente declara novos Pontos de Extensão;
- Depende de um conjunto de outros *Plug-ins*;
- Contém bibliotecas *Java* e outros ficheiros;
- Pode exportar APIs baseadas em *Java*.

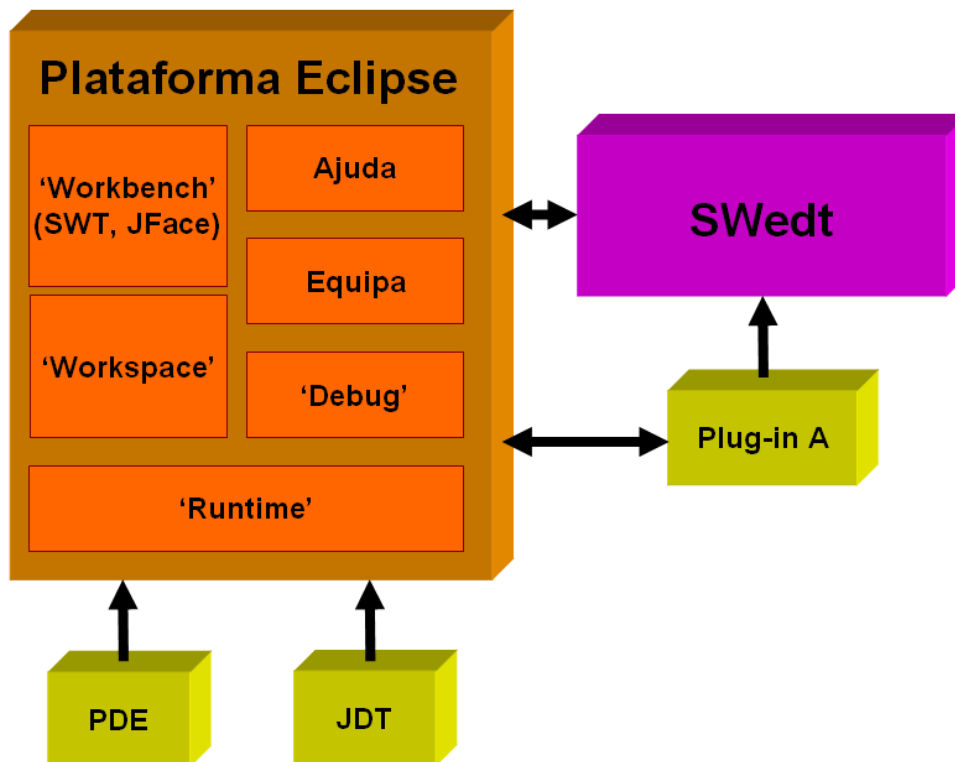


Figura 5.2: A plataforma *Eclipse*.

A figura 5.2 apresenta os principais componentes integrados e/ou associados da plataforma *Eclipse*. Ela é composta por cinco blocos coloridos. O bloco localizado à esquerda e de cor de laranja representa a plataforma *Eclipse* e é por sua vez composto por seis blocos internos que serão alvo de aprofundamento no decorrer desta secção. O bloco localizado à direita de cor roxa

## 5.2. A FERRAMENTA SWEDT

---

representa a ferramenta SWedt. Os dois blocos de cor amarela que se encontram localizados na parte inferior do bloco da plataforma *Eclipse* representam dois *Plug-ins* específicos que se encontram associados à plataforma *Eclipse* (PDE<sup>130</sup> e JDT<sup>131</sup>) e que, também, serão alvo de aprofundamento no decorrer desta secção. Por fim existe um bloco, também, de cor amarela que se encontra localizado na parte inferior do bloco da ferramenta SWedt que representa uma ou mais tecnologias/ferramentas que poderão vir a ser facilmente associados à plataforma *Eclipse* ou à ferramenta SWedt.

A plataforma *Eclipse* é constituída por um conjunto base de componentes, apresentados como blocos internos na figura 5.2, que pode ser associado e facilmente estendido com a associação e integração de novos componentes/*Plug-ins*. A seguir é apresentada uma breve descrição dos principais componentes integrados na plataforma *Eclipse*:

- Componente *Workbench*: Este componente disponibiliza uma estrutura base para a interacção entre a plataforma *Eclipse* e os utilizadores e, também, uma *User Interface* (UI) extensível. O paradigma de IDE do *Eclipse* é baseado em Editores, Vistas e Perspectivas. Do ponto de vista do utilizador, uma janela do *Eclipse* consiste visualmente em Editores e Vistas. As Perspectivas manifestam-se através da apresentação e disposição integrada de arranjos de Editores e Vistas na janela do *Eclipse*:
  - Os Editores permitem que o utilizador abra, edite<sup>132</sup> e guarde objectos. Seguem o ciclo de vida "abrir-guardar-sair". A plataforma *Eclipse* fornece um Editor padrão para recursos do texto, mas outros, e mais específicos, Editores podem ser fornecidos através da utilização de *Plug-ins*;
  - As Vistas fornecem informação sobre objectos presentes num Editor ou em outras Vistas. Uma Vista pode cooperar com um Editor fornecendo mais informação sobre os dados que estão sendo editados. As Vistas possuem um ciclo de vida mais simples

---

<sup>130</sup> *Plug-In Development Environment*.

<sup>131</sup> *Java Development Tooling*.

<sup>132</sup> Inserção e manipulação de objectos, normalmente texto simples.

## 5.2. A FERRAMENTA SWEDT

---

do que Editores: as modificações feitas em uma Vista, tal como a mudança de um valor de propriedade, são conservadas imediatamente e reflectidas em outras partes relacionadas do IDE. A plataforma *Eclipse* fornece diversas Vistas padrão, mas Vistas específicas podem ser fornecidas através da utilização de *Plug-ins*;

- Uma sessão de *Eclipse* pode possuir diversas Perspectivas, mas apenas uma é visível de cada vez. Cada Perspectiva é constituída por um conjunto de Vistas e Editores que podem ser facilmente reajustados na janela da plataforma *Eclipse* pelo utilizador. A plataforma *Eclipse* fornece algumas Perspectivas padrão, mas Perspectivas mais específicas podem ser fornecidas através da utilização de *Plug-ins*.

O paradigma Editores-Vistas-Perspectivas utilizado pelo componente *Workbench* permite integrar facilmente novas ferramentas à plataforma *Eclipse* através da utilização de pontos de extensões na forma de *Plug-in*. Por fim, o componente *Workbench* integra e depende de outros dois componentes:

- Componente *Standard Widget Toolkit* (SWT): É uma biblioteca gráfica associada ao sistema gráfico de janelas mas independente do sistema operativo. Permite e facilita a implementação de gráficos genéricos de baixo nível (botões, listas, texto, menus, árvores, etc.). Toda a plataforma UI do *Eclipse* e os *plug-ins* desenvolvidos para a plataforma *Eclipse* utilizam a biblioteca SWT sempre que necessitam de apresentar informação para o utilizador;
- Componente *JFace*: É uma ferramenta de desenvolvimento estruturado de GUIs, implementada através da biblioteca SWT, com o objectivo de simplificar tarefas comuns relacionadas com o UI. É independente do sistema de janelas e foi projectada para trabalhar em conjunto com o SWT. Inclui as características normais de uma ferramenta de desenvolvimento de UIs dando uma especial relevância a duas delas como sendo as mais interessantes:

\* Acção<sup>133</sup>: O mecanismo de Acção permite que os comandos do utilizador sejam

---

<sup>133</sup>Em língua Inglesa: *Action*.

definidos independentemente da sua localização exacta na UI. Uma Acção representa um comando que pode ser provocado pelo utilizador através de uma tecla, da selecção de um elemento de um menu, ou da selecção de um elemento de uma barra da ferramenta. Cada Acção conhece as suas propriedades de UI (etiqueta, ícone, etc.), as quais são utilizadas para construir componentes gráficos apropriados para apresentar a Acção. Esta separação permite que a mesma Acção possa ser usada em diversos lugares no UI sem haver a necessidade de alterar o seu código;

\* Visor<sup>134</sup>: Os Visores apresentam um comportamento comum e fornecem uma semântica de alto-nível para a manipulação dos componentes gráficos do SWT. Os Visores de listas, árvores, e tabelas suportam o seu preenchimento com elementos do domínio do cliente e mantêm os componentes gráficos sincronizados com posteriores alterações efectuadas nesse domínio. Estes Visores são configurados através de um fornecedor de conteúdo e de um fornecedor de etiqueta. O fornecedor de conteúdo sabe como mapear o elemento de entrada do Visor ao índice interno do Visor. O fornecedor de etiqueta sabe como produzir a etiqueta e o ícone específico de cada elemento presente no índice interno do Visor.

- Componente *Workspace*: O componente *Workspace* permite mapear facilmente a estrutura de directórios e os ficheiros existentes num computador. A plataforma *Eclipse*, através deste componente, permite manipular directórios e ficheiros em forma de um projecto;
- Componente *Runtime*: O núcleo duro<sup>135</sup> da plataforma *Eclipse* é composto por um conjunto de funcionalidades muito pequeno. Todas as restantes funcionalidades existentes e possíveis de desenvolver são disponibilizadas através de *Plug-ins*, incluindo funcionalidades da própria plataforma. Sempre que a plataforma *Eclipse* é iniciada, são detectados de forma automática todos os *Plug-ins* instalados e é realizada a associação de todas as Extensões com os seus respectivos Pontos de Extensão. Mas, apesar de a plataforma

---

<sup>134</sup>Em língua Inglesa: *Viewer*.

<sup>135</sup>Também designado por coração.

## 5.2. A FERRAMENTA SWEDT

---

*Eclipse* poder possuir uma grande quantidade de *Plug-in* instalados, eles apenas serão activados<sup>136</sup> à medida que forem sendo solicitados para utilização;

- Componente *Debug*: O componente *Debug* apresenta um modelo de GUI genérico de depuração de código associado à plataforma *Eclipse*;
- Componente Equipa: O componente Equipa apresenta um modelo de partilha de recursos por uma equipa de desenvolvimento e gestão de versões *Concurrent Versions System* (CVS) [290];
- Componente Ajuda: O componente Ajuda apresenta um modelo para a criação de conteúdos de ajuda baseado na linguagem HTML;
- *Plug-in Java Development Tooling* (JDT): O componente JDT disponibiliza uma API para o desenvolvimento de programas em linguagem de programação *Java* através da utilização da plataforma *Eclipse*;
- *Plug-in Plug-in Development Environment* (PDE): O componente PDE disponibiliza um ambiente de apoio ao desenvolvimento de *Plug-ins* em linguagem de programação *Java* para a plataforma *Eclipse*.

### 5.2.2 Arquitectura da ferramenta SWedt

A ferramenta SWedt foi desenvolvida na forma de um *Plug-in* a ser utilizado na plataforma *Eclipse*. A escolha de desenvolvimento em forma de *Plug-in* prendeu-se com a facilidade de integrar e associar a ferramenta desenvolvida com outras ferramentas também desenvolvidas para o *Eclipse*. Desta forma os utilizadores poderão facilmente utilizá-la em associação com inúmeras ferramentas/*Plug-ins* diferentes. Isto é, apesar de a ferramenta SWedt manipular apenas ficheiros com a extensão XML, outros tipos de ficheiros<sup>137</sup> podem ser, também, facilmente manipulados durante a sua utilização. Para que isso possa acontecer, basta que as extensões desse ficheiros possuam ferramentas/*Plug-ins* associadas e instaladas na plataforma *Eclipse*.

---

<sup>136</sup>Isto é, carregados para a memória RAM.

<sup>137</sup>Como por exemplo, ficheiros *Java*, HTML, XHTML, RDF, OWL, etc.



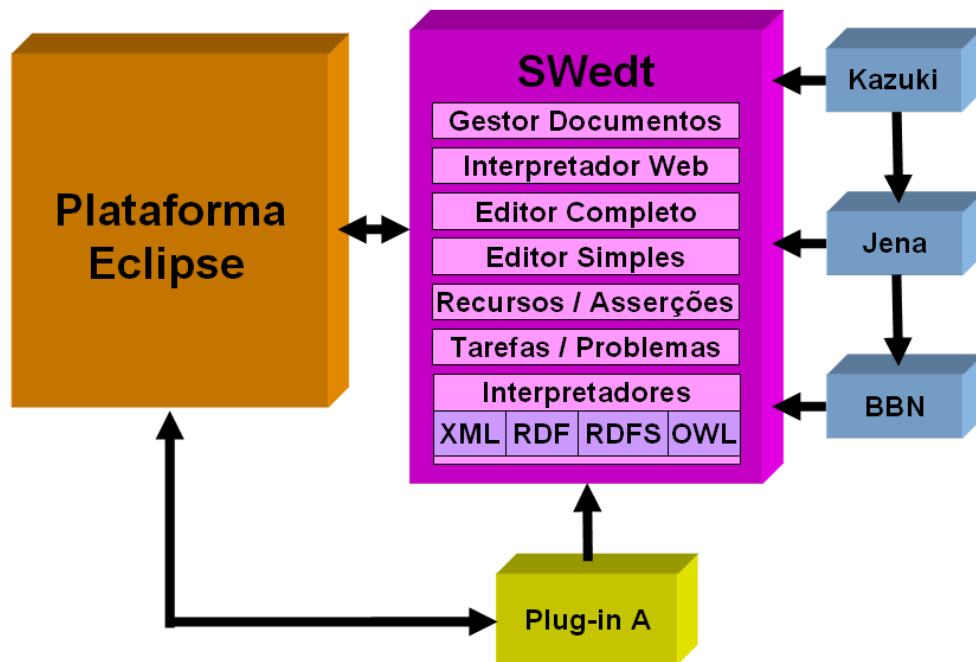


Figura 5.3: Principais componentes da ferramenta SWedt.

Como se pode depreender através da observação da figura 5.3, a qual é composta por cinco blocos coloridos, a ferramenta SWedt utiliza a plataforma *Eclipse* em conjunto com diversas APIs/*Plug-ins*, disponibilizadas pela comunidade científica mundial. O bloco localizado mais à esquerda e de cor de laranja representa à plataforma *Eclipse*, a ferramenta SWedt está representada através do bloco de cor roxa e encontra-se localizado no centro da imagem à direita do bloco cor de laranja e é por sua vez composto por sete blocos internos que serão alvo de aprofundamento no decorrer desta secção. As APIs de nome *Kazuki*, *Jena* e BBN encontram-se representadas na parte direita da imagem através dos blocos de cor azuis e, também, serão alvo de aprofundamento na próxima secção deste capítulo. O bloco de cor amarela que se encontra localizado na parte inferior da imagem representa uma ou mais tecnologias/ferramentas que poderão vir a ser facilmente associados à plataforma *Eclipse* ou à ferramenta SWedt.

A ferramenta SWedt é constituída por um conjunto de componentes, apresentados como blocos internos na figura 5.3, que foram desenvolvidos durante a elaboração desta dissertação e por um conjunto de APIs disponibilizadas pela comunidade científica. De seguida é apresentada

## 5.2. A FERRAMENTA SWEDT

---

uma breve descrição dos principais componentes integrados na ferramenta SWedt:

- Componente Gestor Documento: Representa uma Vista na plataforma *Eclipse* que possibilita a navegação entre os diversos directórios existentes no disco rígido de um computador, criar novos ficheiros, modificar ficheiros e editar ficheiros através da utilização de um editor associado;
- Componente Interpretador Web: Permite a visualização através de um interpretador Web do texto existente no Editor Completo;
- Componente Editor Completo<sup>138</sup>: Representa um Editor que disponibiliza todo o conteúdo de um ficheiro através da utilização das sintaxes RDF/XML, RDFS e OWL;
- Componente Editor Simples: Representa o principal Editor da aplicação. A sua função é a de permitir a inserção de texto, sem qualquer tipo de restrição, por parte dos utilizadores;
- Componente Recursos: Representa uma Vista que permite observar, inserir, alterar e validar associações entre URIs e palavras existentes no Editor Simples;
- Componente Asserções: Representa uma Vista que permite associar entre si os recursos que se encontram definidos na vista Recursos. Por cada associação efectuada nesta vista a ferramenta adiciona, de forma totalmente automática, uma asserção RDF ao conteúdo do ficheiro original;
- Componente Tarefas: É uma Vista que apresenta uma lista de tarefas que pode ser criada e gerida pelos utilizadores;
- Componente Problemas: É uma Vista que apresenta uma lista de problemas detectados pela ferramenta durante a sua utilização;
- Componente Interpretadores: É um componente que disponibiliza interpretadores para as seguintes sintaxes: XML, RDF, RDFS e OWL.

---

<sup>138</sup> Apenas disponibilizado para utilização por parte de utilizadores avançados.

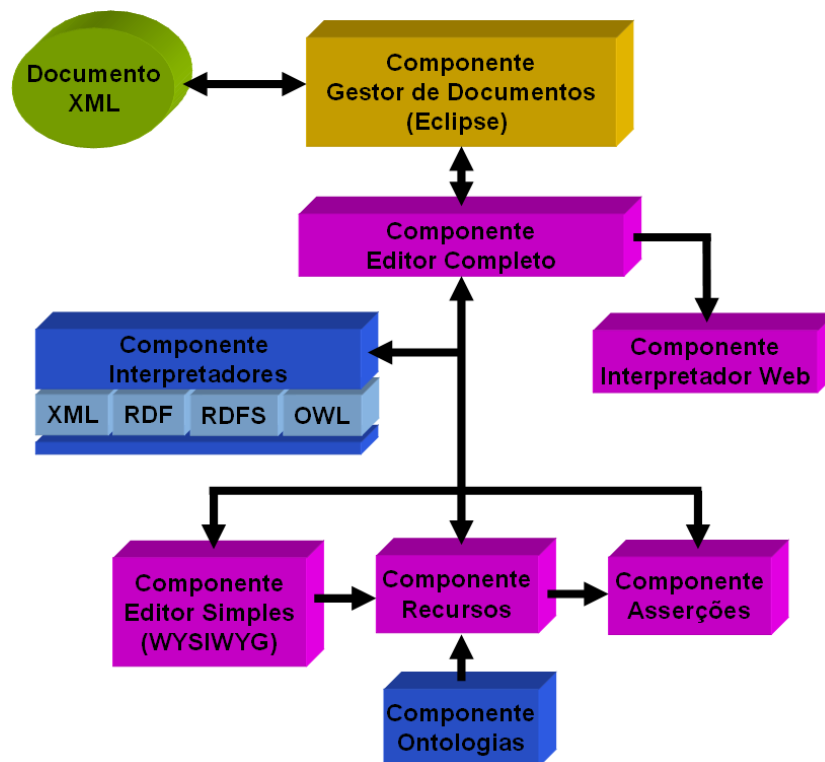


Figura 5.4: Arquitetura da Ferramenta SWedt.

Na imagem da figura 5.4 são apresentadas, através de setas uni e bidireccionais, as interações existentes entre os principais componentes da arquitetura da ferramenta SWedt, representados por oito blocos coloridos. O bloco localizado na parte superior esquerda da imagem em forma oval e de cor verde representa informação em formato XML (Documento XML), o bloco localizados na parte superior da imagem e de cor de laranja representa um componente da plataforma *Eclipse* (mais precisamente Gestor de documentos), os restantes sete blocos de cores roxa e azul existentes na imagem representam componentes que foram desenvolvidos de raiz para a ferramenta SWedt. Os blocos de cor roxa representam componentes da ferramenta em que o utilizador tem uma percepção prática da sua existência, isto é pode interagir com eles, os blocos de cor azul representam componentes da ferramenta que o utilizador não tem percepção prática da sua existência. As interações existentes entre estes componentes vão ser apresentadas e descritas seguindo o fluxo normal de funcionamento da ferramenta, o qual está representado como iniciando na parte superior da imagem, a partir do Documento XML.

## 5.2. A FERRAMENTA SWEDT

---

Os ficheiros em formato XML são a matéria-prima para a ferramenta SWedt. A sintaxe XML é reconhecida pela ferramenta, bem como as sintaxes RDF, RDFS e OWL sempre que estas se encontrem inseridas em ficheiros em formato XML. Através da figura 5.4 podemos verificar que existe um componente, de nome Gestor de Documentos e representado a cor de laranja, responsável por localizar e permitir o acesso da ferramenta a ficheiros em formatos XML. O acesso aos ficheiros em formato XML disponibilizado pelo Gestor de Documentos é bidireccional, isto é, permite que a ferramenta aceda e possa, também, alterar o conteúdo de ficheiro em formato XML.

A partir do momento que o Gestor de Documentos da ferramenta acede a um documento em formato XML, todo o conteúdo desse documento é enviado para o componente Editor Completo, o qual se encontra localizado na parte superior central da imagem através de um bloco de cor roxa. Este editor apresenta sempre a totalidade do conteúdo do documento seleccionado no Gestor de Documentos. O componente Editor Completo permite a manipulação de qualquer parte do documento por parte do utilizador. Todas as alterações efectuadas no conteúdo do documento serão automaticamente reportadas ao Gestor de Documentos.

O Componente Editor Completo representa internamente o documento em formato XML. Desta forma, qualquer alteração implementada por parte da maioria dos outros componentes internos da ferramenta terá uma reflexão no Componente Editor Completo. Assim, na figura 5.4, este componente aparece conectado bidireccionalmente com cinco outros componentes: Gestor de Documentos, Interpretadores, Editor Simples, Recursos e Asserções. Este componente, também, aparece conectado com o componente Interpretador Web, mas apenas de forma unidireccional.

O componente Interpretador Web, representado na parte direita da imagem através de um bloco de cor roxa, permite interpretar e simular a visualização em formato Web do conteúdo presente no componente Editor Completo através da GUI da ferramenta SWedt. Este com-

## 5.2. A FERRAMENTA SWEDT

---

ponente não permite manipular o conteúdo presente no componente Editor Completo e, desta forma, a conexão é representada na imagem de forma unidireccional.

O componente Interpretadores, localizado na parte esquerda da imagem e representado através de um bloco de cor azul, permite interpretar quatro sintaxes diferentes que possam existir em trechos de texto requisitados para interpretação por parte de qualquer um dos seguintes componentes internos da ferramenta: Editor Completo, Editor Simples, Recursos e Asserções. Este componente encontra-se dividido em quatro sub-componentes cada um deles especializado numa sintaxe em particular: XML, RDF, RDFS e OWL.

O componente Editor Simples, localizado na parte inferior esquerda da imagem e representado a cor roxa, encontra-se intimamente ligado ao componente Editor Completo. Enquanto o Editor Completo apresenta todo o conteúdo de documento em formato XML, o Editor Simples em conjugação com o componente Interpretadores "esconde" o código referente às quatro sintaxes reconhecidas pelo ferramenta (XML, RDF, RDFS e OWL) e apenas apresenta o restante conteúdo textual do documento. A informação apresentada pelo Editor Simples pode ser facilmente editada por parte do utilizador e segue a filosofia *What You See Is What You Get* (WYSIWYG) [291]. Sempre que a informação presente num dos editores, Editor Completo ou Editor Simples, da ferramenta é alterada a informação presente no outro editor será automaticamente actualizada pela ferramenta.

Todas as palavras presentes no Editor Simples são recolhidas, organizadas e apresentadas ao utilizador, através de uma tabela, numa Vista da ferramenta. O componente Recursos, localizado na parte inferior central da imagem e representado a cor roxa, é o responsável pela criação dessa tabela. Essa tabela além de apresentar todas as palavras existentes no Editor Simples também procede automaticamente à associação de cada uma delas a conceitos definidos em ontologias reconhecidas pela ferramenta. A lista de ontologias reconhecidas pela ferramenta e a sua exploração é suportada pelo componente Ontologias que se encontra localizado na parte

## 5.2. A FERRAMENTA SWEDT

---

inferior da imagem e está representado a cor azul. O componente Recursos está, também, em constante sincronização com o componente Editor Completo e sempre que o utilizador confirma uma associação de uma palavra do Editor Simples a um conceito de uma ontologia o Editor Completo regista, automaticamente, essa associação no conteúdo do documento XML original.

O componente Asserções, localizado na parte inferior direita da imagem e representado a cor roxa, permite criar associações triplas entre os recursos que se encontrem definidos no documento em formato XML. A ferramenta apenas considera como recurso qualquer registo presente no documento que identifique uma associação entre uma palavra do conteúdo textual e um conceito presente numa ontologia. A lista de triplas de recursos criada pelo utilizador é registada, de forma automática pela ferramenta, no conteúdo do documento em formato XML em utilização.

### 5.2.3 APIs utilizadas

A figura 5.5 apresenta as principais APIs utilizadas pela ferramenta SWedt. O bloco de cor roxa, da figura, representa a ferramenta SWedt e os três blocos de cor azul representam três APIs disponibilizadas pela comunidade científica mundial:

- *Kazuki*: É uma API desenvolvida em Java com o apoio da API *Jena*, que permite a manipulação de informação existente em ontologias do tipo OWL;
- BBN: É um conjunto de *Plug-ins* genéricos da *BBN Technologies*[292] para a plataforma *Eclipse*;
- *Jena*: É a mais importante API utilizada pela ferramenta SWedt. Foi desenvolvida em *Java* e fornece um ambiente de programação para a utilização de modelos RDF, Esquema RDF, OWL, linguagem de interrogação RDQL [210] e um motor de inferência baseado em regras. Pode ser utilizada para criar e manipular grafos RDF como o apresentado na figura 5.6, pois, possui classes de objectos que permitem representar grafos RDF, Recursos, Propriedades, Valores Alfanuméricos<sup>139</sup> e todos os principais conceitos associados à

---

<sup>139</sup>As Interfaces da API que representam Recursos, Propriedades e Valores alfanuméricos denominam-se *Re-*

## 5.2. A FERRAMENTA SWEDT

sintaxe/modelo RDF. O código necessário para criar o grafo, ou modelo, apresentado na figura 5.6 através da utilização da API *Jena*, é apresentado de seguida.

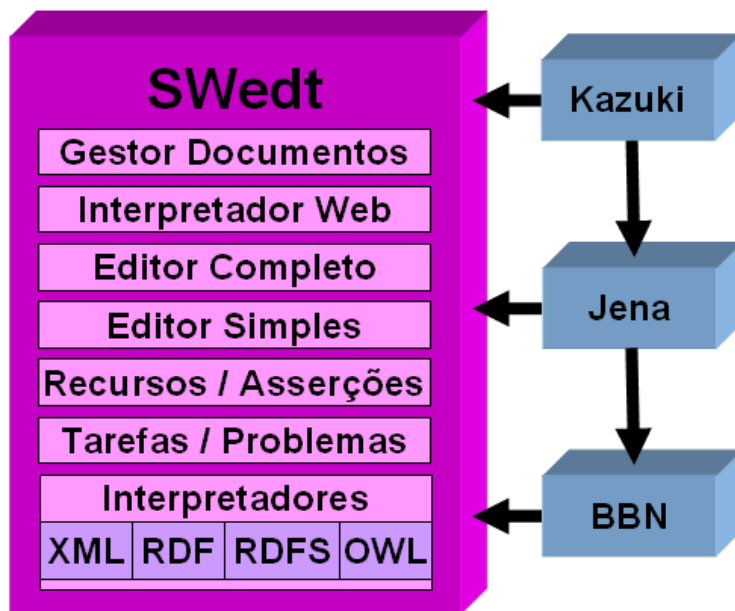


Figura 5.5: APIs utilizadas pela ferramenta SWedt.

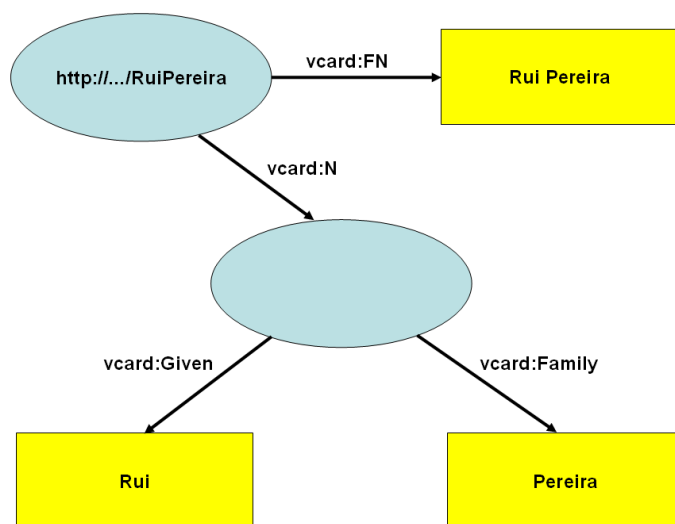


Figura 5.6: Representação RDF do vCard de Rui Pereira.

---

*source*, *Property* e *Literal*. Um grafo RDF é designado por um Modelo RDF e é representado através de um Interface denominada *Model*.

## 5.2. A FERRAMENTA SWEDT

---

```
// algumas definições
String personURI    = "http://somewhere/RuiPereira";
String givenName    = "Rui";
String familyName   = "Pereira";
String fullName     = givenName + " " + familyName;

// criação do modelo
Model model = ModelFactory.createDefaultModel();

// criação dos Recursos e atribuição das suas Propriedades
Resource RuiPereira = model.createResource(personURI)
    .addProperty(VCARD.FN, fullName)
    .addProperty(VCARD.N,
        model.createResource()
            .addProperty(VCARD.Given, givenName)
            .addProperty(VCARD.Family, familyName));
```

Nas primeiras cinco linhas do código fonte apresentado são definidas quatro constantes. Depois é criado um Modelo através da utilização do método *createDefaultModel()* da classe *ModelFactory* da API *Jena*. O modelo apresentado neste exemplo foi implementado na memória, mas a API *Jena* permite utilizar outros tipos de modelos, como por exemplo um que utilize uma BD Relacional. O Recurso Rui Pereira da figura é então criado e associado a duas propriedades. As propriedades são fornecidas por duas constantes da classe *VCARD* que possui objectos que representam todas as definições existentes no Esquema *VCARD* [293, 294]. A API *Jena*, além do Esquema *VCARD*, também fornece Classes associadas aos seguintes Esquemas: *RDF*, *RDFS*, *OWL*, *Dublin Core* e *DAML*. A propriedade *vcard:FN* possui um valor alfanumérico como Objecto, mas a propriedade *vcard:N* possui um Recurso como Objecto, mais precisamente um Recurso Vazio. Desta forma no código é criado um novo recurso como Objecto da propriedade *vcard:N* e são-lhe atribuídas duas propriedades *vcard:Given* e *vcard:Family*. Uma grande



## 5.2. A FERRAMENTA SWEDT

---

vantagem da utilização da API *Jena* prende-se com a possibilidade de, facilmente, converter a representação interna de um grafo RDF na sintaxe RDF/XML. Assim, o grafo da figura 5.6 é facilmente convertido para a sintaxe RDF/XML, como se pode observar no seguinte extracto de código:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#">
  <rdf:Description rdf:about="http://somewhere/RuiPereira">
    <vcard:FN>Rui Pereira</vcard:FN>
    <vcard:N rdf:parseType="Resource">
      <vcard:Given>Rui</vcard:Given>
      <vcard:Family>Pereira</vcard:Family>
    </vcard:N>
  </rdf:Description>
</rdf:RDF>
```

A API, também, possibilita a utilização de outras notações, como a denominada *N-Triples* [137] definida pelo *W3C RDFCore Working Group* e a N3. A API *Jena* é composta por uma diversidade crescente de componentes, dos quais se podem destacar os seguintes:

- Uma API RDF;
- Uma API OWL;
- Possibilidade de criar ou manipular informação definida nos seguintes formatos: RDF/XML, N3 e N-Triples;
- Possibilidade de armazenamento persistente ou não persistente;
- Uma linguagem de interrogação para RDF, denominada RDQL.

### 5.3 Implementação da Ferramenta SWedt

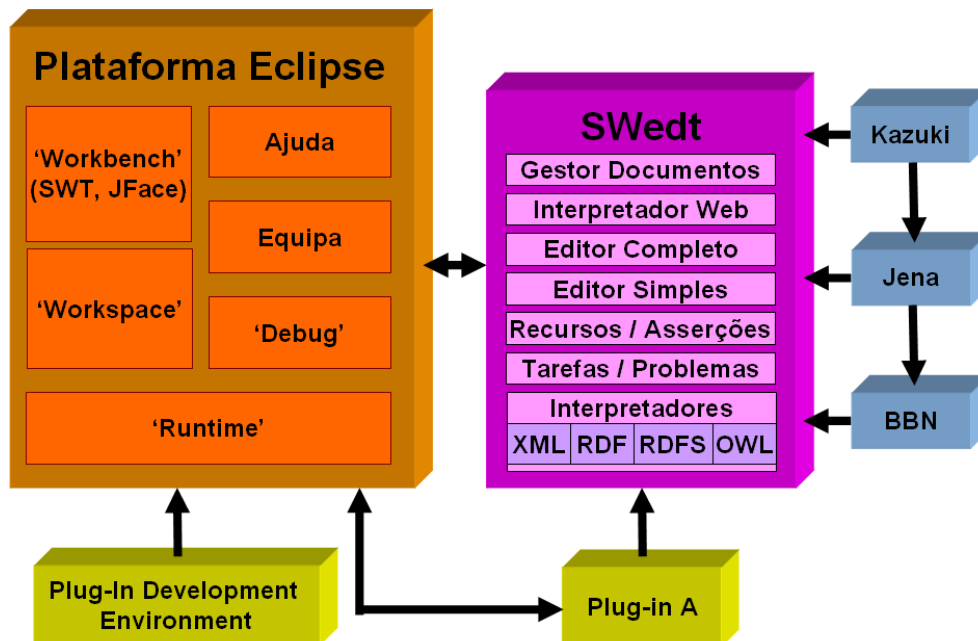


Figura 5.7: Interligação das tecnologias e componentes associados à ferramenta SWedt.

A figura 5.7 apresenta as principais tecnologias relacionadas com a ferramenta SWedt, a qual se encontra representada através do bloco de cor roxa. As relações existentes entre os diversos blocos apresentados na figura estão representadas através de setas unidireccionais e bidireccionais. A Plataforma *Eclipse*, representada na figura através do bloco cor de laranja, foi de extrema importância para a idealização e a materialização da ferramenta SWedt. Mas, em conjugação com a plataforma *Eclipse*, a ferramenta SWedt também utilizou o seguinte conjunto de APIs/componentes/*Plug-ins* disponibilizadas pela comunidade científica mundial:

- PDE. Representado na parte inferior direita da figura através de um bloco de cor amarela;
- API Kazuki. Representada na parte superior direita da figura através de um bloco de cor azul;
- API Jena. Representada na parte central direita da figura através de um bloco de cor azul;

## 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

- API/Componentes BBN. Representada na parte inferior direita da figura através de um bloco de cor azul.

A linguagem de programação escolhida para implementar a ferramenta SWedt foi a linguagem *Java* por utilizar uma metodologia de desenvolvimento orientada aos objectos e ser independente do sistema operativo. Este capítulo apresenta nas próximas secções, o IDE utilizado no desenvolvimento da ferramenta SWedt, bem como as principais entidades e acções implementadas.

### 5.3.1 *Plug-In Development Environment (PDE)*

A ferramenta SWedt foi idealizada para ser implementada como um *Plug-in* para a plataforma *Eclipse*. A plataforma *Eclipse* fornece um ambiente de apoio ao desenvolvimento de *Plug-ins* em linguagem *Java* para a sua plataforma. Esse ambiente de desenvolvimento de *Plug-ins* em linguagem *Java* denomina-se *Plug-In Development Environment (PDE)* e foi o escolhido para ser utilizado no desenvolvimento da ferramenta SWedt.

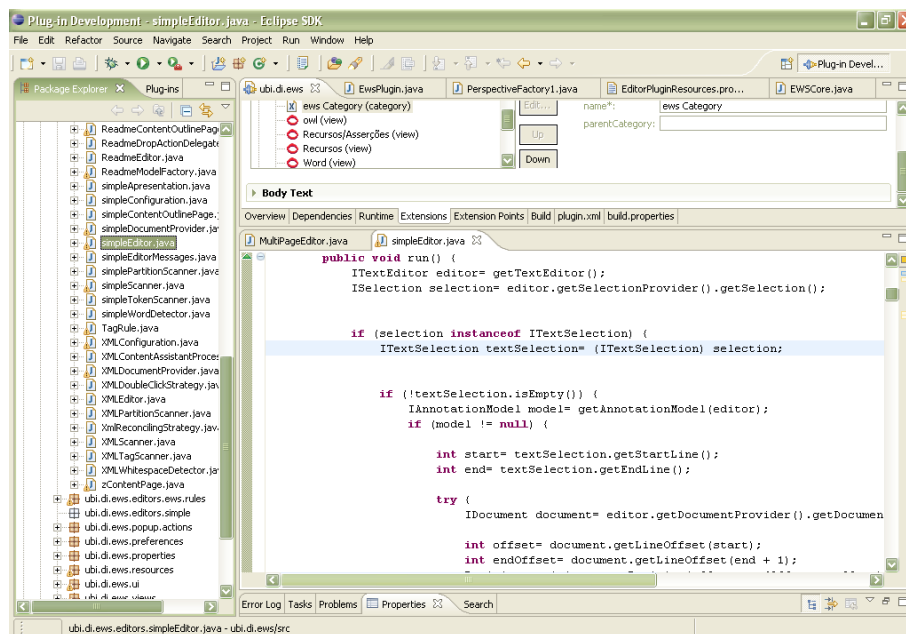


Figura 5.8: *Plug-In Development Environment*.

A figura 5.8 apresenta o PDE da plataforma *Eclipse* durante o desenvolvimento da ferramenta SWedt. De forma muito resumida a imagem pode ser dividida em duas partes, a parte esquerda

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

e a parte direita. A parte esquerda da imagem apresenta a vista *Package Explorer* a qual apresenta os *Packages* e os ficheiros *Java* da ferramenta SWedt. A actual versão da ferramenta SWedt é constituída por 15 *Packages* e mais de 250 classes *Java*, que se encontram representadas através de diagramas de classes em UML no apêndice C. A parte direita da imagem apresenta o Editor Java, o qual permite a manipulação dos ficheiros *Java* pertencentes à ferramenta SWedt. O PDE, também, permite integrar de forma simples os componentes disponibilizados pela plataforma *Eclipse* e as APIs utilizadas pela ferramenta SWedt.

#### 5.3.2 Principais Entidades e Acções Implementadas

O desenvolvimento da ferramenta SWedt foi caracterizado pela definição e implementação de cinco entidades e por cinco acções que relacionam as entidades entre si. As entidades principais definidas e implementadas estão representadas na figura 5.9 através da utilização de cinco blocos coloridos. As acções principais estão representadas na figura através da legendagem das setas que definem as relações existentes entre as entidades principais. Os blocos de cor roxa representam entidades que se encontram relacionadas directamente com a interacção por parte do utilizador. Os restantes blocos, de cores verde e azul, representam entidades que não se encontram relacionados de forma directa com a interacção por parte dos utilizadores. As entidades representadas na figura são:

- Documento XML (Cor verde);
- Texto Simples (Cor roxa);
- Tabela de Associações (Cor roxa);
- Tabela de Asserções (Cor roxa);
- Lista de Ontologias (Cor azul).

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

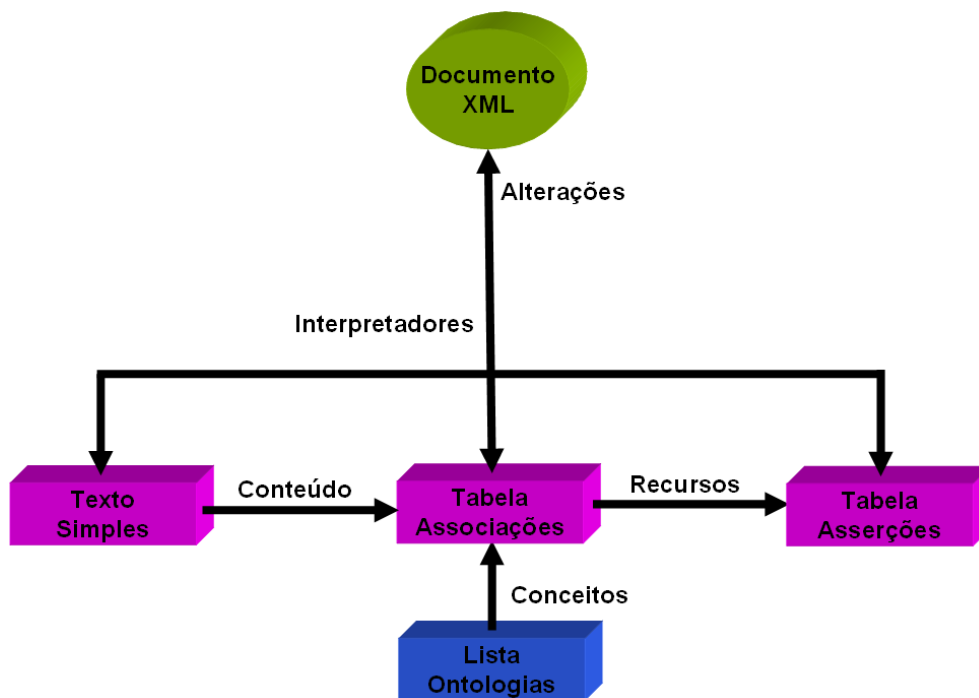


Figura 5.9: Entidades e Acções associadas ao desenvolvimento da ferramenta SWedT.

As acções representadas na figura são:

- Alterações: Representada através da definição de uma relação bidireccional existente entre as entidades: Documento XML, Texto Simples, Tabela de Associações e Tabela de Asserções;
- Interpretadores: Representada através da definição de uma relação bidireccional existente entre as entidades: Documento XML, Texto Simples, Tabela de Associações e Tabela de Asserções;
- Conteúdo: Representada através da definição de uma relação unidireccional existente entre as entidades: Texto Simples e Tabela de Associações;
- Recursos: Representada através da definição de uma relação unidireccional existente entre as entidades: Tabela de Associações e a Tabela de Asserções;
- Conceitos: Representada através da definição de uma relação unidireccional existente entre as entidades: Lista de Ontologias e a Tabela de Associações.

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

Estas entidades e acções são consideradas o núcleo duro da implementação da ferramenta SWedt. Nas próximas sub-secções será apresentado o objectivo e a funcionalidade associada a cada uma destas entidades e acções.

#### **Entidade Documento XML**

A ferramenta SWedt utiliza a API *Jena* para representar internamente através de um Modelo RDF qualquer documento XML que o utilizador possa disponibilizar à ferramenta<sup>140</sup>. Esse Modelo RDF de um documento XML representado internamente pela ferramenta SWedt é também designado por entidade Documento XML.

A entidade Documento XML possui um duplo objectivo. Primeiro, permitir representar virtualmente um ficheiro XML internamente e assim facilitar a interacção existente entre as entidades internas da ferramenta SWedt e os dados existentes no ficheiro XML. Segundo, permitir uma fácil manipulação dos dados existentes no ficheiro XML através da utilização de um Modelo RDF disponibilizado na memória primária do computador.

A criação e gestão da entidade Documento XML é da única responsabilidade da ferramenta SWedt e esta entidade encontra-se associada a três outras entidades da ferramenta: entidade Texto Simples, entidade Tabela de Associações e entidade Tabela de Asserções. A associação existente entre a entidade Documento XML e as três entidades referidas na frase anterior é bidireccional e apresenta dois tipos de acções: Acção Alterações e Acção Interpretadores.

A acção Alterações é bidireccional e permite estabelecer uma actualização entre o conteúdo existente na entidade Documento XML e as três entidades referidas no parágrafo anterior. Sempre que o conteúdo da entidade Documento XML for alterado a ferramenta SWdet, através da acção Alterações, automaticamente actualizará os conteúdos das outras três entidades refe-

---

<sup>140</sup>Independentemente do documento XML disponibilizado possuir, ou não, dados definidos através das sintaxes RDF/XML, RDFS e OWL.

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

ridas de forma a estabelecer a actualidade e integridade do conteúdo apresentado pelos diversos componentes da ferramenta. O inverso também é verdade, isto é, sempre que o conteúdo das três entidade referidas for alterado a ferramenta SWdet, através da acção Alterações, automaticamente actualizará o conteúdo da entidade Documento XML.

A acção Interpretadores, também, é bidireccional. Ela, através da utilização da API *Jena*, permite interpretar e manipular o Modelo RDF da entidade Documento XML. Esta acção é executada sempre que a ferramenta SWedt necessita de actualizar a entidade Documento XML ou de preencher o conteúdo dos dados presentes nas seguintes entidades:

- Entidade Texto Simples: O conteúdo dos dados apresentados nesta entidade é, apenas, constituído pelo conteúdo textual presente no Modelo RDF. Desta forma, todo o conteúdo referente às instruções propriamente ditas das sintaxes XML, RDF, RDFS e OWL existente no Modelo RDF é descartado por parte da actuação da acção Interpretadores;
- Entidade Tabela de Associações: O conteúdo dos dados apresentados nesta entidade é, apenas, constituído pelos Recursos RDF definidos no Modelo RDF e por todas as palavras existentes na entidade Texto Simples. Desta forma, todo o restante conteúdo existente no Modelo RDF é descartado por parte da actuação da acção Interpretadores;
- Entidade Tabela de Asserções: O conteúdo dos dados apresentados nesta entidade é, apenas, constituído pelos Recursos e Asserções RDF existentes no Modelo RDF. Desta forma, todo o restante conteúdo do Modelo RDF é descartado por parte da actuação da acção Interpretadores.

Como a entidade Documento XML representa, de forma interna e actualizada, um ficheiro XML, sempre que a ferramenta SWedt necessite de proceder à sua guarda de forma permanente num dispositivo físico é à própria entidade que será delegada essa tarefa. Assim, a entidade Documento XML, através da utilização da API *Jena*, converte o Modelo RDF do documento actual para um ficheiro XML em sintaxe RDF/XML.

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

#### Entidade Texto Simples

A ferramenta SWedt disponibiliza um editor que permite a manipulação do conteúdo textual presente num ficheiro XML por parte do utilizador. Esse editor é o núcleo duro da entidade Texto Simples. O principal objectivo desta entidade é o de permitir o simples acesso e edição do conteúdo textual presente num ficheiro XML por parte do utilizador da ferramenta SWedt.

Esta entidade permite a manipulação do conteúdo textual de um ficheiro XML através da invocação da acção Interpretador sobre a entidade Documento XML. Assim, através dessa acção, e por sua vez através da utilização da API *Jena*, a entidade acede ao Modelo RDF da ferramenta e extrai todo conteúdo textual do ficheiro XML.

A apresentação do conteúdo textual de um ficheiro XML através de um editor é da responsabilidade da ferramenta SWedt, mas a manipulação desse conteúdo é da responsabilidade do utilizador do editor. Assim, sempre que o conteúdo da entidade Texto Simples for alterado pelo utilizador a ferramenta SWdet, através da acção Alterações, automaticamente actualizará o conteúdo da entidade Documento XML.

Além destas duas acções, Interpretador e Alterações, a entidade Texto Simples também interage com uma acção unidireccional denominada de Conteúdo. Esta última acção é sempre executada quando o conteúdo pertencente à entidade Texto Simples é alterado. A sua função é a de fornecer uma lista de todas as palavras existentes na Entidade Texto Simples à entidade Tabela de Associações.

A entidade Texto Simples apenas apresenta uma versão parcial do conteúdo do ficheiro XML, porque o seu objectivo é o de reduzir ao máximo a necessidade de conhecimento de sintaxes relacionadas com a Web Semântica por parte dos utilizadores desta ferramenta.



### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

#### Entidade Tabela de Associações

A ferramenta SWedt disponibiliza uma tabela com a declaração de todos os Recursos RDF existentes num ficheiro XML. Essa tabela é o núcleo duro da entidade Tabela de Associações. Chama-se entidade Tabela de Associações porque possui associada uma tabela composta por três colunas, sendo cada linha dessa tabela constituída por uma declaração de um Recurso RDF e, por sua vez, cada declaração de um Recurso RDF composta por uma associação entre uma palavra pertencente ao conteúdo da entidade Texto Simples (primeira coluna da tabela) e um URI<sup>141</sup> (segunda coluna da tabela). A terceira coluna da tabela disponibiliza um mecanismo de selecção gráfico que possibilita a validação de associações existentes entre essas palavras e esses URIs por parte do utilizador da ferramenta SWedt. O principal objectivo desta entidade é o de permite a definição de Recursos RDF, através da manipulação de uma simples tabela associada ao documento XML, por parte de utilizadores não familiarizados com a Sintaxe e o Modelo RDF.

Esta entidade preenche a sua tabela através da invocação das seguintes três acções:

- Acção Interpretadores: Esta entidade permite a apresentação e manipulação dos Recursos RDF existentes num ficheiro XML através da invocação da acção Interpretador sobre a entidade Documento XML. Assim, através dessa acção, e por sua vez através da utilização da API *Jena*, a entidade extrai todos os Recursos RDF declarados no Modelo RDF interno da ferramenta e insere-os na tabela de associações;
- Acção Conteúdo: Esta acção fornece uma lista de todas as palavras existentes na Entidade Texto Simples à entidade Tabela de Associações. A entidade Tabela de Associações extrai todas as palavras constituintes dessa lista e insere-as, correspondendo cada palavra a uma linha da sua tabela, na primeira coluna da sua tabela;
- Acção Conceitos: Esta acção cria uma lista de todas as palavras existentes na tabela anterior que não possuam um URI associado, isto é, a segunda coluna da tabela esteja em

---

<sup>141</sup>Conceito definido numa ontologia.

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

branco. Os URI são obtidos através da identificação de conceitos, definidos em ontologias, que seja iguais ou semelhantes as palavras existentes na primeira coluna da tabela. Essa semelhança é verificada sintacticamente e na forma de domínio. A principal função desta acção é a de procurar conceitos, existentes em ontologias, iguais ou semelhantes às palavras existentes na tabela da entidade Tabela de Associações de forma a criar associações entre palavras e URIs de forma automática. Os conceitos definidos em ontologias são disponibilizados pela entidade Lista de Ontologias.

Além da invocação destas três acções, sempre que a tabela é alterada, quer automaticamente pela ferramenta, quer pelo utilizador, esta entidade invoca outras duas acções: Alterações e Recursos. A acção Alterações, sempre que invocada, actualizará, de forma automática, o conteúdo da entidade Documento XML sempre que as alterações efectuadas na tabela exijam o estabelecimento da integridade do conteúdo apresentado por todos os componentes da ferramenta. A acção unidireccional de nome Recursos, sempre que invocada, disponibilizará à entidade Tabela de Asserções uma lista de todos os Recursos existentes na tabela da entidade Tabela de Associações.

A entidade Tabela de Associações apenas apresenta uma versão parcial do conteúdo do ficheiro XML, porque o seu objectivo é o de reduzir ao máximo a necessidade de conhecimento de sintaxes relacionadas com a Web Semântica por parte dos utilizadores da ferramenta SWedt.

#### **Entidade Lista de Ontologias**

A ferramenta SWedt disponibiliza uma lista de ontologias que pode ser actualizada a qualquer momento. A entidade responsável pela gestão dessa lista denomina-se entidade Lista de Ontologias. A sua principal função é a de possibilitar a procura e extracção de conceitos definidos numa lista de ontologia por parte da acção Conceitos. A actividade de navegação através de conceitos definidos em Ontologias foi implementada através da utilização da API *Kazuki*. O funcionamento desta entidade apresenta um comportamento transparente para os utilizadores da ferramenta SWedt.

### 5.3. IMPLEMENTAÇÃO DA FERRAMENTA SWEDT

---

#### Entidade Tabela de Asserções

A ferramenta SWedt disponibiliza uma tabela com a declaração de todas as Asserções RDF existentes num ficheiro XML. Essa tabela é o núcleo duro da entidade Tabela de Asserções. O principal objectivo desta entidade é o de permite a definição de Asserções RDF, através da manipulação de uma simples tabela associada ao documento XML, por parte de utilizadores não familiarizados com a Sintaxe e o Modelo RDF.

Esta entidade preenche a sua tabela através da invocação das seguintes duas acções e por intermédio do utilizador:

- **Acção Interpretadores:** Esta entidade permite a apresentação e manipulação dos Asserções RDF existentes num ficheiro XML através da invocação da acção Interpretador sobre a entidade Documento XML. Assim, através dessa acção, e por sua vez através da utilização da API *Jena*, a entidade extrai todas as Asserções RDF declarados no Modelo RDF interno da ferramenta e insere-os na tabela de asserções;
- **Acção Recursos:** Esta acção fornece uma lista de todos os Recursos RDF existentes na Entidade Tabela de Associações à entidade Tabela de Asserções. A entidade Tabela de Asserções extrai todos os Recursos RDF constituintes dessa lista e insero-os em três combo-listas por forma a poderem ser seleccionados pelos utilizadores;
- **Acção do Utilizador:** Esta entidade fornece uma tabela que ao ser manipulada pelo utilizador permitirá criar ou alterar asserções RDF. A ferramenta disponibiliza apenas Recursos RDF a esta tabela e assim o utilizador ao criar associações tripla entre os recursos RDF disponibilizados estará a criar relações lógicas entre os recursos e sem o saber, estará assim a criar asserções RDF.

Além da invocação destas três acções, sempre que a tabela é alterada, quer automaticamente pela ferramenta, quer pelo utilizador, esta entidade invoca a acção Alterações. A acção Alte-

rações, sempre que invocada, actualizará, de forma automática, o conteúdo da entidade Documento XML sempre que as alterações efectuadas na tabela asserções exijam o estabelecimento da integridade do conteúdo apresentado por todos os componentes da ferramenta.

A entidade Tabela de Asserções apenas apresenta uma versão parcial do conteúdo do ficheiro XML, porque o seu objectivo é o de reduzir ao máximo a necessidade de conhecimento de sintaxes relacionadas com a Web Semântica por parte dos utilizadores desta ferramenta.

### 5.4 Modo e Funcionamento da Ferramenta SWedt

A ferramenta desenvolvida durante a presente dissertação, além de apresentar um modo de funcionamento muito simples, também é visualmente atraente. Neste sub-capítulo serão apresentados os componentes gráficos da GUI, o modo de funcionamento e um exemplo de funcionamento da ferramenta SWedt.

#### 5.4.1 Componentes gráficas da ferramenta

Por forma a homogeneizar o aspecto gráfico e a funcionalidade da ferramenta SWedt, todos os componentes gráficos e ficheiros em formato XML que se apresentem associados à ferramenta na plataforma *Eclipse* apresentarão o ícone da imagem da figura 5.10.



Figura 5.10: Ícone SWedt.

A imagem da figura 5.11 apresenta o GUI da ferramenta SWedt, o qual é constituído por um editor múltiplo associado a seis vistas auxiliares. A plataforma *Eclipse* permite agrupar editores e vistas específicas que trabalhem em sintonia para o mesmo objectivo numa única entidade gráfica denominada Perspectiva. Assim, a actual GUI da ferramenta SWedt, apresentada na figura 5.11, não é mais do que uma perspectiva, de nome SWedt, para a plataforma *Eclipse*.

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

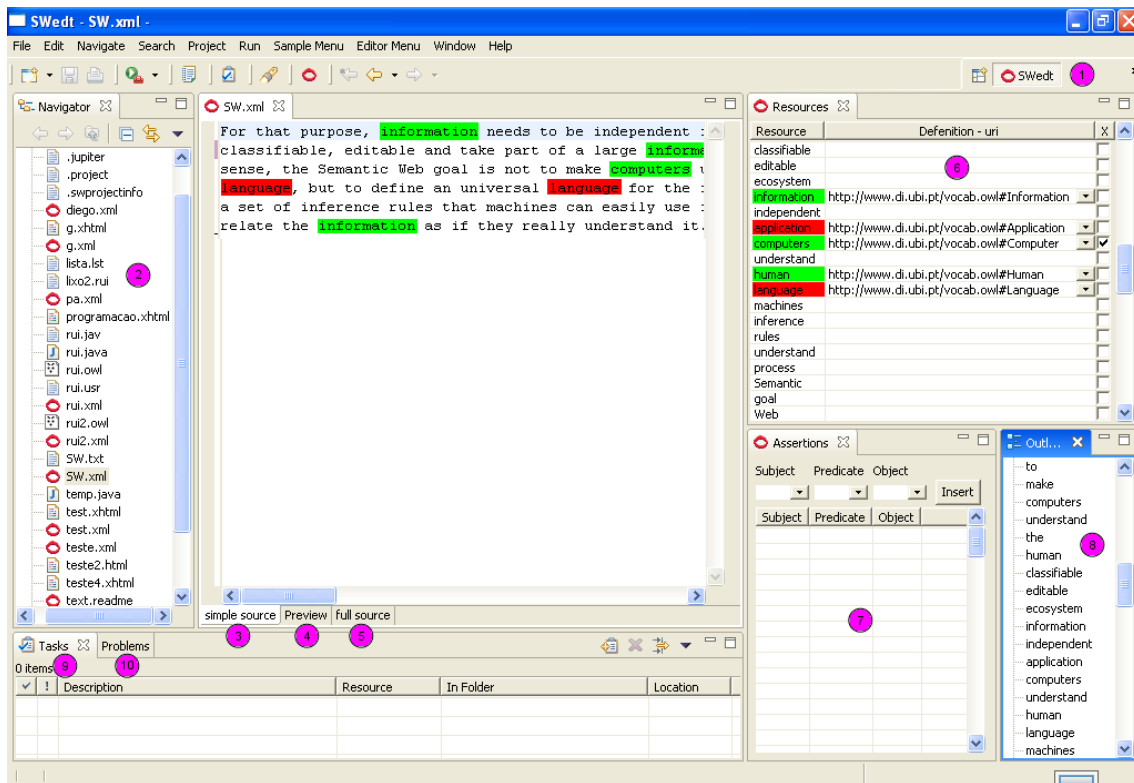


Figura 5.11: Ferramenta SWedt.

De forma a contribuir para uma maior facilidade na identificação dos componentes gráficos constituintes da perspectiva SWedt, é apresentada de seguida, através da utilização de uma legenda para a imagem da figura 5.11, uma breve descrição de cada dos componentes gráficos da ferramenta:

- 1 - Perspectiva seleccionada: Apresenta a perspectiva SWedt como a perspectiva actualmente seleccionada;
- 2 - Vista Gestor de documentos (*Navigator*<sup>142</sup>): Vista que possibilita a navegação entre os diversos directórios existentes no disco rígido de um computador, criar novos ficheiros, modificar ficheiros e editar ficheiros através da utilização de um editor associado;
- 3 - Editor Simples (*simple source*): Principal editor da aplicação. Permite a inserção de texto sem qualquer tipo de restrição;

<sup>142</sup>Tendo em conta que cada vez mais o desenvolvimento de *software* abrange uma amplitude mundial, optei por usar nesta primeira versão da ferramenta termos em língua Inglesa. A língua Portuguesa será contemplada numa próxima versão da ferramenta.

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

- 4 - Interpretador Web (*Preview*): Permite a visualização através de uma página Web interna do conteúdo existente no Editor Simples;
- 5 - Editor Completo (*full source*): Apresenta todo o conteúdo do ficheiro através da utilização das sintaxes XML, RDF, RDFS e OWL;
- 6 - Vista Recursos (*Resources*): Vista que permite observar, inserir, alterar e validar associações entre URIs<sup>143</sup> e palavras existentes no Editor Simples;
- 7 - Vista Asserções (*Assertions*): Vista que permite associar entre si os recursos que se encontram na vista Recursos. Por cada associação efectuada nesta vista a ferramenta adiciona, de forma totalmente automática, uma asserção RDF ao ficheiro em formato XML original;
- 8 - Vista Informação (*Outline*): Vista que apresenta através de uma lista todas as palavras existentes no Editor Simples;
- 9 - Vista Tarefas (*Tasks*): Vista que apresenta e gere uma lista de tarefas a definir pelo utilizador;
- 10 - Vista Problemas (*Problems*): Vista que apresenta e gere uma lista de problemas apresentados durante a execução da ferramenta;

Nas próximas secções serão apresentados, em maior pormenor, cada um dos componentes gráficos constituintes da perspectiva SWedt desenvolvida para a plataforma *Eclipse*.

#### Vista Gestor de Documentos

A Vista Gestor de documentos, figura 5.12, permite a navegação entre os diferentes directórios e ficheiros existentes no disco rígido de um computador, criar novos ficheiros, modificar ficheiros e editar ficheiros através da utilização de um editor associado. Ela apresentada uma estrutura

---

<sup>143</sup>Conceitos definidos em ontologias.

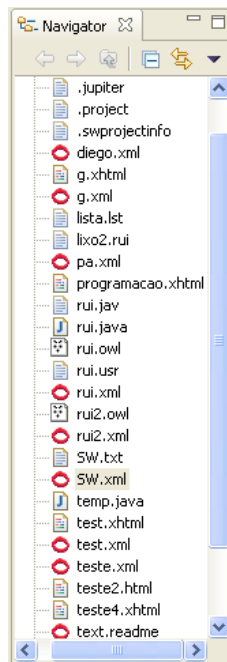


Figura 5.12: Vista Gestor de documentos.

em árvore preenchida com todos os ficheiros e directórios de uma localização do disco rígido em particular. Associado a cada ficheiro está uma pequena imagem<sup>144</sup> que identifica o editor, existente ou instalado na plataforma *Eclipse*, que lhe está associado. No caso concreto da ferramenta *SWedt* o tipo de ficheiro associado é o XML e a imagem associada é a apresentada na figura 5.10.

### Editor Múltiplo

Sempre que o utilizador pretenda editar ficheiros com a extensão XML, a ferramenta *SWedt* disponibiliza um Editor Múltiplo para esse efeito. O Editor apresentado pela ferramenta denomina-se múltiplo porque é composto por três outros elementos:

- Editor Simples;
- Interpretador Web;
- Editor Completo.

---

<sup>144</sup>Normalmente referida como ícone.

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

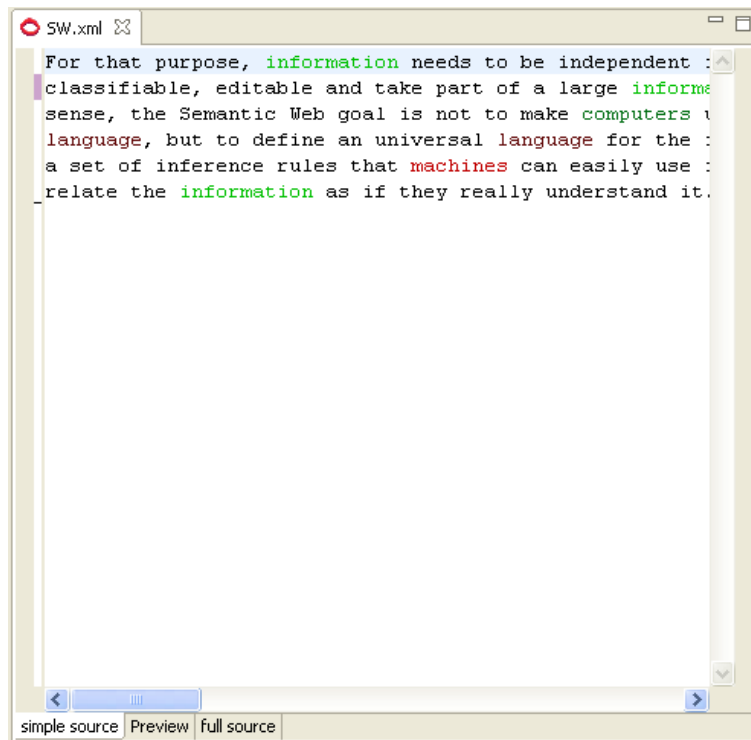


Figura 5.13: Editor Simples da ferramenta SWedT.

O Editor Simples, figura 5.13, é o principal editor da aplicação e tem como característica ser o mais simples possível. O utilizador insere as suas frases neste editor de uma forma livre e sem qualquer tipo de restrições. Por forma a facilitar a interacção com o utilizador, o Editor Simples utiliza uma combinação de cores associadas às palavras existentes no texto de acordo com as seguintes situações:

- Cor verde negrito: Palavra do texto que se encontra associada a um URI e em que essa associação tenha sido confirmada pelo utilizador na Vista Recursos. Por exemplo, a palavra "computers" da figura 5.13;
- Cor verde: Palavra do texto associada a mais do que um URI e em que pelo menos uma dessas associações tenha sido confirmada pelo utilizador na Vista Recursos. Por exemplo, a palavra "information" da figura 5.13;
- Cor vermelha negrito: Palavra do texto associada a um URI e no caso de essa associação ainda não ter sido confirmada pelo utilizador na Vista Recursos. Por exemplo, a palavra



## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

”language” da figura 5.13;

- Cor vermelha: Palavra do texto associada a mais do que um URI e no caso de nenhuma dessas associações ainda ter sido confirmada pelo utilizador na Vista Recursos. Por exemplo, a palavra ”machines” da figura 5.13;
- Cor preta: Restante texto - palavras ainda não associadas a um URI.

A associação entre palavras do texto e URIs específicos é realizada de forma automática pela ferramenta desenvolvida.

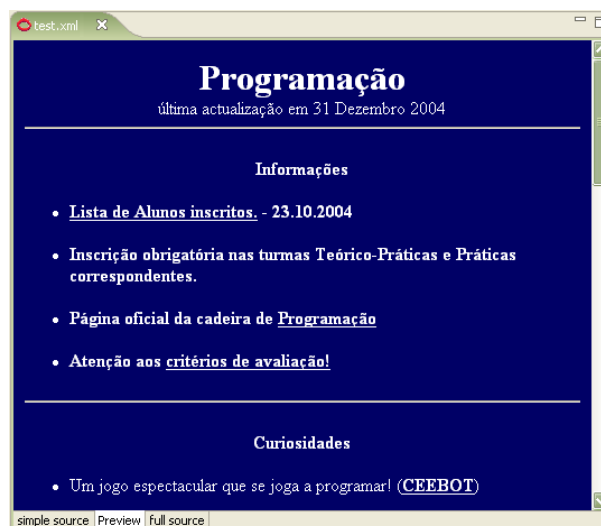


Figura 5.14: Interpretador Web da ferramenta SWedt.

A figura 5.14 mostra como o Interpretador Web apresenta visualmente o texto redigido no Editor Simples. Ele é disponibilizado como sendo uma ”mais valia” para a ferramenta com o objectivo de captar o utilizador. Na verdade, actualmente não existe nenhum editor de páginas web que não possua o seu Visualizador Web interno.

O Editor Completo, apresentado na figura 5.15, é o editor da ferramenta que apresenta todo o conteúdo do ficheiro XML, incluindo desta forma as sintaxes XML, RDF, RDFS e OWL. Ele utiliza o seguinte conjunto de cores associado à sintaxe de cada um dos tipos dos elementos:

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

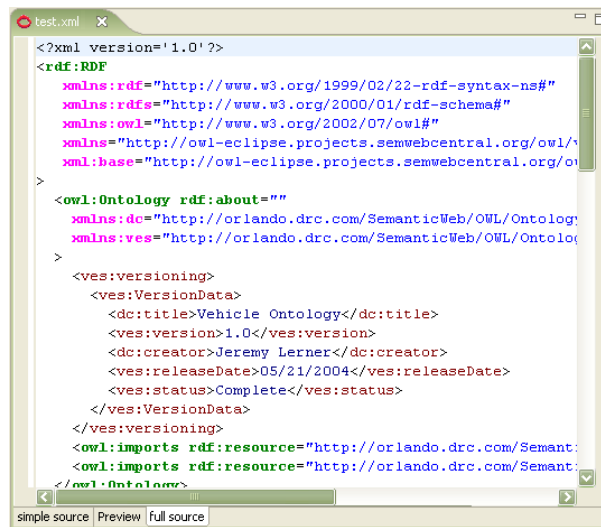


Figura 5.15: Editor Completo da ferramenta SWedt.

- Cor cinzenta: Comentários;
- Cor verde negrito: Elementos RDF, RDFS e OWL;
- Cor azul: Conjunto de caracteres limitado por aspas<sup>145</sup>;
- Cor azul escuro: Valores numéricos;
- Cor vermelha: Elementos XML;
- Cor de rosa: Espaços de Nomes.

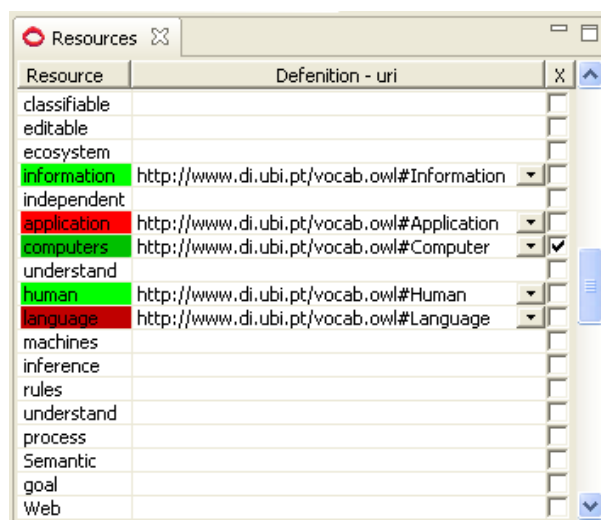
Este editor disponibiliza um auxiliar de escrita de código, isto é, sempre que é inserido o carácter '`<`' o editor apresenta uma lista das classes, propriedades e prefixos dos Espaços de Nomes localmente definidos. Desta forma, o utilizador pode seleccionar um valor da lista em vez de o escrever, evitando assim e até certo ponto o aparecimento de erros na escrita do código. Além disso, este editor também detecta automaticamente erros nas sintaxes XML, RDF, RDFS e OWL. Sempre que um erro de sintaxe é detectado no documento é assinalado através da apresentação do texto incorrecto sublinhado e da colocação de um ícone de erro no lado esquerdo da janela do editor.

<sup>145</sup>Normalmente referidos por *strings*.

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

Apesar de este editor disponibilizar um ambiente agradável e funcional, a sua utilização apenas é aconselhável a utilizadores avançados, isto é, utilizadores familiarizados com as tecnologias da Web Semântica. Mesmo sendo o Editor Completo uma representação interna do documento original em formato XML, que está a ser manipulado pela ferramenta, e estar interligado a todas as acções tomadas nos outros componentes da ferramenta, não existe absoluta necessidade por parte da maioria dos potenciais utilizadores desta ferramenta de utilizar este editor.

### Vista Recursos



Resource	Defenition - uri
classifiable	
editable	
ecosystem	
information	<a href="http://www.di.ubi.pt/vocab.owl#Information">http://www.di.ubi.pt/vocab.owl#Information</a>
independent	
application	<a href="http://www.di.ubi.pt/vocab.owl#Application">http://www.di.ubi.pt/vocab.owl#Application</a>
computers	<a href="http://www.di.ubi.pt/vocab.owl#Computer">http://www.di.ubi.pt/vocab.owl#Computer</a>
understand	
human	<a href="http://www.di.ubi.pt/vocab.owl#Human">http://www.di.ubi.pt/vocab.owl#Human</a>
language	<a href="http://www.di.ubi.pt/vocab.owl#Language">http://www.di.ubi.pt/vocab.owl#Language</a>
machines	
inference	
rules	
understand	
process	
Semantic	
goal	
Web	

Figura 5.16: Vista Recursos da ferramenta SWedt.

A vista Recursos, mostrada na figura 5.16, apresenta uma tabela constituída por três colunas. A primeira coluna, de nome Recurso, apresenta todas as palavras existentes no Editor Simple. A segunda coluna, de nome Definição-URI, apresenta a lista de todos os conceitos definidos em ontologias (URIs) que sejam lexicamente iguais à palavra existente na primeira coluna e/ou todos os conceitos que se encontrem definidos em domínios que estejam associados às restantes palavras que compõem a frase onde a palavra existente na primeira coluna se encontra inserida. A construção da lista de associação entre as palavras existentes no Editor Simple e os conceitos definidos em ontologias é realizada automaticamente por parte da ferramenta, a qual se baseia numa lista geral de ontologias que, também, se encontra associada à ferramenta. Apesar disso, o utilizador pode inserir manualmente outras identificações para os recursos. A terceira coluna

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

possui uma caixa de verificação, a qual poderá ser utilizada pelo utilizador para confirmar a associação existente entre a palavra existente na primeira coluna como o conceito associado a uma ontologia identificado na segunda coluna. Sempre que uma associação deste tipo é confirmada, a ferramenta automaticamente codifica no documento em formato XML original uma associação entre um URI e uma palavra existente no documento.

Como esta tabela tem uma relação muito estreita com o Editor Simples, também utiliza a mesma combinação de cores associadas às palavras existentes no texto de acordo com as situações referidas na secção 5.4.1 - "Editor Múltiplo".

### Vista Asserções

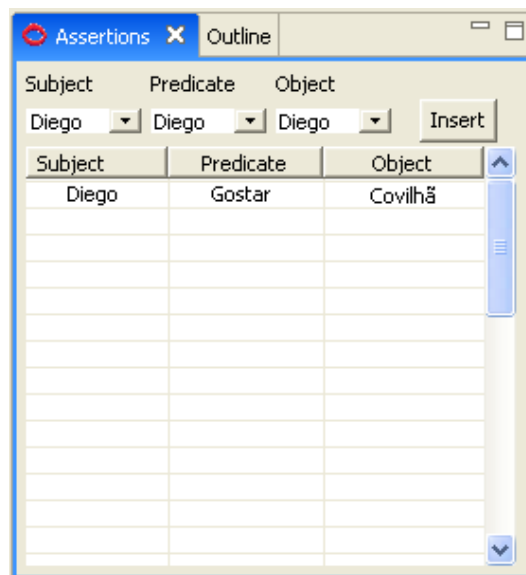


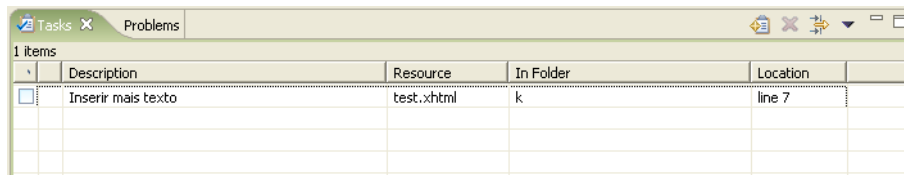
Figura 5.17: Vista Asserções da ferramenta SWedt.

A vista Asserções, mostrada na figura 5.17, apresenta uma tabela, também, constituída por três colunas com os seguintes nomes: Sujeito (*Subject*), Predicado (*Predicate*) e Objecto (*Object*). Esta tabela permite a definição de Asserções RDF de uma forma rápida e simples por parte do utilizador. Em cada uma das colunas existe uma lista com todos os recursos RDF definidos na tabela Recursos anterior. Desta forma é possível criar asserções RDF, bastando para isso que o Sujeito e o Predicado da asserção estejam definidos através de URIs e o Objecto da asserção

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

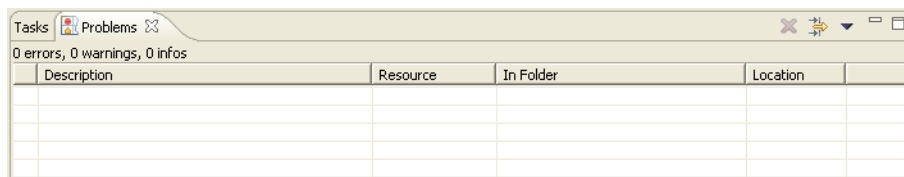
esteja também definido através de um URI ou que seja um simples valor alfanumérico. Sempre que é criada uma asserção RDF, ela é associada, automaticamente pela ferramenta, ao ficheiro em formato XML original.

### Vistas Tarefas e Problemas



	Description	Resource	In Folder	Location
<input type="checkbox"/>	Inserir mais texto	test.xhtml	k	line 7

Figura 5.18: Vista Tarefas da ferramenta SWedt.



Description	Resource	In Folder	Location

Figura 5.19: Vista Problemas da ferramenta SWedt.

As vistas de Tarefas, figura 5.18, e de Problemas, figura 5.19, são utilizadas apenas como um auxílio à edição de texto no Editor Simples e de código fonte no Editor Complexo. A vista Tarefas apresenta uma tabela que pode ser utilizada pelo utilizador para inserir e gerir as tarefas que pretende realizar. A vista Problemas, também, apresenta uma tabela que é actualizada sempre que o Editor Complexo detecta algum erro de sintaxe.

### 5.4.2 Modo de Funcionamento

O esquema da figura 5.20 representa as interacções existentes entre os seis principais componentes gráficos da ferramenta SWedt. Na parte superior da imagem e da esquerda para a direita aparecem os componentes gráficos: Gestor de Documentos, Editor Completo e Interpretador Web. Por sua vez, na parte inferior da imagem e, também, da esquerda para a direita aparecem os componentes gráficos: Editor Simples, Recursos e Asserções. Todos estes componentes

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

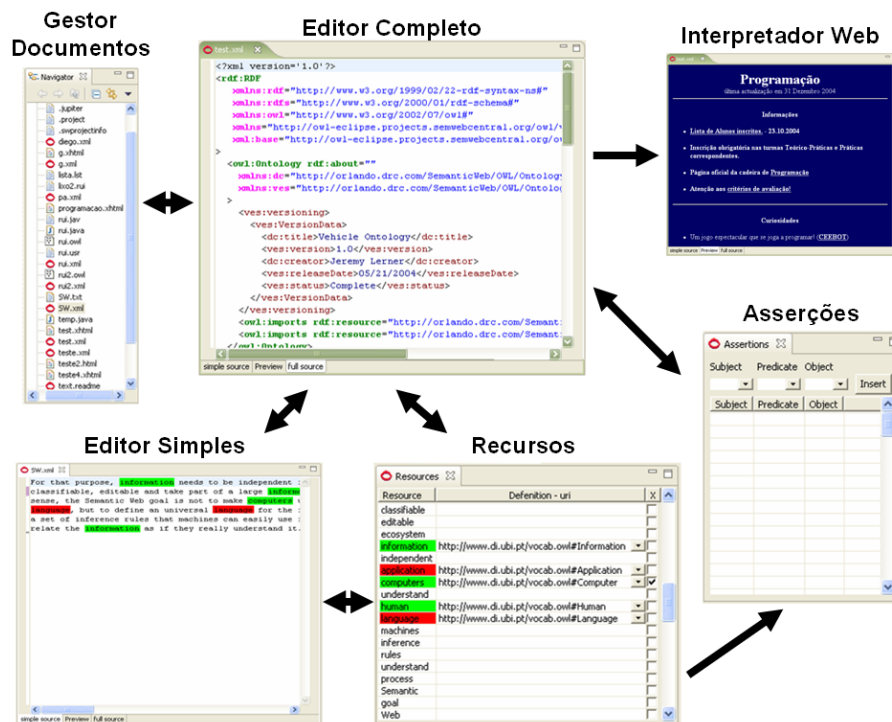


Figura 5.20: Interações existentes entre os componentes gráficos do Editor SWedt.

gráficos foram apresentados na anterior secção deste capítulo. A maioria das interações representadas na imagem são bidireccional, existindo apenas duas interações unidireccional: uma entre o Editor Completo e o Interpretador Web e a outra entre a Vista Recursos e a Vista Asserções.

De uma forma geral, o funcionamento da ferramenta SWedt é simples e transparente na medida em que não existe a necessidade por parte do utilizador de compreender ou manipular directamente os conceitos de URI, Asserções RDF e ontologias. A ferramenta identifica os ficheiros com extensão XML como sendo ficheiros para a Web Semântica e apresenta três fases principais de funcionamento, não obrigatoriamente sequenciais, para a manipulação deste tipo de ficheiros:

- Primeira Fase: A primeira fase é caracterizada pela inserção e edição de texto simples por parte do utilizador. Ela engloba a utilização dos seguintes componentes:
  - Gestor de Documentos: É responsável por localizar e permitir o acesso da ferramenta

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

aos ficheiros existentes no disco rígido do computador. Quando o acesso é solicitado a ficheiros em formato XML, este componente automaticamente envia o todo conteúdo desse ficheiro para o Editor Completo;

- Editor Completo: Representa internamente para a ferramenta o conteúdo do documento em formato XML. Qualquer alteração implementada por parte da maioria dos outros componentes internos da ferramenta terá uma reflexão instantânea neste componente. Neste editor o utilizador pode alterar o conteúdo do documento em formato XML mas a sua utilização é apenas aconselhável para utilizadores conhecedores das tecnologias da Web Semântica; e
  - Editor Simples: Este é o principal editor da ferramenta e por sua vez é o principal interveniente nesta primeira fase de funcionamento da ferramenta. Encontra-se intimamente ligado ao Editor Completo, mas ao contrário dele apenas apresenta o texto simples, "livre" das notações utilizadas pelas sintaxes XML, RDF, RDFS e OWL, que se encontra no documento original. O principal objectivo deste editor é o de permitir a edição de texto de forma simples por parte de qualquer utilizador da ferramenta.
- Segunda Fase: A segunda fase é caracterizada pela associação automática das palavras do texto editado pelo utilizador com URIs. Ela engloba a utilização do seguinte componente:
    - Vista Recursos: Todas as palavras presentes no Editor Simples são recolhidas, organizadas e apresentadas ao utilizador através de uma tabela existente na Vista de Recursos. Essa tabela além de apresentar todas as palavras existentes no Editor Simples também procede automaticamente à associação de cada uma dessas palavras a conceitos definidos em ontologias reconhecidas pela ferramenta, a qual possui uma lista de ontologias que pode ser incrementada a qualquer momento. As associações são efectuadas automaticamente pela ferramenta e podem ser validadas a qualquer momento por parte do utilizador, procedimento este recomendado sempre que aparecem mais do que um URI associado a uma mesma palavra.

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

- Terceira Fase: A terceira fase é caracterizada pela criação de associações entre os recursos por parte do utilizador e que corresponde por sua vez a uma criação automática e interna de asserções RDF por parte da ferramenta. Ela engloba a utilização do seguinte componente:
  - Vista Asserções: É responsável pela criação de associações triplas entre os recursos que se encontrem definidos no conteúdo de um documento em formato XML. A ferramenta apenas considera como recurso um registo que se encontre definido no documento e que esteja representado através de uma associação entre uma palavra do conteúdo textual do documento e um conceito definido numa ontologia.

Além destas três fases principais, o utilizador pode, sempre que o desejar, visualizar o resultado do seu documento XML numa página web, utilizar a Vista Interpretador Web que permite interpretar e simular a visualização em formato Web do conteúdo presente no componente Editor Completo através da GUI da ferramenta SWedt. Este componente não permite contudo manipular o conteúdo presente no componente Editor Completo.

Do ponto de vista dos utilizadores desta ferramenta, os procedimentos a efectuar podem-se resumir a:

- Criar ou abrir um ficheiro em formato XML;
- Editar o conteúdo textual<sup>146</sup> do documento;
- Validar associações entre o conteúdo textual do documento e conceitos associados em ontologias<sup>147</sup>;
- Criar associações triplas entre as associações textuais definidas no ponto anterior.

Como se pode ver através dos procedimentos a efectuar por parte dos utilizadores, apresentados no parágrafo anterior, não existe a necessidade dos utilizadores conhecerem ou dominarem as

---

<sup>146</sup>Texto simples do documento sem qualquer tipo de referências às sintaxes XML, RDF, RDFS e OWL.

<sup>147</sup>O utilizador apenas necessita de validar as associações, a ferramenta é que é responsável pela criação das associações. Não existe uma obrigatoriedade imposta ao utilizador para proceder a essa validação.



tecnologias associadas à Web Semântica. Desta forma, pode-se considerar como atingindo com sucesso os objectivos propostos para o desenvolvimento desta ferramenta.

### 5.4.3 Exemplo de Utilização

A ferramenta SWedt é extremamente simples de utilizar. Afinal, a facilidade de utilização é um dos seus principais objectivos. Mas, o processo de criação de páginas para a Web Semântica não é propriamente simples. Assim, nesta secção é apresentado um exemplo de formatação de dados e informação para a Web Semântica, sub-secção **Dados e Informação na Web Semântica** e uma sub-secção que apresenta um exemplo de utilização da ferramenta SWedt, **Utilização da ferramenta SWedt**.

#### **Dados e Informação na Web Semântica**

Os humanos utilizam uma linguagem comum para comunicarem entre si. Segundo Aaron Swartz [90], a Web pode ser definida como um meio de comunicação, o qual, conseqüentemente é utilizado pelos humanos para comunicarem entre si. Esse meio de comunicação é suportado por computadores e por SPC que, por seu lado, não possuem capacidade suficiente para interpretar a linguagem dos humanos. Desta forma, ao contrário de tentar definir uma nova linguagem universal, a Web Semântica apresenta uma estrutura lógica, evolutiva e universal que permite representar a informação de uma forma compreensível quer pelos humanos quer pelos SPC. Assim, todos os intervenientes deste meio de comunicação poderão compreender a informação e assim proporcionar um serviço de elevada qualidade.

A seguinte frase, retirada de [295], representa uma afirmação expressa em linguagem inglesa:

`Port is traditionally served with Stilton cheese.`

A frase apresentada apenas é totalmente compreendida pelos humanos que falem a língua inglesa, se estes possuírem modelos mentais de conceitos associados às palavras existentes na frase, por exemplo, a palavra **Port** [296] representa o nome de um vinho mundialmente conhecido e

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

a palavra **Stilton** [297] representa a marca de um queijo, também, mundialmente conhecida. Para um humano que desconheça a língua inglesa esta frase não tem qualquer significado, pois ele não consegue identificar as palavras da frase com os seus modelos mentais de conceitos. O mesmo acontece com a generalidade dos SPC, para eles esta frase não passa de uma simples cadeia de símbolos sem qualquer significado, vulgarmente designado por texto simples, isto é, texto que não utiliza qualquer tipo de restrição ou termo predefinido. Actualmente, o texto simples representa a quase totalidade da informação que se encontra disponível na Web. Desta forma, é fácil perceber que a informação presente na Web apenas se encontra disponibilizada com o objectivo de ser compreendida pelos humanos. Mas, a Web Semântica baseia-se num conjunto de tecnologias com o objectivo de inverter esta situação.

Através da utilização da tecnologia XML pode-se, facilmente, estruturar o conteúdo de documentos em texto simples sem alterar o seu significado e sem existir a necessidade de utilizar palavras pré-definidas. De seguida é apresentada a frase **Port is served with Stilton cheese** em sintaxe XML:

```
<frase>
  <vinho>Port</vinho>
  <servido_com>
    <queijo>Stilton</queijo>
  </servido_com>
</frase>
```

Neste exemplo, são utilizadas quatro anotações, denominadas **frase**, **vinho**, **servido\_com** e **queijo**. A anotação, denominada, **frase** engloba toda a frase, as restantes três anotações englobam três palavras específicas existentes na frase. A frase transcrita para a sintaxe XML continua a ser texto simples, mas, agora possui as suas principais palavras estruturadas em forma de uma árvore, como se pode ver na figura 5.21.

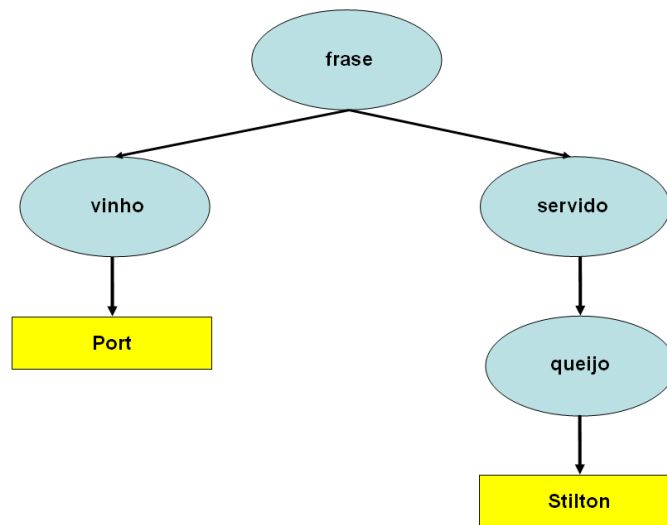


Figura 5.21: Estrutura XML da frase *Port is served with Stilton cheese*.

A utilização de anotações XML não tem qualquer influência na estrutura sintáctica nem no significado da frase original. Uma anotação é composta por uma marcação de início (por exemplo: <vinho>) e uma marcação de fim (por exemplo: </vinho>), sendo o nome que aparece dentro dos sinais < e > o nome da anotação (por exemplo: vinho). O nome a atribuir a uma anotação não está sujeito a nenhum tipo de restrição, podendo ser um qualquer conjunto de caracteres. Geralmente o nome das anotações é utilizado como meta-dados. Neste caso, na frase anteriormente apresentada, a anotação *vinho* está a ser utilizada como meta-dado da palavra *Port*.

Em XML, chama-se elemento XML ao conjunto: marcação de início, marcação de fim e conteúdo delimitado pelas duas marcações anteriores. Neste exemplo, o elemento XML *frase*, informa um SPC que a frase *Port is served with Stilton cheese* se encontra encapsulada no elemento XML de nome *frase*, mas não lhe explica o que é uma frase. O SPC fica apenas a saber que *Port is served with Stilton cheese* é um conjunto de caracteres que está encapsulado no elemento XML *frase*, independentemente do que esse conjunto de caracteres possa querer dizer. Apesar de os SPC não compreenderem o significado da palavra *frase*, sabem que estes dois tipos de dados, os dados propriamente ditos e os respectivos meta-dados, se encontram associados. Assim, um

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

SPC compreende que a palavra **Port** está associada ao meta-dado **vinho** e a palavra **Stilton** está associada ao meta-dado **queijo**, independentemente dos significados que **vinho** e **queijo** possam possuir.

Por vezes existe a necessidade de acrescentar mais informação acerca do conteúdo de um elemento XML. Por exemplo, podemos especificar que a palavra **Port** da frase anterior representa uma marca de vinho em particular. Este tipo de informação auxiliar pode ser fornecida através da utilização de atributos XML nos elementos XML. Um atributo XML possui um nome e um valor, e um elemento XML pode possuir vários atributos XML. Utilizando os atributos denominados **href** e **país**, a frase anterior poderia ser reescrita da seguinte maneira:

```
<frase>
  <vinho país="Portugal"
    href="www.ivp.pt">Port</vinho>
  <servido_com>
    <queijo país="Inglaterra"
      href="www.stiltoncheese.com">
      Stilton</queijo>
  </servido_com>
</frase>
```

Desta forma, foi acrescentada mais informação à frase original, apesar de não se ter utilizado qualquer tipo de termo previamente definido, pois, apesar do termo **href** ser igual a um dos termos utilizados pela linguagem HTML para definir uma localização na Web, na sintaxe XML não possui qualquer significado associado. Desta forma, através da utilização da tecnologia XML, qualquer pessoa pode criar e utilizar os seus próprios meta-dados para estruturar os seus dados. Cada vez mais, os dados estruturados e associados a meta-dados, apresentam um impacto positivo no enriquecimento dos resultados de pesquisa efectuados pela maioria das ferramentas web quando executam pesquisas na Web. Mas quando se pretende utilizar

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

meta-dados a um nível mais abrangente e alargado surge um problema. Neste exemplo, utilizei palavras comuns, como *frase*, *vinho*, *queijo* e *país*, para nomear as minhas anotações, mas, e se alguém utilizar estas mesmas palavras para anotar conceitos diferentes, ou utilizar diferentes palavras para anotar os mesmos conceitos. O problema agrava-se se consideramos que os SPC não conseguirão distinguir facilmente essas situações. Assim, a solução encontrada para evitar o aparecimento destas ambiguidades prende-se com a identificação de forma unívoca dos elementos e atributos XML utilizados. Nesta perspectiva, cada elemento e atributo XML pode ser considerado como um recurso. Desta forma, podemos facilmente identificar os recursos de forma unívoca através da atribuição de um URI a cada um deles<sup>148</sup>. A atribuição consistente de URIs é facilitada através da definição e utilização de **Espaços de Nomes**. Um Espaço de Nomes associa um URI a um conjunto de termos, normalmente designado por vocabulário, associado a uma área de conhecimento específica. A utilização de Espaços de Nomes possibilita que as pessoas possam facilmente criar as suas marcações e combiná-las de forma desambígua com marcações criadas por outras pessoas. Uma possível utilização de um Espaço de Nomes na frase apresentada neste exemplo resultaria da seguinte forma:

```
<frase xmlns:meu="www.di.ubi.pt/meu#">
  <meu:vinho meu:país="Portugal"
    meu:href="www.ivp.pt">Port</meu:vinho>
  <meu:servido_com>
    <meu:queijo meu:país="Inglaterra"
      meu:href="www.stiltoncheese.com">
      Stilton</meu:queijo>
    </meu:servido_com>
  </frase>
```

Na primeira linha do código é definido um Espaço de Nomes, que foi inventado para este exemplo, com o nome *meu* e que foi associado ao URI *www.di.ubi.pt/meu/*. Além de ter sido

---

<sup>148</sup>A atribuição pode ser directa, um novo URI, ou indirecta, através da referência a um URI que identifica um recurso relacionado com o recurso actual.

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

definido, também foi associado aos termos `vinho`, `href`, `servido_com`, `queijo` e `país`. Assim, e ainda no mesmo trecho de código, é utilizado o nome do Espaço de Nomes como prefixo para declarar os elementos e os atributos XML que pertencem a esse Espaço de Nomes, como por exemplo `meu:vinho`. Desta forma, o nome e o URI de um Espaço de Nomes, ao poderem ser definidos por qualquer utilizador, funcionam como um meio simples de eliminar a geração de possíveis ambiguidades durante o processo de atribuição de meta-dados a conceitos.

Não há dúvida de que é muito útil e poderoso o facto de qualquer pessoa poder criar e utilizar os seus próprios URIs nas suas frases ou asserções, mas, o ideal seria se os SPC conseguissem interpretar essas frases ou asserções. A arquitectura RDF apresenta-se como uma possível solução, pois permite expressar asserções que podem ser facilmente interpretadas<sup>149</sup> pelos SPC.

Uma asserção RDF é semelhante a uma simples asserção utilizada na linguagem do dia-a-dia, sendo a principal diferença entre elas o facto de todos os termos de uma asserção RDF serem definidos através de um URI. N-Triples [137] é a linguagem mais simples para escrever uma asserção em RDF. Na prática, ela apenas utiliza uma tripla de URIs ou um conjunto de triplas URIs, como por exemplo:

```
<Port> <served_with> <Stilton>
```

De seguida é apresentada a mesma asserção, mas desta vez utilizando os conceitos definidos de forma unívoca através da utilização de três URIs definidos através do Espaço de Nomes `www.di.ubi.pt/meu`:

```
<www.di.ubi.pt/meu#vinho/href/www.ivp.pt>
```

```
<www.di.ubi.pt/meu#servido_com>
```

```
<www.di.ubi.pt/meu#queijo/href/www.stiltoncheese.com>
```

---

<sup>149</sup>Pretende-se salientar que a interpretação de uma asserção por parte dos SPC não significa que consigam compreender, mas apenas que a conseguem interpretar de uma maneira lógica que dá a sensação que a conseguem compreender.

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

Cada asserção RDF é composta por três URIs: o sujeito, o predicado que relaciona o sujeito com o objecto, e o objecto. Neste exemplo, o sujeito é um URI que representa o vinho do Porto, `www.di.ubi.pt/meu#vinho/href/www.ivp.pt`, o predicado é um URI que representa o termo `servido_com`, `www.di.ubi.pt/meu#servido_com`, e o objecto é um URI que representa o queijo Stilton, `www.di.ubi.pt/meu#queijo/href/www.stiltoncheese.com`. Deste modo, a afirmação RDF deste exemplo exprime:

Vinho do Porto servido com Stilton.

Através da criação e utilização de asserções RDF, qualquer pessoa pode facilmente espalhar asserções interpretáveis por SPC pela Web. Um dos princípios da arquitectura RDF é o permitir a qualquer pessoa dizer o que desejar sobre qualquer assunto. No limite, as asserções em formato RDF até podem ser contraditórias, falsas ou incoerentes, mas esta é uma liberdade que a Web Semântica privilegia com o objectivo de simplificar a sua implementação e utilização por parte de todos os seus utilizadores.

A seguir é apresentada uma possível representação através da sintaxe RDF/XML da asserção RDF anteriormente referida no formado N-Triples:

```
<rdf:RDF
  xmlns:rdf="http://w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:meu="www.di.ubi.pt/meu#">
  <rdf:Description rdf:about=
    "www.di.ubi.pt/meu#vinho/href/www.ivp.pt">
    <meu:servido_com rdf:resource=
      "www.di.ubi.pt/meu#queijo/href/www.stiltoncheese.com"/>
  </rdf:Description>
</rdf:RDF>
```

Através da utilização de documentos em sintaxe RDF/XML, é muitíssimo mais simples e fácil de implementar SPC com capacidade de interpretação da informação. Segundo Aaron Swartz

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

[90], uma fonte rica em informação processável pelos SPC são as BD. Actualmente, a maioria da informação relevante é guardada em BD, o que faz com que existam milhões de BD espalhadas pelo mundo. As BD tradicionais são consideradas óptimas para o armazenamento de informação estruturada. Mas, pelo contrário, não apresentam uma forma clara e versátil para uma sustentação simples e interpretável da enorme quantidade de informação descentralizada e desorganizada que se encontra na Web. A sua principal limitação prende-se com a utilização de um número fixo e pré-determinado de termos admitidos, tornando-as desta forma incompatíveis com a universalidade da Web, na qual todos os dias são inventados e definidos novos termos. Por exemplo, vamos imaginar que a partir de hoje alguém implementava um novo sistema para classificar os artigos científicos numa escala de 1 a 10 em vez de os classificar como por exemplo de "Muito interessante", "Interessante", "Pouco Interessante" e "Desinteressante". Neste caso os SPC baseados no antigo sistema de classificação não seriam capazes de processar a nova informação. É neste campo que a arquitectura RDF apresenta vantagens, pois não é limitada a um número fixo e pré-determinado de termos. A arquitectura RDF não especifica o que os termos ou conceitos significam ou como devem ser usados, apenas permite representá-los através de um modelo e um sintaxe que é facilmente utilizada pelos humanos e interpretável pelos SPC. Desta forma, a arquitectura RDF começa, cada vez mais, a ser utilizada para converter a informação existente nas BD tradicionais em informação processável pelos SPC. Por outro lado, a simples utilização da arquitectura RDF e dos URIs não é suficiente para descrever o significado nem o objectivo da utilização dos recursos. Desta forma, existe a necessidade de conjugar a informação em formato RDF com entidades que descrevam o significado dos recursos, as suas propriedades e as suas inter-relações. A estas entidades dá-se o nome de Ontologias, isto é, estruturas que descrevem de uma forma universal conceitos, as suas propriedades, as suas relações e, também, por vezes, regras de inferência.

Segundo Michael Daconta *et.al.* [146], os humanos criam mentalmente, ao longo da sua vida, representações simbólicas estruturadas sobre os conceitos que os rodeiam, as suas propriedades e as suas inter-relações. Um humano ao ler um documento, apesar de ver apenas simples cadeias



#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

de símbolos, consegue interpretá-los de acordo com o significado que esses símbolos têm nos seus modelos mentais. Por exemplo ao encontrar o conceito **automóvel**, um humano consegue associá-lo a um modelo mental que especifica o que é um automóvel, que tipo de automóveis existem, que funções pode desempenhar, de que tipo de peças é constituído, etc, isto é, um conjunto alargado de informação relacionada com o conceito, e o domínio em que esse conceito se insere, que além de estruturada é extremamente rica em conhecimento. Desta forma, um SPC que aceda a documento web que possua a sua informação associada a uma ontologia, fornecerá uma melhor resposta a qualquer tipo de pesquisa que possa ser submetido, pois poderá processar com um mínimo de ambiguidades a informação existente na Web. Assim, será possível atribuir um significado universal à informação existente em documentos RDF/XML. A informação além de estruturada e facilmente interpretável pelos SPC passará a ter uma definição universal, o que contribuirá de forma extremamente útil para a materialização da Web Semântica.

Neste momento encontram-se em desenvolvimento diversas linguagens ontológicas, sendo a mais recente e importante a OWL. A OWL permite a especificação de conceitos básicos como sub-classe<sup>150</sup>, inversão<sup>151</sup>, etc. e de princípios lógicos que permitirão aos SPC raciocinar através de inferência. Desta forma, como as ontologias podem utilizar princípios lógicos, também podem ser utilizadas para provar a validade de asserções lógicas que possam ser escritas por qualquer pessoa. Apesar de ser difícil de criar provas automaticamente, será muito fácil validá-las automaticamente, pois serão facilmente interpretáveis pelos SPC.

Na Web Semântica, qualquer pessoa pode desenvolver uma ontologia, mas a sua utilização por terceiros estará dependente da confiança que eles depositarem nessa ontologia. Desta forma, começa a ganhar importância a acreditação de ontologias, sendo a DAML umas das organizações que mais contribui para essa acreditação, disponibilizando um sítio web [298] que permite

---

<sup>150</sup>Cada divisão de uma classe.

<sup>151</sup>Ação ou efeito de inverter, de pôr ao contrário.

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

o acesso a uma biblioteca de 282 ontologias<sup>152</sup>. Nessa biblioteca podemos encontrar ontologias sobre os mais variados assuntos, como o cancro<sup>153</sup> [300], a Encefalopatia Espongiforme Bovina (BSE)<sup>154</sup> [301], Armas de Destruição Maciça<sup>155</sup> [302], etc. De entre as várias ontologias OWL disponibilizadas na web, podemos, por exemplo, encontrar uma que descreve Vinhos de Mesa em [303] e uma que descreve Pratos Gastronómicos em [304]. Estas duas ontologias não passam de adaptações mais recentes de uma ontologia DAML [305] desenvolvida inicialmente por Deborah McGuinness.

Assim, continuando a adaptação da frase *Port is served with Stilton cheese* para o modelo da Web Semântica, pode-se, agora, proceder à associação de algum dos termos da frase a conceitos definidos nas ontologias Vinhos de Mesa<sup>156</sup> e Pratos Gastronómicos. Esta associação é extremamente valiosa, pois, além de associar os termos da frase a conceitos definidos dentro de um domínio, os SPC que acedam a esta associação terão também acesso aos modelos que descrevem os domínios em que esses conceitos se encontram definidos. Isto é, apesar da frase apenas referir o termo *Port*, a associação deste termo ao conceito *Port* existente na ontologia Vinhos de Mesa, permitirá a um SPC associar esse termo também aos seguintes conceitos: *RedWine*, *locatedIn:PortugalRegion*, *hasBody:Full*, *hasFlavor:Strong*, *hasSugar:Sweet*, *TaylorPort*, *Vintage*, *WineGrape*, *WhiteWine*, *PotableLiquid*, *hasMaker*, *yearValue*, etc. Um pequeno excerto da ontologia Vinhos de Mesa que se refer directamente ao Vinho do Porto é apresentado de seguida [303]:

---

<sup>152</sup>Existem vários sítios web que disponibilizam ontologias em OWL, como por exemplo o referido em [299].

<sup>153</sup>Doença provocada por uma reprodução descontrolada de células malignas, que são independentes de qualquer tipo de controlo pelo organismo, acabando por invadir tecidos e órgãos e provocar alterações orgânicas.

<sup>154</sup>Também conhecida como a doença das Vacas Loucas.

<sup>155</sup>Incluindo armas químicas, biológicas e radiológicas.

<sup>156</sup>Uma descrição sobre os principais conceitos definidos nesta ontologia pode ser encontrada em [69].

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

(...)

```
<owl:Class rdf:ID="Port">
  <rdfs:subClassOf rdf:resource="#RedWine" />
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#locatedIn" />
      <owl:hasValue rdf:resource="#PortugalRegion" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasBody" />
      <owl:hasValue rdf:resource="#Full" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasFlavor" />
      <owl:hasValue rdf:resource="#Strong" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasSugar" />
      <owl:hasValue rdf:resource="#Sweet" />
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>                                     (...)
```

#### 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

A partir do momento que uma ontologia OWL é disponibilizada na Web, torna-se uma potencial fonte de informação para qualquer SPC desenvolvido para a Web Semântica. Uma nova versão da frase *Port is served with Stilton cheese*, agora com o termos *Port* e *Stilton* associados a conceitos definidos numa ontologia, pode ser assim apresentada:

```
<rdf:RDF
  xmlns:rdf="http://w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:meu="www.di.ubi.pt/meu#">
  <rdf:Description rdf:about=
    "http://www.w3.org/TR/2003/PR-owl-guide-20031209/wine#Port">
    <meu:servido_com rdf:resource=
      "http://www.w3.org/TR/2003/PR-owl-guide-20031209/food#cheese">
      Stilton</meu:servido_com>
    </rdf:Description>
  </rdf:RDF>
```

Este extracto de código, além de ser um pouco complexo, não apresenta, por si só, muitas potencialidades. É verdade que alguns dos seus termos se encontram associados a conceitos definidos em ontologias, mas a sua potencialidade só será demonstrada quando o mesmo for processado e interpretado por um SPC desenvolvido para a Web Semântica. Desta forma, os SPC são uma das peças fundamentais para a materialização da Web Semântica e, desta forma, o seu desenvolvimento está substancialmente a crescer nos últimos anos. Dentro dos SPC actualmente desenvolvidos, existe um que possui como objectivo recomendar os melhores vinhos, que se encontrem disponibilizados na web, para acompanhar pratos gastronómicos específicos. Este SPC chama-se *KSL<sup>157</sup> Wine Agent* [306] e é um tipo de SPC que se interessará, seguramente, pela informação apresentada no anterior extracto de código. Este SPC acede a conceitos e asserções definidas em ontologias, como por exemplo as ontologias OWL Vinhos de Mesa [303] e Pratos Gastronómicos [304], e utiliza um Motor de Inferência chamado JTP [307] na dedução

---

<sup>157</sup> *Knowledge Systems Laboratory.*

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

de estruturas lógicas.

O KSL Wine Agent foi desenvolvido para a Web Semântica, encontra-se disponível na web [306] e é aconselhável utilizá-lo na seguinte situação [69]:

Há a necessidade de se planejar um jantar de cerimónia em que pelo menos um dos convidados é um grande conhecedor de vinhos.

Apesar do anfitrião não ser conhecedor de vinhos, pretende adquirir pela Web e servir os melhores vinhos que combinem com os pratos do cardápio que vai apresentar no jantar de cerimónia.

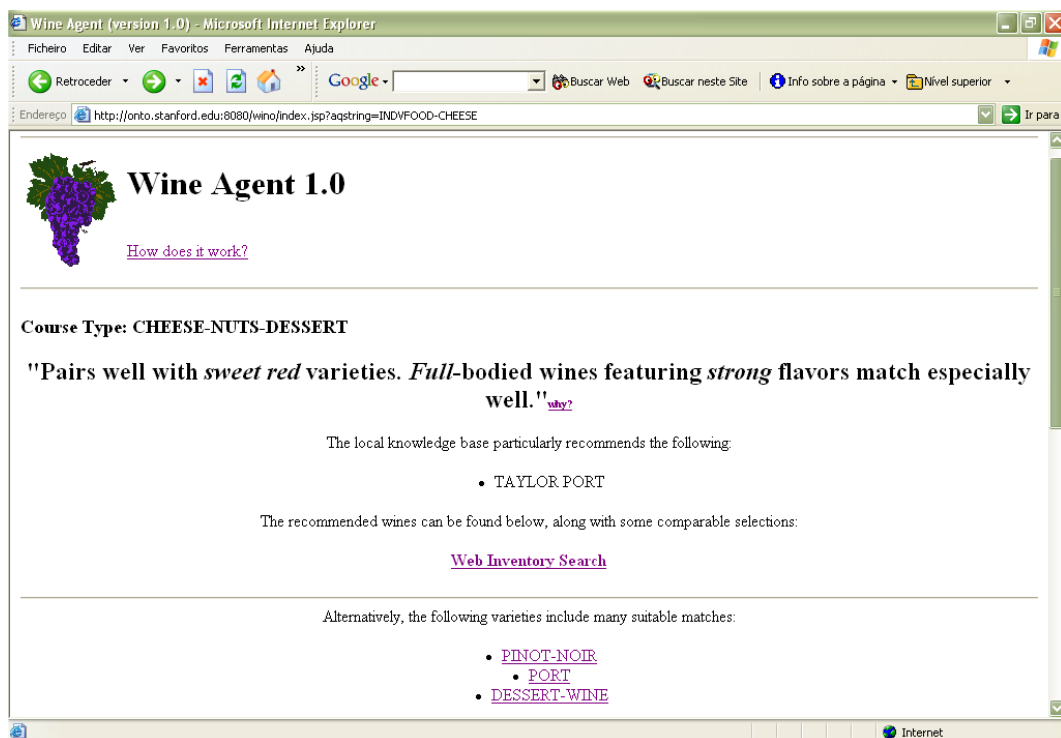


Figura 5.22: KSL Wine Agent: *Course Type: CHEESE-NUTS-DESSERT*.

A título de exemplo, utilizei o KSL Wine Agent para obter o nome do melhor vinho que deveria servir com queijo, e o resultado foi o apresentado na figura 5.22: TAYLOR PORT. Este SPC

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

permite, também, e se o utilizador o desejar, analisar a cadeia de inferência que utilizou para chegar a este resultado, conforme se ilustra na figura 5.23.

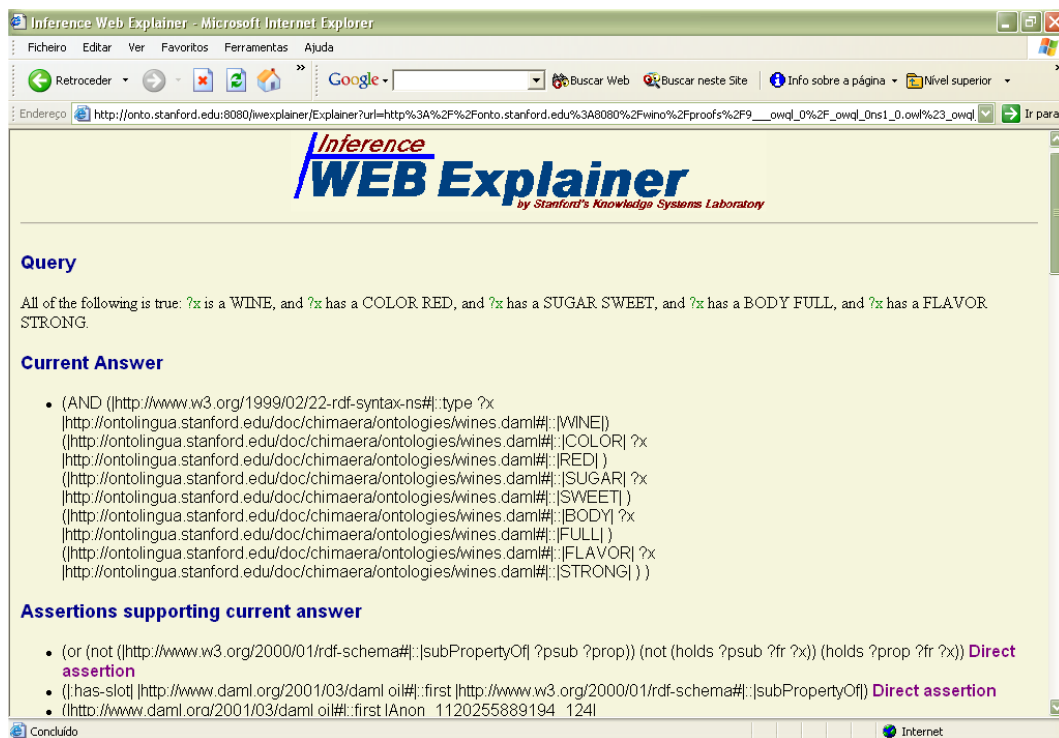


Figura 5.23: *Inference Web Explainer*. Course Type: *CHEESE-NUTS-DESSERT*.

Por fim, como na Web Semântica, qualquer pessoa poderá expressar qualquer comentário acerca de qualquer assunto<sup>158</sup>, a utilização de Assinaturas Digitais<sup>159</sup> associadas à informação será indispensável para a criação de informação de confiança<sup>160</sup>. Desta forma, será possível definir quais as assinaturas digitais que cada SPC deve confiar e qual o grau de confiança associado a cada uma delas. A interligação das políticas de confiança entre diversos SPC, permitirá criar uma Web de confiança. O utilizador final será livre de configurar o grau de confiança que considera aconselhável para a realização de qualquer tarefa em particular. Além disso, o histórico do

<sup>158</sup>Situação que também se verifica na actual Web.

<sup>159</sup>Processo de assinatura electrónica baseado em sistema criptográfico assimétrico composto de um algoritmo ou série de algoritmos, mediante o qual é gerado um par de chaves assimétricas exclusivas e interdependentes, uma das quais privada e outra pública, e que permite ao titular usar a chave privada para declarar a autoria de um documento, e ao destinatário utilizar a chave pública para verificar se a assinatura foi criada mediante o uso da correspondente chave privada e se o documento não foi alterado depois de ter sido assinado.

<sup>160</sup>Asserções RDF/XML assinadas digitalmente.

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

processamento efectuado por um SPC, ou uma cadeia de SPC, na obtenção de um resultado, poderá, também, ser inspeccionado por parte do utilizador.

### Utilização do editor SWedt

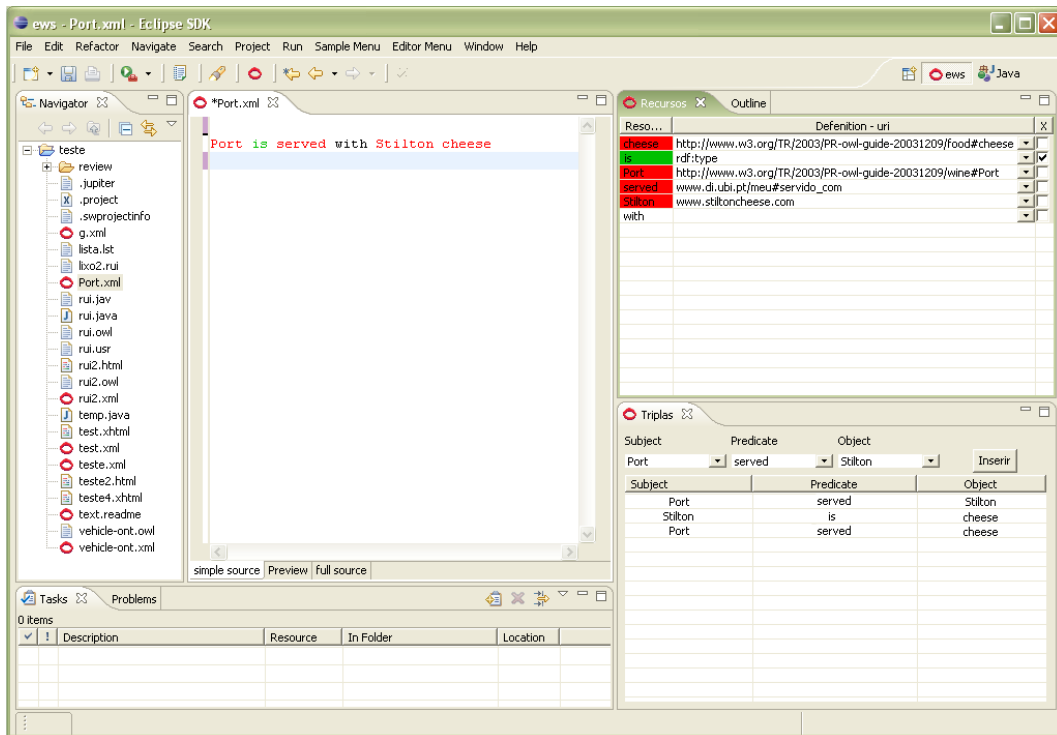


Figura 5.24: SWedt: Port is served with Stilton cheese.

Como vimos através da secção anterior, a criação de páginas para a Web Semântica engloba um grupo alargado de tecnologias e pode tornar-se uma tarefa extremamente complexa. Mas, como podemos ver através da utilização do editor SWedt, ilustrado na figura 5.24, essa tarefa fica extremamente simplificada. O utilizador da ferramenta necessita em primeiro lugar de escrever o seu texto no *simple editor*. No caso da figura, o texto resume-se a *Port is served with Stilton cheese*. À medida que o utilizador vai inserindo o texto a ferramenta apresenta uma lista, na vista *Resources*, com todas as palavras existentes no *simple editor*. Além disso, a ferramenta percorre a sua lista interna de ontologias e verifica se possuem conceitos relacionados com os termos existentes no *simple editor*. Em caso afirmativo, a ferramenta apresenta os seus URIs na lista da vista *Resources*. Em qualquer momento, o utilizador pode confirmar as associações

## 5.4. MODO E FUNCIONAMENTO DA FERRAMENTA SWEDT

---

existentes nessa lista. Por fim, o utilizador pode criar associações triplas entre os recursos existentes na lista da vista **Resources** na vista **Assertions**. No caso das associações apresentadas na figura 5.24 editor SWedt atribuiu ao ficheiro **Port.xml** o seguinte extracto de código:

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:meu="www.di.ubi.pt/meu#">
  <rdf:Description rdf:about=
    "http://www.w3.org/TR/2003/PR-owl-guide-20031209/wine#Port">
    <meu:servido_com rdf:resource=
      "http://www.w3.org/TR/2003/PR-owl-guide-20031209/food#cheese">
      Stilton</meu:servido_com>
  </rdf:Description>
  <rdf:Description
    rdf:about="www.di.ubi.pt/meu#queijo/href/www.stiltoncheese.com">
    <rdf:type rdf:resource=
      "http://www.w3.org/TR/2003/PR-owl-guide-20031209/food#cheese"/>
  </rdf:Description>
</rdf:RDF>
```

Assim, como se pode ver por este simples exemplo, o editor SWedt simplifica o processo de criação de informação para a Web Semântica de tal forma que os seus utilizadores não necessitam de conhecer e dominar os conceitos chaves da Web Semântica: os URI, as ontologias e as Asserções RDF. Um SPC preparado para a Web Semântica que aceda ao ficheiro **Port.xml** já não se limitará a receber uma simples cadeia de símbolos, mas sim, uma quantidade significativa de informação relacionada com Vinho do Porto e queijos, como por exemplo, que o Vinho



do Porto possui diferentes variedades: Ruby, Tawny, Aged Tawny, Vintage, etc., que o próprio autor da frase original pode desconhecer por completo. Além disso, esse SPC poderá possuir a característica de gerar novo conhecimento. Em suma, bem-vindo à Web Semântica.

## 5.5 A Ferramenta SMORE

Tanto quanto era do conhecimento do grupo de investigação onde o trabalho conducente à presente dissertação foi realizado, à data de início dos trabalhos conducentes à elaboração da presente dissertação, não havia qualquer ferramenta para a Web Semântica com as características do SWedt. Contudo, em Agosto de 2005 surgiu a quinta versão da ferramenta SMORE, a qual apresenta características que a assemelham ao editor SWedt. A ferramenta SMORE, apresentado na figura 5.25, permite a criação de documentos HTML associados a ontologias OWL e os seus principais objectivos são [227]:

- Fornecer aos utilizadores da Web um ambiente flexível, onde possam criar as suas páginas web de uma forma simples;
- Facilitar a classificação de informação através da utilização da sintaxe RDF (asserções RDF);
- Facilitar a possibilidade de associar os conceitos definidos nas asserções RDF com conceitos de ontologias existentes na Web;
- Facilitar a criação de novas ontologias a partir de ontologias já existentes;
- Fornecer um meio eficiente e seguro de manipular informação no formato RDF.

Por sua vez, as principais características da ferramenta SMORE são [227]:

- Centrada em OWL: Enquanto as versões anteriores da ferramenta SMORE eram centradas em RDF, a quinta versão foi desenvolvida para permitir a fácil criação de ontologias

## 5.5. A FERRAMENTA SMORE

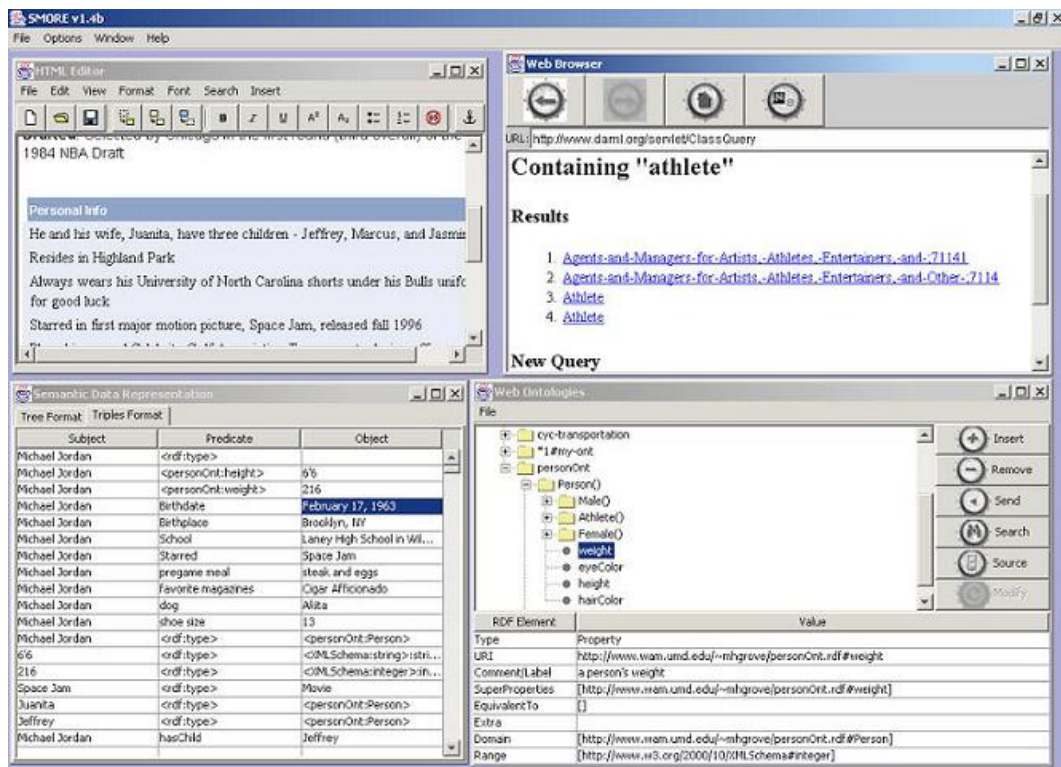


Figura 5.25: A ferramenta SMORE.

- OWL a partir de informação existente em páginas Web. A ferramenta permite importar ontologias já existentes para o interior de uma nova ontologia;
- Criação simples de entidades OWL: O utilizador apenas necessita de seleccionar uma parcela de texto localizada numa página Web ou num documento e clicar no botão correspondente da barra de ferramentas para criar Classes OWL, Propriedades ou Indivíduos;
  - Edição de Triplas OWL: Para os utilizadores familiarizados com a criação de triplas RDF, a ferramenta possibilita a utilização de entidades OWL em triplas;
  - Editor Inteligente de Indivíduos: O editor de Indivíduos utiliza as restrições de domínio e de alcance para apresentar uma lista de alvos elegíveis para todas as propriedades associadas a um Indivíduo;
  - Navegador SWOOP: A ferramenta integra o Navegador de Ontologias SWOOP [229], fornecendo assim uma simples e consistente maneira para procurar e aceder a classes e propriedades em Ontologias;

## 5.5. A FERRAMENTA SMORE

---

- **Prevenção de Erros:** A ferramenta identifica a existência de triplas inválidas na tabela de triplas e alerta o utilizador para corrigir os erros;
- **Editor HTML:** A ferramenta integra o Editor HTML, de nome Ekit [308], o qual permite aos utilizadores efectuarem edições simples nos documentos HTML.

Os principais objectivos da ferramenta SMORE e da ferramenta SWedt são semelhantes. Ambas permitem a criação e a utilização de asserções RDF associadas a ontologias, mas, contudo, apresentam algumas diferenças que as distinguem significativamente. De seguida serão apresentadas as principais vantagens apresentadas por cada uma destas ferramentas sobre a outra.

As principais vantagens da ferramenta SMORE sobre a ferramenta SWedt são:

- **Robustez:** Possui maior robustez em termos de funcionamento. Característica resultante de a ferramenta já se encontrar na sua quinta versão<sup>161</sup>, enquanto a ferramenta SWedt se encontra na sua primeira versão;
- **Receptibilidade:** Possui uma grande receptibilidade por parte dos utilizadores. Situação que pode ser atribuída quer à sua fiabilidade quer ao facto de já se encontrar há quatro anos disponível na Web. Mas, por outro lado, os seus utilizadores necessitam de possuir conhecimentos mínimos sobre as sintaxes XML, RDF, RDFS e OWL e de introduzir a localização das ontologias que necessitem de utilizar<sup>162</sup>.

Por outro lado, as principais vantagens da ferramenta SWedt sobre a ferramenta SMORE são:

- **Simplicidade:** Apresenta uma abordagem muito simples para o desenvolvimento de páginas de acordo com os princípios da Web Semântica, permitindo estabelecer uma ponte entre os investigadores da Web Semântica e os utilizadores da Web;

---

<sup>161</sup>A quinta versão foi lançada a 4 de Agosto de 2005. Sendo a primeira versão de 2002.

<sup>162</sup>Situações que não se verificam com a ferramenta SWedt.

- Integração: Encontra-se integrada numa plataforma aberta que apresenta um GUI e um sistema de manipulação de documentos partilhado com milhares de outras ferramentas, potencializando assim, uma fácil identificação e integração por parte do utilizador com centenas de outras ferramentas;
- Alcance: Permite estabelecer uma ponte entre os actuais investigadores da Web Semântica e os actuais utilizadores da Web.

## 5.6 Resumo

Neste capítulo foi descrita a arquitectura e a implementação da ferramenta SWedt. A especificação e o desenvolvimento da ferramenta SWedt constituiu um dos principais objectivos da presente dissertação.

A actual investigação sobre a Web Semântica prende-se sobretudo com o desenvolvimento de ferramentas especializadas e normalmente associadas a uma das suas camadas da arquitectura. Esta tendência é evidenciada sobretudo pelo desenvolvimento de uma variedade de sistemas e estruturas de *software* informático visando prover editores e integradores de tecnologias presentes nestas camadas. Mas o verdadeiro poder da Web Semântica só será perceptível quando representar um benefício para os seus utilizadores. De facto, actualmente, ainda não existem ferramentas que permitam suportar uma ponte estável entre os investigadores da Web Semântica e os utilizadores da Web. Desta forma, é indispensável desenvolver ferramentas e SPC especializados no apoio aos utilizadores de forma a facilitarem o seu envolvimento e a despertarem o seu interesse pela Web Semântica. Contudo, não deverá ser esquecido que o desenvolvimento destas ferramentas terá de levar em consideração a forma como os utilizadores actualmente interactivam com a Web e, assim, essa interacção deverá ser alterada o mínimo possível.

## 5.6. RESUMO

---

Tendo em conta esta última preocupação e o facto de que uma tecnologia só vir a ser efectivamente acolhida pelos utilizadores se se mostrar útil e simples de utilizar, estabeleceu-se como um dos objectivos desta dissertação, o desenvolvimento de um editor que integrasse as principais tecnologias associadas à arquitectura estratificada da Web Semântica por parte dos actuais utilizadores da Web, e principalmente por aqueles que não estejam familiarizados com essa arquitectura. Por isso, foi desenvolvida uma ferramenta na forma de um *plug-in*, designada por SWedt, que pode ser integrada na plataforma *Eclipse*. A plataforma *Eclipse*, além de ser desenvolvida em Código Aberto, permite, de forma simples e transparente, a integração da ferramenta SWedt com todas as outras ferramentas desenvolvidas, ou que se encontrem em desenvolvimento, para esta plataforma.

A ferramenta SWedt tem como principal objectivo facilitar a criação de páginas para a Web Semântica de uma forma simples e de despertar o interesse da comunidade Web pela Web Semântica. Ela identifica ficheiros com extensão XML como sendo ficheiros para a Web Semântica e apresenta três fases principais para a manipulação deste tipo de ficheiros. O funcionamento da ferramenta é simples, e transparente, na medida em que não existe a necessidade por parte do utilizador de compreender ou manipular directamente os conceitos chaves da Web Semântica, como por exemplo: os URI, as Ontologias e as Asserções RDF.

De entre as principais ferramentas, actualmente desenvolvidas para a Web Semântica, a ferramenta SMORE é a que mais se aproxima da ideia explorada pela ferramenta SWedt. Apesar de a ferramenta SMORE também permitir a criação e utilização de asserções RDF associadas a ontologias, ela, ao contrário da ferramenta SWedt, impõem, necessariamente, que os utilizadores possuam conhecimentos mínimos sobre a sintaxe RDF e que forneçam, obrigatoriamente, a localização das ontologias que necessitem de utilizar. Outras duas vantagens muito importantes da ferramenta SWedt relativamente à ferramenta SMORE prendem-se com o facto de, por um lado, apresentar uma abordagem mais simples para o desenvolvimento de páginas com conteúdos para a Web Semântica, e por outro, de estar integrada numa plataforma que pos-

## 5.6. RESUMO

---

sibilita a sua integração com centenas de outras ferramentas através da utilização do mesmo interface e do mesmo sistema de interacção com os ficheiros.

## Capítulo 6

# Conclusões e Perspectivas de Trabalho Futuro

"Is this rocket science? Well, not really. The Semantic Web, like the World Wide Web, is just taking well established ideas, and making them work interoperability over the Internet. This is done with standards, which is what the World Wide Web Consortium is all about. We are not inventing relational models for data, or query systems or rule-based systems. We are just webizing them. We are just allowing them to work together in a decentralized system - without a human having to custom handcraft every connection."

Tim Berners-Lee

"Business Case for the Semantic Web", Outubro 2001 [309]

### 6.1 Conclusões

A ideia essencial da Web Semântica consiste em aproveitar ao máximo os amplos recursos da actual Web: fundir os milhões de conceitos independentes que abrange numa única e extensa rede de informação auto-descrita que possa ser facilmente manuseada, quer pelas pessoas quer pelos SPC, e simplificar as interacções através do aumento da automação de tarefas. Ela absorveu, alguma inspiração por parte dos actuais modelos de IR e de KR mas, ao contrário deles, apresenta um modelo descentralizado para definir, extrair e contextualizar conceitos pertencentes a várias áreas do conhecimento, de forma universal, desambigua e facilmente interpretável por SPC. Assim, ela encoraja a livre utilização e definição de ontologias por parte de qualquer entidade, pertencendo aos utilizadores finais, pessoas ou SPC, a decisão de confiar ou não nessas ontologias.

A integral definição e implementação da Web Semântica ainda não é uma realidade. A arquitectura estratificada, apresentada no terceiro capítulo desta dissertação, tem de ser encarada como um protótipo para uma futura Web e não como uma solução final. Os actuais investigadores da Web Semântica, devem evitar realizar previsões muito optimistas, uma vez que, poderão revelar-se a curto prazo como erradas, como aconteceu com os primeiros investigadores da IA nas décadas de 50 e 60. Assim, é importante que os investigadores mantenham sempre os pés bem assentes no chão para não levantarem falsas expectativas por parte dos futuros utilizadores. Além disso, devem tomar em consideração que os seres humanos manifestam alguma resistência à mudança e às novas ideias. Por outro lado, a Web Semântica necessita de criar o Efeito em Rede para se tornar uma realidade. Desta forma, é importante referir que uma única página Web Semântica não possui qualquer valor por si só, sendo que a conjugação e integração dinâmica de várias páginas Web Semânticas, ontologias e SPC é que constituirá a verdadeira potencialidade da Web Semântica. Outra característica importante da Web Semântica, resulta do facto de ela transbordar do mundo virtual e abraçar o mundo físico. Nesta nova perspectiva os URIs poderão referenciar qualquer recurso, incluindo entidades físicas, o que permitirá



## 6.1. CONCLUSÕES

---

utilizar a arquitectura RDF e as ontologias para descrever dispositivos como telemóveis, televisores, rádios, frigoríficos, automóveis, etc. Assim, esses dispositivos poderão anunciar as suas funcionalidades, isto é, o que podem fazer e como podem ser controlados, aos SPC através da utilização de uma representação universal, a Web Semântica. Sendo muito mais flexível do que os actuais esquemas de baixo-nível como o *Universal Plug and Play*, este tipo de aproximação semântica abre, também, um novo mundo de potencialidades.

A actual investigação sobre a Web Semântica caracteriza-se sobretudo pelo desenvolvimento de ferramentas especializadas associadas a algumas das camadas da sua arquitectura estratificada de especificações. Esta tendência é evidenciada sobretudo pelo desenvolvimento de uma variedade de sistemas e estruturas de *software* informático que visam prover editores e integradores de tecnologias presentes a essas camadas. Mas o verdadeiro potencial da Web Semântica só será perceptível quando representar um benefício evidente para os seus utilizadores. De facto, actualmente, ainda não existem ferramentas que permitam suportar uma ponte estável entre os investigadores da Web Semântica e os utilizadores da Web. Efectivamente, a principal conclusão que se pode retirar do estudo de categorização das actuais ferramentas para a Web Semântica, apresentado no quarto capítulo desta dissertação, é a de que apesar de existir um grupo alargado de ferramentas para a Web Semântica elas são, na sua maioria, desenvolvidas com o objectivo de apoiar os actuais investigadores da Web Semântica. Assim, existe uma lacuna no desenvolvimento de ferramentas que forneçam a possibilidade de criar, de uma forma bastante simples, páginas web para a Web Semântica, por parte de utilizadores não familiarizados com os seus conceitos e a sua arquitectura. Desta forma, é indispensável desenvolver ferramentas e SPC especializados para o apoio dos utilizadores, de forma a facilitar o seu envolvimento e a despertar o seu interesse pela Web Semântica. Tendo em conta a lacuna apresentada nos anteriores parágrafos e o facto de que uma tecnologia ou ferramenta só ser efectivamente acolhida pelos utilizadores se ela se mostrar útil e simples de utilizar, foi durante a presente dissertação, definida uma arquitectura para uma nova ferramenta, denominada SWedt, para a Web Semântica e, por sua vez, implementada através do desenvolvimento de um *plug-in* para a plataforma

## 6.1. CONCLUSÕES

---

*Eclipse*. A nova ferramenta foi implementada para a plataforma *Eclipse* porque, além de ser desenvolvida em Código Aberto, permite de forma simples e transparente, a sua integração com centenas de outras ferramentas desenvolvidas para a mesma plataforma. A ferramenta SWedt identifica ficheiros com extensão XML como sendo ficheiros para a Web Semântica e apresenta três fases principais para a manipulação deste tipo de ficheiros. A primeira fase corresponde à edição do texto por parte do utilizador, a segunda fase corresponde à associação automática das palavras do texto editado pelo utilizador com conceitos definidos em ontologias, por fim, a terceira fase corresponde à criação de associações entre os conceitos (asserções RDF) por parte do utilizador. O funcionamento da ferramenta é simples e transparente, na medida em que não existe a necessidade por parte do utilizador de compreender ou manipular directamente os conceitos chaves da arquitectura da Web Semântica, como por exemplo: os URI, as ontologias e as asserções RDF.

Apesar de, a Web Semântica, ter sido durante algum tempo considerada como sendo uma obra de ficção ela, cada vez mais, encontra-se mais próxima de se tornar uma realidade. Acima de tudo, ela pode ser considerada como um enorme desafio, como o foi em tempos a implementação da Web. Do ponto de vista do autor da presente dissertação, a materialização da Web Semântica contribuirá de forma positiva para uma nova ordem na disposição de informação na Web podendo, também, ser definida como um passo determinante na aproximação dos sistemas informáticos ao Homem e na evolução do próprio conhecimento humano. O autor desta dissertação acredita que, apesar de a ideologia da Web Semântica ainda não ter recebido a atenção merecida por parte dos actuais investigadores e utilizadores da Web, estará, dentro de poucos anos, de tal maneira incutida no nosso modo de pensar que irá parecer como uma ideia inata, da mesma maneira como actualmente acontece com a ideia de Internet e de Web.

Através da presente dissertação, o autor pretendeu apresentar uma descrição geral do estado da arte sobre as tecnologias disponíveis para a Web Semântica, realizar um trabalho de avaliação e categorização acerca das principais ferramentas a ela associadas e desenvolver uma ferramenta

## 6.2. SUGESTÕES PARA TRABALHO FUTURO

---

que permitisse a criação de páginas web, de acordo com os princípios da Web Semântica, por parte de utilizadores não familiarizados com a sua arquitectura. Como conclusão do trabalho realizado, é com muita satisfação, que o autor assevera ter a convicção de que foram atingidos com sucesso todos os objectivos previamente estabelecidos e que a realização desta dissertação demonstrou-se muito gratificante e enriquecedora, principalmente na conciliação dos diversos conceitos ligados ao mundo da Web em geral e da Web Semântica em particular. Desta forma, o autor augura que muito em breve teremos uma Web diferente, uma Web que terá como base a Web Semântica. Além disso, e na perspectiva do autor, é imprescindível que os engenheiros informáticos e, em geral, os utilizadores da Web se empenhem em estabelecer um contacto urgente com esta nova área e que ela deveria, indubitavelmente, ser incluída como matéria indispensável nos actuais planos curriculares dos cursos de Engenharia Informática. Por fim, o autor espera que a contribuição resultante da elaboração desta dissertação e da disponibilização da ferramenta desenvolvida, venha a ser útil para todos quantos dela possam vir a retirar algo de proveitoso para a sua aprendizagem.

## 6.2 Sugestões para Trabalho Futuro

O desenvolvimento de uma ferramenta para a Web Semântica numa altura em que a sua arquitectura ainda se encontra em fase de investigação, é um desafio em constante transformação e evolução. A ferramenta desenvolvida encontra-se na sua versão inicial, mas como foi desenvolvida na forma de um *Plug-in* para a plataforma *Eclipse*, a inserção de novas funcionalidades bem como a extensão das suas funcionalidades, poderá ser facilmente conseguida. Desta forma, esta versão pode ser considerada como a implementação de uma base robusta para o desenvolvimento de novas e mais poderosas capacidades a inserir em futuras versões desta ferramenta. De qualquer forma, todos os melhoramentos que se venham a implementar na actual ferramenta deverão obedecer à sua principal característica: a criação de páginas web por parte de utilizadores não familiarizados com a arquitectura da Web Semântica.

A seguir é apresentada uma lista de possíveis melhoramentos a integrar numa futura versão da ferramenta SWedt:

- Desenvolver uma funcionalidade que permita assinar digitalmente as asserções RDF desenvolvidas durante a utilização da ferramenta. Esta funcionalidade deverá ser muito simples de utilizar e o mais transparente possível para o utilizador. Ela poderá ser realizada através da implementação de um mecanismo que associe uma Assinatura Digital XML a cada um dos utilizadores da ferramenta;
- Desenvolver uma funcionalidade que possibilite a actualização automática da lista interna de ontologias. Esta funcionalidade poderá ser implementada através da criação e constante actualização de um ficheiro, preferencialmente em notação XML, que possa ser facilmente armazenado numa localização pré-definida na Web e que possuiria uma lista com a localização de todas as ontologias conhecidas. Desta forma, a ferramenta acederia a esse ficheiro e automaticamente actualizaria a sua lista interna de ontologias;
- Aperfeiçoar o Editor Simples da ferramenta. Esta funcionalidade poderá ser implementada através da inclusão de mecanismos que facilitem a inclusão de elementos gráficos, a sua detecção, formatação e validação no Editor Simples. Mas sempre, com o cuidado de não colocar em causa a simplicidade de utilização do editor;
- Implementar uma versão bilingue: Português e Inglês.

# Referências

1. Berners-Lee, Tim & Hendler, James & Lassila, Ora., "*The Semantic Web*", Scientific American. Maio 2001. Informação disponível no endereço da Web: <http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2> (última visita a 10 de Junho de 2005)
2. *W3C Semantic Web Activity*. Informação disponível no endereço da Web: <http://www.w3.org/2001/sw/> (última visita a 10 de Junho de 2005).
3. *Extensible Markup Language XML 1.0* (Terceira versão). *W3C Recommendation* de 4 de Fevereiro de 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/2004/REC-xml-20040204/> (última visita a 15 de Setembro de 2004).
4. *RDF/XML Syntax Specification (Revised)*. *W3C Recommendation* de 10 de Fevereiro de 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/rdf-syntax-grammar/> (última visita a 15 de Setembro de 2004).
5. *Web Ontology Language (OWL)*. Informação disponível no endereço da Web: <http://www.w3.org/TR/owl-ref/> (última visita a 11 de Junho de 2004).
6. Pereira, Rui G. & Freire, Mário M., "Classification of Semantic Web Technologies", in *Encyclopedia of Information Science and Technology*, 2nd Edition, Mehdi Khosrow-Pour (Ed.), Idea Group, Inc., 2006, accepted for publication.
7. Ferramenta *RDF Editor* (RDFedt). Informação disponível no endereço da Web:

- [http://atlas.ucpel.tche.br/~barbosa/compiladores/t1/rdf/editor\\_rdfedt.htm](http://atlas.ucpel.tche.br/~barbosa/compiladores/t1/rdf/editor_rdfedt.htm) (última visita a 10 de Junho de 2005).
8. Ferramenta *RDF Instance Creator* (RIC). Informação disponível no endereço da Web: <http://www.mindswap.org/~mhgrove/RIC/RIC.shtml> (última visita a 10 de Junho de 2005).
  9. Ferramenta Protégé. Informação disponível no endereço da Web: <http://protege.stanford.edu/> (última visita a 10 de Junho de 2005).
  10. Ferramenta OIEd. Informação disponível no endereço da Web: <http://oiled.man.ac.uk/> (última visita a 10 de Junho de 2005).
  11. Ferramenta OntoEdit. Informação disponível no endereço da Web: <http://www.ontoknowledge.org/tools/ontoedit.shtml> (última visita a 10 de Junho de 2005).
  12. Pereira, Rui. G. & Freire, Mário M., "*Semantic Web*", in *Encyclopedia of Multimedia Technology and Networking*, Margherita Pagani (Ed.), Idea Group Reference, ISBN: 1-59140-561-0, Vol. II, pp. 917-924, 2005.
  13. Pereira, R. & Freire, M., "Integration of Ontologies and Semantic Annotations with Resource Description Framework in Eclipse-based Platforms with Editing Features for Semantic Web", in CD ROM Proceedings of The International Conference on Information Networking (ICOIN'2006), Sendai, Japan, January 16, 2006, 10 pages.
  14. Pereira, R. & Freire, M., "Integration of Ontologies and Semantic Annotations with Resource Description Framework in Eclipse-based Platforms with Editing Features for Semantic Web", in *Information Networking*, Ilyoung Chong and Kenji Kawahara (Eds.), *Lecture Notes in Computer Science*, LNCS 3961, Revised Selected Papers of ICOIN 2006, Springer-Verlag, Berlin Heidelberg, 2006, accepted for publication.

15. Pereira, R. & Freire, M., "SWedt: A Semantic Web Editor Integrating Ontologies and Semantic Annotations with Resource Description Framework", in CD ROM Proceedings of International Conference on Internet and Web Applications and Services (ICIW 2006), Gosier, Guadeloupe, French Caribbean, February 22-24, 2006, IEEE Computer Society Press, Los Alamitos CA, February 2006, ISBN-10 0-7695-2522-9, 6 pages
16. Drucker, Peter., "*Beyond the Information Revolution*". Publicado na revista *The Atlantic Monthly* 284(4):47-57. Outubro de 1999.
17. Quarterman, John S. & Carl-Mitchell, Smoot., "*What is the Internet, Anyway?*". Matrix News, Agosto de 1994. Informação disponível em <http://ietf.org/rfc/rfc1935> (última visita a 16 de Agosto de 2005).
18. Tremel, Daniel., "*Todo o conhecimento do mundo*". Publicado na Revista Galileu, São Paulo, Globo, n. 164, p. 36-39, Março de 2005. Informação disponível em <http://revistagalileu.globo.com/Galileu/0,6993,ECT917065-1719-1,00.html> (última visita a 16 de Agosto de 2005).
19. Batstone, David., "*BigBrother.com: The Net War on Privacy*". Publicado na Revista Inc.com, em Fevereiro de 2000. Informação disponível em <http://www.inc.com/articles/2000/02/17442.html> (última visita a 16 de Agosto de 2005).
20. Tak-ho, Fong., "*China targets media's 'evil trend'*". Asia Times Online. Dezembro, 2004. Informação disponível em <http://www.atimes.com/atimes/China/FL08Ad06.html> (última visita a 16 de Agosto de 2005).
21. Lucier, Thomas., "*How To Use The Internet To Perform Due Diligence On Real Property*". 2005. Informação disponível em <http://www.thomaslucier.com/duediligence.html> (última visita a 16 de Agosto de 2005).
22. Williams, Phil., "*Organized Crime and Cybercrime: Synergies, Trends,*

- and Responses*". International Information Programs, U.S. Department of State. Agosto, 2001, Volume 6, Number 2. Informação disponível em <http://usinfo.state.gov/journals/itgic/0801/ijge/gj07.htm> (última visita a 16 de Agosto de 2005).
23. Dicionário *On-line* de Língua Portuguesa da Porto Editora. Informação disponível no endereço da Web: <http://www.portoeditora.pt/dol/> (última visita a 10 de Junho de 2005).
  24. *Federal Networking Council* (FNC). Definição do termo "Internet". Informação disponível no endereço da Web: [http://www.nitrd.gov/fnc/Internet\\_res.html](http://www.nitrd.gov/fnc/Internet_res.html) (última visita a 16 de Junho de 2005).
  25. *Large Scale Networking* (LSN). Informação disponível no endereço da Web: <http://www.nitrd.gov/subcommittee/lsn.html> (última visita a 18 de Junho de 2005)
  26. Elmer, Greg., *"The Economy of Cyberpromotion: Awards on the World Wide Web"*. A. Herman & T. Swiss (Eds.), *"The World Wide Web and contemporary cultural theory"*. Londres. Routledge. Novembro 1998.
  27. Wagner, Jim., *"Father of the Web Wins Millennium Award"*. Internetnews.com. Abril 2004. Informação disponível no endereço da Web: <http://www.internetnews.com/dev-news/article.php/3341741>(última visita a 19 de Junho de 2005).
  28. *Internet World Stats*. Informação disponível no endereço da Web: <http://www.internetworldstats.com/stats.htm> (última visita - 6 de Fevereiro de 2006)
  29. Voiskounsky, Alexander., *"Internet: Diversity or Unification?"*. 48th Annual Conference of the INTERNATIONAL COMMUNICATION ASSOCIATION. Jerusalem, Israel. Julho 1998. Informação disponível no endereço da Web: <http://www.ifi.uib.no/staff/konrad/research/culture/DalexApplicatICA98ica> paper.htm(última visita a 19 de Junho de 2005).



30. Comer, Douglas., *"The Internet Book: Everything You Need to Know About Computer Networking and How the Internet Works"*. 3ª edição. 2000. Prentice Hall (ISBN 0-13-030852-8).
31. Gralla, Preston., *"How The Internet Works"*. 7ª edição. 2003. QUE (ISBN 0-7897-2973-3).
32. Bird, Linda., *"Complete Guide to Using and Understanding the Internet"*. 2003. Prentice Hall (ISBN 0-13-140289-7).
33. Rosenfeld, Louis & Morville, Peter., *"Information Architecture for the World Wide Web: Designing Large-Scale Web Sites"*. 2ª edição. 2002. O'Reilly (ISBN 0596000359).
34. Licklider, J., *"Man-Computer Symbiosis"*. IRE Transactions on Human Factors in Electronics, volume HFE-1, Março 1960. Informação disponível no endereço da Web: <ftp://gatekeeper.research.compaq.com/pub/DEC/SRC/research-reports/SRC-061.pdf> (última visita a 19 de Junho de 2005).
35. Kleinrock, Leonard., *"Information Flow in Large Communication Nets"*. Proposta de Tese de Doutoramento - MIT. Julho 1961. Informação disponível no endereço da Web: <http://www.lk.cs.ucla.edu/LK/Bib/REPORT/PhD/proposal.html> (última visita a 19 de Junho de 2005).
36. Baran, Paul., *"On Distributed Communications Networks"*. IEEE Transactions on Communications, 1964. Informação disponível no endereço da Web: <http://www.rand.org/publications/RM/baran.list.html> (última visita a 19 de Junho de 2005).
37. Berners-Lee, Tim., *"Information Management: A Proposal"*. CERN. Informação disponível no endereço da Web: <http://www.w3.org/History/1989/proposal.html> (última visita a 19 de Junho de 2005).
38. *European Center of Nuclear Research / Laboratório Europeu de Física de Partículas (CERN)*. Informação disponível no endereço da Web: <http://public.web.cern.ch/public>

(última visita - 10 de Junho de 2005).

39. Connolly, Dan., *"A Little History of the World Wide Web"* W3C (MIT, ERCIM, Keio). 2000. Informação disponível no endereço da Web: <http://www.w3.org/History.html> (última visita a 10 de Junho de 2005).
40. Navegador Web denominado Internet Explorer. Informação disponível no endereço da Web: <http://www.microsoft.com/windows/ie/default.msp> (última visita a 10 de Junho de 2005).
41. Navegador Web denominado Netscape Navigator. Informação disponível no endereço da Web: <http://browser.netscape.com/ns8/> (última visita a 10 de Junho de 2005).
42. Navegador Web denominado Opera. Informação disponível no endereço da Web: <http://www.opera.com/> (última visita a 10 de Junho de 2005).
43. Navegador Web denominado Mozilla Firefox. Informação disponível no endereço da Web: <http://www.mozilla.org/> (última visita a 10 de Junho de 2005).
44. Navegador Web denominado Lynx. Informação disponível no endereço da Web: <http://lynx.browser.org/> (última visita a 10 de Junho de 2005).
45. Navegador Web denominado Mosaic, desenvolvido por Marc Andreessen. Informação disponível no endereço da Web: <http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html> (última visita a 10 de Junho de 2005).
46. *National Center for Supercomputing Applications* (NCSA). Informação disponível no endereço da Web: <http://www.ncsa.uiuc.edu/> (última visita a 10 de Junho de 2005).
47. Netscape. Informação disponível no endereço da Web: <http://home.netscape.com/main.adp?fc=1> (última visita a 10 de Junho de 2005).

48. Microsoft. Informação disponível no endereço da Web: <http://www.microsoft.com/> (última visita a 10 de Junho de 2005).
49. Naughton, John., "*Darth Vader meets his Match?*", London Observer. Novembro 1998. Informação disponível no endereço da Web: <http://molly.open.ac.uk/Personal-pages/Pubs/981108.htm> (última visita a 10 de Junho de 2005).
50. Lesk, Michael., "*The seven Ages of Information Retrieval*", IFLANET. Junho 1995. Informação disponível no endereço da Web: <http://www.ifa.org/VI/5/op/udtop5/udtop5.htm> (última visita a 10 de Junho de 2005).
51. Oliveira, Paula., "*Um Modelo baseado em Agentes Móveis para Procura de Informação em Sistemas de Informação Distribuídos*", Tese de Doutoramento, Universidade de Trás-os-Montes e Alto Douro. 2003.
52. Bush, Vannevar., "*As we may think*". The Atlantic Monthly. Julho 1945. Informação disponível no endereço da Web: <http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush.shtml> (última visita a 10 de Junho de 2005).
53. *Text Retrieval Conference* (TREC). Informação disponível no endereço da Web: <http://trec.nist.gov/> (última visita a 10 de Junho de 2005).
54. Motor de Busca denominado Google. Informação disponível no endereço da Web: <http://www.google.com/> (última visita a 10 de Junho de 2005).
55. Motor de Busca denominado Altavista. Informação disponível no endereço da Web: <http://www.altavista.com/> (última visita a 10 de Junho de 2005).
56. Motor de Busca denominado SAPO. Informação disponível no endereço da Web: <http://www.sapo.pt/> (última visita a 10 de Junho de 2005).
57. Adaptação de uma imagem disponível no endereço da Web: <http://www.thefarside.com/> - última visita a 10 de Junho de 2004)

58. Moura, Ana., "*A Web Semântica: Fundamentos e Tecnologias*". Tutorial ministrado no VI Congresso Internacional de Ciencias de la Computación, La Paz, Outubro de 2001. Informação disponível no endereço da Web: <http://www.ipanema.ime.br/ana-moura/bibliometadados.html> (última visita a 10 de Junho de 2005).
59. Huhns, Michael N., "*The Sentient Web*". Internet Computing, IEEE Volume 7, Issue 6, Nov.-Dec. 2003. Page(s):82-84.
60. *HyperText Markup Language* (HTML). Informação disponível no endereço da Web: <http://www.w3.org/MarkUp/> (última visita a 28 de Junho de 2005).
61. *Active Server Pages* (ASP). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Active\\_Server\\_Pages](http://en.wikipedia.org/wiki/Active_Server_Pages) (última visita a 28 de Junho de 2005).
62. *PHP: Hypertext Preprocessor* (PHP). Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/PHP> (última visita a 28 de Junho de 2005).
63. *JavaServer Pages* (JSP). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/JavaServer\\_Pages](http://en.wikipedia.org/wiki/JavaServer_Pages) (última visita a 28 de Junho de 2005).
64. Fensel, D., Hendler J., Lieberman, H. & Wahlster, W. (editores)., "*Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*". Livro publicado pela MIT Press, London. 2003.
65. Berners-Lee, Tim., "*Semantic Web Road map*". W3C (MIT, ER-CIM, Keio). Outubro 1998. Informação disponível no endereço da Web: <http://www.w3.org/DesignIssues/Semantic.html> (última visita a 10 de Junho de 2005).
66. Hawke, Sandro., "*How the Semantic Web Works*". Abril 2002. Informação disponível no endereço da Web: <http://www.w3.org/2002/03/semweb/> (última visita a 10 de Junho de 2005).

67. Koivunen, Marja-Riitta & Miller, Eric., *"W3C Semantic Web Activity"*, Semantic Web Kick-Off in Finland: Vision, Technologies, Research, and Applications, HIIT Publications, Helsinki Institute for Information Technology, Helsinki, Finland. Novembro 2001.
68. Berners-Lee, Tim., *"What the Semantic Web can represent"*. W3C (MIT, ERCIM, Keio). Setembro 1998. Informação disponível no endereço da Web: <http://www.w3.org/DesignIssues/RDFnot.html> (última visita a 10 de Junho de 2005).
69. Smith, Michael., Welty, Chris., McGuinness, Deborah. & Van Harmelen, Frank., *"OWL Web Ontology Language Guide"*. W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/owl-guide/> (última visita em 8 de Junho de 2005).
70. Adaptação de uma imagem disponível no endereço da Web: <http://www.inf.ufsc.br/sbc2002/palestras/lucena.ppt> (última visita a 23 de Junho de 2005).
71. Berners-Lee, Tim & Connolly, Dan & Swick, Ralph R., *"Web Architecture: Describing and Exchanging Data"*. W3C (MIT, ERCIM, Keio). Junho 1999. Informação disponível no endereço da Web: <http://www.w3.org/1999/04/WebData> (última visita a 10 de Junho de 2005).
72. Berners-Lee, Tim & Karger, David R. & Stein, Lynn Andrea & Swick, Ralph R. & Weitzner, Daniel J., *"Semantic Web Development"*. W3C (MIT, ERCIM, Keio). Fevereiro 2000. Informação disponível no endereço da Web: <http://www.w3.org/2000/01/sw/DevelopmentProposal> (última visita a 10 de Junho de 2005).
73. *Knowledge Interchange Format (KIF)*. Informação disponível no endereço da Web: <http://logic.stanford.edu/kif/kif.html> (última visita a 10 de Junho de 2005).

74. *Cycorp* (Cyc). Informação disponível no endereço da Web: <http://www.cyc.com/> (última visita a 10 de Junho de 2005).
75. *Universal Resource Identifiers* (URI). Informação disponível no endereço da Web: [http://www.w3.org/Addressing/URL/URL\\_Overview.html](http://www.w3.org/Addressing/URL/URL_Overview.html) (última visita a 15 de Setembro de 2004).
76. *RDF Model and Syntax Specification. W3C Recommendation* de 22 de Fevereiro de 1999. Informação disponível no endereço da Web: <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/> (última visita a 15 de Setembro de 2004).
77. Felber, P. & Chan, C.Y. & Garofalakis, M.N. & Rastogi, R., "Scalable Filtering of XML Data for Web Services". IEEE Internet Computing, Volume 7, Issue 1, pp 49-57. 2003. (Informação disponível em <http://www.eurecom.fr/felber/publications/IC-03.pdf>)
78. *Extensible Markup Language XML 1.0. W3C Recommendation* de 10 de Fevereiro de 1998. Informação disponível no endereço da Web: <http://www.w3.org/TR/1998/REC-xml-19980210> (última visita a 15 de Setembro de 2004).
79. *EXtensible HyperText Markup Language*. Informação disponível no endereço da Web: <http://www.w3.org/TR/2005/WD-xhtml2-20050527/> (última visita a 15 de Setembro de 2004).
80. *Portable Document Format* (PDF). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Portable\\_Document\\_Format](http://en.wikipedia.org/wiki/Portable_Document_Format) (última visita a 28 de Junho de 2005).
81. *PostScript*. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/PostScript> (última visita a 28 de Junho de 2005).
82. *Scalable Vector Graphics* (SVG). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Scalable\\_Vector\\_Graphics](http://en.wikipedia.org/wiki/Scalable_Vector_Graphics) (última visita a 28 de Junho de 2005).

83. *Management Information Format* (MIF). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Management\\_Information\\_Format](http://en.wikipedia.org/wiki/Management_Information_Format) (última visita a 28 de Junho de 2005).
84. *Printer Command Language* (PCL). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Printer\\_Control\\_Language](http://en.wikipedia.org/wiki/Printer_Control_Language) (última visita a 28 de Junho de 2005).
85. Iannella, Renato & Waugh, Andrew., *"Metadata: Enabling the Internet"*. 1997. Informação disponível em <http://archive.dstc.edu.au/RDU/reports/CAUSE97> (última visita a 28 de Junho de 2005).
86. *Dublin Core Metadata Initiative*. Informação disponível no endereço da Web: <http://dublincore.org/> (última visita a 10 de Junho de 2004).
87. Sowa, John., *"Ontology, metadata, and semiotics"*. B. Ganter & G. W. Mineau, eds. *Conceptual Structures: Logical, Linguistic, and Computational Issues: 8th International Conference on Conceptual Structures*. Darmstadt, Germany. Agosto 2000. (Informação disponível em <http://www.jfsowa.com/ontology/ontometa.htm>)
88. Phipps, Simon., *"Meaning, not Markup,"*. XML Journal, vol. 1, no. 1, p. 66. Fevereiro 2000.
89. Ford, Paul., *"August 2009: How Google beat Amazon and Ebay to the Semantic Web"*. Julho 2002. Informação disponível no endereço da Web: [http://www.ftrain.com/google\\_takes\\_all.html](http://www.ftrain.com/google_takes_all.html) (última visita a 10 de Junho de 2005)
90. Swartz, Aaron., *"The Semantic Web In Breadth"*. Logicerror. Maio 2002. Informação disponível no endereço da Web: <http://logicerror.com/semanticWeb-long> (última visita a 10 de Junho de 2005).
91. *RDF vocabulary description language (RDF Schema)* (RDFS). Informação disponível no endereço da Web: <http://www.w3.org/TR/rdf-schema/> (última visita a 11 de Junho de 2005).

92. *XML Signatures*. Informação disponível no endereço da Web: <http://www.w3.org/Signature/> (última visita a 4 de Fevereiro de 2005).
93. *XML Encryption*. Informação disponível no endereço da Web: <http://www.w3.org/Encryption/2001/> (última visita a 4 de Fevereiro de 2005).
94. Linguagem *JavaScript*. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/JavaScript> (última visita a 28 de Junho de 2005).
95. *Graphics Interchange Format (GIF)*. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Graphics\\_Interchange\\_Format](http://en.wikipedia.org/wiki/Graphics_Interchange_Format) (última visita a 28 de Junho de 2005).
96. *Portable Network Graphics (PNG)*. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Portable\\_Network\\_Graphics](http://en.wikipedia.org/wiki/Portable_Network_Graphics) (última visita a 28 de Junho de 2005).
97. *Cascading Style Sheets (CSS)*. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Cascading\\_Style\\_Sheets](http://en.wikipedia.org/wiki/Cascading_Style_Sheets) (última visita a 28 de Junho de 2005).
98. *eXtensible Stylesheet Language (XSL)*. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Extensible\\_Stylesheet\\_Language](http://en.wikipedia.org/wiki/Extensible_Stylesheet_Language) (última visita a 28 de Junho de 2005).
99. *Standard Generalized Markup Language (SGML)*. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/SGML> (última visita a 28 de Junho de 2005).
100. *eXtensible Stylesheet Language (XSL) Version 1.0. W3C Recommendation* de 15 de Outubro de 2001. Informação disponível no endereço da Web: <http://www.w3.org/TR/2001/REC-xsl-20011015/> (última visita a 15 de Setembro de 2004).



101. *GNU General Public License* (GPL). Informação disponível no endereço da Web: <http://www.gnu.org/copyleft/gpl.html> (última visita a 5 de Maio de 2005).
102. *The Open Source Definition. W3C Recommendation* de 15 de Outubro de 2001. Informação disponível no endereço da Web: <http://www.opensource.org/docs/definition.php> (última visita a 10 de Junho de 2005).
103. *Artificial Intelligence Mark-up Language* (AIML). Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/AIML> (última visita a 10 de Junho de 2005).
104. *Astronomical markup Language* (AML). Informação disponível no endereço da Web: <http://www.cs.odu.edu/~dorton/cs695/assign1/AML.html> (última visita a 10 de Junho de 2005).
105. *Chemical Markup Language* (CML). Informação disponível no endereço da Web: <http://www.xml-cml.org/> (última visita a 10 de Junho de 2005).
106. *Dental Charting Markup Language* (DCML). Informação disponível no endereço da Web: <http://edr.uchc.edu/dcml/default.html> (última visita a 10 de Junho de 2005).
107. *Directory Services Markup Language* (DSML). Informação disponível no endereço da Web: <http://xml.coverpages.org/dsml.html> (última visita a 10 de Junho de 2005).
108. *Gene Expression Markup Language* (GEML). Informação disponível no endereço da Web: <http://xml.coverpages.org/geml.html> (última visita a 10 de Junho de 2005).
109. *Investment Research Markup Language* (IRML). Informação disponível no endereço da Web: <http://xml.coverpages.org/irml.html> (última visita a 10 de Junho de 2005).
110. *Mathematical Markup Language* (MathML). Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/MathML> (última visita a 10 de Junho de 2005).
111. *Medical Markup Language* (MML). Informação disponível no endereço da Web: [http://www.medxml.net/worldwide/What\\_is\\_MML.htm](http://www.medxml.net/worldwide/What_is_MML.htm) (última visita a 10 de Junho de 2005).

2005).

112. *News Markup Language* (NML). Informação disponível no endereço da Web: <http://xml.coverpages.org/nml.html> (última visita a 10 de Junho de 2005).
113. *Weather Observation Definition Format* (OMF). Informação disponível no endereço da Web: <http://www.metnet.navy.mil/Metcast/XML/OMF.html> (última visita a 10 de Junho de 2005).
114. *Petroleum XML* (PetroXML). Informação disponível no endereço da Web: <http://www.petroxml.com/> (última visita a 10 de Junho de 2005).
115. *Synchronized Multimedia Integration Language* (SMIL). Informação disponível no endereço da Web: <http://www.w3.org/TR/REC-smil/> (última visita a 10 de Junho de 2005).
116. *Theological Markup Language* (ThML). Informação disponível no endereço da Web: <http://www.ccel.org/ThML/> (última visita a 10 de Junho de 2005).
117. *Universal Description Discovery Integration* (UDDI). Informação disponível no endereço da Web: <http://www.uddi.org/> (última visita a 10 de Junho de 2005).
118. *Virtual Instruments Markup language* (VIML). Informação disponível no endereço da Web: <http://viml.org/> (última visita a 10 de Junho de 2005).
119. *Wireless Markup Language* (WML). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Wireless\\_Markup\\_Language](http://en.wikipedia.org/wiki/Wireless_Markup_Language) (última visita a 10 de Junho de 2005).
120. *Web Services Description Language* (WSDL). Informação disponível no endereço da Web: <http://www.w3.org/TR/wsdl> (última visita a 10 de Junho de 2005).
121. *Transaction Authority Markup Language* (XAML). Informação disponível no endereço da Web: <http://xml.coverpages.org/xaml.html> (última visita a 10 de Junho de 2005).

122. *Bookmark Exchange Language* (XBEL). Informação disponível no endereço da Web: <http://xml.coverpages.org/xbel.html> (última visita a 10 de Junho de 2005).
123. *Web Services*. Informação disponível no endereço da Web: <http://www.w3.org/2002/ws/> (última visita a 10 de Junho de 2005).
124. Ramalho, J. & Henriques, P., "*XML & XSL: da teoria à prática*". FCA Editora (ISBN 972-722-347-8). Outubro 2002.
125. *Document Type Definition* (DTD). Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Document\\_Type\\_Definition](http://en.wikipedia.org/wiki/Document_Type_Definition) (última visita a 10 de Junho de 2005).
126. Esquema XML (*XML Schema*). Informação disponível no endereço da Web: <http://www.w3.org/XML/Schema> (última visita a 10 de Junho de 2005).
127. Decker, Stefan. & Melnik, Sergey. & Van Harmelen, Frank. & Fensel, Dieter. & Klein, Michel. & Broekstra, Jeen. & Erdmann, Michael. & Horrocks, Ian., "*The Semantic Web: The Roles of XML and RDF*", IEEE Internet Computing. 2000. (Informação disponível em <http://www.ontoknowledge.org/oil/downl/IEEE00.pdf>)
128. *Simple API for XML* (SAX). Informação disponível no endereço da Web: <http://www.saxproject.org/> (última visita a 9 de Junho de 2005).
129. *Document Object Mode* (DOM). Informação disponível no endereço da Web: <http://www.w3.org/DOM/> (última visita a 9 de Junho de 2005).
130. Vaswani, V., "*PHP and the Document Object Model (DOM)*". Pearson Education, Que Publishing. Outubro 2002. Informação disponível no endereço da Web: <http://www.quepublishing.com/articles/article.asp?p=29800&seqNum=7&rl=1> (última visita a 9 de Junho de 2005).
131. Repositório de ferramentas XML que se encontra disponível no endereço da Web: <http://www.garshol.priv.no/download/xmltools/> (última visita a 20 de Junho de 2005).

132. Ferramenta denominada IsaViz. Informação disponível no endereço da Web: <http://www.w3.org/2001/11/IsaViz/> (última visita a 8 de Maio de 2005).
133. Bizer, C. & Westphal, D., "*Developers Guide to Semantic Web Toolkits for different Programming Languages*". 2005. Informação disponível no endereço da Web: <http://www.wiwiss.fu-berlin.de/suhl/bizer/toolkits/> (última visita a 11 de Julho de 2005).
134. Ferramenta denominada Jena. Informação disponível no endereço da Web: <http://www.hpl.hp.com/semweb/jena2.htm> (última visita a 8 de Maio de 2005).
135. *HP Labs Semantic Web Programme*. Informação disponível no endereço da Web: <http://www.hpl.hp.com/semweb/> (última visita a 11 de Junho de 2005).
136. Notation3: trata-se de uma linguagem baseada em RDF para a Web Semântica. Informação disponível no endereço da Web: <http://www.w3.org/DesignIssues/Notation3.html> (última visita a 11 de Junho de 2005).
137. *W3C RDF Core WG Internal Working Draft N-Triples*: é uma formatação em texto simples para a sintaxe RDF. Informação disponível no endereço da Web: <http://www.w3.org/2001/sw/RDFCore/ntriples/> (última visita a 1 de Novembro de 2005).
138. Brickley, Dan. & Guha, R.V., "*RDF Vocabulary Description Language 1.0: RDF Schema*". W3C (MIT, ERCIM, Keio). Novembro 2002. (Informação disponível em <http://www.w3.org/TR/2002/WD-rdf-schema-20021112/>)
139. Imagem de exemplo sobre o Esquema RDF apresentada pelo W3C. <http://www.w3.org/TR/2002/WD-rdf-schema-20021112/intro-example.png> (última visita a 11 de Julho de 2005).
140. Gruber, T., "*A translation approach to portable ontologies*". Knowledge Acquisition 5(2):199-220. Junho 1993. (Informação disponível em [http://ksl-web.stanford.edu/KSL\\_Abstracts/KSL-92-71.html](http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html))

141. Lógica Formal. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Formal\\_logic](http://en.wikipedia.org/wiki/Formal_logic) (última visita a 4 de Fevereiro de 2005).
142. Dicionário *On-line* de Língua Portuguesa da Editora Universal. Informação disponível no endereço da Web: [http://www.priberam.pt/dlpo/definir\\_resultados.aspx](http://www.priberam.pt/dlpo/definir_resultados.aspx) (última visita a 12 de Julho de 2005).
143. McGuinness, Deborah., "Ontologies Come of Age". 6º Capítulo do livro *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. Editado por Fensel, Dieter. & Hendler, James. & Lieberman, Henry. & Wahlster, Wolfgang. MIT Press. 2003.
144. Lógica Fuzzy. Informação disponível no endereço da Web: <http://www-2.cs.cmu.edu/Groups/AI/html/faqs/ai/fuzzy/part1/faq-doc-2.html> (última visita a 11 de Junho de 2005).
145. Redes Neurais. Informação disponível no endereço da Web: <ftp://ftp.sas.com/pub/neural/FAQ.html> (última visita a 11 de Junho de 2005).
146. Daconta, Michael C. & Obrst, Leo J. & Smith, Kevin T., "The Semantic Web: A guide to the future of XML, Web Services, and Knowledge Management". Wiley Publishing, Inc. 2003.
147. *Roget's Thesaurus of English Words and Phrases*. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Roget%27s\\_Thesaurus](http://en.wikipedia.org/wiki/Roget%27s_Thesaurus) (última visita a 3 de Maio de 2005).
148. NISO - *National Information Standards Organization*. "Guidelines for the Construction, Format, and Management of Monolingual Thesauri". NISO Press. 1994. (Informação disponível em <http://www.niso.org/standards/resources/Z39-19.html>)
149. *WordNet Thesaurus*. Informação disponível no endereço da Web: <http://www.cogsci.princeton.edu/wn/> (última visita a 21 de Dezembro de 2004).

150. Tsichritzis, D. & Klug, A., *"The ANSI/X3/SPARC DBMS framework report of the study group on database management systems"*. Information Systems, 3(3):173-191. Março 1978.
151. *Simple HTML Ontology Extensions (SHOE)*. Informação disponível no endereço da Web: <http://www.cs.umd.edu/projects/plus/SHOE/spec.html> (última visita a 11 de Junho de 2004).
152. *XML-based Ontology Exchange Language (XOL)*. Informação disponível no endereço da Web: <http://www.ai.sri.com/pkarp/xol/xol.html> (última visita a 11 de Junho de 2004).
153. Linguagem de Interrogação denominada *Open Knowledge Base Connectivity (OKBC)* . Informação disponível no endereço da Web: <http://www.ai.sri.com/okbc/> (última visita a 5 de Junho de 2005).
154. *Ontology Inference Layer (OIL)*. Informação disponível no endereço da Web: <http://www.ontoknowledge.org/oil/> (última visita a 11 de Junho de 2004).
155. *DARPA Agent Markup Language (DAML)*. Informação disponível no endereço da Web: <http://www.daml.org/> (última visita a 11 de Junho de 2004).
156. *DARPA Agent Markup Language+Ontology Inference Layer (DAML+OIL)*. Informação disponível no endereço da Web: <http://www.daml.org/2001/03/daml+oil-index> (última visita a 11 de Junho de 2004).
157. Edd Dumbill, Edd. *Putting RDF to Work*. Agosto, 2000. Informação disponível no endereço da Web: <http://www.xml.com/pub/a/2000/08/09/rdfdb/> (última visita a 30 de Novembro de 2005).
158. Frauenfelder, Mark. *Sir Tim Berners-Lee*. Outubro, 2004. Informação disponível no endereço da Web: <http://www.technologyreview.com/articles/04/10/frauenfelder1004.asp?p=1> (última visita a 30 de Novembro de 2005).

159. Adaptação de uma imagem disponível no endereço da Web: [http://www-db.stanford.edu/OntoAgents/LINK\\_0\\_D41D8CD98F00B204E9800998ECF8427E.html](http://www-db.stanford.edu/OntoAgents/LINK_0_D41D8CD98F00B204E9800998ECF8427E.html) (última visita a 11 de Junho de 2004).
160. Linguagem de Programação *Java*. Informação disponível no endereço da Web: <http://java.sun.com/docs/books/jls/> (última visita a 11 de Junho de 2005).
161. Linguagem de programação, denominada Python. Informação disponível no endereço da Web: <http://www.python.org/> (última visita a 5 de Maio de 2005).
162. Linguagem de programação funcional, denominada Haskell. Informação disponível no endereço da Web: <http://www.haskell.org/> (última visita a 5 de Maio de 2005).
163. linguagem Oracle PL/SQL. Informação disponível no endereço da Web: <http://www.unix.org.ua/oreilly/oracle/langpkt/> (última visita a 5 de Maio de 2005).
164. *Berkeley Software Distribution* (BSD). Informação disponível no endereço da Web: <http://www.opensource.org/licenses/bsd-license.php> (última visita a 5 de Maio de 2005).
165. *GNU Lesser General Public License* (LGPL). Informação disponível no endereço da Web: <http://www.gnu.org/copyleft/lesser.html> (última visita a 5 de Maio de 2005).
166. *Mozilla Public License* (MPL). Informação disponível no endereço da Web: <http://www.mozilla.org/MPL/MPL-1.1.html> (última visita a 5 de Maio de 2005).
167. *W3C Software License* (W3C license). Informação disponível no endereço da Web: <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231> (última visita a 5 de Maio de 2005).
168. *Apache Software License* (ASL). Informação disponível no endereço da Web: <http://www.opensource.org/licenses/apachepl.php> (última visita a 5 de Maio de 2005).
169. *Massachusetts Institute of Technology License* (MIT License). Informação disponível no endereço da Web: <http://www.opensource.org/licenses/mit-license.php> (última visita a 5

de Maio de 2005).

170. *Nokia Open Source License* (NOKOS). Informação disponível no endereço da Web: <http://wilbur-rdf.sourceforge.net/docs/LICENSE-NOKOS.html> (última visita a 5 de Maio de 2005).
171. Ferramenta denominada SemanticWord. Informação disponível no endereço da Web: <http://mr.teknowledge.com/daml/SemanticWord/SemanticWord.htm> (última visita a 8 de Maio de 2005).
172. Tallis, M. (2003). Semantic Word Processing for Content Authors, In Workshop Notes of the Knowledge Markup and Semantic Annotation Workshop (SEMANNOT 2003), Second International Conference on Knowledge Capture (K-CAP 2003), October 26, 2003, Sanibel, Florida, USA.
173. Ferramenta denominada OntoMat-Annotizer. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/ontomat/> (última visita a 8 de Maio de 2005).
174. Ferramenta denominada PhotoStuff. Informação disponível no endereço da Web: <http://www.mindswap.org/2003/PhotoStuff/> (última visita a 8 de Maio de 2005).
175. Halaschek-Wiener, C. (2004). PhotoStuff overview. Mind-Swap research lab. Informação disponível no endereço da Web: <http://www.mindswap.org/2003/PhotoStuff/talks/psOverview.pdf> (última visita a 5 de Maio de 2005).
176. Golbeck, J., Alford, R., Baker, R., Grove, M., Hendler, J., Kalyanpur, A., Loomis, A., Reck, R., Semantic Web Tools from MIND SWAP. Poster. International Semantic Web Conference (ISWC). June 2002, Sardinia, Italy. Informação disponível no endereço da Web: [http://iswc2002.semanticweb.org/posters/golbeck\\_a4.pdf](http://iswc2002.semanticweb.org/posters/golbeck_a4.pdf) (última visita a 5 de Maio de 2005).



177. Ferramenta denominada Swangler. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/swangle/> (última visita a 8 de Maio de 2005).
178. Ferramenta denominada *AeroText Semantic Web Automated Relation Markup* (AeroSWARM). Informação disponível no endereço da Web: <http://ubot.lockheedmartin.com/ubot/hotdaml/aeroswarm.html> (última visita a 8 de Maio de 2005).
179. Kogut, P., Holmes, W.: AeroDAML: Applying Information Extraction to Generate DAML Annotations from Web Pages. First International Conference on Knowledge Capture (K-CAP 2001). Workshop on Knowledge Markup and Semantic Annotation, Victoria, B.C. October 21, 2001.
180. Protocolo *Simple Object Access Protocol* (SOAP). Informação disponível no endereço da Web: <http://www.w3.org/TR/soap/> (última visita a 4 de Maio de 2005).
181. Ferramenta denominada Microsoft Word. Informação disponível no endereço da Web: <http://www.microsoft.com/office/word/prodinfo/default.mspx> (última visita a 5 de Junho de 2005).
182. Ferramenta denominada OWL Semantic Search Services. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owl-semsearch/> (última visita a 8 de Maio de 2005).
183. Sistema Operativo Linux. Informação disponível no endereço da Web: <http://www.linux.org/> (última visita a 8 de Maio de 2005).
184. Ferramenta denominada *Components for Ontology Processing* (CODIP). Informação disponível no endereço da Web: <http://codip.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
185. Ferramenta denominada HAWK. Informação disponível no endereço da Web: <http://www.cse.lehigh.edu/~zhp2/hawk/readme.html> (última visita a 8 de Maio de 2005).

186. Ferramenta denominada Kazuki. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/kazuki/> (última visita a 8 de Maio de 2005).
187. Ferramenta denominada *Named Graphs API for Jena* (NG4J). Informação disponível no endereço da Web: <http://www.wiwiss.fu-berlin.de/suhl/bizer/ng4j/> (última visita a 8 de Maio de 2005).
188. Ferramenta denominada OWL API. Informação disponível no endereço da Web: <http://sourceforge.net/projects/owlapi> (última visita a 8 de Maio de 2005).
189. Ferramenta denominada *Simple Ontology Framework API* (SOFA). Informação disponível no endereço da Web: <http://sofa.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
190. Ferramenta denominada CARA. Informação disponível no endereço da Web: <http://cara.sourceforge.net/> (última visita a 8 de Maio de 2005).
191. Ferramenta denominada Sparta. Informação disponível no endereço da Web: <http://www.mnot.net/sw/sparta/> (última visita a 8 de Maio de 2005).
192. Ferramenta denominada *Python Rdf Api* (Pyrple). Informação disponível no endereço da Web: <http://infomesh.net/pyrple/> (última visita a 8 de Maio de 2005).
193. Base de Dados denominada MySQL. Informação disponível no endereço da Web: <http://www.mysql.com/> (última visita a 8 de Maio de 2005).
194. *Java Database Connectivity API* (JDBC). Informação disponível no endereço da Web: <http://java.sun.com/products/jdbc/> (última visita a 8 de Maio de 2005).
195. *Named Graphs*. Informação disponível no endereço da Web: <http://www.w3.org/2004/03/trix/> (última visita a 5 de Junho de 2005).
196. Ferramenta denominada OWL-S API. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owl-s-api/> (última visita a 8 de Maio de 2005).

197. Ferramenta denominada PySesame. Informação disponível no endereço da Web: <http://pysesame.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
198. Ferramenta denominada An Entry Sub-ontology of OWL-Time. Informação disponível no endereço da Web: <http://entry-owl-time.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
199. Ferramenta denominada Time Zone Resource in OWL. Informação disponível no endereço da Web: <http://timezone.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
200. *Hypertext Transfer Protocol* (HTTP). Informação disponível no endereço da Web: <http://www.w3.org/Protocols/> (última visita a 14 de Dezembro de 2005).
201. Ferramenta denominada 3Store. Informação disponível no endereço da Web: <http://sourceforge.net/projects/threestore/> (última visita a 8 de Maio de 2005).
202. Ferramenta denominada RDFStore. Informação disponível no endereço da Web: <http://rdfstore.sourceforge.net/> (última visita a 8 de Maio de 2005).
203. Ferramenta denominada *Yet Another RDF Store* (YARS). Informação disponível no endereço da Web: <http://sw.deri.org/2004/06/yars/yars.html> (última visita a 8 de Maio de 2005).
204. Ferramenta denominada Kowari. Informação disponível no endereço da Web: <http://www.kowari.org/> (última visita a 8 de Maio de 2005).
205. Ferramenta denominada Sesame. Informação disponível no endereço da Web: <http://www.openrdf.org> (última visita a 8 de Maio de 2005).
206. Base de Dados denominada Berkeley. Informação disponível no endereço da Web: <http://db.cs.berkeley.edu/> (última visita a 5 de Junho de 2005).
207. Base de Dados denominada PostgreSQL. Informação disponível no endereço da Web: <http://www.postgresql.org/> (última visita a 5 de Junho de 2005).

208. Base de Dados denominada Oracle. Informação disponível no endereço da Web: <http://www.oracle.com/database/index.html> (última visita a 5 de Junho de 2005).
209. Base de Dados denominada Microsoft SQL Server. Informação disponível no endereço da Web: <http://www.microsoft.com/sql/default.msp> (última visita a 5 de Junho de 2005).
210. Linguagem de Interrogação denominada *RDF Data Query Language* (RDQL). Informação disponível no endereço da Web: <http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/> (última visita a 5 de Junho de 2005).
211. Linguagem de Interrogação denominada *Tucana SQL-like query language* (iTQL) . Informação disponível no endereço da Web: <http://www.kowari.org/271.htm> (última visita a 5 de Junho de 2005).
212. Linguagem de Interrogação denominada *Protocol And RDF Query Language* (SPARQL) . Informação disponível no endereço da Web: <http://www.w3.org/TR/rdf-sparql-query/> (última visita a 5 de Junho de 2005).
213. McGuinness, Deborah. & Van Harmelen, Frank. "OWL Web Ontology Language Overview". W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/owl-features/> (última visita em 8 de Junho de 2005).
214. Linguagem de Interrogação denominada *Notation3 Query Language* (N3QL) . Informação disponível no endereço da Web: <http://www.w3.org/DesignIssues/N3QL> (última visita a 5 de Junho de 2005).
215. Ferramenta denominada ParkasW. Informação disponível no endereço da Web: <http://parkasw.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
216. Ferramenta denominada Owl2dig. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owl2dig/> (última visita a 8 de Maio de 2005).
217. Ferramenta denominada Wsdl2owl-s. Informação disponível no endereço da Web: <http://www.daml.ri.cmu.edu/wsdl2owls/> (última visita a 8 de Maio de 2005).

218. Ferramenta denominada *java2owl-s*. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/java2owl-s/> (última visita a 8 de Maio de 2005).
219. Ferramenta denominada *Web Service Ontology Language (OWL-S)*. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/java2owl-s/> (última visita a 8 de Maio de 2005).
220. Formato *Description Logic Interface (DIG)*. Informação disponível no endereço da Web: <http://www.sts.tu-harburg.de/~r.f.moeller/racer/interface1.1.pdf> (última visita a 8 de Maio de 2005).
221. Ferramenta denominada *Owl Converter*. Informação disponível no endereço da Web: <http://www.mindswap.org/~golbeck/code.shtml> (última visita a 8 de Maio de 2005).
222. Ferramenta denominada *owl-s2uddi*. Informação disponível no endereço da Web: <http://owl-s2uddi.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
223. Ferramenta denominada *SWeHG*. Informação disponível no endereço da Web: <http://www.cs.helsinki.fi/group/seco/swehg/> (última visita a 8 de Maio de 2005).
224. Ferramenta denominada *OWL Mode for Emacs*. Informação disponível no endereço da Web: <http://owl-emacs.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
225. *Portable Operating System Interface (POSIX)*. Informação disponível no endereço da Web: <http://www.knosof.co.uk/posix.html> (última visita a 8 de Maio de 2005).
226. *Ontology engineRING ENvironment* (ORIENT). Informação disponível no endereço da Web: <http://apex.sjtu.edu.cn/projects/orient/News.htm> (última visita a 8 de Maio de 2005).
227. Ferramenta denominada *SMORE*. Informação disponível no endereço da Web: <http://www.mindswap.org/2005/SMORE/> (última visita a 8 de Maio de 2005).

228. Ferramenta denominada SWeDE. Informação disponível no endereço da Web: <http://owl-eclipse.projects.semwebcentral.org/> (última visita a 8 de Maio de 2005).
229. Ferramenta denominada SWOOP. Informação disponível no endereço da Web: <http://www.mindswap.org/2004/SWOOP/> (última visita a 8 de Maio de 2005).
230. *Eclipse White Paper. "Eclipse Platform Technical Overview"*. Eclipse Project. Fevereiro 2003. Informação disponível no endereço da Web: <http://www.eclipse.org> (última visita a 8 de Maio de 2005).
231. Ferramenta denominada OWL Filetype Plugin for VIM. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owlvim/> (última visita a 8 de Maio de 2005).
232. VIM: Editor de texto. Informação disponível no endereço da Web: <http://www.vim.org/> (última visita a 8 de Maio de 2005).
233. Editor *Emacs*. Informação disponível no endereço da Web: <http://www.gnu.org/software/emacs/emacs.html> (última visita - 8 de Maio de 2005);
234. Ferramenta denominada DumpOnt. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/dumpont/> (última visita a 8 de Maio de 2005).
235. Ferramenta denominada RDF Gateway. Informação disponível no endereço da Web: <http://www.intellidimension.com/default.jsp?topic=/pages/site/products/rdfgateway.jsp> (última visita a 8 de Maio de 2005).
236. Ferramenta denominada Wilbur. Informação disponível no endereço da Web: <http://wilbur-rdf.sourceforge.net/> (última visita a 8 de Maio de 2005).
237. Ferramenta denominada 4Suite. Informação disponível no endereço da Web: <http://4suite.org/index.xhtml> (última visita a 8 de Maio de 2005).

238. Ferramenta denominada KAON. Informação disponível no endereço da Web: <http://kaon.semanticweb.org/> (última visita a 8 de Maio de 2005).
239. Ferramenta denominada PerIRDF. Informação disponível no endereço da Web: [http://www.gingerall.com/charlie/ga/xml/p\\_rdf.xml](http://www.gingerall.com/charlie/ga/xml/p_rdf.xml) (última visita a 8 de Maio de 2005).
240. Ferramenta denominada RDF Gateway. Informação disponível no endereço da Web: <http://www.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/index.html> (última visita a 8 de Maio de 2005).
241. Ferramenta denominada Redland. Informação disponível no endereço da Web: <http://librdf.org/> (última visita a 8 de Maio de 2005).
242. Base de Dados denominada SQLite. Informação disponível no endereço da Web: <http://www.sqlite.org/> (última visita a 5 de Junho de 2005).
243. Linguagem de Interrogação de BD: SQL, versão 2. Informação disponível no endereço da Web: <http://xarch.tu-graz.ac.at/autocad/docs/sql2.htm> (última visita a 5 de Junho de 2005).
244. *ActiveX Data Objects Data Base* (ADOdb). Informação disponível no endereço da Web: <http://msdn.microsoft.com/data/Default.aspx> (última visita a 5 de Junho de 2005).
245. *Object Linking and Embedding Data Base* (OleDB). Informação disponível no endereço da Web: <http://msdn.microsoft.com/data/Default.aspx> (última visita a 5 de Junho de 2005).
246. DB2. Informação disponível no endereço da Web: <http://www-306.ibm.com/software/data/db2/> (última visita a 5 de Junho de 2005).
247. Linguagem de Interrogação denominada Versa. Informação disponível no endereço da Web: <http://skew.org/mike/4suite-docs/html/Versa.html> (última visita a 5 de Junho de 2005).
248. Linguagem de Interrogação denominada *RDF Data Query Language* (RDFQL). Informação disponível no endereço da Web:

<http://www.intellidimension.com/default.rsp?topic=/pages/rdfgateway/reference/db/default.rsp>  
(última visita a 5 de Junho de 2005).

249. Linguagem de Interrogação denominada *Windows Management Instrumentation Query Language* (WQL). Informação disponível no endereço da Web: [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/wmisdk/wmi/querying\\_with\\_wql.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/wmisdk/wmi/querying_with_wql.asp) (última visita a 5 de Junho de 2005).
250. Ferramenta denominada FaCT++. Informação disponível no endereço da Web: <http://owl.man.ac.uk/factplusplus/> (última visita a 8 de Maio de 2005).
251. Ferramenta denominada F-OWL. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/fowl/> (última visita a 8 de Maio de 2005).
252. Ferramenta denominada Hoolet. Informação disponível no endereço da Web: <http://owl.man.ac.uk/hoolet/> (última visita a 8 de Maio de 2005).
253. Ferramenta denominada EulerMoz. Informação disponível no endereço da Web: <http://sourceforge.net/projects/eulermoz> (última visita a 8 de Maio de 2005).
254. Ferramenta denominada Instance Store. Informação disponível no endereço da Web: <http://instancestore.man.ac.uk/> (última visita a 8 de Maio de 2005).
255. Ferramenta denominada Euler. Informação disponível no endereço da Web: <http://eulerssharp.sourceforge.net/> (última visita a 8 de Maio de 2005).
256. Ferramenta denominada Swish. Informação disponível no endereço da Web: <http://www.ninebynine.org/RDFNotes/Swish/Intro.html> (última visita a 8 de Maio de 2005).
257. Ferramenta denominada Metalog. Informação disponível no endereço da Web: <http://www.w3.org/RDF/Metalog/> (última visita a 8 de Maio de 2005).



258. Ferramenta denominada OWLJessKB. Informação disponível no endereço da Web: <http://edge.cs.drexel.edu/assemblies/software/owljesskb/> (última visita a 8 de Maio de 2005).
259. Ferramenta denominada Pychinko. Informação disponível no endereço da Web: <http://www.mindswap.org/%7Ekatz/pychinko/> (última visita a 8 de Maio de 2005).
260. Ferramenta denominada RdfLib. Informação disponível no endereço da Web: <http://www.semanticplanet.com/library?pagename=RdfLib.HomePage> (última visita a 8 de Maio de 2005).
261. Linguagem de Interrogação denominada *Pseudo Natural Language* (PNL). Informação disponível no endereço da Web: <http://www.w3.org/RDF/Metalog/docs/pnl.html> (última visita a 8 de Maio de 2005).
262. Linguagem de Interrogação denominada Jess. Informação disponível no endereço da Web: <http://herzberg.ca.sandia.gov/jess/> (última visita a 8 de Maio de 2005).
263. *XML Schema Definition* (XSD). Informação disponível no endereço da Web: <http://www.w3.org/XML/Schema> (última visita a 8 de Maio de 2005).
264. Ferramenta denominada Pellet. Informação disponível no endereço da Web: <http://www.mindswap.org/2003/pellet/index.shtml> (última visita a 8 de Maio de 2005).
265. Ferramenta denominada FlinkCommands. Informação disponível no endereço da Web: <http://flink.dcc.ufba.br/en/software/commands.html> (última visita a 8 de Maio de 2005).
266. Ferramenta denominada OWL-S IDE. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owl-s-ide/> (última visita a 8 de Maio de 2005).
267. Ferramenta denominada Visual Variable-Depth Info Display. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/vivid/> (última visita a 8 de Maio de 2005).

268. Ferramenta denominada HyperDAML. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/hyperdaml/> (última visita a 8 de Maio de 2005).
269. Ferramenta denominada Object Viewer. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/objectviewer/> (última visita a 8 de Maio de 2005).
270. Ferramenta denominada OWL-P. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owlp/> (última visita a 8 de Maio de 2005).
271. Ferramenta denominada Semantic Discovery Service. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/sds/> (última visita a 8 de Maio de 2005).
272. Sistema Operativo MacOS. Informação disponível no endereço da Web: [http://en.wikipedia.org/wiki/Mac\\_OS](http://en.wikipedia.org/wiki/Mac_OS) (última visita a 8 de Maio de 2005).
273. Ferramenta denominada *Closed World Machine* (Cwm). Informação disponível no endereço da Web: <http://www.w3.org/2000/10/swap/doc/cwm> (última visita a 8 de Maio de 2005).
274. Ferramenta denominada Lehigh University Benchmark. Informação disponível no endereço da Web: <http://swat.cse.lehigh.edu/projects/lubm/index.htm> (última visita a 8 de Maio de 2005).
275. Ferramenta denominada *OWL-S Service Retrieval Test Collection* (OWLS-TC). Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/owls-tc/> (última visita a 8 de Maio de 2005).
276. Ferramenta denominada TAP. Informação disponível no endereço da Web: <http://tap.stanford.edu/> (última visita a 8 de Maio de 2005).

277. Linguagem de Interrogação denominada GetData. Informação disponível no endereço da Web: <http://tap.stanford.edu/tap/docs/client.html> (última visita a 8 de Maio de 2005).
278. Linguagem de Interrogação denominada *OWL Query Language* (OWL-QL). Informação disponível no endereço da Web: (<http://projects.semwebcentral.org/projects/owl-ql/> - última visita - 8 de Maio de 2005).
279. . Linguagem de Interrogação denominada *Rule Extension of OWL Mobile Commerce Lab* (ROWL). Informação disponível no endereço da Web: (<http://projects.semwebcentral.org/projects/rowl/> - última visita - 8 de Maio de 2005).
280. Ferramenta denominada OWL Validator. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/vowlidator/> (última visita a 8 de Maio de 2005).
281. Ferramenta denominada SWRL Validator. Informação disponível no endereço da Web: <http://projects.semwebcentral.org/projects/swrl-val/> (última visita a 8 de Maio de 2005).
282. Ferramenta denominada ConsVISor. Informação disponível no endereço da Web: <http://68.162.250.6/index.html> (última visita a 8 de Maio de 2005).
283. *Semantic Web Rule Language* (SWRL). Informação disponível no endereço da Web: <http://www.daml.org/2003/11/swrl/> (última visita a 8 de Maio de 2005).
284. Ferramenta denominada VisioOWL. Informação disponível no endereço da Web: <http://web.tampabay.rr.com/flynn/VisioOWL/VisioOWL.htm> (última visita a 8 de Maio de 2005).
285. *Graph Stylesheets* (GSS). Informação disponível no endereço da Web: <http://www.w3.org/2001/11/IsaViz/gss/gssmanual.html> (última visita a 8 de Maio de 2005).

286. Ferramenta denominada SVG-OWL Viewer. Informação disponível no endereço da Web: [http://www.mindswap.org/aditkal/svg\\_owl.shtml](http://www.mindswap.org/aditkal/svg_owl.shtml) (última visita a 8 de Maio de 2005).
287. Betamax. Modelo de cassetes de video. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/Betamax> (última visita a 8 de Maio de 2005).
288. Processador de texto denominado AmiPro. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/AmiPro> (última visita a 8 de Maio de 2005).
289. Processador de texto denominado WordPerfect. Informação disponível no endereço da Web: <http://en.wikipedia.org/wiki/WordPerfect> (última visita a 8 de Maio de 2005).
290. *Concurrent Versions System* (CVS). Informação disponível no endereço da Web: <http://www.nongnu.org/cvs/> (última visita a 8 de Maio de 2005).
291. Definição de *What You See Is What You Get* (WYSIWYG). Informação disponível no endereço da Web: [http://searchsmb.techtarget.com/sDefinition/0,290660,sid44\\_gci213392,00.html](http://searchsmb.techtarget.com/sDefinition/0,290660,sid44_gci213392,00.html) (última visita a 1 de Dezembro de 2005).
292. BBN Technologies. Informação disponível no endereço da Web: <http://www.bbn.com/> (última visita a 8 de Maio de 2005).
293. Iannella, R., *Representing vCard Objects in RDF/XML*. Fevereiro 2001. Informação disponível no endereço da Web: <http://www.w3.org/TR/vcard-rdf> (última visita a 1 de Novembro de 2005).
294. Código do Esquema VCARD. Informação disponível no endereço da Web: <http://www.w3.org/2001/vcard-rdf/3.0> (última visita a 1 de Novembro de 2005).
295. IntoWine - Sítio da Web sobre vinhos de todo o Mundo: <http://www.intowine.com/> (última visita a 3 de Dezembro de 2005).

296. Instituto dos Vinhos do Douro e Porto - Sítio da Web sobre o vinho do Porto: <http://www.ivp.pt/index.asp> (última visita a 3 de Dezembro de 2005).
297. Sítio da Web sobre o queijo Stilton: <http://www.stiltoncheese.com/> (última visita a 3 de Dezembro de 2005).
298. *DAML Ontology Library* - Biblioteca de ontologias da DAML. Informação disponível no endereço da Web: <http://www.daml.org/ontologies/> (última visita a 3 de Dezembro de 2005).
299. Sítio Web com ontologias em OWL: <http://protege.stanford.edu/plugins/owl/owl-library/> (última visita a 3 de Dezembro de 2005).
300. Ontologia relacionada com o cancro. Informação disponível no endereço da Web: <http://www.mindswap.org/2003/CancerOntology/nciOncology.owl> (última visita a 11 de Junho de 2005).
301. Ontologia relacionada com a Encefalopatia Espongiforme Bovina (BSE), também conhecida como a doença das Vacas Loucas. Informação disponível no endereço da Web: [http://www.cs.man.ac.uk/~horrocks/OWL/Ontologies/mad\\_cows.owl](http://www.cs.man.ac.uk/~horrocks/OWL/Ontologies/mad_cows.owl) (última visita a 11 de Junho de 2005).
302. Ontologia relacionada com as Armas de Destruição Maciça. Informação disponível no endereço da Web: <http://reliant.teknowledge.com/DAML/WMD.owl> (última visita a 11 de Junho de 2005).
303. Ontologia, muito simples, relacionada com vinhos de mesa. Informação disponível no endereço da Web: <http://www.w3.org/TR/2004/REC-owl-guide-20040210/wine.rdf> (última visita a 3 de Dezembro de 2005).
304. Ontologia, muito simples, relacionada com pratos gastronómicos. Informação disponível no endereço da Web: <http://www.w3.org/TR/2004/REC-owl-guide-20040210/food.rdf> (última visita a 3 de Dezembro de 2005).

305. McGuinness, Deborah. Knowledge Systems Laboratory, Stanford University. Classic wines knowledge. Informação disponível no endereço da Web: <http://www.daml.org/ontologies/76> (última visita a 3 de Dezembro de 2005).
306. *KSL Wine Agent 1.0*. Informação disponível no endereço da Web: <http://www.daml.org/ontologies/76> (última visita a 3 de Dezembro de 2005).
307. JTP - *An Object-Oriented Modular Reasoning System*. Informação disponível no endereço da Web: <http://www.ksl.stanford.edu/software/JTP/> (última visita a 3 de Dezembro de 2005).
308. Editor HTML, de nome Ekit. Informação disponível no endereço da Web: <http://www.hexidec.com/ekit.php> (última visita a 30 de Outubro de 2005).
309. Berners-Lee, T., *Business Case for the Semantic Web*. Outubro 2001. Informação disponível no endereço da Web: <http://www.w3.org/DesignIssues/Business> (última visita a 3 de Dezembro de 2005).
310. Grupo de Trabalho *Web Ontology Working Group*. Informação disponível no endereço da Web: <http://www.w3.org/2001/sw/WebOnt/> (última visita - 8 de Maio de 2005).
311. Dean, Mike., Schreiber, Guus., Bechhofer, Sean., Van Harmelen, Frank., Hendler, Jim., Horrocks, Ian., McGuinness, Deborah., Patel-Schneider, Peter. & Stein, Lynn., *OWL Web Ontology Language Reference*. W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/owl-ref/> - última visita em 8 de Junho de 2005.
312. Patel-Schneider, Peter., Hayes, Patrick. & Horrocks, Ian., *OWL Web Ontology Language Semantics and Abstract Syntax*. W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/owl-semantics/> - última visita em 8 de Junho de 2005.
313. Carroll, Jeremy. & Roo, Jos., *OWL Web Ontology Language Test Cases*. W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web:

<http://www.w3.org/TR/owl-test/> - última visita em 8 de Junho de 2005.

314. Heflin, Jeff., *OWL Web Ontology Language Use Cases and Requirements*. W3C (MIT, ERCIM, Keio). 2004. Informação disponível no endereço da Web: <http://www.w3.org/TR/webont-req/> - última visita em 8 de Junho de 2005.
315. Miller, E. e Hendler, J., *Web Ontology Language (OWL)*. 2005. Informação disponível no endereço da Web: (<http://www.w3.org/2004/OWL/> - último acesso: 5 de Maio de 2005).





# Apêndice A

## Sintaxe RDF/XML

Este apêndice apresenta de uma forma prática os principais elementos definidos na sintaxe RDF/XML. A sintaxe RDF/XML permite definir quatro tipo de valores:

- Valores Simples;
- Valores Estruturados;
- Colecções de Objectos;
- Asserções baseadas em asserções.

De seguida serão apresentados alguns exemplos para os quatro tipos de valores que são possíveis de definir através da utilização desta sintaxe:

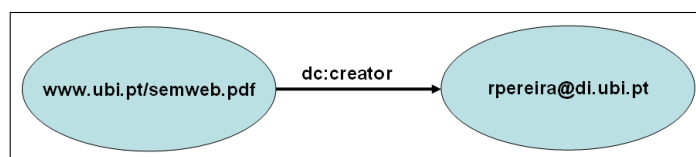


Figura A.1: Rui Pereira é o autor do documento semweb.pdf.

- Valores Simples: Uma entidade simples é utilizada quando se deseja mencionar apenas uma característica sobre um valor de um determinado recurso como, por exemplo, o grafo

apresentado na figura A.1. A representação do grafo da figura anterior é expressa através da sintaxe RDF/XML da seguinte forma:

```
<?xml version= "1.0"?>
<rdf:RDF xmlns:rdf=
    "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
    about="www.di.ubi.pt/semweb.pdf">
    <dc:creator>rpereira@di.ubi.pt</dc:creator>
  </rdf:Description>
</rdf:RDF>
```

Esta representação utiliza dois Espaços de Nomes identificados pelo conjunto de caracteres 'xmlns'. O primeiro Espaço de Nomes utiliza o prefixo 'rdf' e o URI "http://www.w3.org/1999/02/22-rdf-syntax-ns", e referencia o esquema de definições básicas do modelo RDF; o segundo utiliza o prefixo 'dc' e o URI "http://purl.org/dc/elements/1.1/", e referencia o esquema dos elementos definidos no vocabulário *Dublin Core* [86]. A asserção RDF<sup>163</sup> da figura A.1, associa o URI da pessoa "Rui Pereira" ao URI de um documento do tipo PDF. É utilizado o termo 'Description' para definir uma nova asserção RDF e o termo 'about' para identificar o recurso que se comporta como Sujeito da asserção. Os termos 'Description' e 'about' pertencem à sintaxe RDF e encontram-se especificados no Espaço de Nomes "http://www.w3.org/1999/02/22-rdf-syntax-ns". O Predicado da asserção RDF, termo 'publisher', encontra-se especificado no Espaço de Nomes "http://purl.org/dc/elements/1.1/". O Objecto da asserção RDF, neste caso um recurso (rpereira@di.ubi.pt), é associado ao Predicado no interior da declaração da asserção RDF.

- Valores Estruturados: Uma entidade estruturada é utilizada quando se deseja mencionar

---

<sup>163</sup>Uma asserção RDF é constituída por três elementos: o Sujeito, o Predicado e o Objecto.

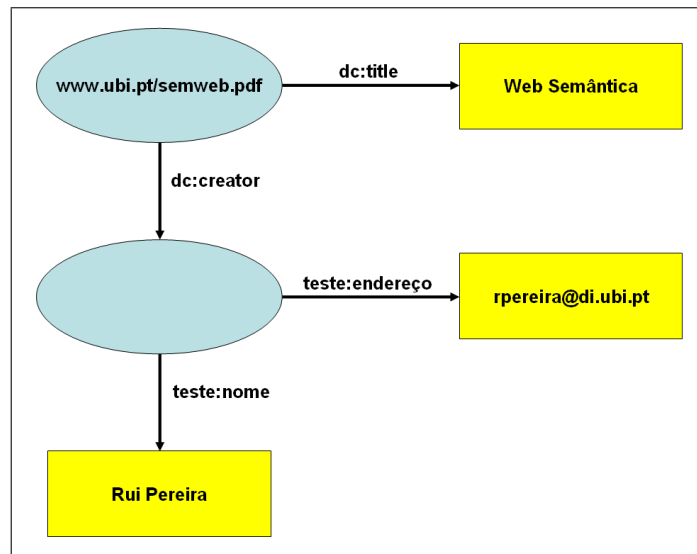


Figura A.2: Utilização de "recursos anónimos" em asserções RDF.

mais do que uma característica sobre um mesmo valor de um recurso RDF. Ela pode ser representada como um novo recurso. Quando esse novo recurso não possui um nome é conhecido por "recurso anónimo" e a sua representação é feita através de uma elipse vazia. Na figura A.2 é apresentada uma estrutura RDF criada a partir da seguinte declaração:

A pessoa cujo nome é Rui Pereira e o seu e-mail é rpereira@di.ubi.pt, é o autor do recurso `www.di.ubi.pt/semweb.pdf` intitulado de Web Semântica.

A representação utilizando a sintaxe RDF/XML é dada por:

```
<?xml version= "1.0"?>
<rdf:RDF xmlns:rdf=
  "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  xmlns:teste="www.di.ubi.pt/teste">
  <rdf:Description
    about="www.di.ubi.pt/semweb.pdf">
    <dc:titulo>Web Semântica</dc:titulo>
```

```

<dc:creator>
  <teste:nome>Rui Pereira</teste:nome>
  <teste:endereco>rpereira@di.ubi.pt</teste:endereco>
</mod:creator>
</rdf:Description>
</rdf:RDF>

```

Apenas a título de exemplo, utilizou-se um novo Espaço de Nomes, denominado por "teste" e definido pelo URI "www.di.ubi.pt/teste".

- Coleções de Objectos: Possibilidade de criar uma coleção de valores associada a uma propriedade/recurso. A sintaxe RDF define três tipos de Coleções de objectos:

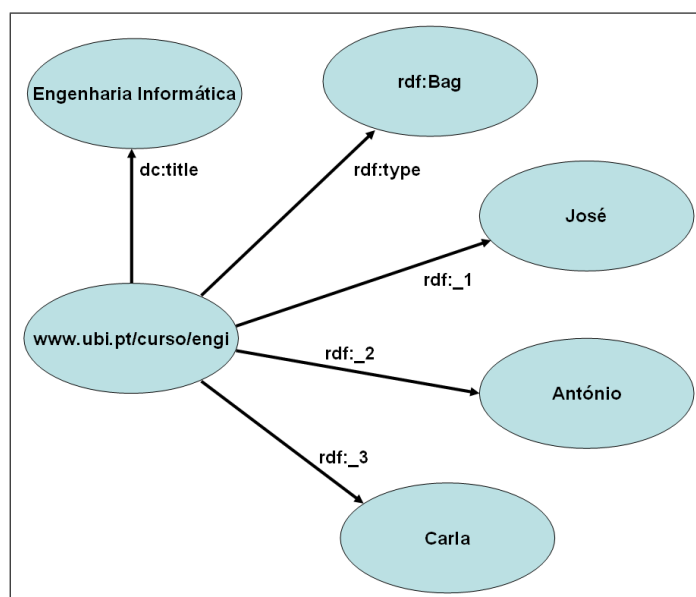


Figura A.3: Coleção de Objectos do tipo Bag.

- 'Bag'<sup>164</sup>: Define uma lista de recursos, ou de valores, não ordenados, permitindo a existência de valores duplicados, que pode ser associada a uma propriedade/recurso. Por exemplo, a figura A.3 apresenta e especifica que um curso representado pelo

<sup>164</sup>Em língua Portuguesa: saco.

recurso [www.ubi.pt/curso/EI](http://www.ubi.pt/curso/EI) é constituído pelos alunos José, António e Carla. A propriedade 'rdf:type' especifica o tipo de Coleção de Objectos<sup>165</sup> que está a ser utilizado. Cada elemento da colecção é nomeado através de propriedades numeradas como "\_1", "\_2", "\_3" e assim sucessivamente. De seguida é apresentada a mesma representação na sintaxe RDF/XML:

```
<?xml version= "1.0"?>
<rdf:RDF xmlns:rdf=
  "http://www.w3.org/1999/02/22-rdf-syntax-ns"
  xmlns:teste="www.di.ubi.pt/teste">
<rdf:Description about=
  "www.ubi.pt/curso/engi">
  <dc:titulo>Engenharia Informática</dc:titulo>
  <teste:alunos>
    <rdf:Bag>
      <rdf:_1 resource=
        "www.ubi.pt/alunos/José"/>
      <rdf:_2 resource=
        "www.ubi.pt/alunos/António"/>
      <rdf:_3 resource=
        "www.ubi.pt/alunos/Carla"/>
    </rdf:Bag>
  </teste:alunos>
</rdf:Description>
</rdf:RDF>
```

---

<sup>165</sup>Neste caso do tipo: 'rdf:Bag'.

- 'Sequence'<sup>166</sup>: Define uma lista de recursos, ou valores, idêntica à Colecção de Objectos 'Bag', só que neste caso, ela pode ser ordenada alfabeticamente ou numericamente.
  - 'Alternative'<sup>167</sup>: É uma lista de recursos, ou valores, alternativos e de escolha livre que podem ser associados a uma propriedade/recurso.
- Asserções baseadas em asserções: o modelo RDF permite também descrever as próprias asserções. Para isso, ele utiliza as entidades Sujeito, Predicado, Objecto e Tipo (*type*). Por exemplo, a seguinte frase é constituída por duas asserções:

O artigo "A rota do queijo" cujos autores são Carlos Pereira e Fernando Serra, foi aceite para publicação no jornal Terra da Covilhã.

Esta frase pode ser facilmente representa na sintaxe RDF, para isso basta definir:

- Sujeito: Define o recurso que é descrito pela primeira asserção, neste caso **A rota do queijo**;
- Predicado: Identifica a propriedade associada ao Sujeito da primeira asserção, neste caso **autor**;
- Objecto: Identifica o valor da propriedade associada ao Sujeito da primeira asserção, neste caso pode ser definida como uma Colecção de Objectos do tipo Bag: **Carlos Pereira e Fernando Serra**;
- Tipo: Identifica uma asserção como um novo recurso. Todas as asserções deste tipo são instâncias de `rdf:statement`. Um recurso do tipo `rdf:statement` é, por definição, composto pelas propriedades `rdf:subject`, `rdf:predicate` e `rdf:object`.

---

<sup>166</sup>Em língua Portuguesa: sequência.

<sup>167</sup>Em língua Portuguesa: alternativa.

De seguida, é apresentada a frase anterior, mas agora através da sintaxe RDF/XML:

```
<?xml version= "1.0"?> <rdf:RDF xmlns:rdf=
    "http://www.w3.org/1999/02/22-rdf-syntax-ns"
    xmlns:teste="www.di.ubi.pt/teste">
  <rdf:Description>
    <rdf:Statement>
      <rdf:subject rdf:resource=
        "www.escriptor.carlosP/rotaQueijo.pdf"/>
      <rdf:predicate rdf:resource="teste:autor"/>
      <rdf:object>
        <rdf:bag>
          <rdf:_1 resource="Carlos Pereira"/>
          <rdf:_2 resource="Fernando Serra"/>
        </rdf:bag>
      </rdf:object>
      <rdf:type rdf:resource=
        "http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
    </rdf:Statement>
    <teste:aceite_por>
      "http://www.jornal.terra.Covilhã.pt"
    </teste:aceite_por>
  </rdf:Description>
</rdf:RDF>
```





# Apêndice B

## OWL

De entre as várias linguagens ontológicas actualmente desenvolvidas, a *Web Ontology Language* (OWL) é aquela que mais se aproxima da camada Ontologia apresentada na arquitectura da Web Semântica. Ela foi desenvolvida com o objectivo de contribuir, de uma forma simples e universal, para a criação e partilha de ontologias na Web Semântica.

É baseada na linguagem DAML+OIL<sup>168</sup> e tal como ela pode ser considerada como um vocabulário extensivo<sup>169</sup> do RDF. Ao contrário da maioria da linguagens ontológicas que foram desenvolvidas para serem usadas em comunidades específicas, a OWL foi desenvolvida para ser compatível com a arquitectura da Web de uma forma geral e com a arquitectura da Web Semântica de uma forma particular. Desta forma, ela fornece uma estrutura para a gestão, partilha e reutilização de informação através da Web independentemente do tipo de ferramentas utilizado e, além disso, encontra-se aprovada pelo W3C como uma **Recomendação W3C** desde o dia 10 de Fevereiro do ano de 2004 [5].

A linguagem OWL possibilita que a informação contida em documentos possa ser realmente

---

<sup>168</sup>Linguagem desenvolvida por uma equipa internacional fundada pela DARPA e pelo grupo *Information Science Technologies* (IST) da Comissão Europeia.

<sup>169</sup>A linguagem OWL é implementada a partir do Esquema RDF, ao qual adiciona mais vocabulário para descrever classes e propriedades: relações entre classes (e.g. *disjointness*), cardinalidade (e.g. *"exactly one"*), igualdade, definições mais ricas de tipos de propriedades, características das propriedades (e.g. simetria) e classes enumeradas.

processada por aplicações e não apenas utilizada para ser apresentada de forma visual às pessoas. Ela pode ser utilizada para representar explicitamente o significado dos termos existentes em vocabulários e as relações existentes entre esses termos.

Segundo Jim Hendler e Guus Schreiber, directores do Grupo de Trabalho *Web Ontology Working Group* [310], a OWL constitui um gigantesco passo em frente na representação e organização do conhecimento da Web. Ela equilibra de forma simples e eficiente as necessidades exigidas quer pela indústria (uma linguagem compatível com os modelos actuais utilizados na Web) quer pelas restrições inerentes ao desenvolvimento de uma linguagem ontológica que incorpore os resultados das mais recentes investigações da área. Este grupo de trabalho é constituído por mais de cinquenta membros que representam a actual investigação, quer a nível industrial, quer a nível académico da OWL.

Durante o seu funcionamento, o *Web Ontology Working Group* desenvolveu seis documentos sobre a OWL que foram considerados a base que contribuiu e possibilitou a obtenção da **Recomendação W3C** por parte da W3C:

- *OWL Web Ontology Language Overview* [213]: É um documento que apresenta, de forma geral, as características da linguagem OWL;
- *OWL Web Ontology Language Guide* [69]: É um documento que apresenta, de forma mais detalhada, as características da linguagem OWL bem como alguns exemplos de utilização;
- *OWL Web Ontology Language Reference* [311]: É um documento que serve como referência para os detalhes de cada uma das características da linguagem OWL;
- *OWL Web Ontology Language Semantics and Abstract Syntax* [312]: É um documento que apresenta a semântica da linguagem OWL e detalhes sobre a conversão de informação de OWL para RDF;
- *OWL Web Ontology Language Test Cases* [313]: É um documento que apresenta vários

casos de testes práticos efectuados durante a utilização da linguagem OWL;

- *OWL Web Ontology Language Use Cases and Requirements* [314]: É um documento que apresenta exemplos práticos de utilização da linguagem OWL.

A linguagem OWL baseia-se na utilização do conceito de URI, na sintaxe RDF/XML e na linguagem Esquema RDF para gerar ontologias com as seguintes características [315]:

- Facilidade de serem manipuladas de forma distribuída através de diferentes sistemas;
- Versáteis às necessidades da Web;
- Compatíveis com as normalizações web referentes à acessibilidade e à internacionalidade;
- Utilização de uma especificação aberta e facilmente extensível.

Segundo Deborah McGuinness e Frank Van Harmelen, a linguagem OWL pode ser dividida em três sub-linguagens [213]:

- *OWL Lite*: Deverá ser utilizada quando se desejar implementar uma classificação hierárquica simples<sup>170</sup> e que necessite apenas de ser apoiada em simples restrições<sup>171</sup>;
- *OWL Description Logic* (DL): É uma linguagem um pouco mais complexa que a *OWL Lite*. Deverá ser utilizada quando se pretender associar o máximo de expressividade à informação descrita, mas de uma forma que não coloque em causa a sua correcta interpretação e atempada execução por parte dos SPC. A linguagem OWL DL implementa todos os construtores da linguagem OWL, apesar de alguns deles apenas poderem ser utilizados de acordo com certas restrições<sup>172</sup>.

---

<sup>170</sup>Por exemplo, uma Taxonomia ou um Thesaurus.

<sup>171</sup>O exemplo de uma restrição é por exemplo: "Apenas serem permitidas associações de cardinalidade de valor 0 ou 1."

<sup>172</sup>Por exemplo: "Apesar de uma classe poder ser uma sub-classe de várias classes, ela nunca poderá ser considerada uma instância de uma outra classe."

- *OWL Full*: Apenas deverá ser utilizada quando se pretender associar o máximo de expressividade da informação independentemente da existência ou não de garantias da sua correcta interpretação e atempada execução por parte dos SPC. Por exemplo, na linguagem *OWL Full*, uma classe pode ser simultaneamente referida como uma colecção de indivíduos ou apenas como um indivíduo específico. Além disso permite que os vocabulários predefinidos, RDF, Esquema RDF e OWL possam ser estendidos, o que poderá levantar problemas de compatibilidade entre diferentes interpretações por parte de diferentes SPC.

A sub-linguagem *OWL Full* pode ser vista como uma extensão da sintaxe RDF, por outro lado, as sub-linguagens *OWL Lite* e *OWL DL* podem ser vistas como extensões restritas da sintaxe RDF. Qualquer documento OWL pode ser considerado um documento RDF e qualquer documento RDF pode, por sua vez, ser considerado um documento *OWL Full*. No entanto, só alguns documentos RDF é que podem ser considerados documentos *OWL Lite* ou *OWL DL*.

A seguir é apresentada uma breve descrição dos termos definidos na linguagem OWL. Todos os termos apresentados que já existam na sintaxe RDF ou sintaxe Esquema RDF serão precedidos dos seguintes prefixos `rdf:` e `rdfs:` [311]:

- Termos utilizados que se encontram relacionados com o Esquema RDF:
  - **Class**: Uma Classe define um grupo de indivíduos que compartilham algumas propriedades em comum. As classes podem ser organizadas numa hierarquia através do uso do termo `subClassOf`. Existe uma definição geral de classe denominada **Thing**, a qual é a classe de todos os indivíduos e é super-classe de todas as classes OWL. Há também uma definição específica de classe denominada **Nothing**, a qual, por sua vez, não possui nenhuma instância associada e é sub-classe de todas as classes OWL;
  - `rdfs:subClassOf`: Este termo é utilizado para construir hierarquias de classes através da declaração que determinadas classes são sub-classes de outras classes;

- **rdf:Property**: As propriedades podem ser usadas para definir relacionamentos que possam existir, quer entre indivíduos, quer entre os indivíduos e os seus valores. Exemplo de propriedades OWL: **hasChild**, **hasRelative**, **hasSibling** e **hasAge**.
  - **rdfs:subPropertyOf**: As hierarquias de propriedades podem ser criadas fazendo uma ou mais indicações, através do termo **rdfs:subPropertyOf**, de que uma propriedade é um sub-propriedade de uma ou mais propriedades.
  - **rdfs:domain**: Um domínio de uma propriedade limita os indivíduos a que a propriedade pode ser aplicada. Se uma propriedade relacionar um primeiro indivíduo a um segundo indivíduo, e a propriedade tiver uma classe como um de seus domínios, então o segundo indivíduo deve pertencer a essa classe.
  - **rdfs:range**: Uma propriedade pode limitar, através do termo **rdfs:range**, o intervalo de indivíduos que a propriedade pode ter como seu valor. Se uma propriedade relacionar um indivíduo com um outro indivíduo e a propriedade tiver uma classe específica no seu intervalo de indivíduos, então o outro indivíduo deve obrigatoriamente pertencer a essa classe específica.
  - **Individual**: Os indivíduos são instâncias de classes e propriedades que podem ser usadas para relacionar os indivíduos entre si.
- Termos utilizados que se encontram relacionados com a Igualdade e a Desigualdade:
    - **equivalentClass**: Duas classes podem ser declaradas como sendo equivalentes/ sinónimas.
    - **equivalentProperty**: Duas propriedades podem ser declaradas como sendo equivalentes/ sinónimas.
    - **sameAs**: Dois indivíduos podem ser declarados como sendo o mesmo indivíduo. O uso deste termo é útil para associar diferentes **nomes** a um mesmo indivíduo.
    - **differentFrom**: Um indivíduo pode ser declarado como sendo diferente de outros indivíduos.

- AllDifferent: Os elementos de um grupo de indivíduos podem ser declarados mutuamente distintos;
- Termos utilizados que se encontram relacionados com Propriedades:
  - inverseOf: Uma propriedade pode ser declarada como sendo a inversa de uma outra propriedade. Se a propriedade P1 for declarada como sendo a inversa da propriedade P2 e se X for associado a Y pela propriedade P2 então Y ficará associado a X pela propriedade P1.
  - TransitiveProperty: As propriedades podem ser declaradas como sendo transitivas. Se uma propriedade P for transitiva e se os pares  $(x,y)$  e  $(y,z)$  forem instâncias da propriedade P, então o par  $(x,z)$  será também uma instância de P.
  - SymmetricProperty: As propriedades podem ser declaradas como sendo simétricas. Se uma propriedade P for simétrica e se o par  $(x,y)$  for uma instância da propriedade P, então o par  $(y,x)$  também será uma instância da propriedade P.
  - FunctionalProperty: Se uma propriedade for declarada como FunctionalProperty então apenas pode possuir zero ou um valor para cada um dos seus indivíduos associados. Esta propriedade funciona como uma referência rápida para declarar que o cardinalidade mínima da propriedade é zero e que a cardinalidade máxima é 1.
  - InverseFunctionalProperty: Se uma propriedade for declarada como InverseFunctionalProperty então a sua inversão é FunctionalProperty.
- Termos utilizados que se encontram relacionados com Propriedades Restritivas:
  - allValuesFrom: Este termo é utilizado quando se pretende associar um intervalo de restrições, de carácter obrigatório, a uma propriedade P em relação a uma classe em particular. Desta forma só se poderá associar um outro indivíduo a esta classe através da propriedade P se ele for uma instância do mesmo intervalo de restrições associado à propriedade P.

- **someValuesFrom**: Este termo é utilizado quando se pretende associar um intervalo de restrições, sendo pelo menos uma obrigatória, a uma propriedade P em relação a uma classe em particular. Desta forma só se poderá associar um outro indivíduo a esta classe através desta propriedade P se ele for uma instância de pelo menos uma das restrições do intervalo de restrições associado à propriedade P.
- Termos utilizados que se encontram relacionados com Restrições de Cardinalidade:
  - **minCardinality**: A cardinalidade é associada a uma propriedade em relação a uma classe em particular. Se for associada uma cardinalidade mínima (**minCardinality**) de valor um a uma propriedade em relação a uma classe em particular, todas as instâncias dessa classe estarão relacionadas com pelo menos um indivíduo através dessa propriedade. A utilização deste termo é o mesmo que declarar uma restrição em termos da quantidade mínima de indivíduos que podem ser associados a uma propriedade.
  - **maxCardinality**: Inverso do termo **minCardinality**.
  - **cardinality**: Permite definir um valor único, quer para o mínimo, quer para o máximo de cardinalidade, que pode ser associada a uma propriedade em relação a uma classe em particular.
- Termos utilizados que se encontram relacionados com a Intersecção de Classes:
  - **intersectionOf**: Este termo permite relacionar classes com restrições;
- Termos utilizados que se encontram relacionados com outras Características:
  - **oneOf**: (classes enumeradas): Uma classe pode ser descrita através da enumeração dos todos os indivíduos que constituem a classe.
  - **hasValue** (valores de propriedades): Uma propriedade pode ser especificada como possuindo um determinado indivíduo como seu valor;

- `disjointWith`: As classes podem ser declaradas como não possuindo qualquer tipo de relação entre si.
- `unionOf`, `complementOf`, `intersectionOf` (Combinações lineares): Combinação booleanas entre classes e restrições;

Apesar de a linguagem OWL ser uma *Recomendação W3C* há relativamente pouco tempo, o número de ferramentas que de alguma forma a utilizam ou a interpretam é enorme. Por exemplo, do estudo apresentado no capítulo 4, quarenta e nove das oitenta ferramentas apresentadas utilizam de alguma forma a linguagem OWL ou ontologias OWL.



# Apêndice C

## Diagramas de Classes em UML da ferramenta SWedt

Este apêndice apresenta diagramas de classes da ferramenta SWedt através da *Unified Modeling Language* (UML). A UML é uma linguagem de modelagem usada para especificar, visualizar, construir e documentar entidades<sup>173</sup> de um sistema de software orientado a objectos. Basicamente, permite visualizar as várias entidades de um projecto através de diagramas normalizados. Devido ao elevado número de classes definidas na ferramenta SWedt serão apresentados doze diagramas de classes em UML.

O diagrama UML da figura C.1 apresenta as principais classes Java definidas na ferramenta SWedt. A maioria das classes encontra-se agrupadas em Pacotes<sup>174</sup>. Na figura, identificam-se cinco classes: SWedtCore, ISWedtModel, SwedtPlugin, ISwedtConstants, ISWedtModel e cinco Pacotes: perspective, editors, views, dom, resources. Em suma, este diagrama engloba todas as classes definidas durante o desenvolvimento da ferramenta, sendo os seguintes onze diagramas especificações deste primeiro diagrama. No caso dos Pacotes DOM (figura C.2), RDF (figura C.3), RDFS (figura C.4) e OWL (figura C.5), as classes definidas correspondem aos termos utilizados pelos respectivos interpretadores.

---

<sup>173</sup>Neste caso em particular, classes.

<sup>174</sup>Do termo em inglês: *Packages*.

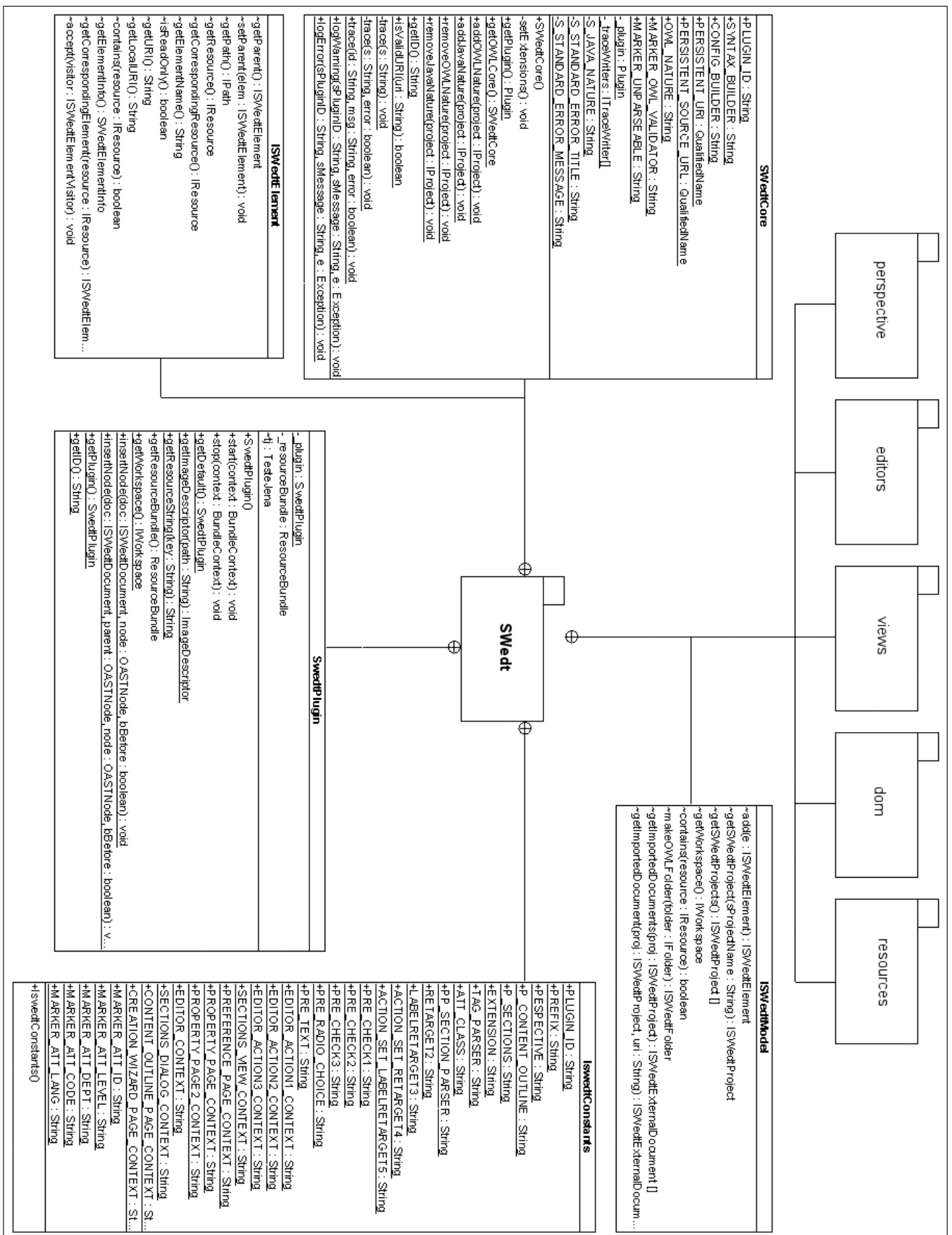


Figura C.1: Diagrama de Classes da ferramenta SWedt.

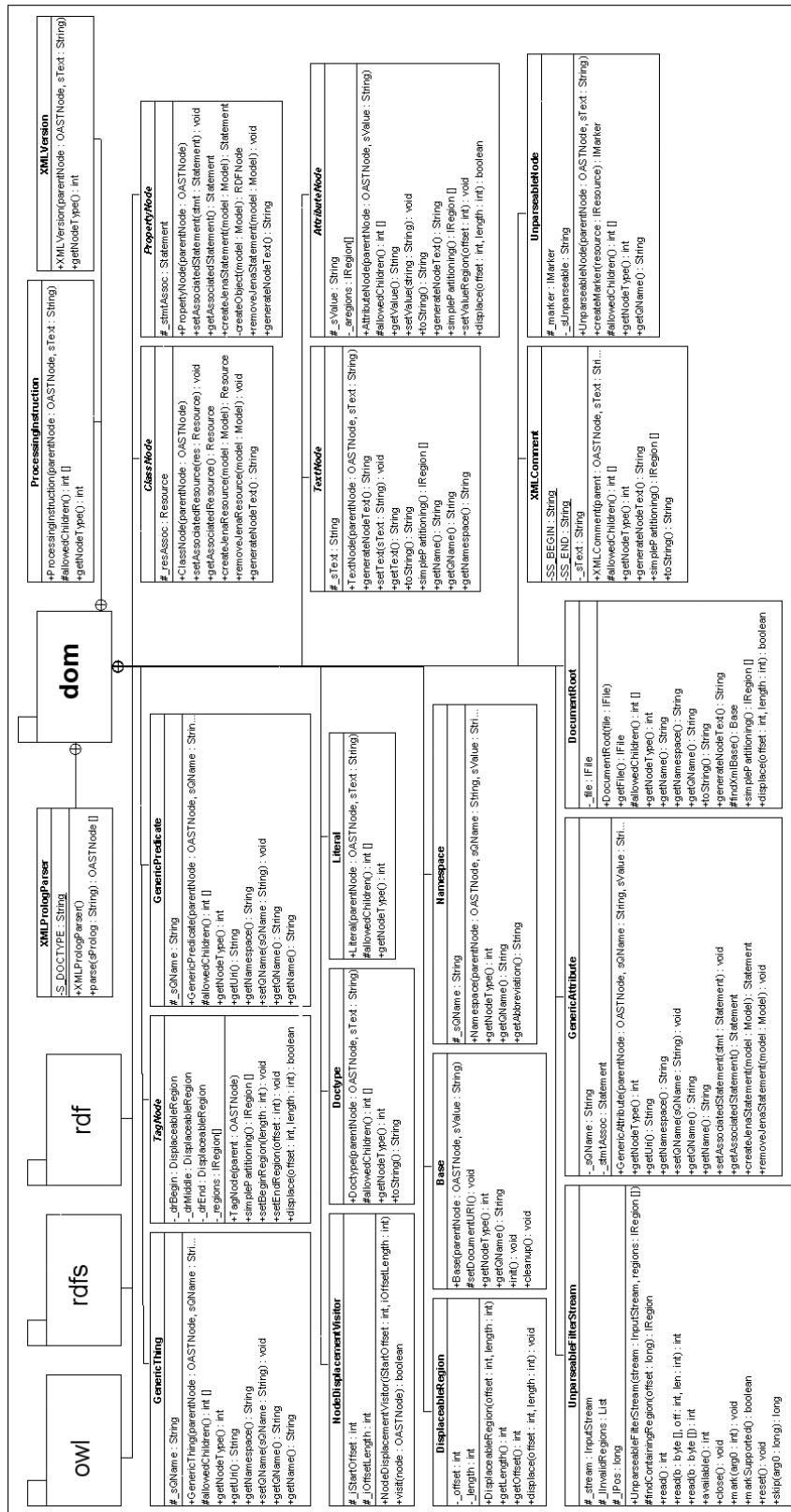


Figura C.2: Diagrama de Classes do Pacote DOM.

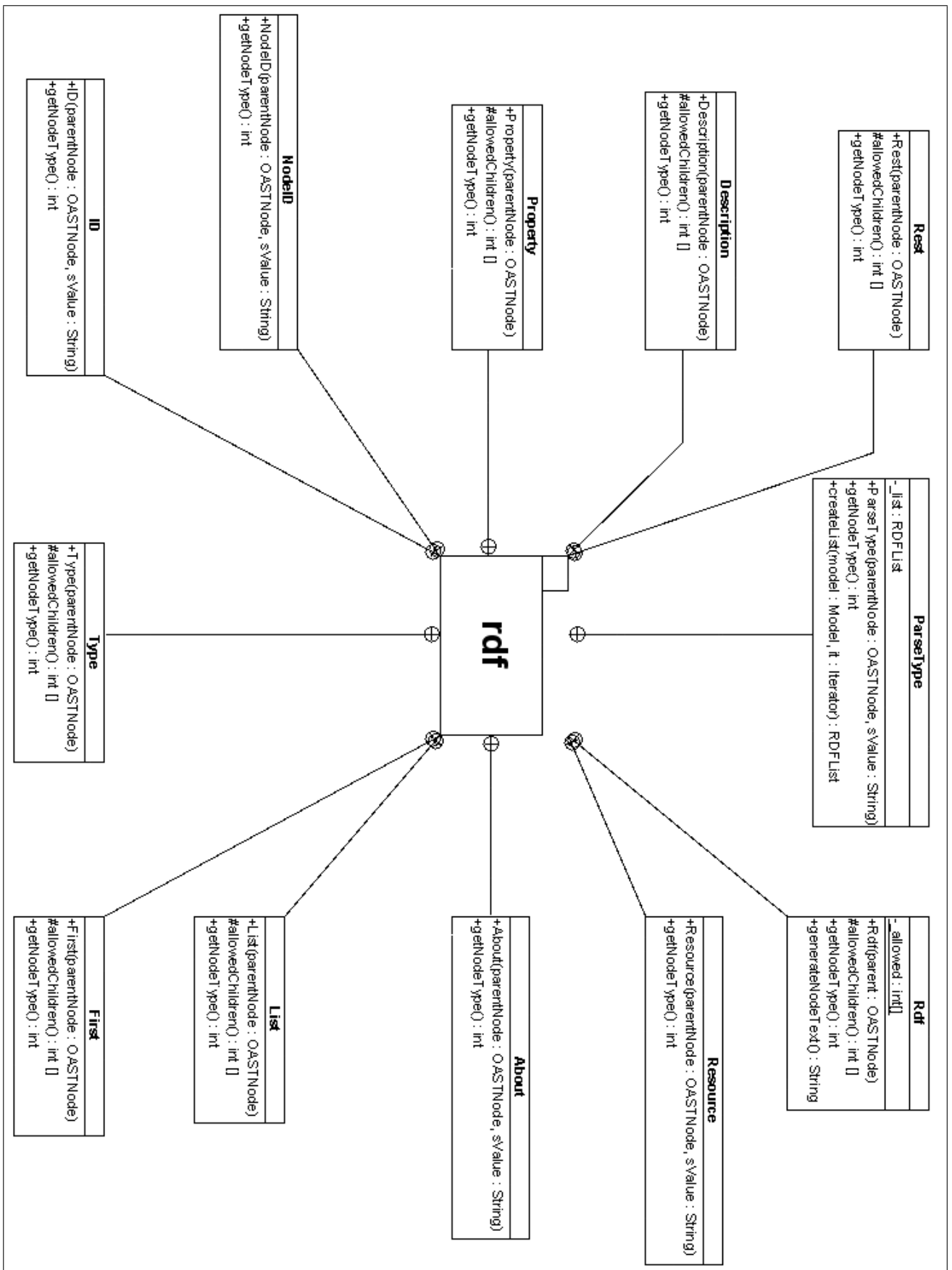


Figura C.3: Diagrama de Classes do Pacote RDF.

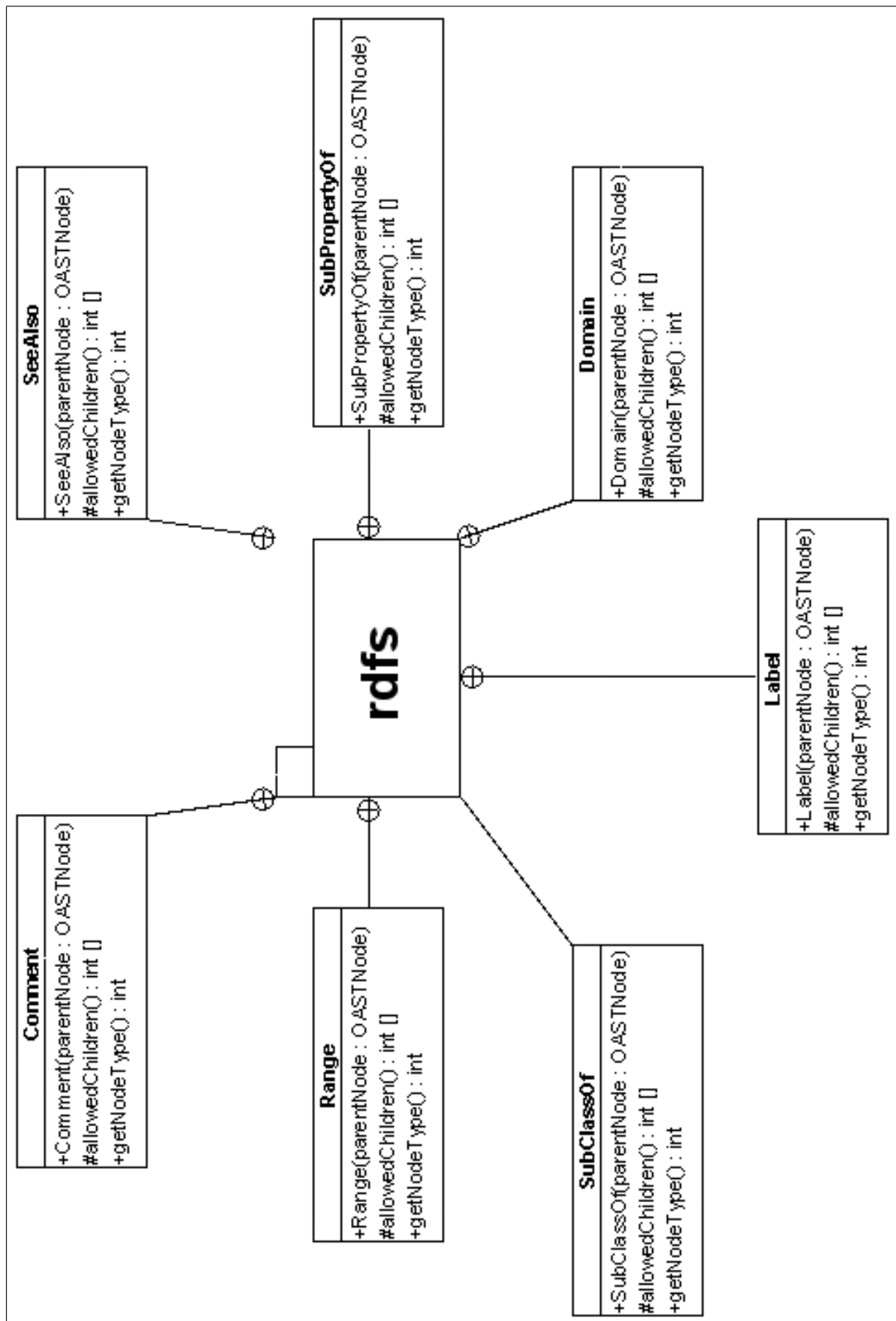


Figura C.4: Diagrama de Classes do Pacote RDFS.



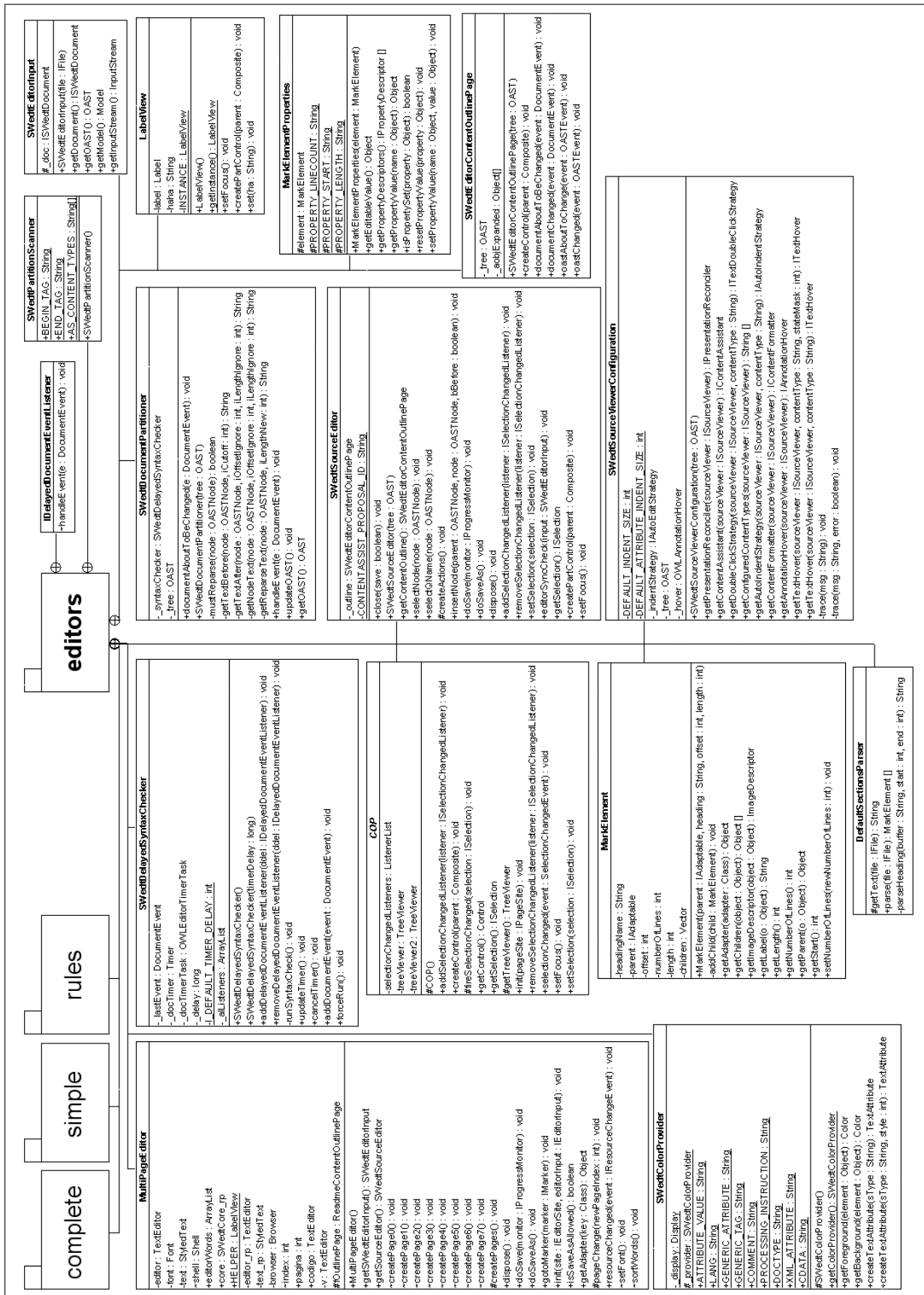


Figura C.6: Diagrama de Classes do Pacote Editors.

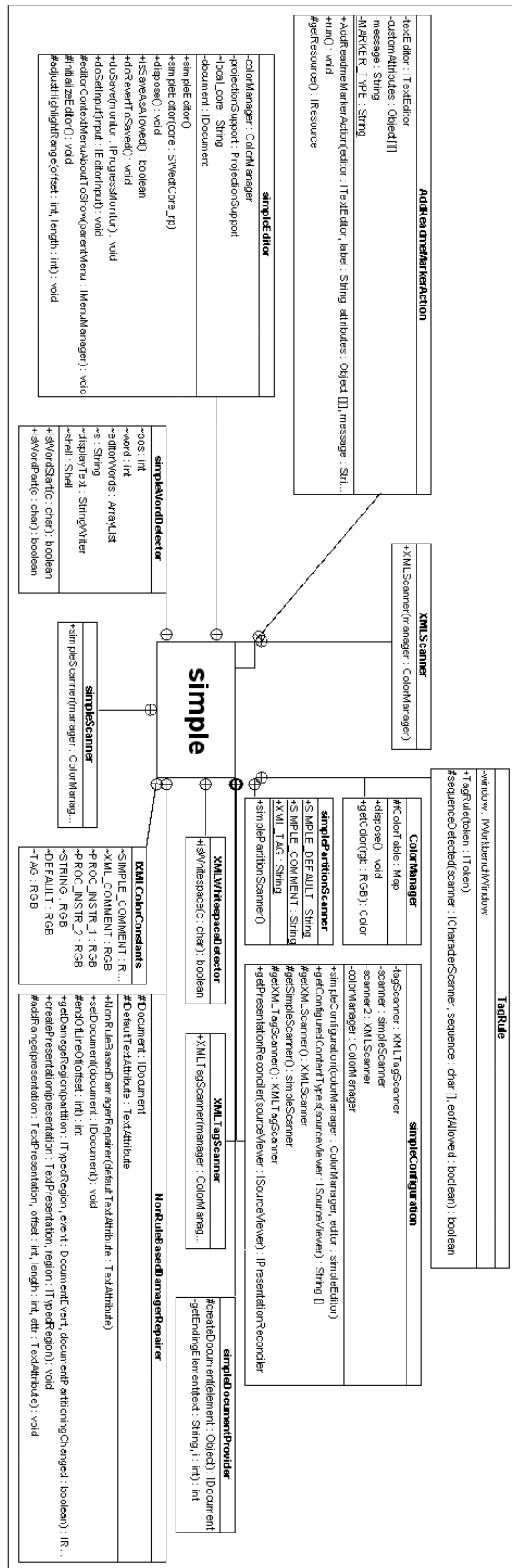


Figura C.7: Diagrama de Classes do Pacote Simple.



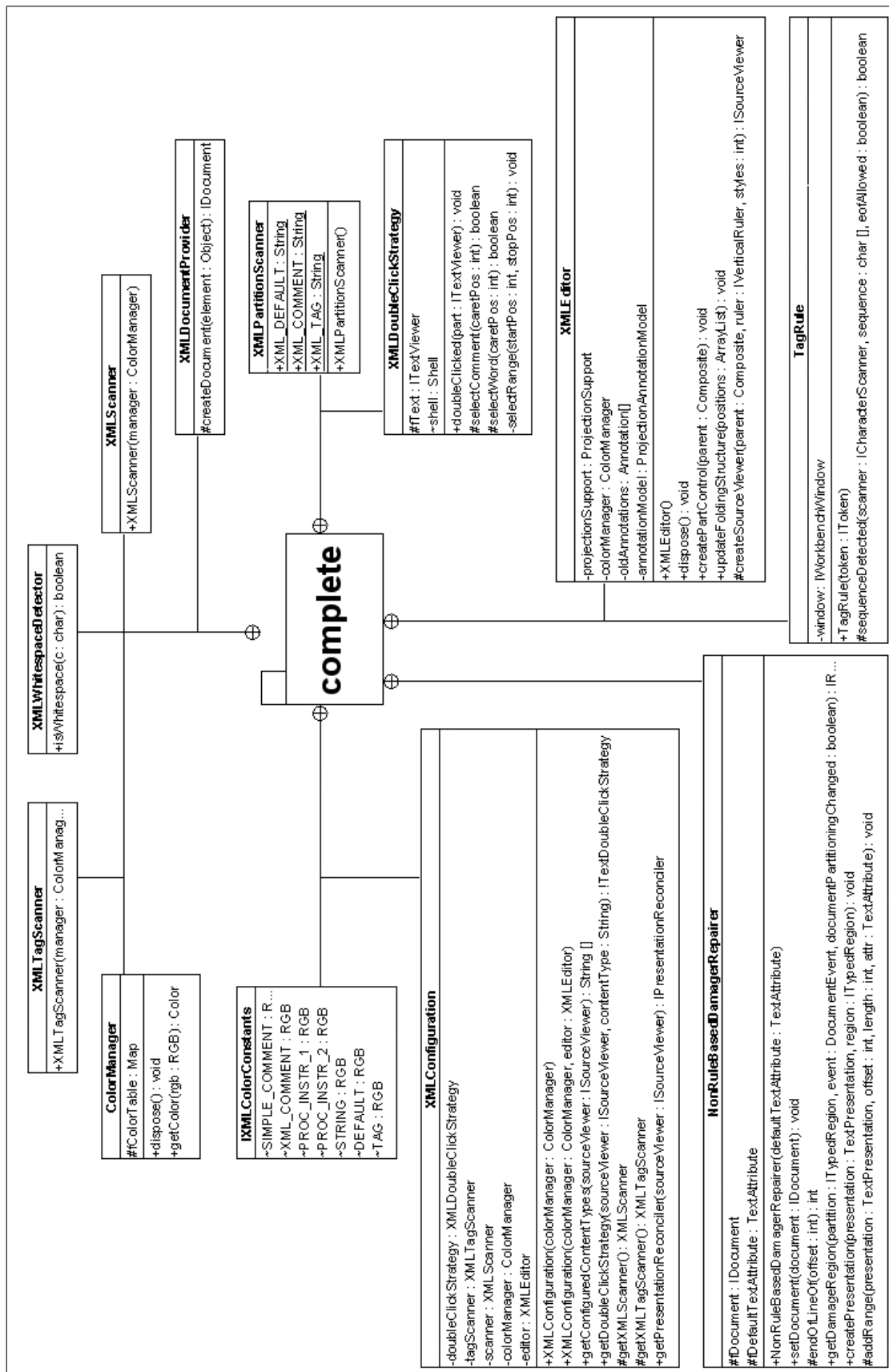


Figura C.8: Diagrama de Classes do Pacote Complete.

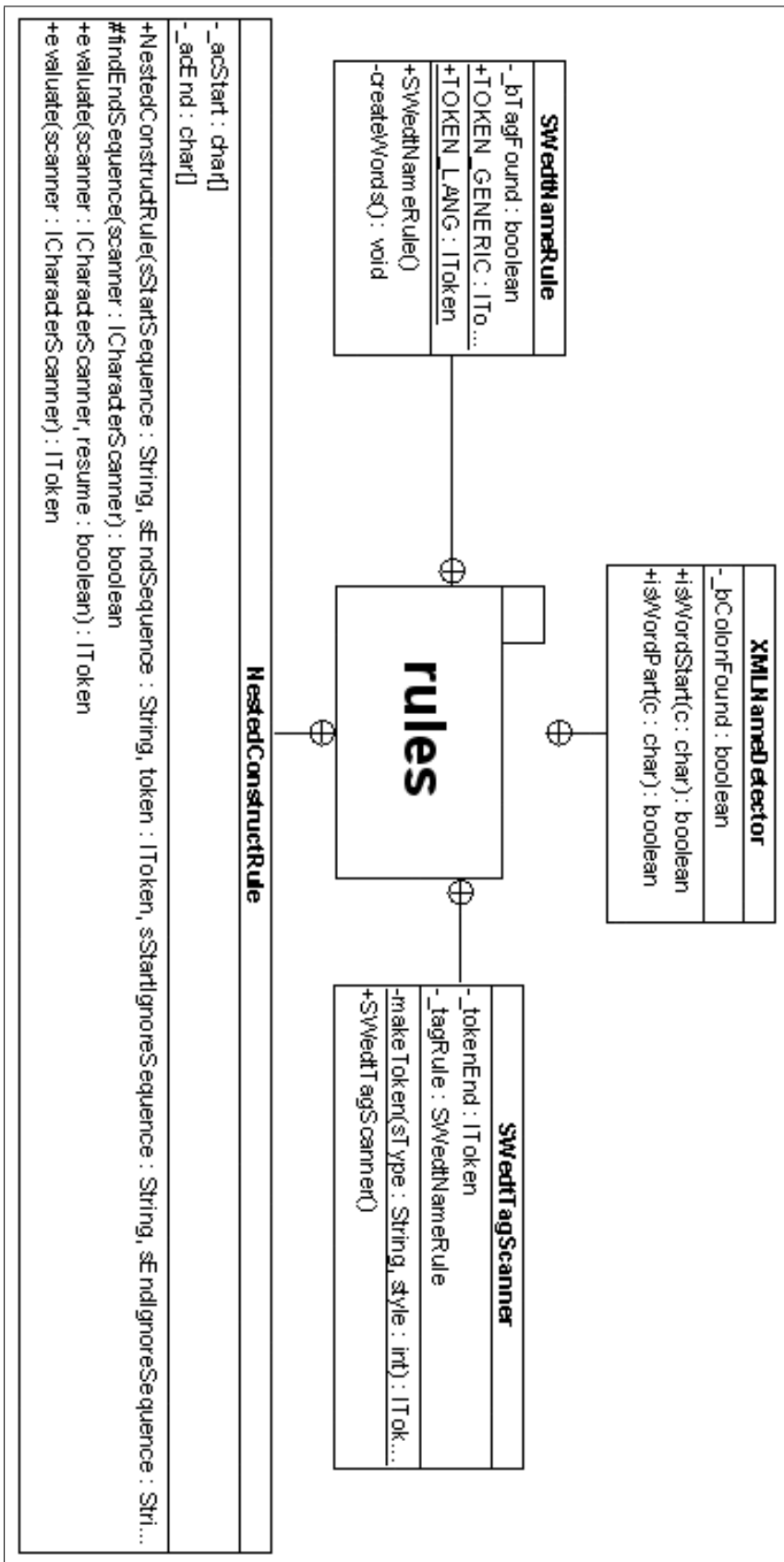


Figura C.9: Diagrama de Classes do Pacote Rules.

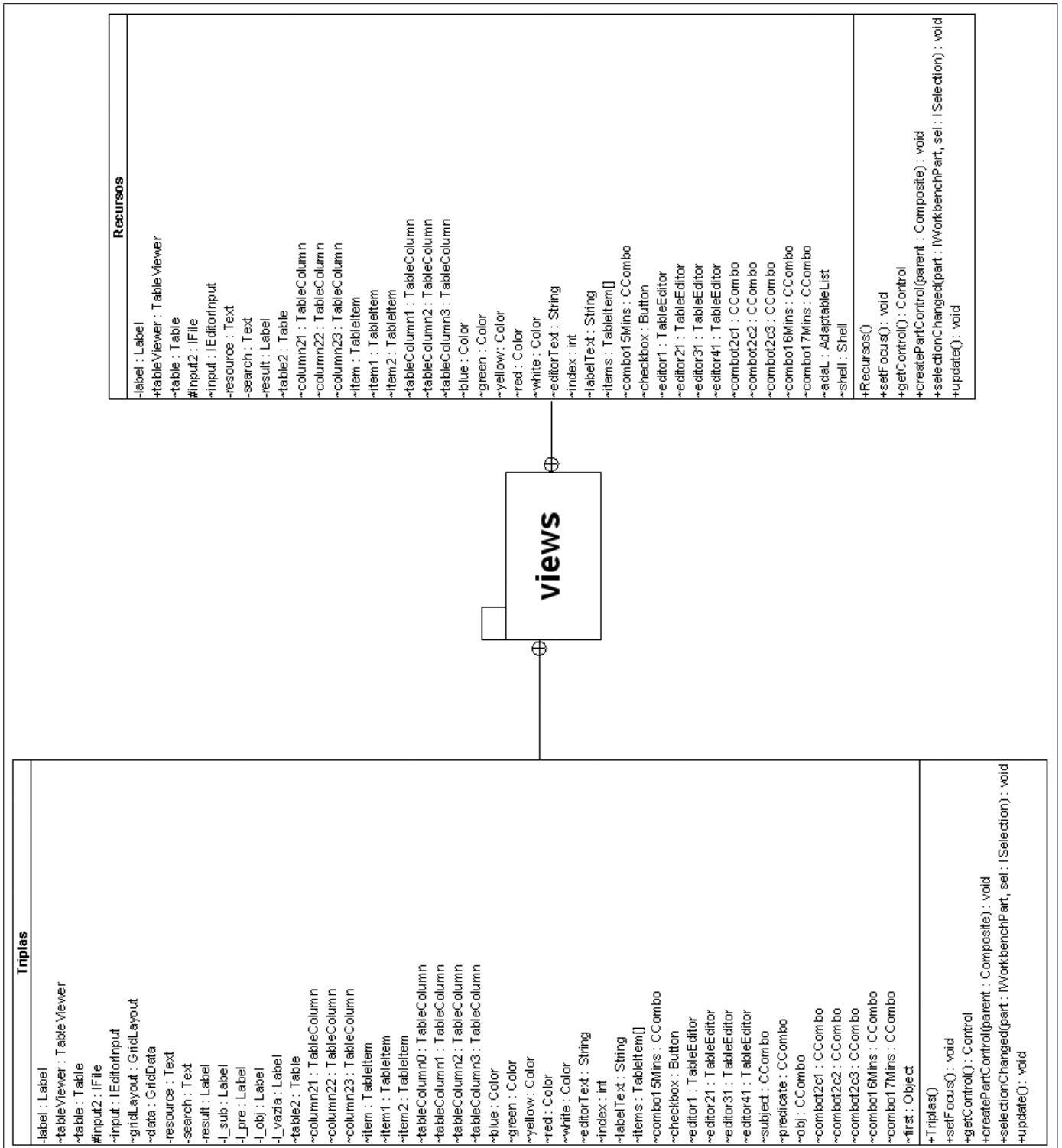


Figura C.10: Diagrama de Classes do Pacote Views.

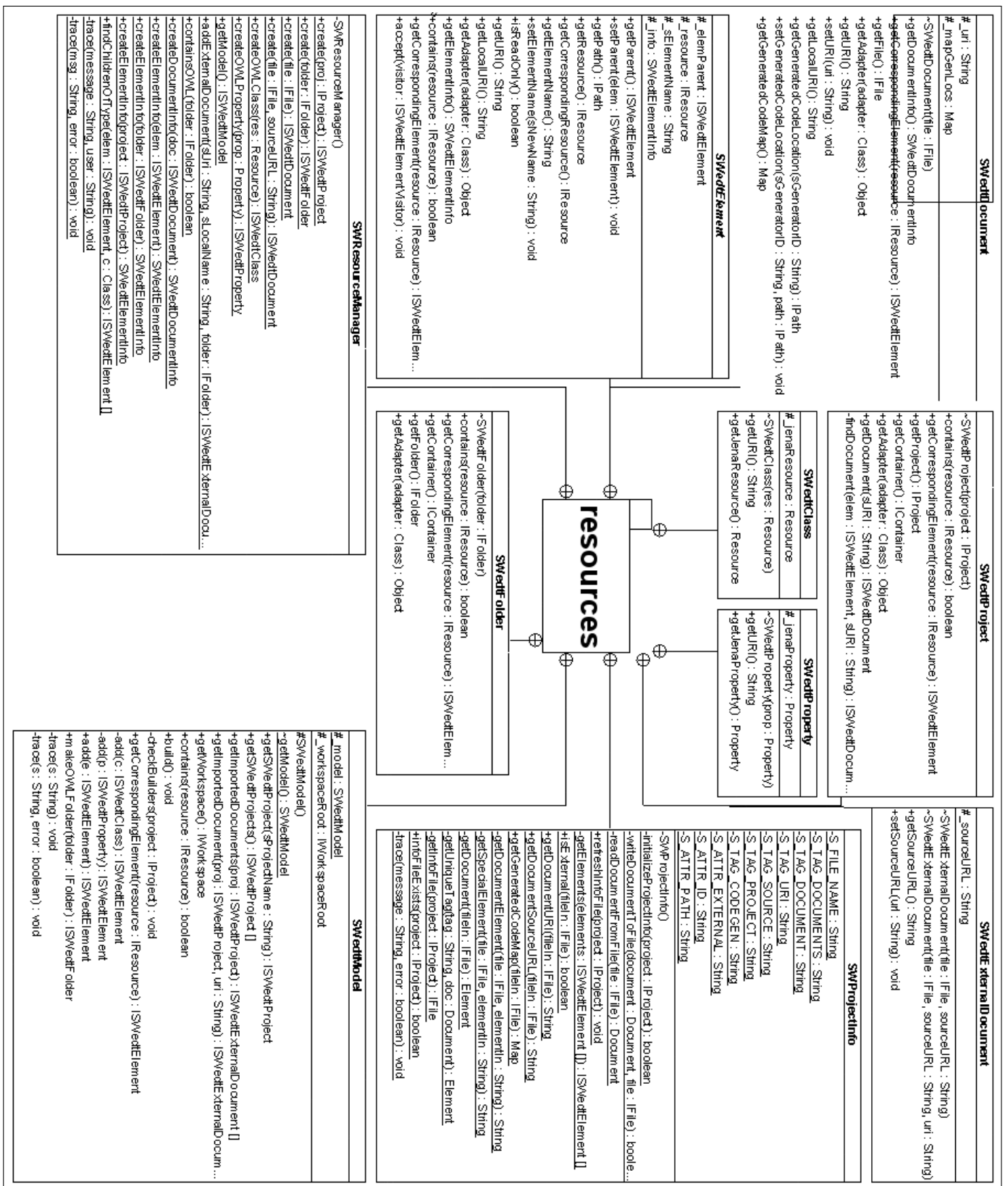


Figura C.11: Diagrama de Classes do Pacote Resources.

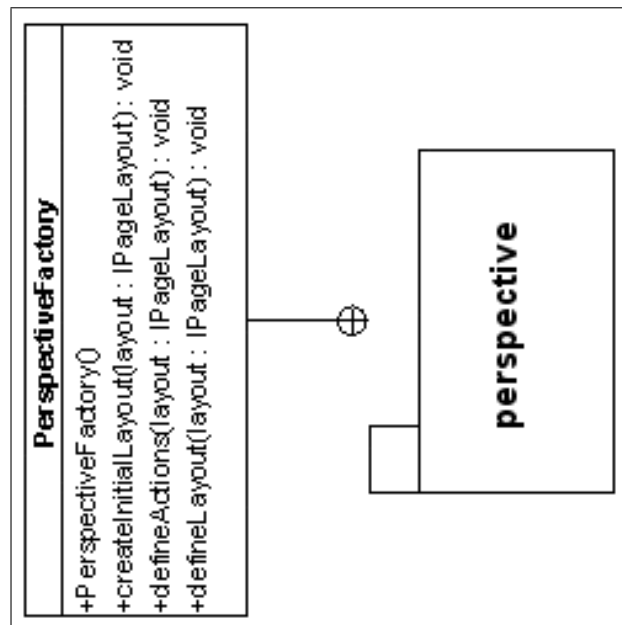


Figura C.12: Diagrama de Classes do Pacote Perspective.