

Rumo a uma Web mais inteligente

Vitor Manuel Barrigão Gonçalves

Escola Superior de Educação – Instituto Politécnico de Bragança
vg@ipb.pt

Eurico Manuel Carrapatoso

Faculdade de Engenharia da Universidade do Porto / INESC Porto
emc@fe.up.pt



Resumo:

A *World Wide Web* foi criada com a visão de que seria um espaço onde a informação teria um significado bem definido, facilitando a cooperação e a comunicação entre as pessoas e os agentes computacionais [Berners-Lee et al. 2001]. No entanto, essa visão não foi ainda alcançada, uma vez que a Web se encontra apenas organizada na perspectiva dos humanos, tornando a busca de informação nesta rede de documentos uma tarefa árdua. Para que a Web contemple a perspectiva das máquinas, torna-se necessário reduzir a ambiguidade da linguagem natural e aumentar a precisão na recuperação de informação, através da implementação de agentes computacionais inteligentes.

Exige-se o aparecimento de motores de busca capazes de reunir informação de documentos dispersos, processá-la e inferir novos conteúdos de acordo com a pesquisa submetida e de modo a responder eficientemente aos requisitos dos utilizadores. Por conseguinte, é necessário que a Web seja suportada por uma semântica que atribua um significado bem definido à informação disponibilizada. É neste contexto que surge a denominada Web Semântica ou Inteligente (*Semantic Web*) [W3C 2003] como um cenário desejável para o futuro da Web.

O objectivo deste ensaio teórico é traçar o estado da arte da Web Semântica dando a conhecer um conjunto de tecnologias da informação e do conhecimento, que tem vindo a ser considerado como base para o desenvolvimento de aplicações que tornem a Web mais inteligente.

Palavras-chave:

Web Semântica, Linguagem de Marcação Estendida (XML), Metadados, Arquitectura de Descrição de Recursos (RDF), Ontologias, Agentes Web Inteligentes.

KeyWords:

Semantic Web, eXtensible Markup Language (XML), Metadata, Resource Description Framework (RDF), Ontologies, Intelligent Web Agents.

1- Introdução

Cada vez mais as Tecnologias de Informação e Comunicação (TICs) possibilitam a criação de ambientes baseados na Internet que visam a recolha, armazenamento, disseminação, recuperação e apresentação de informação acerca de qualquer área do conhecimento.

A Internet, mais concretamente a *World Wide Web* (WWW ou simplesmente Web), constitui a infra-estrutura ideal para a publicação e disseminação de informação (conteúdos), nomeadamente devido à independência de plataforma, simplicidade dos protocolos e facilidade de uso e acesso [Bieber 1997]. Estas condições impulsionaram o aparecimento de vários sítios de informação (*sites Web*), desde as Páginas Web Pessoais até aos mais complexos Sistemas de Informação para a Web.

A Web tornou-se numa biblioteca de documentos à escala mundial, constituindo actualmente o maior repositório de informação que dispo-

nibiliza conteúdos em páginas estáticas ou dinâmicas e oferece serviços para praticamente todos os sectores da sociedade.

Face à proliferação de conteúdos na Web, surgiram mecanismos de pesquisa, cuja principal função é facilitar a exploração e recuperação de informação, quer através de directórios (por exemplo: Yahoo! ou Lycos), onde a informação se encontra catalogada de acordo com as descrições das páginas fornecidas por quem as submete, quer através de motores de busca de informação (por exemplo: Google ou Altavista), que baseiam a busca de informação em palavras-chave, em robots que realizam o *download* das páginas Web, em indexadores que criam índices extraindo automaticamente os termos chave das páginas ou em interfaces de consulta que comparam os termos recebidos com a base de termos indexados [Ramalho 2002]. Contudo, o constante aumento do número de servidores e de páginas Web faz com que a teia de informação assuma dimensões gigantescas, tornando difícil encontrar aquilo que se deseja ou procura, mesmo recorrendo aos agentes dos motores de busca, pois o número de resultados devolvidos é exageradamente grande e impreciso.

Actualmente, este problema tem vindo a agravar-se significativamente face à evolução da Web para um estado onde os conteúdos são cada vez mais criados pelos utilizadores. Falamos mais concretamente da evolução que tem vindo a ocorrer na Web desde 2004 e que ficou conhecida como Web 2.0 (termo derivado de uma afirmação utilizada pela *O'Reilly Media* na conferência *MediaLive Internacional* em 2004). Esta nova realidade pretende identificar a mais recente geração de serviços baseados na Internet, cujo expoente máximo pode ser encontrado nas redes sociais colaborativas *on-line* formadas por blogs, wikis e demais ferramentas de comunicação interactiva. Assim, o utilizador abandona uma atitude passiva de mero consumidor de conteúdos para se tornar um utilizador activo, produtor e distribuidor de conteúdos que alimentam a rede. Ou seja, assume-se plenamente como o verdadeiro EMEREC (o receptor é simultaneamente emissor, assumindo um papel activo na comunicação) anunciado por Jean Cloutier (1975).

Imagine que procurava informações acerca da atleta “Rosa Mota” na Internet. Como resultado da pesquisa obteria diversas páginas Web acerca de cores (cor-de-rosa), de flores (estudo ou venda de rosas), de “postais electrónicos com rosas”, da “Banda de Música Pop-Rock Rosa Mota”, da “Cidade de Santa Rosa (Califórnia)”, do “Condado de Santa Rosa (Flórida)”, dos poetas “António Ramos Rosa” ou “Rosa Alice Branco”, para além de muitas páginas onde aparecem os termos “Atleta Rosa Mota” ou “Pavilhão Rosa Mota”, entre muitas outras respostas sobre rosas e motas que dificilmente corresponderão ao sonho cor-de-rosa que tinha tido.

Embora os motores de busca, através dos seus robots e agentes, nos

libertem da tarefa monótona, repetitiva e difícil de procurar os sites Web que contêm informação acerca da temática que pretendemos, ainda nos restam tarefas de leitura, interpretação, compreensão, comparação e selecção dos resultados da pesquisa. Isto acontece porque os agentes computacionais não são capazes de processar e entender os dados, não distinguem os poetas dos atletas, as cores das flores, os nomes comuns dos nomes próprios ou as revistas digitais dos serviços de comércio electrónico. Se bem que devamos acrescentar ou retirar termos que permitam aperfeiçoar a pesquisa e indicar outros parâmetros na pesquisa avançada, tais como o título, a data de publicação, a língua, o país, o domínio ou partes da URL, a verdade é que na grande maioria das pesquisas o resultado continua a incluir dezenas ou centenas de páginas que não correspondem ao resultado pretendido, tornando difícil separar o trigo do joio.

Embora a mente humana consiga distinguir facilmente as diferenças entre uma página Web da atleta Rosa Mota e uma página Web da Banda de Pop-Rock Rosa Mota ou entre uma busca de livros de um autor e uma busca de livros acerca desse mesmo autor, as máquinas ainda não o conseguem fazer. Não basta pensar a informação na óptica dos humanos, é necessário pensá-la também na óptica das máquinas (*Machine-understandable Information*). Por conseguinte, as máquinas (ou mais propriamente os agentes computacionais) exigem uma linguagem baseada em categorias semânticas que lhes permita filtrar eficientemente, num universo de dados disponíveis, as informações que interessam ao utilizador num determinado contexto.

A possibilidade dos computadores serem não só capazes de apresentar a informação contida nas páginas Web, como também de entenderem sintáctica, semântica e morfológicamente essa mesma informação, levou vários investigadores (Tim Berners-Lee, James Hendler, Ora Lassila, John Davies e Michael Daconta, entre outros) a unirem esforços nestes últimos anos. Fruto destes esforços surgiu a Web Semântica (*Semantic Web*) como uma possível solução para tornar a Web mais inteligente.

Este ensaio teórico enquadra-se numa investigação de doutoramento que visa avaliar o impacto que esta nova visão da Web vai originar, na sociedade em geral e na Educação em particular, com vista a melhorar os mecanismos de pesquisa de informação educativa na Internet, recorrendo às tecnologias para a Web Semântica e ao desenvolvimento de agentes computacionais inteligentes no contexto educativo.

A Web constitui a maior base de conteúdos educativos para professores e alunos oferecendo oportunidades de geração de novo conhecimento que nem sempre é possível concretizar dadas as dificuldades de localização das fontes de informação. A representação do conhecimento em metadados, a construção de ontologias e o desenvolvimento de

mecanismos de inferência para o contexto educativo facilitarão não só o acesso e recuperação de informação de diversos *Websites* mas também a gestão do conhecimento nos sistemas de *e-Learning*.

Assim, após esta introdução, que teve como finalidade perspectivar a revolução que está a acontecer em prol do utilizador final (consumidor ávido de informação), este artigo aborda, no segundo capítulo, o conceito de Web Semântica e refere as principais motivações que têm originado diversos projectos de investigação. No terceiro capítulo é apresentada a arquitectura sugerida para a Web Semântica e referidas as principais camadas e correspondentes normas, linguagens, tecnologias e ferramentas. Finalmente, no quarto capítulo são tecidas algumas considerações sobre o impacto da Web Semântica, evidenciando a evolução da Web e dos motores de busca e os benefícios da adopção desta nova visão no desenvolvimento de conteúdos educativos.

Por conseguinte, após referir algumas ideias e conceitos sobre esta visão e a arquitectura correspondente, este artigo tem como objectivo abordar os aspectos mais relevantes na construção de aplicações para a Web Semântica: Como representar o conhecimento?; Qual a utilidade dos metadados?; Porque é que é necessário um modelo de descrição de recursos?; Qual a importância das ontologias neste contexto?; Para que servem os agentes inteligentes?; e Que tipos de tecnologias, linguagens e arquitecturas são utilizadas?.

2- Web Semântica

O termo “Semântica”, utilizado pela primeira vez pelo linguista francês Michel Bréal (1883), designava a ciência que se dedicava ao estudo das significações. Genericamente, podemos definir semântica como a parte da linguística que se ocupa da significação das palavras e da evolução do seu sentido. Sendo a Web uma das mais amplas e importantes formas de comunicação e de disseminação de informação das diversas áreas de conhecimento, torna-se indispensável procurar soluções de integração, troca e compreensão semântica dessa informação. Tais soluções procuram, através da criação de normas, arquitecturas de metadados, serviços de inferência e ontologias, entre outras, a melhor forma de tornar a informação não só compreensível pelos humanos mas também compreensível pelas máquinas [Moura 2002].

Com vista a usar a Web de uma forma mais inteligente, no final dos anos 90 surgiu o projecto Principia Cybernetica que se baseava na analogia “ligações entre páginas Web // ligações entre neurónios”. Isto é, modelada a actividade neural, conclui-se que, se existisse um conceito A, relacionado fortemente com um conceito B que, por sua vez, se relacionasse com um conceito C, então, A relacionar-se-ia com C. Transpondo esta visão para o software foi possível interligar páginas que à partida

não tinham qualquer relação. Desta forma, a pesquisa de informação deixou de estar dependente apenas de palavras-chave, sendo possível usar também conceitos relacionados [Thompson 2004].

A necessidade de recuperar informação armazenada em grandes repositórios de informação disponíveis na Internet e de responder com maior precisão às pesquisas efectuadas directamente pelos cibernautas, tem motivado a conjugação de esforços ao nível do W3C (*World Wide Web Consortium*) no sentido de adicionar informação semântica às páginas Web. Procura-se, assim, aumentar a eficiência e a selectividade dos motores de busca e de outros tipos de ferramentas de processamento automático de documentos.

As propostas de incorporação de informação semântica em sistemas de busca aplicam abordagens distintas, enfatizando um ou outro aspecto da análise linguística e utilizando diferentes métodos de organização de bases de conhecimento [Ataíde 2001].

A Web Semântica não é uma nova Web ou uma Web à parte, mas sim uma extensão da actual em que a informação tem um significado bem definido, possibilitando aos computadores e às pessoas trabalharem em cooperação.

A maior parte dos conteúdos que hoje aparecem na Internet foram projectados para serem lidos pelas pessoas e não para poderem ser manipulados por agentes de software [Berners-Lee e al. 2001]. Para que isto seja possível, não será obrigatório dotar os computadores de uma inteligência artificial, bastará que os conteúdos disponibilizados pelas páginas Web sejam previamente produzidos através de “editores de Web semânticos” ou descritos de modo a serem entendidos pelos agentes de software.

Um dos desafios que se coloca neste sentido é fornecer uma linguagem que expresse dados e regras para raciocínio sobre esses dados, tendo em vista a possibilidade de deduzir novos dados e regras e permitir que regras existentes em sistemas de conhecimento possam ser exportadas para a Web [Farias e Rosário 2002].

A Web Semântica tem como finalidade estruturar e atribuir semântica aos dados, o que conduzirá à construção de aplicações para aceder à informação de uma forma mais eficiente e precisa e à consequente diminuição ou anulação da grande maioria das limitações referidas.

A Web, enquanto meio de acesso a documentos para as pessoas, passará a ser um sistema de dados e informação processáveis automaticamente. Por exemplo, será possível que os agentes de software, que vagueiam de uma página a outra, possam realizar tarefas mais sofisticadas para os utilizadores, tais como as realizadas por agentes de viagens ou agentes imobiliários, que nos auxiliam de forma especializada a encontrar aquilo que melhor satisfaz as nossas necessidades. Aquilo que aparentemente é

um pequeno passo para o Homem será um grande passo para as máquinas na compreensão do significado da informação e do conhecimento.

O desenvolvimento da Web Semântica será impulsionado pelo uso intensivo de metadados, favorecendo o acesso automatizado à informação, com base no processamento semântico dos dados e em heurísticas, feito por máquinas [Berners-Lee et al. 2001].

Segundo a visão da Web Semântica, um qualquer utilizador, quer navegando num catálogo ou directório, quer utilizando um mecanismo de pesquisa, perante a indicação do assunto que pretende procurar, seria auxiliado por agentes inteligentes que, através de mecanismos e linguagens de inferência e orientados por ontologias articuladas, percorreriam páginas Web anotadas e repositórios de metadados com vista a recuperar os documentos ou recursos de informação sobre o assunto em questão (em vez de todos os documentos que contenham as palavras-chave indicadas). Ou seja, já não será a sintaxe a determinar o resultado do processo de recuperação de informação, mas sim a semântica.

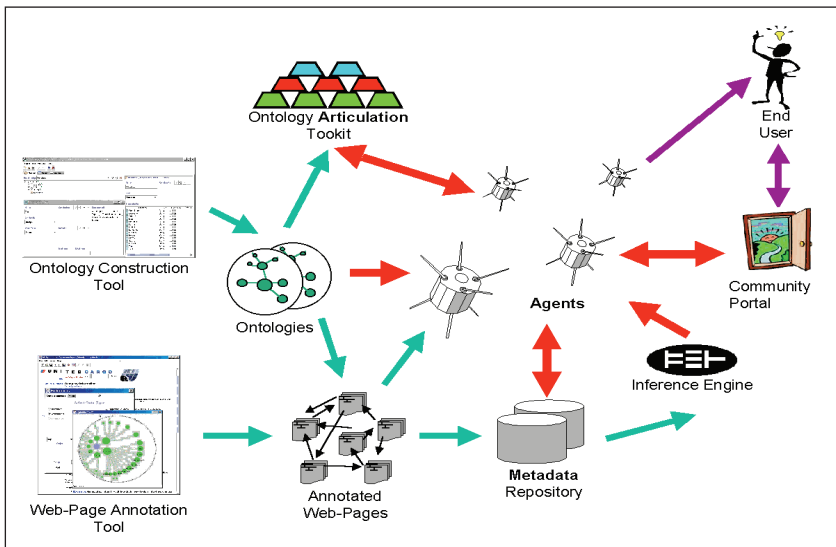


Figura 1 - Web Semântica

Aspectos tais como interoperabilidade e cooperação entre múltiplas fontes de informações, descrevendo diferenças sintáticas, semânticas e estruturais entre sistemas são também importantes para a Web Semântica [Cunha 2002]. A integração das tecnologias *eXtensible Markup Language* (XML), *Resource Description Framework* (RDF), arquiteturas de metadados, ontologias, agentes computacionais, entre outras, favorecerá o aparecimento de serviços Web que garantam a interoperabilidade e cooperação.

3- Arquitectura da Web Semântica

Em primeira instância, a Web Semântica é uma iniciativa ou projecto, liderado pelo *World Wide Web Consortium (W3C)*, que pretende criar um meio universal para troca de informação, atribuindo significado (semântica) ao conteúdo dos documentos da Web, de modo a que esse significado seja compreendido não só por humanos, mas também por máquinas. Em última instância, a Web Semântica é uma extensão da Web actual que permitirá encontrar, partilhar e (re)combinar informação mais facilmente, fruto da colaboração entre o utilizador e os agentes de software.

Em 2000, o W3C divulgou a primeira arquitectura para a Web Semântica baseada em camadas sobrepostas, em que cada camada ou tecnologia deveria obrigatoriamente ser complementar e compatível com as camadas inferiores, sem que tivesse de estar dependente de qualquer camada superior, tal como se pode verificar na figura Figura 2 – Arquitectura para a Web Semântica (versão 2000).

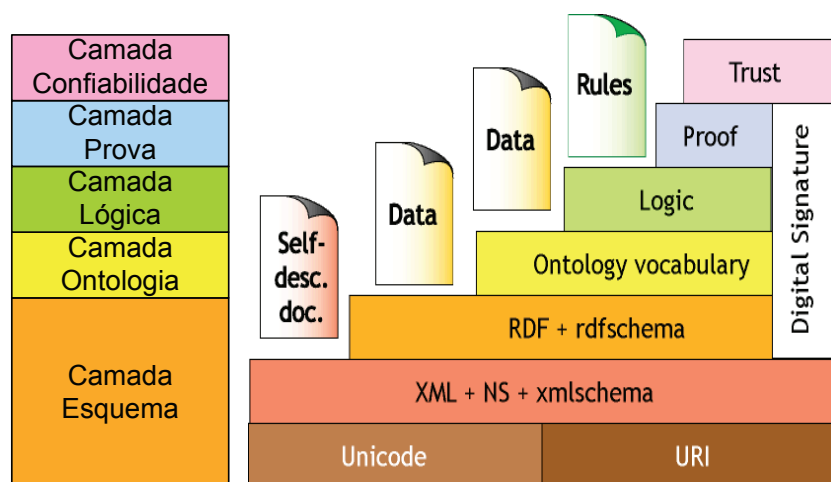


Figura 2 – Arquitectura para a Web Semântica (versão 2000)

A arquitectura sugerida para a Web Semântica é composta genericamente por três camadas principais:

- **Camada Esquema (Schema Layer):** responsável por estruturar os dados e definir o seu significado para que se possa elaborar um raciocínio lógico. Para além da definição de que todos os documentos, objectos e eventos são recursos identificados por URIs (*Uniform*

Resource Identifiers) e de que a sua leitura universal é garantida pela norma Unicode, XML(S) e RDF(S) constituem os principais elementos-chave desta camada no que diz respeito à estruturação e atribuição de significado aos dados. Usámos os termos XML(S) e RDF(S) para a tecnologia XML e o seu esquema XML *Schema* e para a tecnologia RDF e ao seu esquema RDF *Schema*, respectivamente.

- **Camada Ontologia (*Ontology Layer*):** constituída pelo vocabulário ontológico que fornece o significado dos termos e das teorias de determinada área do conhecimento compartilhados por todos os participantes do domínio, incrementando a expressividade semântica fornecida pela RDF e o seu esquema RDFS. As ontologias, descritas como documentos ou ficheiros que formalmente definem as relações entre os termos a usar na comunicação entre agentes computacionais e destes com as pessoas, assumem-se como outro dos elementos chave da Web Semântica.
- **Camada Lógica (*Logic Layer*):** responsável pela declaração de regras lógicas que podem ser executadas pelos agentes de software para realizar inferências automáticas e verificar o nível de coerência lógica dos recursos.

Os esforços de investigação dos últimos anos concentraram-se essencialmente na implementação de linguagens e ferramentas para a camada ontológica. Consequentemente, em 2004, principalmente como resultado do grupo de trabalho *Web Ontology Working Group*, foi publicada a *Web Ontology Language* (OWL) como linguagem recomendada pelo W3C para o desenvolvimento de ontologias. Em 2005, foi publicada a versão mais recente da arquitectura para a Web Semântica (Figura 3 – Arquitectura para a Web Semântica (versão 2005)).

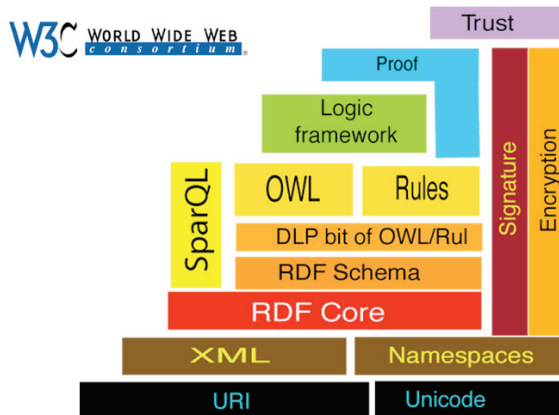


Figura 3 – Arquitectura para a Web Semântica (versão 2005)

3.1- Camada Esquema

O objectivo principal da Web Semântica não é, pelo menos para já, treinar as máquinas para que se comportem como pessoas, mas sim desenvolver tecnologias e linguagens que tornem a informação legível para as máquinas.

O primeiro passo rumo à Web Semântica passa pela adopção de uma linguagem que permita a exportação dos conteúdos Web atendendo às regras de qualquer sistema de representação do conhecimento.

O projecto Web Semântica é um sucessor do projecto Metadados (dados sobre dados na Web), mas que inclui categorias semânticas. Recorrendo a uma linguagem particularmente flexível, a XML, para fornecer a sintaxe e estruturar a informação, aos metadados para descrever essa informação e às linguagens RDF(S) para descrever e atribuir significado a essa informação, a Web pode tornar-se num espaço onde a informação terá um significado bem definido, facilitando a cooperação e a comunicação entre as pessoas e os agentes computacionais.

3.1.1- Linguagem XML

Com o crescimento exponencial da Web a indexação de informação tornou-se uma tarefa árdua. A linguagem HTML (*HyperText Markup Language*) é a linguagem básica de marcação que é usada na grande maioria das páginas Web. As principais limitações desta linguagem residem no facto de não admitir adicionar significado à informação e de não permitir estruturar os dados de um documento HTML, ou seja, não há distinção entre a codificação dos seus componentes básicos (conteúdo, estrutura e apresentação dos dados). As *tags* HTML não oferecem qualquer descrição acerca dos dados, apenas descrevem como uma página deve ser exibida (preocupação estética na apresentação da informação), e os dados são apresentados na linguagem natural.

XML, sendo uma linguagem de marcação extensível, permite a definição de marcas e atributos. Podemos então dizer que XML é um conjunto de regras para definir etiquetas semânticas que dividem um documento em partes e identificam as diferentes partes desse documento [Harold 1999]. É a linguagem de marcação apropriada para a representação de dados, documentos e demais entidades, cuja essência se fundamenta na capacidade de agregar informação [Pimentel et al. 2000]. Com XML conseguimos não só garantir um melhor controlo em relação à interface, mas também uma maior flexibilidade na pesquisa, devido essencialmente à separação entre os rótulos de descrição e a apresentação do conteúdo.

Genericamente, os principais componentes da XML são: as declarações, os elementos, os atributos, as entidades, os comentários, as instruções de processamento e as regras definidas num DTD (*Document Type Definition*) ou num esquema XML ou *XML Schema*.

```

<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="exemplo.xsl"?>
<!DOCTYPE BIBLIOTECA SYSTEM "exemplo.dtd">
<BIBLIOTECA>
  <LIVRO>
    <CODIGO>001</CODIGO>
    <TITULO>Revista EduSer</TITULO>
    <VALOR>5 Euros</VALOR>
  </LIVRO>
  <LIVRO>
    <CODIGO>002</CODIGO>
    <TITULO>Revista Toques Formativos</TITULO>
    <VALOR>2 Euros</VALOR>
  </LIVRO>
</BIBLIOTECA>

```

Figura 4 - Exemplo de XML

Um DTD não é mais do que um ficheiro que especifica um conjunto de regras para restringir e validar a estrutura do documento XML, ou seja, define a estrutura de um documento através da especificação dos seus elementos básicos, relacionamentos, atributos, entidades e o conjunto de *tags* associadas para marcar o documento. A utilização de um DTD facilita o processamento de documentos XML por aplicações, para além de controlar a coerência das marcações de XML. De acordo com o documento de XML visualizado anteriormente, o DTD correspondente está representado na Figura 5.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!ELEMENT BIBLIOTECA (LIVRO+)>
<!ELEMENT LIVRO (CODIGO,TITULO,VALOR)>
<!ELEMENT CODIGO (#PCDATA)>
<!ELEMENT TITULO (#PCDATA)>
<!ELEMENT VALOR (#PCDATA)>

```

Figura 5 - Exemplo de DTD

Uma vez que o DTD tem uma sintaxe particular, tentou encontrar-se uma forma de escrever em XML a estrutura dos documentos XML. O XML *Schema* é, portanto, uma alternativa ao DTD baseada em XML.

Por fim, recorrendo a folhas de estilo distintas em linguagem XSL (*eXtensible Style Language*), os dados de um documento XML podem ser apresentados em vários suportes (computador, telemóvel, TV, etc.).

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/
Transform" version="1.0">
<xsl:output method="html"/>
<xsl:template match="/BIBLIOTECA">
<html>
  <head>
    <title>Biblioteca</title>
  </head>
  <body bgcolor="#FFFFFF">
    <h1>Revistas da ESEB</h1>
    <table border="1" cellspacing="0" cellpadding="5">
      <xsl:apply-templates/>
    </table>
  </body>
</html>
</xsl:template>
<xsl:template match="LIVRO">
  <tr>
    <xsl:apply-templates/>
  </tr>
</xsl:template>
<xsl:template match="CODIGO">
  <xsl:call-template name="COLUNA"/>
</xsl:template>
<xsl:template match="TITULO">
  <xsl:call-template name="COLUNA"/>
</xsl:template>
<xsl:template match="VALOR">
  <xsl:call-template name="COLUNA"/>
</xsl:template>
<xsl:template name="COLUNA">
  <td>
    <xsl:value-of select="."/>
  </td>
</xsl:template>
</xsl:stylesheet>

```

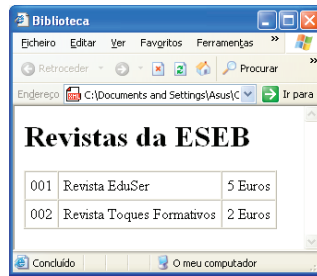


Figura 6 - Exemplo de XSL

A orientação do HTML para a apresentação implica o recurso a palavras-chave aquando da busca de informações. Com a linguagem XML é dado o primeiro passo para que a interpretação semântica dos documentos seja efectuada também pelas máquinas, uma vez que os recursos podem ser relacionados e os conteúdos descritos de forma clara, facilitando a busca de informações ao recorrer a metadados.

A XML assume-se como uma linguagem através da qual o significado

pode ser comunicado independentemente da plataforma, facultando a criação de uma estrutura arbitrária nos documentos, mas nada diz acerca do significado dessa estrutura. Esta tarefa é deixada para a linguagem RDF que permite descrever a informação de forma não ambígua para depois ser processada pelas máquinas [Thompson 2004].

3.1.2- Arquitectura de Descrição de Recursos RDF

Se a ideia é fazer com que os computadores entendam os conteúdos Web, o primeiro passo é a organização e estruturação da informação e o segundo é a adição de semântica a essa informação, para que os agentes computacionais a possam compreender.

A *Resource Description Framework* (RDF) é uma *framework*, tecnologia ou linguagem para representar informação na Web [Lassila e Swick 1999]. Sendo uma das aplicações da XML, usa a notação XML como sintaxe de codificação e de descrição dos metadados. **Metadados**, não são mais do que “dados sobre dados”, ou dados sobre os sistemas que manipulam esses dados.

O esquema de metadados mais popular é o Dublin Core Metadata (não esquecendo o MARC 21 no âmbito das bibliotecas e centros de documentação), enquanto que o LOM (Learning Object Metadata) é o mais utilizado no campo da Educação. Contudo, a RDF é a proposta que mais se destaca no contexto particular da Web Semântica, uma vez que é uma recomendação W3C e permite expressar os conjuntos de elementos de metadados dos esquemas DCM e LOM, permitindo uma maior eficiência da descoberta de recursos de informação na Web.

Uma vez que os metadados são utilizados para descrever as características de um recurso e as suas relações, o objectivo principal da arquitectura RDF é definir um mecanismo para descrever recursos não vinculados a um domínio específico de aplicação, ou seja, um mecanismo que sirva para descrever recursos de qualquer área do conhecimento.

A RDF é uma linguagem que permite a codificação, o intercâmbio e a reutilização de metadados estruturados [Daconta et al. 2003]. Basicamente, a RDF fornece uma forma de associar propriedades a recursos, através de *statements*. Um *statement* (declaração) tem três componentes: *resource* (recurso), *property* (propriedade) e *value* (valor), através dos quais é possível descrever os recursos de informação disseminados na rede, definindo recursos e identificando as suas propriedades e valores.

Uma propriedade é qualquer característica que possa descrever um recurso. Mas uma propriedade pode representar também o relacionamento entre recursos. Para definir o significado, as características e as relações entre recursos recorreremos ao Esquema RDF (*RDF Schema* ou RDFS), tal como veremos mais adiante.

Uma declaração corresponde a um determinado recurso, às suas

propriedades e aos respectivos valores. Podemos resumir uma declaração como “o recurso (*subject*) possui a propriedade (*predicate*) com o valor (*object*). Por exemplo, a declaração <“http://www.vgportal.ipb.pt”, “autor”, “Vitor Gonçalves”> teria o significado: A página http://www.vgportal.ipb.pt (recurso) tem como autor (propriedade) Vitor Gonçalves (valor).

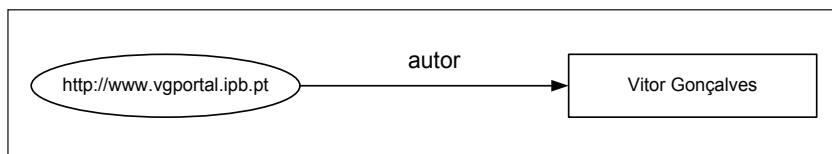


Figura 7 – Grafo de um Modelo de Dados RDF básico

As declarações representam um grafo que vai do nó sujeito para o nó objecto e o arco tem o nome da propriedade, tal como se pode verificar na representação gráfica da sentença inerente ao exemplo. Para apoiar a representação gráfica podemos recorrer a ferramentas tais como a DLG (Directed Label Graphs). Se aplicarmos a sintaxe RDF baseada em XML ao grafo da Figura 7 – Grafo de um Modelo de Dados RDF básico obtemos o seguinte código:

```
<?xml:namespace ns = http://www.w3.org/RDF/RDF/"
prefix="rdf" ?>
<?xml:namespace ns = http://purl.org/dc/elements/1.1/"
prefix="dc" ?>

<rdf: RDF>

  <rdf:Description about="http://www.vgportal.ipb.pt">
    <dc:Creator>Vitor Gonçalves</dc:Creator>
  </rdf:Description>

</rdf: RDF>
```

Figura 8 – Exemplo básico de RDF

O elemento *Description* permite agrupar diversos *statements* (decla-

rações) para o mesmo recurso. O atributo *about* deste elemento indica o recurso ao qual todas essas declarações se aplicam. O elemento *dc:Creator* é um elemento da especificação para metadados *Dublin Core* e permite declarar o valor “Vitor Gonçalves”. Caso pretendêssemos fornecer o título do URI deste exemplo, bastaria recorrer ao elemento *dc:Title* de metadados *Dublin Core*, tal como veremos na secção seguinte. Para além da sintaxe XML, o código acima necessita das facilidades oferecidas pelos XML *namespaces*, que mais não são do que conjuntos de nomes, identificados por URIs, para usar em documentos XML como nomes de atributos e tipos de elementos [Bray et al. 1999].

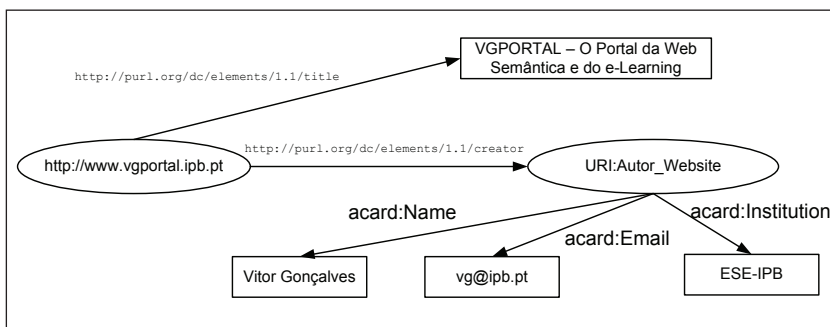


Figura 9 – Grafo de um Modelo de Dados RDF

Num diagrama de modelo de dados, um objecto pode corresponder a outro recurso identificado por um URI, uma *string* ou outro qualquer tipo de dados definido em XML. Por exemplo, se necessitarmos fornecer mais informação sobre o autor (e-mail e instituição), o valor “Vitor Gonçalves” teria que ser substituído por um recurso (identificado por um URI) com as propriedades (nome, e-mail e instituição). A diferença é que, no primeiro exemplo, temos uma declaração com uma única propriedade (o literal “Vitor Gonçalves”) e no segundo, passou a existir uma propriedade estruturada que possui três propriedades (nome, email e instituição). Assim, o grafo do modelo de dados da Figura 9 – Grafo de um Modelo de Dados RDF expressa os metadados através da sintaxe de serialização RDF/XML da Figura 10.

```

    <?xml:namespace ns = http://www.w3.org/RDF/RDF/" prefix="rdf"
?>
    <?xml:namespace ns = http://purl.org/dc/elements/1.1/"
prefix="dc" ?>
    <?xml:namespace ns = "http://www.vgportal.ipb.pt/acard/"
prefix="acard" ?>
    <rdf:RDF>
        <rdf:Description about="http://www.vgportal.ipb.pt">
            <dc>Title>VGportal - O Portal da Web Semântica e do
e-Learning</dc>Title>
            <dc:Creator>Vitor Gonçalves</dc:Creator>
        </rdf:Description>
        <rdf:Description ID = "http://www.vgportal.ipb.pt/vg">
            <acard:Name> Vitor Gonçalves </acard:Name>
            <acard:Email> vg@ipb.pt </acard:Email >
            <acard:Institution> ESE-IPB </acard:Institution>
        </rdf:Description>
    </rdf:RDF>

```

Figura 10 – Exemplo de RDF

Por conseguinte, as *tags* RDF de uma página Web contêm recursos que descrevem uma ou mais propriedades para cada recurso. Cada propriedade possui um valor ou aponta para outro recurso. Assim, as máquinas sabem como usar o valor armazenado em cada propriedade, uma vez que se encontra definido publicamente e armazenado através dos denominados *namespaces*. De referir que o prefixo *acard* corresponde a um *namespace* específico definido uma única vez na terceira linha do código RDF/XML.

Mas, mesmo assim, há a necessidade de especificar que deverá ser indiferente usar como identificação “Gonçalves, Vitor”, “Vitor Gonçalves” ou “Vitor Barrigão Gonçalves”, pois o mecanismo de busca interpretá-los-á correctamente, já que cada propriedade corresponde a um tipo definido globalmente. Um exemplo similar é a confusão entre datas do tipo UK (dia/mês/ano) e do tipo US (mês/dia/ano) [Thompson 2004].

A resolução das ambiguidades e a obtenção de maior precisão na recuperação de informação disponível na Web constituem duas das principais preocupações dos projectos da WS. A ambiguidade pode ser solucionada atribuindo-se um URI diferente a cada conceito de uma

palavra. Assim, os agentes dos motores de busca poderão encontrar páginas que se refiram a conceitos específicos e não todas as páginas nas quais a palavra ambígua é utilizada. Mas durante o processo de busca, o agente também pode encontrar dois ou mais URIs diferentes para o mesmo conceito. Para resolver este problema, existem as ontologias que vão fornecer o vocabulário necessário para a comunicação entre os agentes computacionais e as páginas Web e mostrar as relações entre os conceitos, fornecendo mais significado que as descrições RDF/XML. Não obstante, podem ocorrer outros tipos de ambiguidades no conteúdo de documentos disponíveis na Web que interfiram na precisão da recuperação da informação.

A RDF não fornece a informação suficiente para que uma máquina (aplicação ou agente de software) perceba que o valor de uma propriedade que representa um autor é uma referência a uma pessoa e não a uma instituição, localidade ou veículo. Necessitamos de um esquema que valide as afirmações expressas em RDF. Necessitamos do esquema RDFS ou de outra proposta ontológica. Em suma, necessitamos de mais uma camada semântica: a Camada Ontológica.

3.2- Camada Ontológica

As ontologias já são usadas no âmbito da Inteligência Artificial há alguns anos, geralmente associadas à inferência lógica e técnicas similares. Mais recentemente, começaram também a ser usadas para a localização e recuperação de informação no âmbito das Ciências da Computação e das Ciências da Informação [Garshol 2004]. Em primeira instância, uma ontologia é um modelo de dados que representa um determinado domínio do conhecimento.

Os dicionários fornecem as definições dos conceitos, enquanto que os índices fornecem as indicações para os locais onde cada conceito é referenciado. A taxonomia clarifica os relacionamentos hierárquicos entre os conceitos, criando uma estrutura de classes/subclasses, enquanto que o tesouro amplia esta estrutura com algumas relações predefinidas. Finalmente, uma ontologia descreve os conceitos relevantes existentes num determinado domínio do conhecimento, definindo quaisquer relações binárias entre eles que se julgue interessante citar. Portanto, há uma forte conexão entre os diversos instrumentos de representação do conhecimento, tal como se pode verificar na Figura 11 – Relação entre Dicionário, Índice, Taxonomia, Tesouro e Ontologia [Librelotto 2005].

As ontologias representam o culminar da exposição evolutiva e cumulativa que fizemos dos instrumentos convencionais para a representação do conhecimento. Ao fornecerem um vocabulário aberto para descrever objectos de informação, as ontologias assumem-se como fulcrais no âmbi-

to da Web Semântica, uma vez que as restantes abordagens representam linguagens com vocabulários fixos.

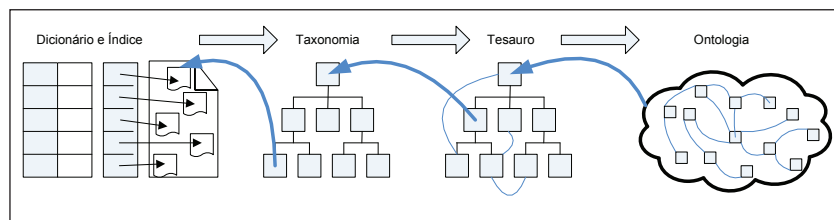


Figura 11 – Relação entre Dicionário, Índice, Taxonomia, Tesouro e Ontologia

Uma ontologia é uma forma de representar o conhecimento através de hierarquias elementares [Hendler 2001]. Pode ser vista como uma taxonomia formada por classes e subclasses de objectos, relacionadas entre si, à qual juntamos mais um conjunto de propriedades e regras de inferência. Segundo [Pereira 2003] uma ontologia é uma descrição explícita e precisa de conceitos e relações que existem num domínio (área de conhecimento) em particular, permitindo a troca e reutilização do conhecimento.

A utilização de ontologias permite lidar com conceitos, representando-os formalmente; evita problemas inerentes ao vocabulário da linguagem natural (homonímia, sinonímia, metonímia, etc.); favorece a reutilização do conhecimento e facilita a formulação de consultas tendo em conta a semântica do domínio, entre outras características.

Imagine que tentava procurar informação sobre a “Família do Escritor José Saramago” usando, por exemplo, o motor de busca *Google*; o resultado da pesquisa traduzir-se-ia em mais de 11 mil páginas Web, uma vez que o mecanismo de busca não efectua a procura com base no significado, mas sim em palavras-chave desprovidas de significado e contexto. Um mecanismo de busca semântico é capaz de percorrer a informação representada em ontologias e responder mais precisamente a questões tais como: “Quem são os netos de José Saramago?”; “Quem é o genro de José Saramago?”; “Qual a nacionalidade da sua actual mulher?” ou mesmo “Qual a sua relação com o compositor italiano Azio Corghi?”, uma vez que, após verificar que José Saramago pertence à classe dos humanos, o mecanismo de busca perceberá que a informação que procuramos incide nas suas relações com outros humanos.

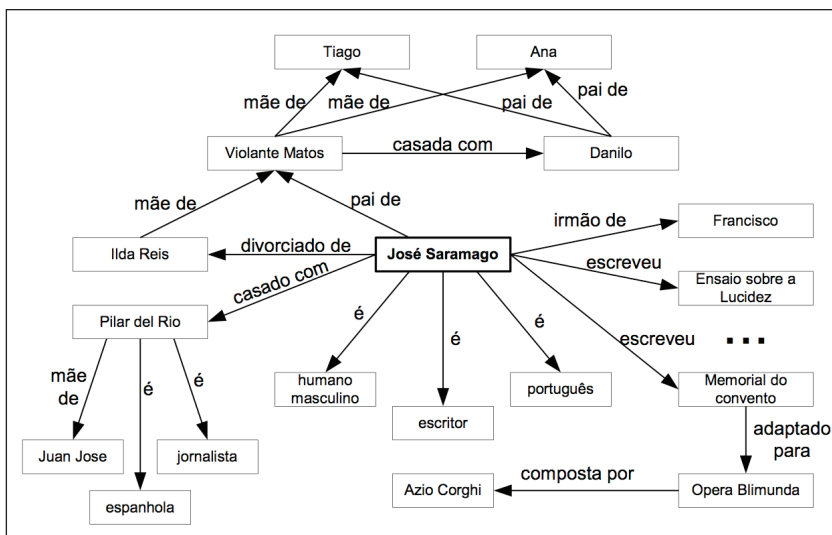


Figura 12 - Parte de uma possível ontologia para José Saramago

Em suma, a vantagem da utilização das ontologias está na habilidade de deixar de lado muitas características indesejáveis [Moura 2002]. Ou seja, ao usar ontologias a velocidade e qualidade das pesquisas melhorarão significativamente, uma vez que os agentes computacionais pesquisarão apenas páginas que se referem à informação pretendida, em vez de pesquisarem todas as páginas usando palavras-chave ambíguas.

A RDF por si só não permite criar ontologias; no entanto, ao permitir descrever vocabulários que representam o conhecimento sob a perspectiva das redes semânticas, constitui a base para outras linguagens com essa finalidade, tais como *RDF Schema* (RDFS) e *OWL (Web Ontology Language)* que mais não são do que tecnologias que definem as primitivas para a criação de ontologias [Davies 2003].

OWL é baseada na lógica descritiva para fornecer uma semântica formal, permite a representação de hierarquias e facilita um maior nível de interpretação dos conteúdos Web por parte da máquina do que aquele que é proporcionado por XML e RDF (ou mesmo RDFS).

Por conseguinte, o sucesso da Web Semântica depende fortemente da proliferação de ontologias, pelo que se requer uma construção rápida das mesmas [Maedche et al. 2001]. Actualmente, existem algumas propostas para desenvolver métodos de criação e manutenção automática de ontologias. No entanto, há ainda muito trabalho por fazer nomeadamente ao nível das linguagens de definição de ontologias, das metodologias e ferramentas de construção e manipulação de ontologias (OILED, OntoEdit, Kaon, Protégé, etc.), da integração de ontologias, da

aprendizagem de ontologias, das ferramentas de anotação de páginas (OntoMat, OntoAnnotate, UBOT, etc.), dos motores de inferência (LoPix, TRIPLE, etc.), ou mesmo dos agentes inteligentes e dos *Web Services*, entre outras ferramentas para o desenvolvimento de aplicações para a Web Semântica (Jena da HP, Wilbur da Nokia, etc.).

3.3- Camada Lógica

Com a descrição da estrutura dos dados em XML, a representação da semântica simples desses dados em RDF(S) e a representação formal comumente aceita sobre o que significam esses dados em ontologias mais expressivas como OWL estão criadas as condições para que os agentes possam raciocinar sobre dados com significado bem definido. Não obstante, para que os agentes de software possam raciocinar sobre as estruturas de dados, tirando partido das relações entre os objectos dessas estruturas esclarecidas através de ontologias, pode ser necessário definir previamente regras para os mecanismos de inferência.

A Camada Lógica é composta por um conjunto de regras de inferência que fornecem aos agentes computacionais a capacidade de raciocinar sobre as estruturas de dados, ou seja, para realizar inferências sobre os dados.

As regras de inferência fornecem aos agentes computacionais o poder de raciocinar sobre as estruturas de dados definidas na Camada Esquema (*Schema Layer*) tirando partido da relação entre os objectos dessas estruturas definidas na Camada Ontológica (*Ontology Layer*) [Afonso 2001]. Tipicamente, em cada sistema de representação do conhecimento existe uma distinção clara entre os dados e as regras de inferência. As regras controlam as etapas de inferência que o mecanismo de inferência realiza sobre os dados [Berners-Lee 2000].

As regras de inferência são também um elemento chave para a visão da WS, uma vez que permitem que os dados provenientes de diversas fontes sejam integrados, derivados e transformados de uma forma distribuída, transparente e escalável.

A iniciativa RuleML (*Rule Markup Initiative*) tem colaborado com o W3C na normalização de regras para a Web, nomeadamente no âmbito dos Sistemas Distribuídos, Web Semântica, Computação Autónoma e Arquitecturas Orientadas a Serviços (*Service Oriented Architecture*) activados semanticamente. O objectivo desta iniciativa é desenvolver a RuleML (*Rule Markup Language*) como uma linguagem canónica para regras baseada na XML, na semântica formal e em implementações eficientes. A RuleML cobre todo o espectro das regras: regras derivadas, regras transformadas e regras reactivas. Com esta linguagem é possível realizar consultas e inferências em ontologias Web, alinhamentos de ontologias, etc [Boley et al. 2005].

A SWRL - *Semantic Web Rule Language* [Horrocks et al. 2004] e a WRL - *Web Rule Language* [Angele et al. 2005] são as linguagens de regras para a Web que mais se têm destacado. SWRL é uma linguagem baseada nas cláusulas de Horn e na combinação das sub-linguagens OWL DL e OWL Lite com as sub-linguagens *Datalog* unárias/binárias da linguagem RuleML. Enquanto que WRL é uma linguagem baseada na programação lógica e com um nível de expressividade muito próximo da *Description Logic*.

Para realizar a inferência pode ser usada a linguagem SPARQL (*RDF Data Query Language*). No entanto, há situações em que as ontologias não expressam todo o conhecimento necessário para realizar a inferência. Assim, as regras de inferência, definidas por exemplo através da linguagem SWRL, proporcionam uma camada lógica básica para ampliar as capacidades dos agentes inteligentes na localização e recuperação de informação e na geração de conhecimento.

Esta secção não teve como objectivo apresentar completamente a Camada Lógica, uma vez que essa camada, bem como as restantes camadas superiores da arquitectura para a Web Semântica ainda se encontram numa fase embrionária de desenvolvimento, não existindo por enquanto tecnologias recomendadas.

3.4 - Agentes Inteligentes

Os Agentes Inteligentes são sistemas automáticos (hardware ou software) que contêm mecanismos de IA, e são capazes de tomar decisões e aperfeiçoar o seu desempenho [Berners-Lee et al. 2001].

Os Agentes Inteligentes devem possuir as seguintes características [Cunha 2002]:

- Autonomia: trabalhar sem intervenção humana;
- Habilidade social: saber interagir com humanos ou outros agentes;
- Reactividade: poder receber estímulos do ambiente e responder em tempo útil;
- Pró-actividade: ter comportamento direccionado a um objectivo, tomando a iniciativa da acção sem precisar de esperar por estímulos;
- Portabilidade: facilidade de mobilidade para outros ambientes;
- Continuidade temporal: funcionar continuamente.

Na Internet, os agentes inteligentes (*Intelligent Web Agents*) servem principalmente para explorar serviços na Web e entender regularidades geradas pela Web.

Quando existem vários Agentes Inteligentes a actuar de forma integrada e cooperativa, o sistema é chamado Multi-Agente. Geralmente,

cada agente inteligente possui conhecimentos próprios e diferentes. Estes agentes interagem entre si, partilhando informação e conhecimento para soluções de problemas mais complexos, os quais dificilmente seriam resolvidos por qualquer um dos agentes de maneira isolada.

Resumindo, a Web está prestes a tornar-se na maior base de dados do mundo – uma verdadeira biblioteca universal. A infra-estrutura para a Web Semântica deverá fornecer a sintaxe para representar metadados (esquemas XML e RDF) e os vocabulários para exprimir esses metadados (ontologias) [Harmelen 2004].

Considerando que os Metadados são o primeiro passo rumo à Web Semântica e que a RDF constitui os alicerces para processar esses metadados, torna-se necessário construir esquemas RDF que correspondam aos esquemas de metadados XML, contribuindo para a indexação dos recursos distribuídos na Web. Um esquema RDF permite expressar as relações entre diferentes recursos e propriedades para que possam ser usados por diferentes aplicações. Assim, colecções de recursos e suas relações são definidas nas ontologias.

Os agentes de software assumem-se como a componente que concretiza a visão da Web Semântica, permitindo que a Internet evolua de um espaço estático, cuja utilização requer uma grande intervenção humana, para um espaço mais complexo e dinâmico, caracterizado por um crescente recurso a agentes de software e, conseqüentemente, pela diminuição do número de horas despendidas em tarefas pouco criativas e monótonas por parte do utilizador final.

Os agentes devem ter capacidade para: compreender o significado e a relação entre os objectos, baseando-se no vocabulário ontológico (*ontology vocabulary*); raciocinar sobre eles, utilizando as regras de inferência definidas na camada lógica (*logic layer*); e ainda devem ser capazes de trocar dados, resultados do processamento de dados e provas de inferência (*proofs*) de forma eficiente e eficaz, uma vez que são eles que representam valor para os utilizadores da Web Semântica.

A existência de vários esquemas RDF (por exemplo na área da Educação) possibilitará a interoperabilidade. Mas, a RDF e as ontologias associadas podem conter dados incorrectos (por descuido ou não!), pelo que muito há ainda a fazer no que diz respeito à segurança desta infra-estrutura.

Por conseguinte, torna-se necessário implementar um sistema de certificação através de assinaturas digitais (*digital signatures*) para garantir a autenticidade e a confiabilidade das fontes, uma vez que passamos a ter agentes raciocinando sobre os dados, o que afectará significativamente a acção humana. O recurso a assinaturas digitais possibilitará aos agentes inteligentes aferirem o grau de certeza associado ao resultado derivado do seu raciocínio, podendo mesmo ignorar a informação caso a fonte não seja fidedigna.

4 - Considerações Finais

As Tecnologias da Informação e Comunicação, das quais se destacam o computador e a Internet, constituem tecnologias educativas poderosas que exigem uma base educacional e cultural consistente, promovendo uma verdadeira comunicação e cooperação entre pessoas e máquinas, para a utilização do seu verdadeiro potencial e da sua plena capacidade.

Aquando da sua apresentação ao mundo em Agosto de 1989, no âmbito do Laboratório Europeu de Física de Partículas (CERN), não se previa que a *World Wide Web* tivesse o impacto que teve na sociedade em geral.

A primeira geração da rede (ênfase na infra-estrutura tecnológica da Internet) permitia apenas a troca de dados entre máquinas distintas. A segunda geração (ênfase nas aplicações Web) provocou uma revolução por disponibilizar uma vasta gama de aplicações e informação para as pessoas, tornando também possível o comércio electrónico entre clientes e empresas.

Da criação manual de páginas Web rapidamente se passou à geração automática de páginas baseadas em base de dados. Porém, como as páginas só possuem informação léxica, mesmo os agentes e/ou robots mais avançados encontram um ambiente hostil para a realização das suas tarefas, porque tanto o conteúdo das páginas como o relacionamento entre elas é difícil de ser compreendido pelas máquinas, por se encontrarem geralmente em linguagem natural. Mesmo que os conteúdos sejam gerados a partir de uma base de dados, o significado dessa estrutura não é suficientemente claro para que um agente computacional possa realizar inferências sobre os dados.

A próxima geração da rede tem vindo a ser designada rede semântica ou inteligente. O maior desafio é transformar os dados e aplicativos em elementos úteis, legíveis e compreensíveis para os agentes inteligentes, facilitando-lhes a comunicação e a cooperação dinâmicas. Ou seja, por um lado é necessário facilitar a pesquisa e a interpretação através de normas que definam de forma universal o significado do conteúdo das páginas Web. Por outro, é urgente desenvolver aplicações que consigam simultaneamente partilhar informação de diversas fontes e inferir informação ou conhecimento adicional.

A cada uma destas gerações correspondem diferentes gerações de motores de busca. Na primeira geração podemos identificar motores de busca caracterizados por basear as suas respostas em informação intrínseca - informação ou palavras-chave contidas no próprio documento ou página Web. À segunda geração correspondem motores de busca que, para além das informação intrínseca, utilizam informação extrínseca - informação não incluída no documento ou página Web que é gerada a partir da sua estrutura de links (*link analysis*) ou da sua popularidade em relação aos outros documentos (*usage analysis*). A tendência dos motores

de busca de terceira geração aponta para a capacidade de inferir dados através da colaboração e interoperação entre agentes Web inteligentes [Rezende 2002].

O conceito de Web Semântica possibilitará aos utilizadores da próxima geração da Internet obter informação mais precisa, devido à utilização de agentes computacionais inteligentes nos motores de busca, os quais permitirão encontrar exactamente o que procurarmos. Em vez de obtermos como resposta a uma consulta algumas dezenas ou centenas de possibilidades, a Web será suficientemente inteligente para perceber exactamente aquilo que pretendemos.

Por exemplo, utilizando o Google (páginas em Português) na Web actual, se efectuarmos uma pesquisa do tipo “livros acerca de José Saramago” surgem mais de 580 hipóteses, entre as quais muitas relativas a livros escritos pelo autor e não a livros que falem sobre o autor. Com a Web Semântica, os agentes serão suficientemente inteligentes para decifrar o sentido da palavra “acerca” que os conduzirá a procurar livros sobre o autor e não livros escritos pelo autor. Esta situação justifica a evolução de uma rede de documentos para uma rede de dados.

Através da representação do conhecimento em XML (transporte dos dados), RDF (representação da estrutura dos dados) e ontologias (representação semântica dos dados através OWL), permitindo a exploração de consultas em metadados, torna-se possível esta evolução, favorecendo a comunicação entre máquinas e a colaboração com os humanos e, consequentemente, atinge-se a visão de Berners-Lee.

A Web Semântica poderá ajudar na evolução do conhecimento humano em geral [Berners-Lee et al. 2001] ao facilitar, por exemplo, a colaboração entre equipas investigadoras de um mesmo tema que se encontrem em pontos distintos do planeta. Tal como refere o filósofo francês Pierre Lévy (2003), a Web Semântica é o instrumento necessário ao desenvolvimento da inteligência colectiva.

Por conseguinte, é normal que exista um grande interesse, quer a nível comercial, quer a nível académico, por tornar a Web Semântica uma realidade, uma vez que se pensa que pode ser uma peça crucial no desenvolvimento da actual Sociedade da Informação e do Conhecimento. Muitas são as empresas, nomeadamente as de telecomunicações, que têm vindo a investir nesta área por considerarem que as redes 3G constituirão a plataforma ideal para o desenvolvimento da Web Semântica.

Este ensaio teórico tentou descrever o esforço que tem vindo a ser levado a cabo no sentido de desenvolver a infra-estrutura de suporte à Web Semântica e as aplicações que a viabilizam, entre outras soluções e ideias que impulsionam esta visão inovadora. Paralelamente, tentou sensibilizar os produtores de conteúdos (incluindo professores e alunos) para os benefícios destas novas tecnologias.

Finalmente, escusado será afirmar que os processos de aprendizagem, a investigação académica e a relação professor-aluno só terão a beneficiar com a aplicação da Web Semântica à Gestão da Informação e do Conhecimento (pesquisas inteligentes e recuperação de informação relevante) e ao Acesso à Informação (motores de busca munidos de agentes inteligentes) ou mesmo aos sistemas de *e-Learning* (favorecendo a compreensão dos dados e eliminando a ambiguidade).

Num futuro muito próximo, professores e alunos passarão mais tempo a ler, ver e ouvir material didáctico multimédia útil para a sua investigação, e muito menos tempo a filtrar essa informação de entre um conjunto vasto de resultados devolvidos pelos motores de busca actuais [Williamson e Miller 2003].

Referências Bibliográficas

- AFONSO, M. M. R. **Semantic Web**. Porto: Universidade do Porto, 2001. Disponível em: <http://www.fe.up.pt/~mgi00014/ari/SW.doc>
- ANGELE, J., Boley, H., Bruijn, J. d., Fensel, D., Hitzler, P., Kifer, M., et al. (2005). Web Rule Language (WRL). *W3C Member Submission*. Acesso em: 16 de Novembro de 2006, Disponível em: <http://www.w3.org/Submission/2005/SUBM-WRL-20050909/>
- ATAÍDE, Isabel. **Semantic Web**. Covilhã: Universidade da Beira Interior, 2003. Disponível em: http://www.di.ubi.pt/~api/semantic_web.pdf
- BERNERS-LEE, T. (2000). Axioms of Web Architecture - Rules and Facts: Inference engines vs Web. *World Wide Web Consortium*. Acesso em: 20 de Outubro de 2004, Disponível em: <http://www.w3.org/DesignIssues/Rules.html>
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The semantic web – a new form of the Web content that is meaningful to computer will unleash a revolution of new possibilities**. Scientific American, Maio, 2001. Disponível em: http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21.
- BIEBER, M. e Vitali, F. **Toward support for Hypermedia on the World Wide Web**, Computer, 1997.
- BOLEY, H., Dean, M., Grosz, B., Kifer, M., Tabet, S., & Wagner, G. (2005a). RuleML Position Statement, *W3C Workshop on Rule Languages for Interoperability Position Paper*. Washington, D.C., USA: The RuleML Initiative - WWW Consortium.
- BOURDA, Yolaine; DOAN, Bich-Liên. **The Semantic Web for learning Resources**. IEEE Computer Society, 2003. Dispo-

- nível em: <http://csdl.computer.org/comp/proceedings/icalt/2003/1967/00/19670322.pdf>
- BRAY, T., Hollander, D., & Layman, A. (1999). Namespaces in XML. *World Wide Web Consortium*. Acesso em: 16 de Janeiro de 2006, Disponível em: <http://www.w3.org/TR/1999/REC-xml-names-19990114/>
- CUNHA, Luiz. **Web Semântica: Estudo Preliminar**, Sao Paulo: Embrapa, 2002. Disponível em: <http://www.cnptia.embrapa.br/publica/2002/doc18.pdf>
- GARSHOL, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all. *Ontopia*. Acesso em: 28 de Fevereiro de 2006, Disponível em: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>
- DAVIES, J., Fensel, D., & Harmelen, F. v. (2003). *Towards The Semantic Web – Ontology-Driven Knowledge Management*. Chichester, England: John Wiley & Sons, Ltd.
- DCMI. **Dublin Core Metadata Initiative**, 2002. Disponível em: <http://www.dublincore.org>
- DORI, Dov. **Object-Process Methodology as a basis for the Visual Semantic Web**. IEEE Computer Society, 2003. Disponível em: <http://csdl.computer.org/comp/proceedings/dexa/2003/1993/00/19930617abs.htm>
- FARIAS, C. G. de; ROSÁRIO, G. **Uma análise da web semântica e suas implicações no acesso à informação**. Disponível em: <http://maae.deinf.ufma.br/Ensino/IA/Uma%20An%C3%A1lise%20da%20Web%20Sem%C3%A2ntica%20e%20suas%20Implica%C3%A7%C3%B5es%20no%20Acesso%20%C3%A0%20Informa%C3%A7%C3%A3o.PDF>
- GOÑI, J.; FERNANDES, C. e LUCENA, C.. **E-learning e a Web Semântica**. VI Congresso da RIBIE, Pontifícia Universidade Católica do Rio de Janeiro, 2002. Disponível em: <http://lsm.dei.uc.pt/ribie/docfiles/txt2003729185022paper-212.pdf>
- HARMELEN, Frank. **The Semantic Web: What, Why, How, and When**. IEEE Intelligent Systems. Março, 2004.
- HAROLD, Elliotte Rusty. **XML Bible**. - 1. ED.- Chicago: IDG Books, 1999.
- HENDLER, J. (2001). Agents and the Semantic Web. *IEEE Intelligent Systems Journal*. Acesso em: 09 de Maio de 2004, Disponível em: <http://www.cs.umd.edu/users/handler/AgentWeb.html>
- HORROCKS, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosf, B., & Dean, M. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Member Submission*. Acesso em: 14 de Novembro de 2006, Disponível em: <http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/>
- HORROCKS, I., Patel-Schneider, P. F., & Harmelen, F. V. (2003). From

- SHIQ and RDF to OWL: The Making of a Web Ontology Language. *Journal of Web Semantics*, 1(1), 7-26.
- LASSILA, O. e SWICK, R. R.. **Resource Description Framework (RDF) Model and Syntax Specification**. W3C Recommendation 22 de Fevereiro de 1999. Disponível em: <http://www.w3.org/TR/REC-rdf-syntax>.
- LÉVY, Pierre. **Semantic Web: A tool for Education and Human Development**. University of Ottawa, Junho de 2003. Disponível em: <http://www.ricardoneves.eti.br/Docs/Pierre%20Levy%20Slides%20Web%20Semantica%20portugu%EA.s.ppt#1>
- LIBRELOTTO, G. R. (2005). *Topic Maps: da Sintaxe à Semântica, Dissertação de doutoramento*. Universidade de Braga, Braga.
- MAEDCHE, Alexander and STAAB, Steffen. **Learning Ontologies for the Semantic Web**. In: Semantic Web 2001 (at WWW10). May 1, Hongkong, China. Disponível em: <http://www.aifb.uni-karlsruhe.de/WBS/sst/Research/Publications/semweb2001.pdf>.
- MOURA, Ana M. de C. **A web semântica: fundamentos e tecnologias**. Rio de Janeiro: IME, 2002. Disponível em: <http://ipanema.ime.eb.br/~anamoura/publicacoes.html>
- PEREIRA, Rui. **Semantic Web**. Covilhã: Universidade da Beira Interior, 2003.
- PIMENTEL, M. G.C., TEIXEIRA, C. A. C., SANTANCHE, A. **XML: Explorando suas Aplicações na Web**. Congresso da Sociedade Brasileira de Computação, Curitiba 2000. Actas da XIX Jornada de Atualização em Informática.
- RAMALHO, Franklin. **Web Semântica**. I Workshop de Dissertações da COPIN, Universidade Federal de Campina Grande, Agosto 2002. Disponível em: <http://www.dsc.ufcg.edu.br/~copin/eventos/IWDCOPIN/apresentacoes/WebSemantica.ppt>
- REZENDE, B. e BAX, M.. **Projecto Indexa: ferramenta de auxílio à divulgação de informações na Web**, Universidade Federal de Minas Gerais, 2002. Disponível em: http://cuba.paradigma.com.br/paradigma/artigos/artigos_04.pdf
- STAAB, S., MAEDCHE, A.. **knowledge Portals - Ontologies at Work**. Institute AIFB, University of Karlsruhe, Ontoprise Gmb H. Disponível em: <http://www.aifb.unikarlsruhe.de/WBS/Publ/pub2001.html>.
- THOMPSON, Jon. Features: **Semantic Web – Straight Talking**. PC Pro, Maio 2004.
- WILLIAMSON Ben e MILLER Libby. **The semantic web: A touch of intelligence for the internet?**. The Guardian, 21 de Junho 2003. Disponível em: <http://education.guardian.co.uk/print/0,3858,4695866-108699,00.html>

Mais informação em:

W3C. **Semantic Web**, 2003. Disponível em: <http://www.w3.org/2001/sw/>

W3C. **XML**, 2003. Disponível em: <http://www.w3.org/xml/>

W3C. **RDF**, 2003. Disponível em: <http://www.w3.org/rdf/>

W3C. **Ontologias**, 2003. Disponível em: <http://www.w3.org/2001/sw/WebOnt/SemanticWeb.org>. **The Semantic Web Community**

Portal, 2003. Disponível em: <http://www.semanticweb.org>