

# 1 Introdução

## 1.1 A Web Semântica

A Web Semântica é definida por seus idealizadores como uma extensão da Web atual, onde as informações recebem um significado bem definido, permitindo maior cooperação entre pessoas e computadores (Berners-Lee et al., 2001). As argumentações e prol da Web Semântica se baseiam no fato de que, atualmente, os dados disponíveis na WWW possuem apenas informações de apresentação dos mesmos. A linguagem HTML utilizada para publicação das páginas Web permite, por exemplo, dizer que um texto deve ser apresentado em negrito, como um cabeçalho, ou em uma cor diferente. Entretanto, o crescimento da Web tem demonstrado que tais informações não bastam para executarmos nossas tarefas da forma mais eficiente, tirando maior proveito da capacidade de processamento de nossos computadores.

Imagine o cenário no qual uma pessoa deseja obter informações na Web sobre a famosa escultura da Vênus de Millo. Uma busca pela palavra “Vênus” nos mecanismos de busca mais populares retornará uma lista de resultados com páginas que falam sobre a referida escultura, mas também páginas sobre o planeta Vênus, a deusa da mitologia romana, a pintura de Botticelli e muitos outros tópicos sem relação alguma com o que se procurava. Cabe então à pessoa o trabalho de separar o joio do trigo. Isso acontece, porque o mecanismo de busca não consegue fazer a distinção entre os diferentes domínios das ocorrências da palavra “Vênus”. O código HTML que ele interpreta simplesmente não contém nenhuma informação que possa ajudá-lo nesse sentido.

A idéia da Web Semântica é justamente dotar o conteúdo da Web de informações adicionais na forma de metadados. Esses metadados permitiriam, por exemplo, a criação de mecanismos de busca que identifiquem a diferença entre informações sobre uma escultura, um planeta, uma figura mitológica, uma pintura, etc. Ao se representar os metadados de forma que computadores também possam

entender estas informações, o poder analítico das máquinas pode ser aproveitado com mais eficiência (Berners-Lee, 2000). Podem ser criados programas que não apenas buscam, transmitem e exibem as informações da Web, mas também as utilizam para automação e integração em diversas aplicações (W3C Semantic Web, 2004). Os primeiros passos nesse sentido já foram dados, com a especificação por parte do W3C (*World Wide Web Consortium*), de um conjunto de padrões que constituem os pilares fundamentais da futura Web Semântica.

A base de todo o trabalho é o padrão RDF (*Resource Description Framework*), um modelo de dados onde os itens de interesse podem ser tratados como recursos, identificados por um URI (*Uniform Resource Identifier*), e descritos com base em propriedades simples e valores de propriedades. Esta abordagem também permite que as informações descritas em RDF sejam representadas como um grafo, com nós e arcos representando os recursos, propriedades e seus valores.

Foram definidos outros dois padrões que especificam vocabulários para definição de classes e propriedades de recursos RDF. O primeiro deles, o padrão RDFS (*RDF Schema*) permite a definição de classes e propriedades, e inclui uma semântica para criação de hierarquias de generalização destas classes e propriedades. O segundo padrão é a linguagem OWL (*Web Ontology Language*), que utiliza RDF e RDFS como base, e define um vocabulário mais rico para descrição de classes e propriedades, introduzindo semântica adicional, como por exemplo, classes disjuntas, propriedades simétricas e cardinalidade de propriedades. Estes dois padrões podem ser empregados para a criação de vocabulários específicos, e a apresentação de uma informação, utilizando classes e propriedades destes vocabulários, é uma forma de ao mesmo tempo estabelecer e informar qual deve ser a interpretação utilizada. De agora em diante, utilizaremos a expressão “modelo RDF” ou “modelo de dados RDF” para denotar as representações em RDF de informações sobre um dado domínio que estejam de acordo com os padrões RDFS ou OWL.

A seguir é apresentado um exemplo de como a informação de que *Miguel de Cervantes é o autor de Don Quixote* pode ser apresentada em RDF com uma sintaxe em XML.

```

<ex:Book rdf:ID="http://www.example.org/DonQuixote">
  <dc:title>Don Quixote</dc:title>
  <dc:creator>
    <foaf:Person rdf:ID="http://www.example.org/Cervantes">
      <foaf:name>Miguel de Cervantes</foaf:name>
    </foaf:Person>
  </dc:creator>
</ex:Book>

```

Quadro 1 - Exemplo de declaração de dados em RDF/XML

Neste exemplo o livro Don Quixote é tratado como um recurso identificado pelo URI `http://www.example.org/DonQuixote` e Cervantes é tratado como outro recurso identificado pela URI `http://www.example.org/Cervantes`. Em resumo, são utilizadas propriedades e classes dos vocabulários *Example* (prefixo *ex:*), *Dublin Core* (prefixo *dc:*) e *Friend of a Friend* (prefixo *foaf:*), para declarar que o recurso `http://www.example.org/DonQuixote` é um livro (*ex:Book*), cujo título (*dc:title*) é “Don Quixote”, e que foi criado (*dc:creator*) pelo recurso `http://www.example.org/Cervantes`, que é uma pessoa (*foaf:Person*) cujo nome (*foaf:name*) é “Miguel de Cervantes”. A figura abaixo mostra a mesma informação representada como um grafo RDF.

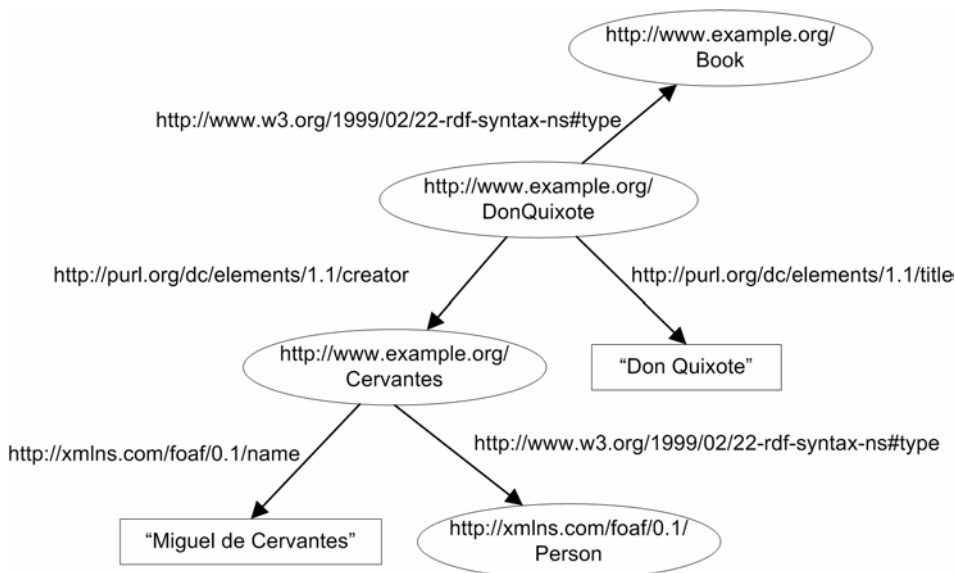


Figura 1 - Exemplo de um grafo RDF

Os padrões definidos até o momento estão voltados para solução da primeira parte do desafio de tornar a Web Semântica uma realidade: Acrescentar significado às informações disponíveis na Web, fazendo isso de forma que estas

informações possam ser entendidas por computadores. A segunda parte do desafio é o uso efetivo dos padrões, na criação dos programas que através da interpretação destas informações permitirão a automação e integração de aplicações. Esta é uma tarefa que envolve a solução de uma série de problemas menores, podendo-se tomar os seguintes como exemplo: a construção de ferramentas para autoria de conteúdo semântico, a definição de mecanismos efetivos para busca e reuso de vocabulários, a criação de tecnologias para armazenamento, processamento e busca de informações semânticas, e a exploração destas informações por seres-humanos para realização de alguma tarefa. A maior parte dos trabalhos ligados a Web Semântica até o presente enfocam primordialmente a questão da viabilização do processamento da informação por programas que trocam dados entre si. O enfoque do trabalho aqui apresentado, no entanto, são as aplicações interativas na Web Semântica, onde seres humanos exploram e interagem com a informação.

## **1.2 Explorando dados na Web Semântica**

Considerando-se que um dos objetivos finais da Web Semântica é permitir uma maior cooperação entre pessoas e computadores, ou colocando de outra forma, fazer com que computadores permitam às pessoas tirar maior proveito das informações disponíveis na Web, as aplicações da Web Semântica podem ser vistas como parte de um time homem-máquina empenhado na realização de uma tarefa (Rossi, 1996). Sob esta ótica, a execução de uma tarefa é caracterizada por um cenário no qual uma pessoa explora e analisa informações, para então instruir a máquina a executar uma ação específica. Ao se considerar este cenário no contexto atual da Web Semântica, fica visível a necessidade de se criar maneiras que permitam a exploração, por seres-humanos, das informações representadas em RDF, formatadas para máquinas.

Nossa proposta para o tratamento desta questão é a apresentação dessas informações como uma aplicação hipermídia, desenvolvida segundo o método SHDM (*Semantic Hypermedia Design Method*) (Lima, 2003). Nessa abordagem, a aplicação hipermídia em questão define uma visão navegacional sobre dados da Web Semântica, que representam o modelo conceitual da aplicação. Esta visão por sua vez é especificada com base na análise de cenários e casos de uso

descrevendo as tarefas a serem desempenhadas pelos usuários, permitindo a identificação das informações e operações envolvidas nessas tarefas. A separação entre o modelo conceitual e o modelo navegacional da aplicação vem do reconhecimento de que, assim como no caso de aplicações tradicionais para WWW, em aplicações para a Web Semântica, usuários com perfis distintos, ou desempenhando tarefas diferentes, podem necessitar de diferentes visões navegacionais sobre um mesmo conjunto de informações, e que as abstrações fornecidas pelo modelo conceitual não são suficientes para especificar aplicações que incorporam a metáfora do hipertexto, como discutido em Rossi et al. (1999).

Em sua maioria, outras abordagens encontradas para aplicações interativas na Web Semântica não incorporam o conceito de visão navegacional, apresentando diretamente todos os dados do grafo RDF. Algumas são literalmente visualizadores gráficos do grafo RDF, apresentando as informações em um formato muito semelhante ao da representação mostrada na Figura 1. Exemplos do uso dessa abordagem são o IsaViz<sup>1</sup> e o Fentwine<sup>2</sup>, que também funcionam como ferramentas para autoria de conteúdo RDF. Outras soluções apresentam as informações do grafo RDF como páginas HTML, fazendo isso através de navegadores/editores RDF como o BrownSauce<sup>3</sup>, SWOOP<sup>4</sup> e Foaf Explorer<sup>5</sup> (específico para dados definidos segundo a ontologia *Friend of a Friend*), ou através da criação de *websites* para um domínio específico como os projetos Esperonto<sup>6</sup> e Mindswap<sup>7</sup>, ou ainda com o uso de transformações XSLT para converter RDF representado na sintaxe RDF/XML diretamente em páginas XHTML.

Também existem abordagens que empregam o princípio de um modelo navegacional distinto do modelo conceitual como, por exemplo, o OntoWeaver (Lei et al., 2002, 2003). Nossa solução é baseada na experiência prévia com o método OOHDM (*Object Oriented Hypermedia Design Method*), do qual o

---

<sup>1</sup> - <http://www.w3.org/2001/11/IsaViz/>

<sup>2</sup> - <http://www.w3.org/2001/sw/Europe/events/foaf-galway/papers/pp/fentwine/>

<sup>3</sup> - <http://brownsauce.sourceforge.net/>

<sup>4</sup> - <http://www.mindswap.org/2004/SWOOP/>

<sup>5</sup> - <http://xml.mfd-consult.dk/foaf/explorer/>

<sup>6</sup> - <http://www.esperonto.net>

<sup>7</sup> - <http://www.mindswap.org>

método SHDM é uma evolução, e cujas abstrações já se mostraram, na prática, adequadas para o desenvolvimento de aplicações hipermídia.

### **1.3 Organização da dissertação**

Os capítulos a seguir estão organizados da seguinte forma. O capítulo 2 apresenta a evolução do método SHDM, enfocando as principais alterações e extensões introduzidas, para que fosse possível desenvolver uma aplicação hipermídia com quaisquer dados da Web Semântica. No capítulo 3 é apresentada em detalhes a forma de declaração da visão navegacional. O capítulo 4 contém um exemplo completo da especificação de do modelo conceitual e do mapeamento navegacional de uma aplicação. No capítulo 5 é apresentada a arquitetura de implementação desenvolvida para a solução proposta. Finalmente no capítulo 6 são apresentadas as conclusões finais e sugestões de trabalhos futuros.