

1 Introdução

A Web Semântica é uma visão [W3C, 2001b]: uma idéia de termos dados na Web definidos e conectados de modo a serem utilizados por máquinas não só com objetivo de apresentação, mas também para automação, integração e reuso de dados através de várias aplicações. Atualmente, a Web Semântica tem sido o centro de atenções de muitos esforços, tanto da área acadêmica quanto da industrial, uma vez que é considerada o próximo passo evolutivo da *World Wide Web* (WWW). O objetivo principal desta Web do futuro é alcançar um estágio onde uma imensa quantidade de dados estará disponível com seus metadados para auxiliar tanto pessoas quanto máquinas a encontrar e processar recursos úteis [Berners-Lee, 2000], e também trocar dados entre aplicações [W3C, 2002].

Uma grande ênfase tem sido dada no processo de localizar recursos úteis na Web Semântica, e diversas propostas têm sido feitas para organização e busca. Mas sabemos que só isto não basta. Encontrar um documento (ou recurso qualquer) é apenas uma parte do problema (a primeira parte, é verdade!). Mas uma vez que o usuário encontra um documento que deseja, quais são as possíveis tarefas que ele executará em seguida? Ele irá navegar para outro documento com o mesmo critério de seleção do documento corrente? Ou preferirá ver mais informações a respeito de um tópico específico encontrado no documento inicial? As respostas para estas questões podem ser obtidas através de um espaço navegacional bem definido, com base na análise das tarefas e perfis dos usuários. E nosso objetivo é justamente projetar as aplicações que oferecerão suporte a este espaço navegacional, ou seja, suporte às navegações desejadas.

Para muitas pessoas a Web é importante devido à funcionalidade oferecida por aplicações Web, através das quais pode-se fazer acesso e processar informações armazenadas na própria Web. Portanto, muito do que é almejado para a Web Semântica só será possível alcançar se as aplicações Web puderem tirar proveito da meta-informação disponível. Assim, além dos metadados descrevendo os recursos da Web, precisamos disponibilizar também metadados a respeito de aplicações.

Para descrever os metadados das aplicações, desenvolvemos um método chamado SHDM (*Semantic Hypermedia Design Method*) [Lima & Schwabe 2002a, 2002b, 2003a, 2003b], tema principal desta tese. O método é baseado em nossa experiência em projeto de aplicações Web e hipermídia com o método OOHDM (*Object Oriented Hypermedia Design Method*) [Rossi, 1996, Rossi et al., 1999, Schwabe & Rossi, 1998, Schwabe et al. 2001b]. Durante esta pesquisa estamos interessados em usar o conhecimento representado em ontologias e nosso maior interesse é em processar este conhecimento, tirando proveito da nova infra-estrutura que está sendo construída para a Web Semântica.

O método SHDM usa uma abordagem iterativa baseada em modelos para projeto de aplicações Web, organizada em cinco etapas: Levantamento de Requisitos, Modelagem Conceitual, Modelagem Navegacional, Projeto de Interface Abstrata e Implementação. Cada etapa faz foco em um aspecto particular e produz modelos, descrevendo detalhes a respeito da aplicação que será executada na Web. O método SHDM mantém a separação entre projeto conceitual e projeto navegacional, um importante destaque do método OOHDM. Ao separarmos explicitamente os projetos conceitual e navegacional, endereçamos diferentes aspectos das aplicações Web. Enquanto o projeto conceitual deve expressar objetos e comportamentos do domínio da aplicação, o projeto navegacional tem como objetivo organizar o espaço navegacional, levando em consideração perfis de usuários e suas tarefas. O projeto navegacional é uma atividade chave na implementação de aplicações Web, e nós incentivamos uma separação explícita do projeto conceitual [Rossi et al., 1999]. No SHDM, a etapa de projeto navegacional produz modelos ainda mais expressivos, capazes de representar aplicações Web e até mesmo famílias de aplicações Web.

Apresentamos nossa abordagem para o desafio de criar modelos conceituais e navegacionais para domínios baseados em ontologias. Atualmente, as linguagens DAML+OIL [Connolly et al., 2001] e OWL [Smith et al., 2002] (com RDF [Lassila & Swick 1999] e RDF(S) [Brickley & Guha, 2000]) vêm sendo utilizadas para construir ontologias em domínios complexos, através da descrição de classes e propriedades. Além do uso destas linguagens mencionadas, enriquecemos a etapa de projeto navegacional fornecendo primitivas para especificar navegações facetadas, que expressam de forma concisa as abstrações para estruturas de acesso facetadas e contextos facetados. Estes novos recursos permitem o projeto de estruturas de acesso mais ricas, fornecendo ao usuário formas mais flexíveis de alcançar conjuntos de

objetos que são relevantes para a tarefa a ser executada. Oferecemos também primitivas para permitir ao projetista da aplicação descrever concisamente hierarquias de metadados facetados.

Atualmente estamos utilizando o método SHDM em um ambiente de “mundo fechado” com definições de recursos locais; perfis de usuários e evolução serão abordados em trabalhos posteriores. Antevemos o uso do método em um modo semi-automático, com uma evolução modular da aplicação, tirando proveito da capacidade de inferência das linguagens de definição escolhidas.

1.1. A Web Atual

A Web é cada vez mais uma realidade em nossas vidas. A cada dia aumenta o número de usuários e a quantidade de dados disponíveis para acesso. Já presenciamos a primeira geração da Web com suas páginas estáticas confeccionadas individualmente à mão. Em seguida, vivenciamos a segunda geração composta de páginas dinâmicas geradas automaticamente através de *templates* e fazendo acesso a bases de dados. Estas gerações tiraram proveito da simplicidade da linguagem HTML, mas também tiveram que enfrentar suas limitações e buscar inovações para resolver questões para as quais a linguagem não oferecia suporte.

Surgiu então uma nova linguagem, XML (*eXtensible Markup Language*) [Bray et al., 2000] que, devido a maior flexibilidade, causou grande impacto e popularidade. Observamos, no entanto, que apenas XML não é o suficiente, aliás XML é um ótimo ponto de partida... de um longo caminho que ainda estamos percorrendo.

A Tabela 1 descreve a evolução da Web em diferentes aspectos: no eixo Y, vemos as diferentes linguagens de codificação; as diferentes formas de criação das páginas e o uso destas páginas por pessoas e/ou aplicações; a evolução dos paradigmas; a evolução da forma do usuário pensar; a evolução das aplicações; exemplos das áreas de domínio comuns e as principais aplicações. No eixo X, podemos observar as diferentes gerações da Web e também a linha do tempo.

Vale comentar que na época da elaboração deste trabalho as questões transacionais ocupavam posição de destaque em ambientes empresariais e

também em projetos de pesquisa. O tópico que tratamos, a Web Semântica, tem como data prevista, o ano de 2005.

	Estática	Dinâmica	Transacional	Semântica
Codificação	HTML, CGI, Perl, ...	+ SGDBRs, JSP, ASP, Java,...	+XML, S2EE, .NET, ...	+RDF, OWL, ?
Criação	Criação manual por pessoas Uso por pessoas	<i>Templates</i> criados por pessoas Uso por pessoas	Criação por aplicações baseadas em esquemas fixos Uso por aplicações e pessoas	Criação por aplicações baseadas em modelos Uso por aplicações, dispositivos e pessoas
Paradigma	Informação, um grande jornal	Jornal se transforma em catálogo	Catálogo se transforma em plataforma transacional	Plataformas se conectam
"Pensamento"	Navegar	Recuperar/Atualizar	Interagir	Interoperar
Ex. de Aplic	Marketing	Vendas	Serviços	Integração
Principais Aplics	• Navegador	• Busca • Gerenciamento de conteúdo • Servidores de Aplicação Web	• Portais • Processos de integração • Serviços Web	• Agentes pessoais • Aplicações IP • <i>Advisors</i> • Ferramentas Cognitivas
	1995	2000	2005	

Tabela 1 - A Evolução da Web [TopQuadrant 2003]

1.2. A Web Semântica

Tim Berners-Lee et al. [2001] afirmam que a Web Semântica é uma extensão da Web atual, onde um significado é fornecido a cada informação disponível, permitindo que computadores e pessoas trabalhem em cooperação. De acordo com [W3C 2001b], a Web Semântica pode ser definida como a representação abstrata dos dados na WWW, com base nos padrões RDF e em outros padrões a serem definidos. Esta nova Web está sendo desenvolvida sob a coordenação do Consórcio W3C (*World Wide Web Consortium*) [W3C 2001a] em colaboração com um grande número de pesquisadores e parceiros industriais.

Esforços para disponibilizar metadados na Web têm se tornado prioritários para diversas comunidades. A Web só alcançará seu potencial total se chegar a ser um local onde dados possam ser compartilhados e processados, tanto por ferramentas automáticas, quanto por pessoas. Para a Web evoluir, os futuros programas devem ser capazes de compartilhar e processar dados, mesmo que estes programas tenham sido projetados de forma totalmente independente.

Neste caminho rumo à Web Semântica, observamos diferentes investigações visando a disponibilização dos dados da Web como um grande sistema de representação do conhecimento, utilizando linguagens de definição de ontologias como DAML+OIL e a nova linguagem ainda em desenvolvimento pelo consórcio W3C, chamada *Web Ontology Language* ou apenas OWL. Também podemos observar que pesquisas a respeito de Serviços Web (*Web Services*) estão expandindo rapidamente, devido ao aumento da necessidade de interoperabilidade entre aplicações.

Até recentemente a Web Semântica era muito pouco além de um nome para a próxima geração de infra-estrutura da Web, como imaginado por um de seus criadores, Tim Berners-Lee. Com o surgimento de XML e RDF, e também de novas propostas como RDF(S), DAML+OIL e OWL, a Web Semântica está rapidamente tomando forma.

1.3. Contribuições

Nesta tese, diversas contribuições são apresentadas, tanto contribuições relativas a Web Semântica, quanto contribuições relativas a evolução de métodos para projeto de aplicações na Web (ex.: método OOHDM).

Vale destacar que os pilares essenciais da abordagem OOHDM foram mantidos, sendo eles: (1) existe um modelo navegacional diferente do modelo conceitual e há um mapeamento entre os dois; (2) no modelo navegacional temos contextos de navegação e (3) existe uma interface abstrata. Podemos afirmar que a principal contribuição do OOHDM para o desenvolvimento de aplicações na Web é justamente a identificação de que modelos navegacionais não são modelos conceituais, isto é, não devemos navegar em nosso modelo conceitual, conforme detalhado em [Rossi et al., 1999]. O SHDM manteve este princípio importante, que não vem sendo utilizado pelas primeiras aplicações desenvolvidas para a Web Semântica, como Esperonto [Esperonto, 2003] OntoWebber [Jin et al., 2001] e Mindswap [Mindswap 2003] (veja subseção 7.1).

Nestas aplicações os usuários navegam no modelo conceitual, pois não existe modelo navegacional.

As principais contribuições desta tese são:

- um método com abordagem abrangente, considerando as diversas etapas de projeto de aplicações Web, integrado com os formalismos da Web Semântica. O método SHDM (*Semantic Hypermedia Design Method*) apresentado é na verdade, uma evolução da versão anterior chamada OOHDM (*Object Oriented Hypermedia Design Method*);
- um conjunto de primitivas de modelagem a ser utilizado em cada uma das etapas de projeto. O SHDM apresenta novas abstrações, que são mais ricas e permitem especificação de um conjunto maior de aplicações Web de forma concisa e não ambígua;
- a especificação dos metadados de aplicações Web através de extensões de linguagens propostas pelo consórcio W3C, como XML, RDF, RDF(S), DAML+OIL e OWL;
- um modelo navegacional conciso e com maior poder expressivo, devido ao uso de uma linguagem de consulta apropriada, que permite que a definição de primitivas navegacionais seja feita com base tanto nos dados quanto nos metadados. Pelo fato de poder consultar metadados, é possível definir-se famílias de aplicações com características comuns;
- uma nova primitiva para a especificação de estruturas de acesso, através da incorporação do conceito de facetas. Esta primitiva permite que se construa índices que dão suporte à seleção incremental de objetos de navegação utilizando múltiplos critérios simultaneamente;
- uma arquitetura de software que utiliza a própria especificação da aplicação, descrita nos formalismos propostos, para implementá-la, gerando páginas dinamicamente. Esta arquitetura garante a aderência da implementação à especificação, o que não ocorria anteriormente.

1.4. Estrutura da Tese

Esta tese apresenta o método SHDM (*Semantic Hypermedia Design Method*), uma evolução do método OOHDM (*Object Oriented Hypermedia Design Method*), que visa modelar aplicações da futura Web Semântica.

No Capítulo 2 revisamos brevemente os principais fundamentos utilizados das diversas áreas relacionadas. A Web Semântica é uma área de pesquisa

recente, e vem sendo construída com base em novas tecnologias para Web, como novas linguagens de definição de ontologias e de consulta, e também novas arquiteturas de implementação. Neste mesmo capítulo fazemos um resumo do método OOHDM.

No Capítulo 3 apresentamos um resumo dos principais conceitos do novo método SHDM e suas etapas, mostrando como é possível tirar proveito de primitivas de modelagem mais expressivas, ao estender o método OOHDM.

Duas etapas do SHDM são detalhadas nos capítulos seguintes. A etapa de Projeto Conceitual é descrita no Capítulo 4, enquanto que o Projeto Navegacional é abordado no Capítulo 5.

Uma Arquitetura de Implementação é descrita no Capítulo 6 e também são feitos comentários a respeito de outras implementações possíveis no ambiente da Web Semântica do futuro.

As considerações finais deste trabalho são apresentadas no Capítulo 7, juntamente com os trabalhos relacionados e os trabalhos futuros.

O Apêndice expõe dois exemplos de aplicações modeladas com o método, uma ontologia do domínio de Artes e outra do domínio de Vinhos, ilustrando como as novas primitivas podem ser utilizadas. Os Anexos apresentam duas tabelas comparativas elaboradas por outros autores e citadas ao longo do texto.