

InfoArch: uma ontologia para modelar o domínio da Arquitetura da Informação para Web

Marckson Roberto Ferreira de Sousa^{*}

Edilson Leite da Silva^{**}

Guilherme Ataíde Dias^{***}

Maria Amélia Teixeira da Silva^{****}

Frederico Luiz Gonçalves de Freitas^{*****}

Ryan Ribeiro de Azevedo^{*****}

Resumo Neste trabalho é apresentada uma ontologia para modelar o domínio da Arquitetura da Informação para Web (AI para web), de acordo com os preceitos definidos por Morville e Rosenfeld no livro *Information Architecture for the Word Wide Web*, versão 2006. Objetiva-se estruturar o conhecimento relacionado ao domínio de AI para Web, formalizando o mesmo, bem como auxiliar o ensino dos conceitos e relacionamentos do domínio da área de AI para Web. A pesquisa realizada é de caráter teórico e qualitativo, e classifica-se como descritiva e exploratória. A modelagem foi realizada mediante a utilização da linguagem *Ontology Web Language* (OWL) e do *framework Protégé* 3.4.1, seguindo os passos da metodologia 101. Os resultados mostram a InfoArch, uma ontologia que representa os conceitos e relacionamentos, além de possibilitar responder a questionamentos sobre o domínio. Considera-se que InfoArch

* Doutor em Engenharia Elétrica pela Universidade Federal da Paraíba (UFPB) e professor adjunto do Departamento de Ciência da Informação (DCI) da UFPB. Endereço postal: UFPB, Centro de Ciências Sociais Aplicadas, DCI, Cidade Universitária, Castelo Branco, Campus I, João Pessoa, Paraíba, CEP: 58051-900. Tel. (83) 3216-7264 e e-mail marckson.dci.ufpb@gmail.com.

** Mestre em Ciência da Computação pela Universidade Federal de Pernambuco (UFPE) e professor do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) e do Centro Universitário de João Pessoa (Unipê). Endereço postal: Unipê, BR 230, Km 22, Água Fria, João Pessoa, Paraíba, CEP. 58053-000. Tel. (83) 2106-9274 e e-mail els3@cin.ufpe.br.

*** Doutor em Ciência da Comunicação pela Universidade de São Paulo (USP) e professor adjunto do Departamento de Ciência da Informação (DCI) da UFPB. Endereço postal: UFPB, Centro de Ciências Sociais Aplicadas, DCI, Cidade Universitária, Castelo Branco, Campus I, João Pessoa, Paraíba, CEP. 58051-900. Tel. (83) 3216-7264 e e-mail guilherme@dci.ccsa.ufpb.br.

**** Mestranda em Ciência da Informação na UFPB e bolsista da Capes pelo Programa de Pós-Graduação em Ciência da Informação da UFPB. Endereço postal: UFPB, Centro de Ciências Sociais Aplicadas, DCI, Cidade Universitária, Castelo Branco, Campus I, João Pessoa, Paraíba, CEP. 58051-900. Tel. (83) 3216-7483 e e-mail melteixeiraufpb@gmail.com.

***** Doutor em Engenharia Elétrica pela Universidade Federal de Santa Catarina (UFSC) e professor adjunto do Centro de Informática (CIN) da UFPE. Endereço postal: UFPE, CIN, Av. Luz Freire, s/n, Cidade Universitária, Recife, Pernambuco, CEP. 50740-560. Tel. (81) 2126-8430 e e-mail fred@cin.ufpe.br.

***** Doutorando em Ciência da Computação pela UFPE e professor assistente do Departamento de Sistemas de Informação da Universidade Federal do Piauí (UFPI). Endereço postal: UFPE, Centro de Informática, Av. Luz Freire, s/n, Cidade Universitária, Recife, Pernambuco, CEP. 50740-560. Tel: (81) 2126-8430 e e-mail rra2@cin.ufpe.br.

traz contribuições principalmente relativas as questões de ensino, pesquisa e extensão, pois servirá como fonte de informação para pesquisadores, professores e equipes de desenvolvimento de *sites* que trabalhem com Arquitetura da Informação para *Web*.

Palavras-chave Arquitetura da Informação para *Web*; ontologia; desenvolvimento de *site*; *web* semântica

InfoArch: an ontology for modeling the field of Information Architecture for the World Wide Web

Abstract This paper presents an ontology to model the field of Information Architecture for Web (Web IA), according to the precepts defined by Morville and Rosenfeld in the book *Information Architecture for the World Wide Web*, version 2006. It aims to structure the related knowledge in the field of IA for the Web, formalizing this area and helping to teach the concepts and relationships in the domain of the IA for the Web. The research is theoretical and qualitative, and is classified as descriptive and exploratory. The modeling was performed using the language of Ontology Web Language (OWL) and the Protégé framework 3.4.1, by following the steps of the methodology 101. The results show the InfoArch, an ontology that represents concepts and relationships, and enables finding answers to questions about the area. It is considered that InfoArch brings contributions especially on issues of teaching, research and extension, and thus will serve as a source of information for researchers, teachers and staff of developing sites that work with Information Architecture for the World Wide Web.

Keywords information architecture for the World Wide Web; ontology; site development; semantic web.

Introdução

As informações disponíveis na *web* devem ser melhor organizadas para que os usuários possam recuperá-las de forma mais rápida e significativa. As transformações que normalmente acontecem nas páginas *web*, com o objetivo de torná-las mais interativas e dinâmicas, por meio da inserção de informações sem critérios de organização, podem gerar desconforto aos usuários, que nem sempre conseguem localizar o que desejam. Tal problema resulta em perdas significativas de informação e gera insatisfação nos usuários dos respectivos espaços informacionais, que nem sempre conseguem localizar o que desejam.

Segundo Souza e Alvarenga (2006), a *web* semântica oferece várias contribuições aos mais diversos campos de estudos relativos ao tratamento da informação. Tais contribuições se dão por meio de sua proposta de estruturar o conteúdo da *web* (BERNERS-LEE, 2001; BREITMAN, 2005; CARLAN, 2006). Dessa forma, a partir da organização das informações, é possível que agentes de software percorram a rede, página a página, para realizar tarefas consideradas “trabalhosas” para o usuário com o objetivo de recuperar a informação desejada. Para facilitar a

realização dessas tarefas, é fundamental o uso de ontologias, que fornecem um vocabulário compartilhado para modelar determinado domínio (área do conhecimento), seus conceitos, propriedades, relações, restrições (GRUBER, 1993; BERNERS-LEE, 2001). As ontologias são a principal camada da *web* semântica, existindo vários tipos que podem ser classificados de acordo com a sua finalidade.

Já a Arquitetura da Informação para *Web* define as formas básicas para disponibilizar informações nos *websites* por meio dos seus sistemas de organização, navegação, rotulação e busca, também denominados de “componentes da arquitetura da informação” (MORVILLE; ROSENFELD, 2006, p. 49, tradução nossa). Para que tais sistemas possam melhorar a estrutura desses *websites*, é necessário que sejam aplicadas durante o seu desenvolvimento metodologias com recomendações para cada etapa do processo, com base em estudos relativos à usabilidade, estudo do usuário e interface humano-computador (SAPIENT, 2000 apud MORROGH, 2003; MORVILLE; ROSENFELD, 2006; CAMARGO, 2010).

Segundo Morville e Rosenfeld (2006), existe uma pesquisa da *Vividence*, feita com 69 *websites*, apontando que 5 dos 7 principais problemas estão relacionados à Arquitetura de Informação, sendo estes: resultados de busca organizados inadequadamente, arquitetura de informação pobre, *home page* com informações confusas, rótulos confusos e navegação inconsistente.

Tanto a Arquitetura da Informação para *Web*, quanto as ontologias são propostas muito discutidas pelas comunidades de estudos relativos a essas temáticas. Em ambas, é possível conceber aplicações práticas que funcionem tanto no nível acadêmico quanto profissional, não existindo contudo uma integração entre essas propostas, uma vez que seus estudos e resultados são realizados de forma independente, embora ambas visem proporcionar facilidades para a recuperação da informação na *web*.

Muitas empresas perdem clientes porque os mesmos não conseguem encontrar os produtos, serviços ou informações que procuram, fazendo com que os mesmos abandonem o *site* e procurem outros, com a perspectiva de não retornar mais, senão quando extremamente necessário. Isso gera insatisfação dos clientes/usuários, os quais passam a procurar o que precisam em outros fornecedores, e assim, a empresa perde clientes e recursos monetários, conduzindo a prejuízos muitas vezes irreparáveis.

Este trabalho visa à integração das propostas da Arquitetura da Informação para *Web* e das ontologias, formalizando o domínio da AI para *web* usando ontologia, de forma a auxiliar no processo de desenvolvimento de *sites* que proporcionem facilidades de uso, e sobretudo de recuperação da informação.

A pesquisa parte da hipótese de que, ao modelar o conhecimento do domínio da Arquitetura da Informação para *Web* utilizando uma ontologia, é possível facilitar o uso da AI para *web* pelos desenvolvedores, os quais poderão utilizá-la na criação de seus *sites* bem como para auxiliar no ensino de disciplinas relacionadas à área. Para alcançar tal proposta, durante a pesquisa foi necessária à realização de atividades intermediárias, como:

- Formalizar o domínio da Arquitetura da Informação para *Web*, a partir dos seus principais conceitos, os componentes (sistemas) e a metodologia;
- Permitir o desenvolvimento de outras ontologias a partir deste trabalho;

- Possibilitar a utilização da ontologia para auxiliar no ensino da disciplina Arquitetura da Informação para *Web*, adotada em programas de pós-graduação em áreas de estudos relacionados à recuperação da informação;
- Apresentar cenários de utilização da ontologia, bem como as respectivas contribuições para estes cenários.

A pesquisa realizada é de caráter teórico e qualitativo, e classifica-se como descritiva e exploratória. Ela é teórica, por ter como objetivo definir e estruturar sistemas e modelos teóricos, desvendando quadros de referência sobre conceitos de Arquitetura da Informação para *Web* e Ontologias. É também qualitativa em virtude de buscar fornecer uma descrição rica em detalhes sobre a realidade específica de um domínio. É descritiva porque vem como complemento à abordagem qualitativa, descrevendo fatos observados, registrados, analisados, classificados e interpretados, sem interferência do pesquisador. Pode ser considerada ainda como exploratória, pois de acordo com o entendimento de Gil, toda pesquisa exploratória busca constatar algo num organismo ou num fenômeno, identificando fatores determinantes para a ocorrência dos fenômenos (GIL, 2000).

A ontologia desenvolvida para modelar o domínio da Arquitetura da Informação para *Web*, recebeu o nome de InfoArch.

Arquitetura da Informação para *Web*

O termo Arquitetura de Informação foi criado por Richard Saul Wurman em 1976, com o intuito de tornar a informação mais compreensível para todos, tendo em vista os problemas advindos da grande oferta de informação, que provocava nos usuários uma sensação de distanciamento entre o que se compreende e o que se deveria compreender (WURMAN, 1991). Seus conceitos foram inicialmente aplicados na organização de materiais gráficos, como guias, mapas e atlas, expandido-se posteriormente para a organização de *layout* de museus e estruturação de imagens radiográficas para uso médico.

No *design* de *websites* a Arquitetura da Informação para *Web* foi inicialmente abordada por Morville e Rosenfeld que, dentre outras definições, referem-se a AI para *web* como a combinação dos esquemas de organização, rotulação, busca e navegação dentro de *websites* e intranets. Complementando a definição anterior, o *Information Architecture Institute* (2002, p. 1), define a AI para *web* como “a arte e ciência de organizar e rotular *websites*, intranets, comunidades online e software, para suportar usabilidade”.

Nessa perspectiva,

A Arquitetura da Informação refere-se ao desenho das informações: como textos, imagens e sons são apresentados na tela do computador, a classificação dessas informações em agrupamentos de acordo com os objetivos do site e das necessidades do usuário, bem como a construção de estrutura de navegação e de busca de informações, isto é, os caminhos que o usuário poderá percorrer para chegar até a informação. (STRAIOTO, 2002 apud VIDOTTI; SANCHES, 2004, p. 2)

Segundo Morville e Rosenfeld (2006), a Arquitetura de Informação para *web* busca compreender e atender a três dimensões de variáveis para organizar a informação, sejam elas: o usuário, o conteúdo e o contexto, conforme mostrado na figura 1.

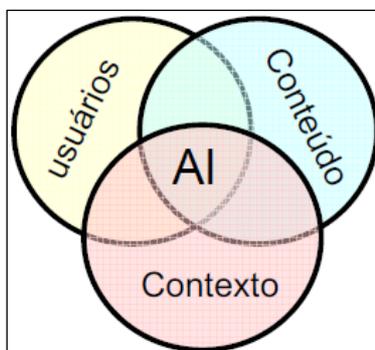


Figura 1: Três variáveis que dimensionam a organização da informação

Fonte: Adaptado de Morville e Rosenfeld (2006)

Conforme exposto, a primeira dimensão são os usuários, dos quais são analisados suas necessidades, hábitos e comportamentos; na segunda dimensão são analisadas as características do conteúdo que será apresentado (volume, formato, estrutura, governança, dinamismo, etc.); na terceira dimensão por sua vez, são verificadas as especificidades do contexto de uso do sistema de informação (objetivo do *website*, cultura e política da empresa, restrições tecnológicas, localização, etc.).

Para Morville e Rosenfeld (2006) os sistemas da Arquitetura da Informação para *Web*, são denominados como sistemas interdependentes, cada um com suas recomendações e aplicações próprias. São eles:

- Sistemas de Organização (*Organization System*) – agrupa e categoriza o conteúdo informacional;
- Sistema de Navegação (*Navegation System*) – determina a maneira de navegar, de mover-se pelo espaço informacional e hipertextual;
- Sistema de Rotulação (*Labeling System*) – determina as formas de representação e apresentação da informação, definindo cada um dos elementos informativos;
- Sistemas de Busca (*Search System*) – determina as perguntas que os usuários podem fazer e as respostas que podem obter;

Esses sistemas integram os componentes da AI para *web* e devem estar apoiados pelas Estruturas de Representação (*Structures of Representation*) que são os metadados, vocabulários controlados e tesouros. Neste trabalho também se considera que as ontologias fazem parte das estruturas de representação, porque também podem ser usadas para representar conhecimento.

Neste contexto, a Arquitetura da Informação para *Web* utilizando esses sistemas, objetiva: organizar; facilitar a navegação e a recuperação; rotular e representar a informação; e atribuir significado às informações disponíveis nos *websites*, de modo a facilitar a recuperação destas por parte dos usuários dos referidos espaços informacionais.

Metodologia para Arquitetura da Informação para Web

Existem várias propostas de metodologias para o desenvolvimento da Arquitetura da Informação para *Web*. Algumas, utilizadas por profissionais envolvidos no desenvolvimento de *websites*, outras, existem apenas como proposta teórica, ou seja, sem aplicação prática.

Neste trabalho a metodologia utilizada baseou-se na proposta de Morville e Rosenfeld (2006), considerando-se que as demais são bem semelhantes, e diferenciam-se apenas quanto ao grau de detalhamento das fases. Assim, ela foi utilizada tanto para descrever as fases comuns as principais metodologias existentes, quanto para os conceitos usados na ontologia.

A referida metodologia é composta por cinco fases: pesquisa, concepção, documentação, implementação e avaliação, detalhadas como:

- Pesquisa – Esta fase objetiva entender o problema e definir o escopo do trabalho a ser realizado para desenvolver o *site*. Nela, é feito um levantamento que contempla: a visão, missão e objetivos da empresa; o conteúdo do *site*; o público alvo e suas principais características; além da análise dos sistemas já existentes na empresa, bem como levantamento de requisitos que possam limitar a usabilidade.
- Concepção – Nesta fase são definidas as regras de organização e apresentação dos conteúdos e serviços disponibilizados no *site* e o modelo de interação de forma a atender: as boas práticas de usabilidade; as necessidades do público alvo; os requisitos do projeto; e os objetivos do negócio. Aqui são definidos os quatro sistemas da Arquitetura da Informação para *Web*, que são os sistemas de organização, navegação, rotulação e busca.
- Documentação – Nesta fase registram-se todos os documentos elaborados e utilizados pelos profissionais envolvidos no desenvolvimento do *site*, a exemplo dos *sitegramas*, fluxo de navegação, *wireframes*, *card sorting* e *blue print*.
- Implementação – Fase na qual os profissionais responsáveis pelo desenvolvimento (programadores, *webdesigners*, etc), por meio dos documentos elaborados na fase de documentação, fazem a implementação (codificação) do *site*, mediante o uso das tecnologias que melhor atendam a sua necessidade.
- Avaliação – Nesta fase verifica-se se foram contemplados todos os requisitos levantados na fase inicial (fase da pesquisa), e faz-se o acompanhamento do uso do *site* para verificar a satisfação do usuário, o que servirá de análise para possíveis mudanças a serem realizadas. Nesta fase deve acontecer a aprovação da versão criada, bem como a proposição de mudanças quando necessário.

Ontologias

A palavra ontologia vem do grego *ontos* (ser) + *logos* (palavra). Para a filosofia, é a ciência do ser, dos tipos de estrutura dos objetos, propriedades, eventos, processos e relacionamentos em todas as áreas da realidade (BREITMAN, 2005).

O termo ontologia pode ter vários significados dependendo da realidade em que esteja sendo aplicado, visto que é usado para descrever domínios e estes domínios podem abordar um determinado assunto muito específico da sua realidade. Nesse sentido, o termo ontologia foi utilizado inicialmente na filosofia, passando contudo a ser utilizado também em outras ciências.

Por ser usado em diversas ciências, existem muitas definições para ontologia que, em alguns casos podem se complementar. No entanto, a encontrada com mais frequência na literatura da *web* semântica, é a proposta por Gruber, que define ontologia como uma especificação formal e explícita de uma conceitualização compartilhada, em que conceitualização representa um modelo abstrato de algum fenômeno que identifica os conceitos relevantes para si mesmo; explícita, significa que os elementos e suas restrições são claramente definidos; formal, na medida em que deve ser passível de processamento automático (por máquinas); e compartilhada, por capturar o conhecimento consensual, aceito por um grupo de pessoas (GRUBER, 1993).

Classificação de ontologias

Cada classificação de ontologia é fundamentada de acordo com o tratamento dado à informação sobre a qual a ontologia está sendo aplicada. Na figura 2, é ilustrada uma classificação de acordo com o nível de generalidade da ontologia.

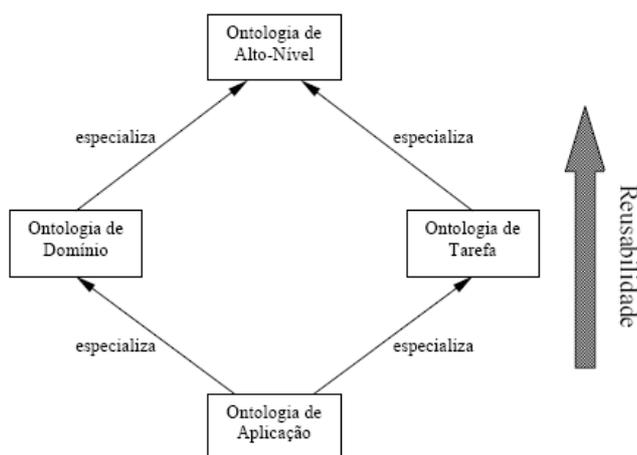


Figura 2: Classificação das ontologias quando à generalidade e seus relacionamentos

Fonte: Guimarães (2002, pág. 57)

As ontologias de nível superior (alto nível) descrevem conceitos mais gerais, tais como espaço, tempo e eventos. São a princípio independentes de domínio e podem ser reutilizadas em novas ontologias. As ontologias de domínio descrevem o vocabulário relativo a um domínio específico por meio da especialização de conceitos das ontologias de alto nível. As ontologias de tarefas descrevem o vocabulário relativo a uma tarefa genérica ou atividade por meio da especialização de conceitos das ontologias de alto nível. As ontologias de aplicação, por sua vez, são ontologias mais específicas, que correspondem, de maneira geral, a papéis desempenhados por entidades do domínio para realizar uma tarefa.

Construção de ontologias

Para desenvolver ontologias, existem vários ambientes, metodologias e linguagens que podem ser utilizados dependendo do tamanho e complexidade do domínio a ser modelado. Neste trabalho foi utilizada a linguagem *Ontology Web Language (OWL)*, o *framework Protégé 3.4.1*, seguindo os passos da metodologia 101¹.

A linguagem utilizada na modelagem foi a OWL, lançada pelo *World Wide Web Consortium (W3C)*, projetada para atender as necessidades das aplicações para *web* semântica. Segundo Freitas (2003) a OWL representa conceitos e seus relacionamentos na forma de uma ontologia e possui três linguagens em ordem de expressividade: *OWL Lite* que suporta a criação de hierarquias simplificadas de classificação e suas restrições, e tem por objetivo fornecer a migração de tesouros e taxonomias para ontologias. Não suporta estruturas sofisticadas, só restrições simples; *OWL-DL*, mais expressiva que a *OWL Lite*, é baseada na Lógica de Descrição (DL), que pode ser construída por união, interseção e complemento, pela enumeração de instâncias e podem ter disjunções. *OWL-DL* permite a verificação de satisfatibilidade de conceitos e classificação de hierarquias. Garante completude, decidibilidade e toda a expressividade da lógica de descrição; e *OWL Full*, que representa a linguagem OWL mais expressiva. Essa linguagem é utilizada quando a expressividade do conhecimento for mais importante que garantias de computabilidade.

O ambiente usado para modelagem foi o *framework Protégé 3.4.1*, uma ferramenta desenvolvida pelos pesquisadores do *Stanford Medical Informatics* da Universidade de Stanford no ano de 1997. Por ser um ambiente integrado para construção e edição de sistemas baseados em conhecimento, o *framework Protégé 3.4.1* possui a possibilidade de carregar, editar e salvar ontologias em vários formatos, tais como, *Resource Description Framework (RDF)*, *Extensible Markup Language (XML)*, OWL, entre outros. A decisão de utilizar esta ferramenta foi devida ao fato de ela ser consolidada, testada e validada pela comunidade, com ampla variedade de material disponível na Internet.

¹ A metodologia 101 é uma metodologia usada para construção de ontologias, que descreve uma seqüência de sete passos para o processo de construção, são eles: Passo 1 - determinar o domínio e o escopo da ontologia; Passo 2 - considerar o reuso de outras ontologias; Passo 3 - enumerar os termos importantes da Ontologia; Passo 4 - definir classes e Hierarquia das classes; Passo 5 - definir as propriedades das classes; Passo 6 - definir os valores das propriedades; e Passo 7 - criar Instâncias. (NOY; MCGUINNESS, 2001).

Segundo Azevedo (2008), o *framework Protégé* é acessível aos mecanismos de inferências para lógica descritiva, possuindo algumas vantagens, tais como: a geração automática de códigos *Java Beans* das ontologias, ferramenta em constante atualização, e existência de documentação que explica o uso da ferramenta.

A metodologia seguida na modelagem foi a 101, proposta por Noy e McGuinness (2001). A metodologia 101 envolve uma modelagem prática, seguindo uma seqüência de passos não linear, podendo ser realizadas inúmeras alterações até conseguir um modelo adequado, para ser validado. Os sete passos da metodologia 101 são os seguintes: determinar o domínio e o escopo da ontologia (passo 1); considerar o reuso de outras ontologias (passo 2); enumerar os termos importantes da ontologia (passo 3); definir classes e a hierarquia de classes (passo 4); definir as propriedades das classes (passo 5); definir os valores das propriedades (passo 6) e criar as instâncias (passo 7).

Processo de modelagem da InfoArch

Na seqüência serão descritos os passos necessários para a modelagem da ontologia ao utilizar o Método 101, considerando os principais conceitos do domínio da Arquitetura da Informação para *Web*.

PASSO 1: Determinar o domínio e o escopo da ontologia. Nesse primeiro passo, foram definidas algumas atividades a serem seguidas para determinar o domínio e o escopo da ontologia. São elas:

- Definição do domínio – Foi definido o domínio da Arquitetura da Informação para *Web*, partindo-se de dois conceitos fundamentais da AI para *web*, que são os sistemas e a metodologia. Os quatro sistemas (organização, navegação, rotulação e busca) além das estruturas de representação e a metodologia que organiza em uma seqüência de fases o desenvolvimento da AI para *web*;
- Propósito da ontologia – O principal objetivo foi modelar os principais conceitos do domínio da Arquitetura da Informação para *Web*, tais como seus sistemas e etapas das metodologias utilizadas, bem como as principais categorias de *sites*, os profissionais envolvidos no processo de desenvolvimento de *sites* e seus respectivos papéis. Esta modelagem objetiva auxiliar no ensino da disciplina Arquitetura da Informação para *Web* e no processo de desenvolvimento de *sites*;
- Questões de competência da ontologia – São questões para as quais a ontologia deve fornecer as respostas para auxiliar no processo de desenvolvimento de *sites* e/ou no auxílio ao ensino da AI para *web*. A modelagem feita é capaz de responder a questões, tais como: Quais os componentes da AI que devem compor os *sites*? Qual a função dos sistemas de organização, navegação, rotulação e busca nos *sites*? Em qual etapa da metodologia são definidos os componentes da AI para *web*? Quais as pessoas envolvidas no desenvolvimento da Arquitetura da Informação para *Web*? Entre outras. Tais perguntas podem ser realizadas por meio de consultas feitas diretamente na ontologia e/ou aplicações desenvolvidas para acessá-la.

- Usuários da ontologia – Todos os profissionais da informação que direta ou indiretamente tenham interesse em conhecer e discutir o assunto, tais como professores, tanto em programas de Pós-Graduação que já oferecem a disciplina, quanto em cursos ministrados por profissionais e/ou empresas que desenvolvem *sites*.

PASSO 2: Considerar o reuso de outras ontologias. A modelagem aqui proposta poderá ser reutilizada para a criação de novas ontologias relacionadas à Arquitetura da Informação para *Web*, a exemplo das ontologias de contexto. Poderá ser utilizada ainda, para modelar as características específicas de *sites* de acordo com a categoria destes. Isto é possível porque a modelagem foi realizada utilizando um *framework* conhecido e usado internacionalmente, utilizando a linguagem OWL e seguindo uma metodologia simples e bastante difundida. Além disso, todo o código OWL será disponibilizado mediante o contato com os autores.

PASSO 3: Enumerar os termos importantes da ontologia. Os termos importantes da ontologia foram obtidos por meio de artigos, livros e conversas com profissionais e pesquisadores do domínio. Nesse contexto os principais termos definidos foram: Arquitetura da Informação para *Web*; arquiteto da informação; sistemas da AI para *web*; metodologia; papel e *site*; a partir dos quais são relacionados e derivados vários outros, ligados por meio de hierarquias e relacionamentos, que serão abordados nos passos seguintes.

PASSO 4: Definir classes e hierarquia das classes. Todas as classes definidas no *Protégé* herdam de uma classe nativa a *thing* direta ou indiretamente. Na figura 3 é mostrada a hierarquia das classes, algumas delas, como, categorias, sistemas, conceitos, etapas e papéis, possuem subclasses, que herdam diretamente delas e indiretamente da classe *thing*.



Figura 3: Hierarquia de classes da ontologia

Fonte: Elaborado pelos autores

PASSO 5: Definir as propriedades das classes. A ontologia possui 34 propriedades que representam os relacionamentos entre as classes. As propriedades podem ser de dois tipos: *Object Properties* ou *Datatype Properties*. Na figura 4 é ilustrada a principal classe da InfoArch (“Componentes”), suas subclasses e relacionamentos.

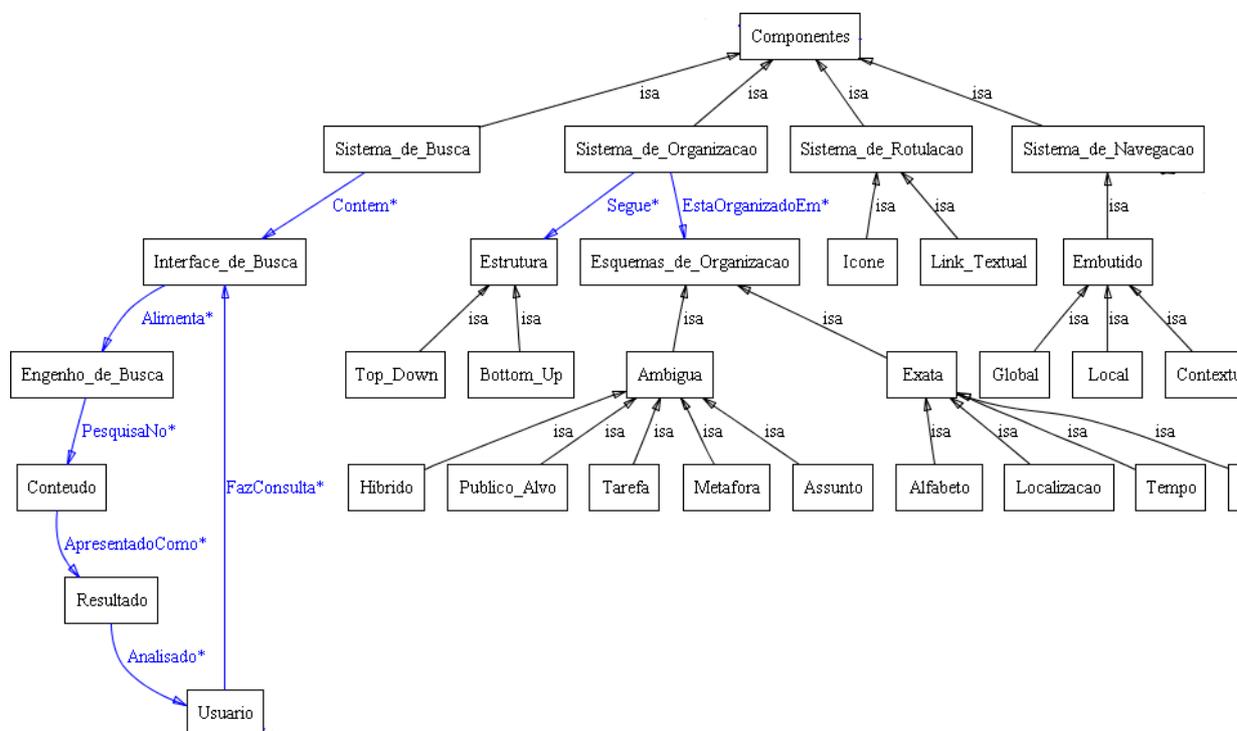


Figura 4: Componentes - a principal classe da ontologia com suas subclasses e relacionamentos

Fonte: Elaborado pelos autores

A partir da classe (“Componentes”), são especializadas e relacionadas à maioria das subclasses da InfoArch. Na seqüência será feita uma descrição de algumas dessas subclasses e seus relacionamentos a fim de facilitar o seu entendimento, possuindo as demais classes uma descrição semelhante.

O componente da modelagem referente ao sistema de busca (representado pela classe “Sistema_de_Busca”), descreve o processo de uma busca realizada internamente em um determinado *site*. Esta classe possui um relacionamento (denominado “Contem”) com a classe “Interface_de_Busca”, para indicar que no *site* em questão o usuário (classe “Usuario”) poderá realizar sua consulta (explicitado pelo relacionamento “FazConsulta”, que relaciona a classe “Usuario” com a classe “Interface_de_Busca”).

Feito o passo anterior, a informação fornecida pelo usuário será repassada para a classe “Engenho_de_Busca” por meio do relacionamento “Alimenta”. De forma similar, na seqüência é feita a busca sobre a informação no conteúdo do *site*, sendo o resultado encontrado pelo engenho

de busca apresentado ao usuário para ser analisado, fechando o ciclo da consulta realizada, como mostrado na figura 5.

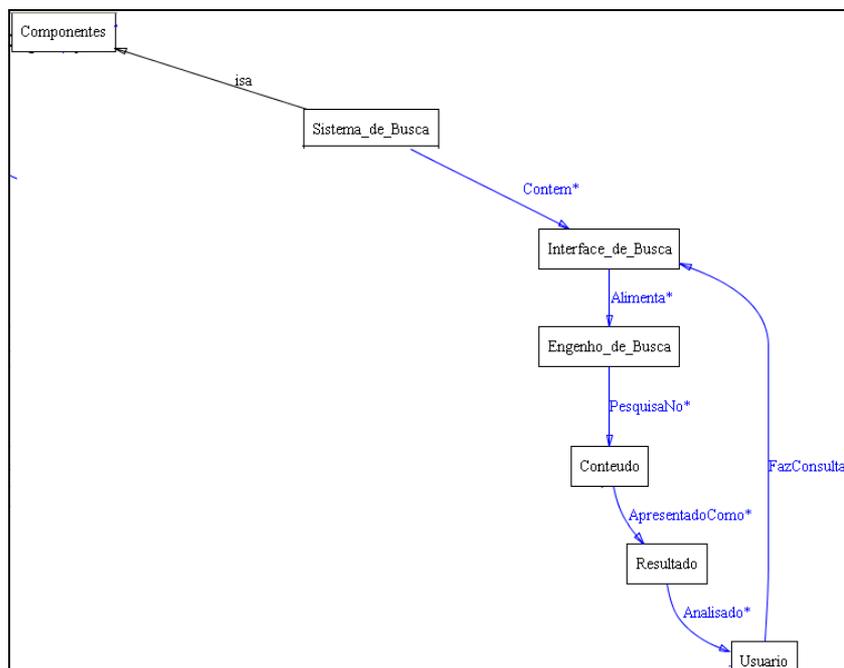


Figura 5: Classe “Sistema_de_Busca” com suas subclasses e relacionamentos

Fonte: Elaborado pelos autores

Já o componente da modelagem referente ao sistema de rotulação mostra como devem ser representados os rótulos nos *sites*. Na ontologia essa representação é feita por meio do relacionamento (“Define”), relacionando a classe (“Sistema_de_Rotulação”) a classe (“Rotulo”) que por sua vez possui as suas subclasses (“Icône”), correspondente às figuras que representam informação e (“Link_Textual”) onde os termos que possuem links internos ou externos ao *site* representam uma informação para o usuário. Esta representação é mostrada na figura 6.

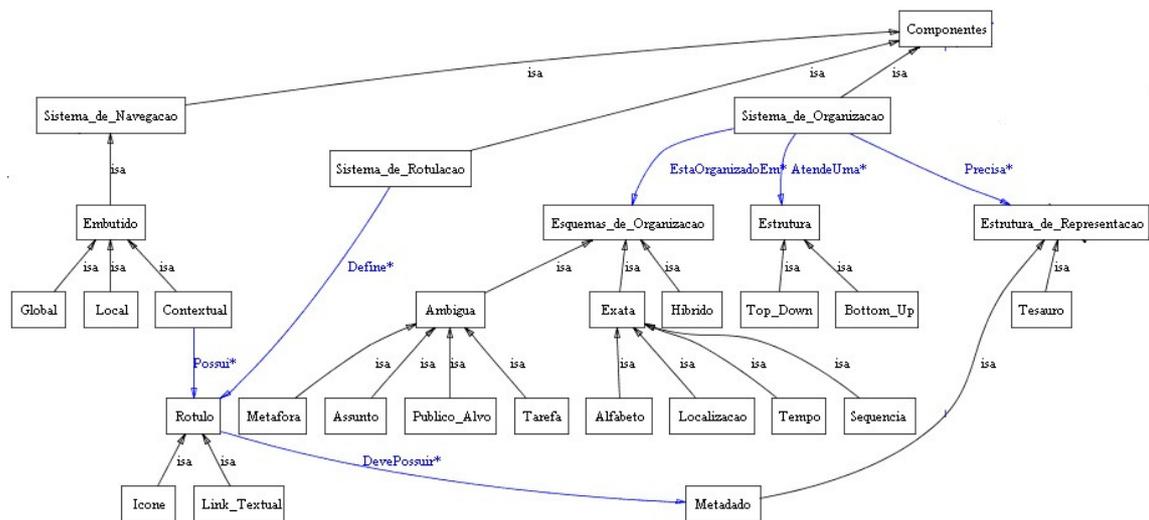


Figura 6: A classe “Sistema_de_Rotulacao” com seu relacionamento com a classe

Fonte: elaborado pelos autores

PASSO 6: Definir os valores das propriedades. As propriedades podem assumir diferentes valores de acordo com a expressividade da linguagem. Cada propriedade deve ter uma descrição diferente como o tipo do valor (funcional, inverso, simétrico ou transitivo), a cardinalidade, e outros recursos que o campo pode ter.

Na InfoArch foram definidas as cardinalidades das propriedades *Datatype Properties* usadas para construir os relacionamentos utilizados nas consultas que respondem às questões de competência. As cardinalidades das propriedades são representadas entre parênteses ao lado das instâncias que são ligadas por meio dessas propriedades.

A partir da utilização da ontologia é possível obter respostas para vários questionamentos que envolvam aspectos relacionados à Arquitetura da Informação para Web, podendo-se utilizar o próprio *Protégé*, ou por meio de aplicações que venham a ser desenvolvidas para acessá-la.

A consulta abaixo, por exemplo, traz a relação dos cinco componentes definidos na fase de concepção da metodologia recomendada para definir os componentes da Arquitetura da Informação para Web.

```
SELECT *
WHERE { ?OsComponentes:SaoDefinidos ?Concepcao }
```

Já esta outra consulta traz informações específicas sobre o componente sistema de busca especificando que o mesmo é definido na Concepção; que possui interface de busca; e que esta interface alimenta os engenhos de busca, realizando consultas no conteúdo do *site*.

```
SELECT *
WHERE { ?Componentes:SaoDefinidos ?Concepção }
```

{?Sistemas_de_Busca:Contem ?Interface de Busca }

{?Interfaces_de_Busca:Alimenta ?Engenhos de Busca }

{?Engenhos_de_Busca:PesquisaNo ?Conteúdo }

PASSO 7: Criar Instâncias. Nesta última fase são criadas instâncias individuais para as classes, para que por meio dos relacionamentos entre essas classes possam ser definidas as consultas que retornam resultados como respostas a questões de competência da ontologia. Nesta etapa foram definidas as instâncias que respondem às principais questões de competência, como: Componentes; Estruturas_de_Representacao; Sistemas_de_Busca; Sistemas_de_Organizacao; Sistemas_de_Rotulacao; Sistemas_de_Navegacao; Conceitos; Icontexto; Iconteudo; Engenhos_de_Busca; Esquema_de_Organizacao; Estruturas; Etapa; Iconcepcao; Interfaces_de_Busca; Resultados; Rótulos, dentre outras.

Vale salientar que para criar as instâncias foi usado o mesmo padrão utilizado para criar as classes, ou seja, letra inicial maiúscula e as demais minúsculas, além de separar nomes por *underline* “_” quando estes forem compostos. Como o *Protégé* não permite que os nomes das classes e instâncias sejam exatamente iguais aos das classes, convencionou-se diferenciar pelo singular e plural, como por exemplo, para a classe “Componentes”, criou-se a instância “Componente”. Quando não for conveniente usar esta padronização, optou-se por acrescentar um “I” (inicial de instância), no início do nome, por exemplo, classe “Concepcao”, instância “Iconcepcao”.

Ressalta-se que tal convenção não causa prejuízo para a ontologia, visto que as consultas definidas para trazer respostas às questões de competência já são montadas observando-se estes detalhes e as consultas são feitas por meio dos relacionamentos, que podem ser os mesmos para relacionar tanto as classes quanto as instâncias.

Resultados

A InfoArch possui 17 classes principais a partir das quais são herdadas outras 52 classes, totalizando 69 classes e 34 propriedades dos tipos *object properties* e/ou *datatype properties*. Algumas dessas classes foram instanciadas para, juntamente com suas propriedades, proporcionar a realização de consultas sobre o domínio, que responderão a questionamentos nos seus cenários de utilização.

Cenários de utilização

A InfoArch foi proposta para modelar o domínio da Arquitetura da Informação para *Web*, a fim de permitir o compartilhamento do conhecimento entre pesquisadores e/ou profissionais que trabalham e pesquisam o tema.

A InfoArch pode ser usada como facilitadora para o reuso, padronização e compartilhamento do conhecimento sobre Arquitetura da Informação para *Web*. Dessa forma, poderá a princípio ser utilizada em três cenários, sejam eles: ensino da disciplina de Arquitetura da Informação para *Web*; reuso para criação de outras ontologias e/ou aplicações; e desenvolvimento de *sites*.

Para exemplificar a utilização da InfoArch, seguem alguns cenários que ilustram a forma como a mesma poderá proporcionar aos seus usuários um auxílio de extrema relevância.

Cenário 1: ensino da disciplina de Arquitetura da Informação para Web

A disciplina Arquitetura da Informação para *Web* já é lecionada em alguns cursos de Pós-Graduação. Um exemplo disso pode ser visto no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba. A disciplina é lecionada também em alguns cursos de curta duração, ministrados tanto por profissionais de empresas que trabalham com desenvolvimento de *sites*, quanto por professores e pesquisadores de universidades que trabalham o tema em suas pesquisas.

A InfoArch pode ser usada para auxiliar nas aulas referentes tanto aos cursos quanto à própria disciplina. Com a ontologia é possível mostrar os conceitos e os relacionamentos entre eles no contexto da Arquitetura da Informação para *Web*, além de responder a questões sobre este domínio por meio da representação gráfica, que pode ser dada pela interação direta com a ontologia ou por meio de aplicações que venham a ser desenvolvidas para acessá-la.

A partir do semestre 2010.2 a InfoArch será utilizada para auxiliar no ensino da disciplina de Arquitetura da Informação para *Web* do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, e nos estudos e práticas do grupo de pesquisa Arquitetura da Informação e Usabilidade (ARIUS) formado por docentes e discentes de Graduação e Pós-Graduação.

Cenário 2: criação de novas ontologias e/ou aplicações

Este cenário pode acontecer quando houver a necessidade de criação de novas ontologias para domínios semelhantes e/ou aplicações que podem fazer reuso desta ontologia. Neste caso pode-se citar como exemplo a criação de ontologias para descrever a Arquitetura da Informação para *Web* para categorias específicas de *sites*.

Ultimamente estão sendo realizados alguns estudos relacionados à aplicação de princípios de Arquitetura da Informação para *Web* na interface de dispositivos móveis, a exemplo de *e-books*. Nessa perspectiva a ontologia aqui proposta contribuirá para a criação de outras ontologias específicas para este domínio.

As novas ontologias a serem criadas podem ser estendidas a partir da classe “Componentes”, para fazer a descrição dos componentes gerais da Arquitetura da Informação para *Web*, e da

classe “Categorias” (para fazer a representação das características específicas, inerentes a cada categoria do *site*).

Cenário 3: desenvolvimento de sites

Ao desenvolver um *site*, a equipe que deseja adotar as recomendações da Arquitetura da Informação para *Web* poderá consultar a InfoArch para ter acesso às informações nela contidas, bem como à sua metodologia, que descreve as etapas e os artefatos de cada uma delas.

Para proporcionar tal facilidade, a InfoArch traz a classe “Componentes” com suas subclasses e relacionamentos, que descrevem o uso dos componentes da AI para *web*, e mais um conjunto de classes contendo uma metodologia com as suas etapas (pesquisa, concepção, documentação, implementação e manutenção), e o que deve ser realizado em cada uma delas.

A partir da InfoArch e com a criação de novas ontologias do cenário 2, é possível criar um ambiente para o desenvolvimento de *sites* orientados por ontologias e que utilize os conceitos da Arquitetura da Informação para *Web*.

Considerações finais

As ontologias são formas de armazenar dados sobre um determinado domínio com o objetivo de realizar a padronização dos mesmos, proporcionando um melhor entendimento sobre os conhecimentos armazenados e facilitando o acesso às informações. A utilização de ontologias em várias áreas do conhecimento tem se tornado uma atividade muito útil, pois são capazes de armazenar, compartilhar e reutilizar conhecimentos e estruturas de informação referentes a um determinado domínio, permitindo a origem de análises e consultas relacionadas aos conhecimentos.

Já a Arquitetura da Informação para *Web* define os componentes de busca, rotulação, navegação, organização e estruturas de representação como sugestão para estruturar os *websites*, melhorando a usabilidade e conseqüentemente facilitando a recuperação da informação na *web*. A Arquitetura da Informação para *Web* também recomenda metodologias para que tais componentes (sistemas) sejam aplicados no desenvolvimento de *websites*, com base em estudos relativos à usabilidade, estudo do usuário e interface humano-computador.

Este trabalho realizou a modelagem dos principais conceitos e relacionamentos do domínio da Arquitetura da Informação para *Web*, servindo como base para o compartilhamento e reuso do conhecimento deste domínio por profissionais e pesquisadores que discutem e trabalham a temática aqui tratada. Dessa forma, como resultado, pode-se considerar as seguintes contribuições:

- Servir como base para o conhecimento do domínio da Arquitetura da Informação para *Web*, possibilitando o reuso e compartilhamento dos conceitos e seus relacionamentos. A Arquitetura da Informação para *Web* é um tema ainda novo e

pouco conhecido, com pouca literatura disponível, sobretudo em língua portuguesa. Assim, a InfoArch também servirá de fonte de consultas podendo inclusive ilustrar de forma gráfica seus conceitos e relacionamentos, o que vem a facilitar o seu entendimento, principalmente em estudos feitos por iniciantes no assunto;

- Ajudar no desenvolvimento de *sites* que contemplem os componentes da Arquitetura da Informação para *Web*, possibilitando aos usuários destes terem mais facilidade para encontrar as informações que desejam. Neste caso, a ontologia também representa mais uma fonte de consulta, podendo ser utilizada pela equipe de desenvolvimento para sanar dúvidas sobre como utilizar os componentes da Arquitetura da Informação para *Web*, do que até então só seria capaz por meio da literatura disponível e dos arquitetos da informação;
- Auxiliar no ensino da disciplina de Arquitetura da Informação para *Web*, que já faz parte de Programas de Pós-Graduação relacionados a áreas como Ciência da Informação e Design Gráfico. Como também, em cursos ministrados por professores pesquisadores do assunto e/ou por empresas de desenvolvimento de *sites*, tanto para treinar seus desenvolvedores quanto para agregar os componentes (sistemas) da Arquitetura da Informação para *Web* em seu processo de desenvolvimento.

Como é possível observar, a InfoArch pode contribuir diretamente com os vários profissionais que já trabalham ou venham a trabalhar com Arquitetura da Informação para *Web*, tanto na área acadêmica em pesquisa, ensino ou extensão, quanto na área corporativa com as pessoas e empresas responsáveis pelo desenvolvimento de *sites*.

Artigo recebido em 14/02/2011 e aprovado em 28/02/2011.

Referências

AZEVEDO, R. Ribeiro de. CoreSec: uma ontologia para o domínio de segurança da informação. 2008. 137f. Dissertação (Programa de Pós-Graduação em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2008.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web. Scientific Am, May 2001, pp. 34–43.

BREITMAN, Karin Koogan. Web Semântica: a internet do futuro. Rio de Janeiro: LTC: 2005.

CAMARGO, Liriane Soares Araújo de. Metodologia de desenvolvimento de ambientes científicos digitais baseada em princípios da Arquitetura da Informação. Marília, 2010, 322f. Tese (Programa de Pós-Graduação em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2010. Disponível em:

<http://www.marilia.unesp.br/Home/PosGraduacao/CienciadaInformacao/Dissertacoes/camargo_lsa_do_mar.pdf>. Acesso em: 03 ago. 2010.

CARLAN, Eliana. Ontologia e Web Semântica. Disponível em: <<http://eprints.rclis.org/archive/00007140/01/ECarlan.pdf>>. Acesso em: 12 ago. 2010.

FREITAS, F. Ontologias e a Web Semântica. Disponível em: <http://www.inf.ufsc.br/~gauthier/EGC6006/material/Aula%203/Ontologia_Web_semantica%20Freitas.pdf>. Acesso em: 17 ago. 2010.

GIL, Antonio Carlos. Como elaborar projetos de pesquisa. 3. ed. São Paulo: Atlas, 2000.

GRUBER, T.R. A translation approach to portable ontology specifications. Knowledge Acquisition, 1993. n. 5, p. 199-220.

MORROGH, E. Information Architecture: An Emerging 21st Century Profession. New Jersey: Prentice Hall, 2003.

MORVILLE, P.; ROSENFELD, L. Information Architecture for the World Wide Web. 3. ed. Sebastopol: O'Reilly, 2006. 526p.

NOY, N. F.; MCGUINNESS, D. L. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

SOUZA, Renato Rocha; ALVARENGA, Lídia. A Web Semântica e suas contribuições para a ciência da informação. Ciência da Informação, Brasília, v. 33, n. 1, p. 132-141, jan./abril 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n1/v33n1a16.pdf> >. Acesso em: 25 nov. 2006.

THE INFORMATION ARCHITECTURE INSTITUTE. What is Information Architecture?, 2002. Disponível em: <http://iainstitute.org/documents/learn/What_is_IA.pdf >. Acesso em: 02 ago. 2010.

VIDOTTI, Silvana A. B. G. ; SANCHES, Silviane. A. S. Arquitetura da Informação em web sites. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 2004. Anais eletrônicos. Campinas: Unicamp, 2004. Disponível em: <www.libdigi.Unicamp.br?document/?down=8302>. Acesso em: 14 ago. 2010.

WURMAN, Richard Saul. Ansiedade de Informação. São Paulo SP: Cultura Editores Associados, 1991.