

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-8 – Informação e Tecnologia

COMPONENTES DE REPRESENTAÇÃO DA INFORMAÇÃO EM AMBIENTES DE INFORMAÇÃO DIGITAL: ESTUDO DO SISTEMA DE ORGANIZAÇÃO DO SOFTWARE TAINACAN

INFORMATION REPRESENTATION COMPONENTS IN DIGITAL INFORMATION ENVIRONMENTS: A STUDY OF THE TAINACAN SOFTWARE ORGANIZATION SYSTEM

Luciana Candida da Silva – Universidade Federal de Goiás

José Eduardo Santarem Segundo - Universidade Estadual Paulista Júlio de Mesquita Filho

Modalidade: Trabalho Completo

Resumo: Considerando que os componentes de representação da informação são um aporte ao sistema de organização de ambientes de informação digital, parte-se de uma revisão teórica sobre os temas arquitetura da informação, componentes de representação e a interação desses componentes com o sistema de organização para, a seguir, analisar o sistema de organização do software livre de gestão de acervos digitais Tainacan, segundo sua estrutura e esquema. Os resultados apresentaram que o sistema de organização da arquitetura da informação das coleções analisadas do software Tainacan caracteriza-se como esquema híbrido e estrutura principal hierárquica com o uso complementar das estruturas de modelo de banco de dados relacional e hipertexto. Identificou que os principais elementos de representação adotados pelo Tainacan foram os metadados, sistema de classificação e taxonomia; os elementos não identificados foram as listas, os anéis de sinônimo e o tesouro. Sendo assim, sugere-se a adoção dos referidos elementos para que haja uma arquitetura da informação que apresente a informação facilitada e compreensiva aos usuários.

Palavras-Chave: Representação da informação; Sistemas de organização da informação; Arquitetura da informação; Software Tainacan.

Abstract: Considering that the components of information representation are a contribution to the system of organization of digital information environments based on a theoretical revision on the themes of information architecture, representational components and the interaction of these components with the organizational system, and then analyze the organizational free system of the digital collections management software Tainacan, according to its structure and scheme. The results show that the informational architecture system of Tainacan software is characterized as a hybrid scheme and a hierarchical main structure with the complementary use of relational and hypertext database model structures. However, it was identified that the main elements of representation adopted by Tainacan were the metadata, classification system and taxonomy; The unidentified elements were the lists, synonym rings and thesaurus. Therefore, it is suggested to adopt the aforementioned elements so that there is a simple informational architecture that presents comprehensive information to the users.

Keywords: Information representation; Information organization system; Information Architecture; Tainacan software.

1 INTRODUÇÃO

O avanço das tecnologias da informação acelerou o processo de produção da informação e do conhecimento proporcionando um imenso volume de informações disponíveis no meio eletrônico. Diante desse avanço observa-se um aumento no número de ambientes de informação digital, tais como bibliotecas digitais, sites, repositórios institucionais, periódicos eletrônicos, entre outros. Para Oliveira, Vidotti e Bentes (2015, p.32), “os ambientes de informação digital são lugares que armazenam e facilitam o acesso às informações de natureza digital, sobretudo no contexto da internet e da *Web*”. Porém, conforme alerta Oliveira (2014), a inadequada organização da informação em ambientes digitais e analógicos pode comprometer a experiência de interação nesses produtos tecnológicos.

Diante do aumento desordenado da informação e da inadequada organização da informação no meio digital, surgem os elementos da arquitetura da informação digital como alternativa para tornar a informação mais compreensível aos usuários.

Nesse contexto, em uma abordagem sistêmica, Rosenfeld, Morville e Arango (2015) expõem que a arquitetura da informação combina os sistemas que são compartilhados pela maioria dos ambientes interativos de informação: organização, rotulagem, navegação e busca. Além desses sistemas, os autores discutem os componentes de representação: metadados, vocabulários controlados e tesauros como sistemas invisíveis que moldam o ambiente de informação nos bastidores.

Ainda segundo Rosenfeld, Morville e Arango (2015), para que um ambiente de informação digital possa oferecer serviços de informação e acesso a uma vasta gama de conteúdos, bem como promover a sua disseminação, os componentes dos sistemas de arquitetura da informação se interligam com dependências complexas. Dessa forma, o sistema de organização é considerado a base dos demais sistemas. Os metadados e os vocabulários controlados se apresentam como um elo entre os sistemas.

Diante da importância dos componentes de representação na arquitetura da informação em ambientes de informação digital, este estudo se propõe a destacar a relevância dos componentes de representação da informação em um ambiente de informação digital e identificar quais desses componentes estão sendo aplicados ao sistema de organização da informação do software livre de gestão de acervos digitais Tainacan para favorecer os usuários no acesso a informação relevante de forma compreensível.

Para tanto, buscou-se apoio teórico na área da Ciência da Informação acerca dos temas arquitetura da informação e componentes de representação da informação, bem como a sua interação com o sistema de organização. Na sequência, identificou os elementos de representação que estão sendo aplicados no sistema de organização do software livre de gestão de acervos digitais Tainacan como aporte aos usuários.

2 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa é de natureza qualitativa do tipo descritiva-bibliográfica e busca identificar o uso dos componentes de representação da informação no sistema de organização da informação do software livre Tainacan para facilitar a compreensão da informação disponibilizada.

O referencial teórico sobre arquitetura da informação, componentes de representação da informação e a interação desses componentes com o sistema de organização foi construído a partir de artigos científicos recuperados nas bases de dados Scopus e Brapci, bem como em livros e sites. E para análise das coleções adotou o site institucional do software Tainacan, o qual permitiu o direcionamento para a página das coleções. O recorte temporal refere-se ao período de 1997 a 2015 no que se refere à Arquitetura da Informação, de 1983 a 2015 aos componentes de representação da informação, de 1996 a 2015 para a integração dos componentes de representação da informação com o sistema de organização.

Em seguida, desenvolveu-se a análise do software Tainacan, a partir da coleção da Biblioteca Modelo, da Universidade Federal de Goiás (UFG) e da coleção Retratos do Império, do Museu Histórico Nacional (MHN). A escolha das coleções partiu da observação do maior número de elementos de análise entre as demais coleções que adotam o Tainacan para o gerenciamento e publicação de seus acervos.

A análise teve como base os elementos **esquema** e **estrutura** do modelo de sistema de organização da informação proposto por Rosenfeld, Morville e Arango (2015). Após identificar os elementos presentes nas coleções selecionadas, passou-se a observar como ocorrem as suas interações com o sistema de organização como benefício ao usuário.

O software Tainacan é resultado do projeto de gestão social de acervos digitais de cultura desenvolvido pelo Laboratório de Políticas Públicas Participativas (L3P), do Núcleo de Pesquisas em Gestão, Políticas e Tecnologias de Informação (NGPTI/UFG), da Universidade

Federal de Goiás (UFG). O Tainacan é considerado um software livre brasileiro que tem por objetivo servir como suporte as novas ações que promovam a digitalização de objetos culturais e a novas formas de organização e disponibilização de itens já digitalizados (TAINACAN, 2015).

A coleção da Biblioteca Modelo é um projeto coordenado pelo curso de Biblioteconomia da UFG que tem como uma de suas premissas desenvolver uma formação curricular específica de tratamento da informação na área de bibliotecas escolares.

A coleção Retratos do Império faz parte do acervo museológico do Museu Histórico Nacional (MHN), a qual apresenta pinturas de Dom João VI a Dom Pedro II. O MHN possui uma representativa coleção de retratos da família imperial brasileira, realizados por diversos artistas entre o século XVIII e século XX (MUSEU HISTÓRICO NACIONAL, 2019).

3 ARQUITETURA DA INFORMAÇÃO

O termo arquitetura da informação foi introduzido pelo arquiteto Richard Saul Wurman, em 1976, com o trabalho sobre *The architecture of information* (A arquitetura da informação) apresentado na conferência da American Institute of Architects (AIA, Instituto Americano de Arquitetos em português), primeira conferência a tratar da temática. O teórico Wurman, em 1997, apresenta uma das principais definições ligadas às funções realizadas pelos profissionais da arquitetura da informação,

1) o indivíduo que organiza os padrões inerentes aos dados, tornando o complexo claro; 2) a pessoa que cria a estrutura ou mapa da informação, que permite aos outros encontrar seus próprios caminhos na direção do conhecimento; 3) a atividade profissional que surge no século 21 apontando para as necessidades da época, como foco na clareza, na compreensão humana e na ciência da organização da informação (WURMAN, 1997, capa apud ROBREDO, 2008, p.119).

No entendimento de Rosenfeld, Morville e Arango (2015, p. 24) a arquitetura da informação pode ser compreendida nas seguintes perspectivas:

1) Projeto estrutural de ambientes de informação compartilhados; 2) A combinação dos sistemas de organização, rotulagem, busca e navegação dentro dos sistemas digitais; 3) A arte e a ciência de estruturar produtos e experiências de informação para apoiar a usabilidade e a encontrabilidade; 4) uma disciplina emergente e comunidade de prática focada em trazer princípios de design e arquitetura para o cenário digital.

Rosenfeld, Morville e Arango (2015, p.82-83) afirmam que a arquitetura da informação envolve quatro elementos:

- (1) sistemas de organização – maneira como o conteúdo de um site pode ser agrupado;
- (2) sistema de rotulagem – forma como é representada cada unidade de informação do site;
- (3) sistema de navegação – ferramentas auxiliares que permitem ao usuário folhear ou navegar através dessas unidades de informação; e
- (4) sistema de busca – permite ao usuário realizar consultas no todo informacional dentro do site.

Diante da aplicação das novas tecnologias da informação e comunicação em ambientes digitais, Vidotti, Cusin e Corradi (2008), observaram que os autores Peter Morville e Louis Rosenfeld, em 2006, adicionaram os elementos metadados, tesauro, e vocabulário controlado para representação descritiva e temática dos conteúdos presentes na rede informacional, conforme descrição abaixo:

- Metadados: descreve documentos, sites, imagens, software e arquivos de áudio e vídeo e outros conteúdos com o intuito de prover recursos e recuperação da informação;
- Tesauro: melhora o processo de recuperação e navegação da informação; rede semântica de conceitos interligando as palavras e seus sinônimos, homônimos, termos gerais e específicos e termos relacionados; e,
- Vocabulário controlado: qualquer subconjunto de linguagem natural, na sua forma mais simples, é uma lista de termos equivalentes sob a forma de um sinônimo ou de uma lista de termos preferidos. (VIDOTTI; CUSIN; CORRADI, 2008, p.174)

Robredo (2008) analisa a evolução da linha de pensamento de Morville e Rosenfeld e relata que a partir da segunda edição da obra *Information Architecture for the Web* (Arquitetura da Informação para a Web), esses autores vêm apresentando uma visão mais prática da arquitetura da informação dando maior ênfase à representação de conteúdos (metadados, vocabulários controlados, tesauros, relações semânticas, hierarquias, etc.).

Nesse sentido, Vidotti, Cusin e Corradi (2008) apresentam uma definição contemplando os sistemas da arquitetura da informação associados aos elementos do ciclo informacional, trabalhados no âmbito da Ciência da Informação:

[...] a Arquitetura da Informação enfoca [a] organização de conteúdos informacionais e as formas de armazenamento e preservação (sistemas de organização), representação, descrição e classificação (sistema de rotulagem, metadados, tesauro, vocabulário controlado), recuperação (sistema de busca), objetivando a criação de um sistema de interação (sistema de navegação) no qual o usuário deve interagir facilmente (usabilidade) com autonomia no acesso e uso do conteúdo (acessibilidade) no ambiente hipermídia informacional digital. (VIDOTTI; CUSIN; CORRADI, 2009, p. 182).

Após contextualizar a arquitetura da Informação, a partir dos principais autores da área, ressalta-se que os ambientes de informação digital devem ser dinâmicos e possuem necessidades e oportunidades específicas de acordo com o contexto, usuário e conteúdo. O contexto está relacionado com a missão, objetivo, estratégia, equipe, processos e procedimentos, aspectos físicos e tecnológicos de infraestrutura, orçamento e cultura de cada organização. O usuário está ligado às necessidades de informação e comportamento de busca. O conteúdo está relacionado à comunicação técnica. Muitos sistemas digitais são textuais, muitas vezes, a comunicação é construída com palavras e frases que tentam transmitir um significado (ROSENFELD; MORVILLE; ARANGO, 2015).

Além das três áreas para construção do modelo de um projeto eficaz de arquitetura da informação, aliam-se os componentes de representação, os quais influenciam em como os usuários usam os produtos e serviços.

4 COMPONENTES DE REPRESENTAÇÃO DA INFORMAÇÃO

Os componentes de representação da informação – metadados, vocabulários controlados e os tesouros - têm como objetivo possibilitar a interação entre os sistemas da arquitetura da informação proporcionando ao usuário uma boa compreensão da informação disponibilizada no ambiente digital.

Rosenfeld, Morville e Arango (2015) denominam os metadados, os vocabulários controlados e os tesouros como componentes invisíveis por se manifestarem nos bastidores de um site e os usuários raramente interagirem com eles. Estes componentes, muitas vezes alimentam outros componentes para melhorar a busca. Oliveira, Vidotti e Bentes (2015), denominam os componentes invisíveis de sistema de representação. Neste estudo, serão chamados de componentes de representação da informação, por aproximar das definições de Rosenfeld, Morville e Arango (2015) e denominação de Oliveira, Vidotti e Bentes (2015).

No contexto da Ciência da Informação, “a representação da informação, pode ser entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico” (BRASCHER; CAFÉ, 2008, p.5).

Para Novellino (1996, p.38) “a principal característica do processo de representação da informação é a substituição de uma entidade linguística longa e complexa - o texto do documento - por sua descrição abreviada”. Novellino (1996, p. 38) aponta que a

representação da informação “funciona como um artifício para enfatizar o que é essencial no documento considerando a sua recuperação, sendo a solução ideal para organização e uso da informação”.

Segundo Silva, Pinho Neto e Dias (2013) os metadados, os vocabulários controlados e os tesauros são estruturas da representação da informação que auxiliam na construção dos demais sistemas da arquitetura da informação, mediante a escolha e categorização dos termos que melhor representam determinado conteúdo dentro de um contexto, bem como as hierarquias e relacionamentos existentes.

Neste estudo, os tesauros serão descritos como um tipo de vocabulário controlado. Dessa forma, apresentam-se nos itens seguintes os componentes de representação da informação: Metadados e os Vocabulários Controlados.

4.1 METADADOS

O termo metadados possui funções de acordo com a área em que é utilizado. No “contexto da Ciência da Informação tem como objetivo principal a descrição da informação para sua representação, busca e recuperação” (FUSCO, 2011, p.41).

Para Alves e Santos (2013, p. 41 apud ROSETTO, 2003, p. 59)

[...] os metadados são um conjunto de atributos, mais especificamente dados referenciais, metodologicamente estruturados e codificados, conforme padrões internacionais, para localizar, identificar e recuperar pontos informacionais de textos, documentos e imagens disponíveis em meios digitais ou em meios convencionais.

Para os diversos recursos informacionais, os metadados possuem padrões que orientam e auxiliam na forma de descrição de cada tipo de informação.

Para Alves (2010, p. 47), “os padrões de metadados são estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados”.

São exemplos de padrões de metadados: Dublin Core (utilizado na descrição de recursos digitais), MARC 21 (registros bibliográficos), TEI (material textual em ambientes digitais), dentre outros.

As “tags de metadados são usadas para descrever documentos, páginas, imagens, software, vídeo e arquivos de áudio e outros objetos de conteúdo para fins de navegação e

recuperação melhorada” (RONSENFELD; MORVILLE; ARANGO, 2015, p. 270, tradução nossa).

Os ambientes de informação digital vêm exigindo o desenvolvimento de uma infraestrutura de informação que assegure a recuperação, o acesso e o intercâmbio de informação. Nesse contexto, Alves e Santos (2013, p. 15) reforçam

[...] os metadados, nessa infraestrutura, se destacam como elementos intrínsecos e fundamentais para os sistemas e ambientes informacionais digitais, pois englobam aspectos tecnológicos e aspectos representacionais específicos de domínio, promovendo a representação, a individualização, o intercâmbio, a interoperabilidade entre sistemas, o acesso e a recuperação de recursos informacionais.

Para Fusco (2011), os metadados tem a finalidade de facilitar a localização e recuperação das informações para os usuários e, para isso, utiliza-se de procedimentos técnicos de indexação e classificação de conteúdos informacionais, possibilitando a integração de fontes diversificadas e heterogêneas de informação.

A ANSI/NISO Z39.19 (2005) ressalta que vários conjuntos de elementos de metadados foram criados para apoiar as comunidades particulares ou domínios de informação. Estes conjuntos de elementos especificam regras semânticas para cada elemento, que pode incluir o uso recomendado ou exigido de um vocabulário controlado para os valores atribuídos a um elemento. Nesse sentido, os metadados podem ser utilizados com o apoio de vocabulários controlados como fonte para termos permitidos para um determinado elemento de metadados.

4.2 VOCABULÁRIOS CONTROLADOS

Os vocabulários controlados são listas de termos elaboradas para identificar o assunto de um documento para fins de recuperação rápida e eficaz. De acordo com a ANSI/NISO Z39.19 (2005) um vocabulário controlado é utilizado para melhorar a eficácia do armazenamento, navegação e recuperação da informação na *Web*, bem como em outros ambientes que buscam identificar e localizar a informação desejada, por meio de algum tipo de descrição utilizando linguagem controlada. O objetivo do vocabulário controlado é obter consistência na descrição do conteúdo dos objetos e facilitar a sua recuperação.

Para Lancaster (2004, p.19), o vocabulário controlado inclui uma forma de estrutura semântica. Essa estrutura destina-se, especialmente, a:

- controlar sinônimos, optando por uma única forma padronizada, com remissivas de todas as outras;
- diferenciar homógrafos; e

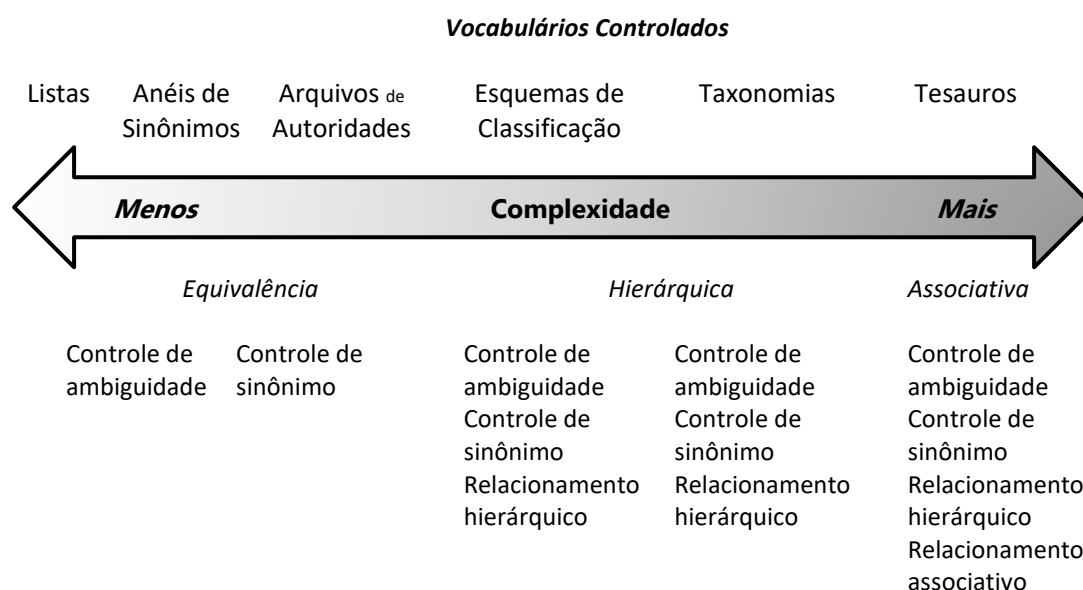
- reunir ou ligar termos cujos significados apresentem uma relação mais estreita entre si.

A linguagem empregada nos vocabulários para descrever os assuntos dos documentos é chamada linguagem de indexação, linguagem documentária ou linguagem descritora.

Neste contexto, Piedade (1983) considera dois métodos fundamentais de indexar o tema de documentos: 1 – indexação por palavra; e 2 – indexação por termo. A primeira emprega palavras encontradas no próprio documento, usa-se uma linguagem natural, não há controle de sinônimos ou homógrafos. A indexação por conceito pressupõe a análise do conteúdo temático do documento, a decisão sobre os conceitos presentes no texto e a tradução do observado em linguagem apropriada, com a qual rotulam-se os documentos e os seus registros bibliográficos. Ainda segundo Piedade (1983, p. 10), “na indexação por conceito determinam-se os cabeçalhos a empregar, distinguem-se homônimos, controlam-se sinônimos, preveem-se ligações hierárquicas e emprega-se uma linguagem controlada”.

Dito isso, destaca-se na figura 1 os principais tipos de vocabulários controlados utilizados em ambientes de informação digital, apresentados a partir de uma linha de complexidade das suas estruturas semânticas. Essa complexidade é proporcional à quantidade de tipos de relações semânticas existentes entre os termos.

Figura 1: Complexidade dos tipos de vocabulários controlados



Fonte: Adaptado de ANSI/NISO Z39.19 (2005) e Rosenfeld, Morville e Arango (2015).

As relações semânticas responsáveis por reunir os termos são: equivalência ou sinonímia, hierárquica e associativa. Para Robredo (2008), as relações de equivalência ou sinonímia são quando o mesmo conceito pode ser representado por dois ou mais termos, sendo que um dos mesmos deve ser escolhido como o preferido. As expressões ‘use’ e ‘usado por’ indicam o termo preferido na lista. A relação hierárquica põe em evidência as relações de subordinação do tipo genérico-específico dos termos. Para indicar as relações hierárquicas são usadas as expressões ‘termo mais amplo’, ‘genérico’ e ‘específico’. A relação associativa, leva um termo a outro que seja correlato e é, geralmente, expressa pelo indicador ‘Veja também’ ou ‘Termo Relacionado’.

5 INTERAÇÃO DOS COMPONENTES DE REPRESENTAÇÃO DA INFORMAÇÃO COM O SISTEMA DE ORGANIZAÇÃO

O modo como é realizada a organização da informação em um *website* influencia na maneira como as pessoas compreendem a informação disponibilizada. Os meios digitais proporcionam ambientes flexíveis para organizar a informação de forma que o usuário possa navegar e encontrar resposta satisfatória para suas buscas.

Os meios digitais proporcionam também facilidade para as pessoas publicarem e organizarem as suas próprias informações em uma variedade de formatos, ocasionando ambiguidades e heterogeneidade. Esses fatores dificultam à compreensão e o acesso a informação em ambientes digitais.

Diante de tal dificuldade Rosenfeld, Morville e Arango (2015) apresentam o sistema de organização em esquemas e estruturas de organização da informação. O esquema define as características comuns de itens de conteúdo e influencia o agrupamento lógico desses itens; a estrutura define os tipos de relações entre itens de conteúdo e grupos, além das principais maneiras pelas quais os usuários podem navegar.

Os esquemas de organização se dividem em exatos (alfabética, cronológica e geográfica), ambíguos (tópico, tarefa, público e metáfora) e híbrido. Os esquemas de organização exatos, dividem a informação em categorias bem definidas e mutuamente exclusivas com regras claras para incluir novos itens. Indicados quando o usuário sabe exatamente o que está procurando. Os esquemas de organização ambíguos, dividem a informação em categorias subjetivas. Baseiam-se na ambiguidade inerente a língua e na subjetividade humana. Não possui regras claras de como incluir novos itens. Indicados

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

quando o usuário não sabe exatamente o que está procurando, no entanto podem fazer melhores combinações de assuntos e obterem um melhor resultado durante as suas buscas. Rosenfeld, Morville e Arango (2015) garantem que o sucesso de um esquema de organização ambígua depende da qualidade de como os itens são inseridos no sistema, além da avaliação constante da satisfação dos usuários que utilizam o sistema.

As subdivisões dos esquemas exatos e ambíguos são:

Quadro 1: Esquemas de Organização

		Indicações de Uso
Esquemas	Exatos	Alfabético
		Cronológico
		Geográfico
		Tópico
	Ambíguo	Tarefa
		Público
		Metáfora

Fonte: Adaptado de Rosenfeld, Morville e Arango (2015) e Reis (2007).

O esquema híbrido reúne dois ou mais esquemas de organização, ou seja, é um esquema composto por diferentes elementos de organização. Pela diversidade de conteúdos na *Web* em uma variedade de formatos, observa-se que o esquema de organização híbrida é de uso frequente entre os sites. Para Reis (2007), a diversidade de conteúdos na *Web* dificulta a criação de um sistema de organização único que atenda a todo o *Website*. Ao se utilizar diferentes esquemas de organização Rosenfeld, Morville e Arango (2015) lembram da importância de preservar a integridade de cada esquema para manter a capacidade de um modelo mental para os usuários.

Os componentes de representação interagem com a camada estrutural definindo os

tipos de relações entre o conteúdo dos itens e dos grupos.

Entre as funções da estruturação de um site, a partir dos componentes de representação, incluem a eliminação da ambiguidade, controle de sinônimos e estabelecimento de relacionamentos semânticos entre conceitos.

As principais estruturas de organização que se aplicam a arquitetura de informação incluem a hierarquia, o modelo de banco de dados orientado e hipertexto. Cada estrutura de organização possui pontos fortes e fracos. Em alguns casos, faz sentido usar um ou o outro. Em muitos casos, faz sentido usar todos os três de uma forma complementar (ROSENFELD, MORVILLE e ARANGO, 2015).

As relações hierárquicas fornecem uma maneira simples e familiar para organizar as informações. Os usuários podem rapidamente compreender ambientes de informação que utilizam modelos de organização hierárquica.

As estruturas com base em modelo de banco de dados, ou *bottom-up*, utiliza bases de dados relacional e metadados. Para Morville e Rosenfeld (2006), ao utilizar esta estrutura de banco de dados em conjunto com metadados e vocabulários controlados é possível ao arquiteto da informação ativar: geração automática de índices alfabéticos, apresentação dinâmica de *links* tipo ver também, pesquisa por campo, filtros e ordenação avançados dos resultados de busca.

A estrutura hipertextual é composta pelos componentes principais: os itens ou pedaços de informação. E, ainda que este tipo de estrutura proporcione grande flexibilidade, tem grande potencial para complexidade e confusão do usuário, devido a particularidades das associações.

6 ANÁLISE DO SISTEMA DE ORGANIZAÇÃO DO SOFTWARE TAINACAN

A ferramenta Tainacan é flexível e permite a gestão e publicação de coleções digitais estabelecendo diferentes filtros de buscas, com mais de um tipo de esquema, o que caracteriza sua organização como híbrida. Apesar disso, as coleções possuem autonomia para definir os melhores filtros, a partir das necessidades de seu público.

A ferramenta possibilita a construção de uma estrutura de categorias com níveis e subníveis permitindo criar uma hierarquia de coleções trabalhando com o conceito de

herança entre elas. As coleções segmentam o acervo de acordo com sua tipologia institucional e documental.

Apresenta-se a seguir a análise do sistema de organização do software livre Tainacan, a partir do estudo das coleções: Biblioteca Modelo da UFG e Retratos do Império do MHN. A análise teve como base os esquemas exatos e ambíguos, bem como a análise estrutural que engloba os componentes de representação da informação.

6.1 SISTEMA DE ORGANIZAÇÃO DO SOFTWARE TAINACAN: ANÁLISE DOS ESQUEMAS EXATOS E AMBÍGUOS

Os sistemas de organização apresentados pelas coleções Biblioteca Modelo da UFG e Retratos do Império do MHN reúnem os esquemas exatos e ambíguos como elementos de organização, porém cada qual com sua metodologia de classificação das categorias. Estas categorias constituem os filtros para a realização das buscas.

A coleção Biblioteca Modelo da UFG adota a classificação das obras por níveis de leitura, sendo: leitor iniciante, leitor habilidade e o leitor fluente. As obras indicadas para o leitor iniciante englobam livros que contém somente imagens e algumas palavras, geralmente são voltadas para o público infantil de 0 a 4 anos de idade. O nível habilidade se constitui em imagens e pequenas frases com letras grandes. Para o leitor considerado fluente a indicação são os livros com textos extensos.

A partir dessa primeira classificação estabeleceram-se os rótulos por palavras definindo as categorias e subcategorias que são utilizadas como filtro de busca. Entende-se que a metodologia de classificação por nível de leitura direciona as obras para o seu público-alvo, demonstrando que a coleção apresenta o esquema ambíguo público.

O acervo da Biblioteca Modelo da UFG é composto também por livros juvenis que são classificados a partir de seu conteúdo, como literatura brasileira, literatura portuguesa, literatura inglesa, dentre outras. Essa organização por assunto é definida por Rosenfeld, Morville e Arango (2015) como esquema ambíguo do tipo tópico.

O esquema exato cronológico é visto na organização da coleção Biblioteca Modelo da UFG ao apresentar a categoria ano de publicação da obra como filtro de busca. Este item deve ser considerado em função das obras serem indicadas a premiações em seu ano de publicação, sendo assim o público pode realizar buscas pelo ano de publicação.

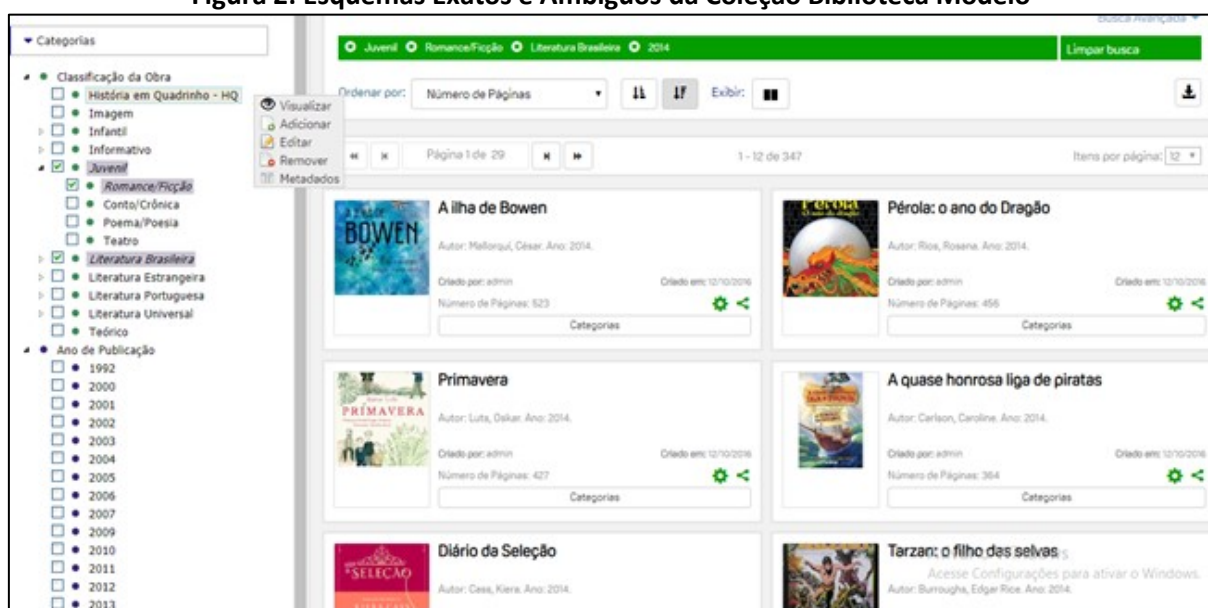
XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

O esquema exato alfabético pode ser observado na coleção Biblioteca Modelo na categoria classificação da obra. Nesta categoria, os assuntos são ordenados alfabeticamente.

A coleção Biblioteca Modelo possui o esquema ambíguo tarefa destinado ao público visualizar, editar, remover e adicionar metadados. Para garantir a integridade e veracidade dos dados, esta tarefa está sujeito a aprovação dos gestores da coleção.

Os esquemas exatos e ambíguos adotados na coleção da Biblioteca Modelo, da UFG podem ser visualizados na figura 2.

Figura 2: Esquemas Exatos e Ambíguos da Coleção Biblioteca Modelo



Fonte: Tainacan (2019).

Na coleção Retratos do Império foram identificados os esquemas exatos alfabético e cronológico e o esquema ambíguo tópico e tarefa. O esquema exato alfabético é estabelecido na organização dos autores e o esquema exato cronológico na organização das datas de produção das pinturas.

O esquema ambíguo do tipo tópico pode ser visualizado para a organização da informação das pinturas por assunto, tipo de material (por exemplo: guache, madeira, marfim, óleo, etc) e técnica de produção da pintura (por exemplo: óleo e guache sobre marfim). Estes tópicos oferecem filtros de buscas que permitem o agrupamento lógico dos itens possibilitando a recuperação da informação útil.

O esquema exato geográfico e o esquema ambíguo metáfora não foram identificados nas coleções analisadas. Na coleção Retratos do Império, do MHN também não foram

identificados os esquemas ambíguos tarefa e público. Acredita-se que os esquemas ausentes sejam irrelevantes para os usuários destas coleções.

6.2 SISTEMA DE ORGANIZAÇÃO DO SOFTWARE TAINACAN: ANÁLISE ESTRUTURAL

As hierarquias definidas pelas coleções analisadas apresentam certo equilíbrio entre amplitude e profundidade. A coleção Biblioteca Modelo apresenta até três níveis hierárquicos, por exemplo, ‘Classificação da obra – Juvenil – Teatro’, sendo necessários três cliques para encontrar a informação mais específica, conforme pôde ser visualizado na figura 2. A coleção Retratos do Império necessita de dois cliques ‘Autor – Auguste Petit’. Esta última coleção não requer maior aprofundamento para encontrar e os objetos informacionais publicados.

Na definição das coleções adotou-se o princípio geral da classificação ao dividir as categorias em subcategorias do geral para o mais específico permitindo, dessa forma, o controle da ambiguidade e sinônimos entre os termos. Este princípio é facilmente identificável na coleção Biblioteca Modelo que possui maior aprofundamento hierárquico.

Esta organização foi realizada a partir das informações preenchidas em formulários de metadados no padrão Dublin Core ou outros padrões de metadados reconhecidos, pois de acordo com o Manual do Usuário do Tainacan (2015), o software “permite a criação de coleções oriundas, por exemplo, de um álbum de fotos do Flickr, vídeos do Youtube, imagens do Facebook, objetos digitais da Europeana e de outros repositórios digitais que permitam coletar dados em padrões de metadados reconhecidos” (TAINACAN, 2015).

Durante a descrição dos documentos em metadados, o Tainacan permite interligar semanticamente uma coleção de livro com uma coleção de pesquisadores, por exemplo. Esse modelo de banco de dados relacional possibilita combinar informações durante a busca avançada. Na busca avançada os usuários podem combinar campos de metadados como assunto, autor, título ou descrição, tipo de documento, formato etc.

O Tainacan utiliza a estrutura hipertexto para interligar itens dentro da mesma coleção ou em ambientes de informação da Web. O uso de hipertextos favorece uma navegação mais dinâmica e estimula o usuário a explorar o ambiente de informação. Apesar disso, não foi possível visualizar na coleção Biblioteca Modelo o uso de hipertexto. Nesse sentido, sugere-se o uso de hipertexto das obras catalogadas inclusive com as indicações e

prêmios recebidos. Enquanto que na coleção Retratos do Império foi identificado o uso de hipertexto no metadado 'condições de reprodução', em que os direitos reservados para reprodução são direcionados para o *link* de outra página do Museu Histórico Nacional.

O ambiente de informação digital orientado por metadados e vocabulários controlados proporciona ao usuário uma experiência mais uniforme e satisfatória. Os metadados possuem um conjunto de elementos para descrever um documento de forma a individualizá-lo entre muitos outros para facilitar a sua recuperação no meio digital. Para tanto, para garantir a uniformidade dos dados, a sua utilização conjunta com regras de descrição e vocabulários controlados apresenta-se como solução.

Sendo assim, analisou nas duas coleções a padronização de preenchimento dos metadados, constatou-se que o acervo da Biblioteca Modelo da UFG adota um nível simples de descrição física com elementos essenciais para identificação dos livros (autor, título, tradutor, ilustrador, edição, local, editora, ano, número de página e ISBN) e segue as regras de descrição do Código de Catalogação Anglo-Americano (AACR2r). A coleção Retratos do Império apresenta a descrição dos campos de metadados para denominação, descrição, número de registro, título, autor, técnica (por exemplo: óleo sobre porcelana), material (por exemplo: óleo, porcelana), forma de aquisição, fonte de aquisição, referência de aquisição, local de produção, data de produção, classe, estado de conservação, altura, largura, termos de indexação, nome da exposição (coleção), condições de reprodução. Apesar da descrição completa e seguir uma certa uniformidade no preenchimento dos campos, não foi possível identificar o uso de padrão bibliográfico para descrição das pinturas.

A classificação encontrada no Tainancan segue o princípio da divisão, segundo as suas semelhanças e diferenças, na organização dos documentos. Para organização física dos documentos nas estantes, a biblioteca Modelo adota um esquema de classificação adaptado da Classificação Decimal Universal (CDU) e Classificação Decimal de Dewey (CDD) elaborada para bibliotecas especializadas em literatura infantil, juvenil e informativa, porém não foi identificado nenhum campo de metadado para o registro da classificação numérica. Para a coleção Retratos do Império não foi possível identificar o uso de sistema de classificação decimal.

A indexação da coleção Biblioteca Modelo é apresentada nas categorias e subcategorias seguindo a mesma metodologia da classificação por leitores. No entanto, não há evidências de uso de listas de cabeçalhos de assunto ou tesauros. O campo de metadado

destinado ao assunto é denominado de classificação da obra. Na descrição da coleção Retratos do Império foi definido o campo de metadado, denominado de termos de indexação; para descrever o assunto, foram identificados entre cinco a vinte termos de indexação, porém não foi identificado o uso de lista de cabeçalhos de assunto ou tesauros para indexação do assunto das pinturas.

O Tainacan não adota um tesouro padrão para as suas coleções. Esse fato pode ser explicado pela diversidade das áreas que utilizam este software para gerenciar as suas coleções digitais e/ou analógicas. Contudo, cada instituição gestora de coleção tem autonomia para adotar um tesouro específico para sua área. Ainda assim, seria interessante a adoção ou o planejamento de um tesouro para a área de cultura, a qual representa o propósito inicial do software – Gestão de Acervos da Cultura Digital Brasileira. Da mesma forma, recomenda-se aos gestores das coleções analisadas o uso de tesouro para o efetivo estabelecimento de relacionamentos equivalentes, hierárquicos e associativos entre os termos, além de controlar os problemas comuns da linguagem natural como ambiguidades, heterogeneidade, homografia e a polissemia.

O Tainacan permite o estabelecimento de taxonomias facetadas, pois define uma hierarquia distinta para cada metadado ao criar um filtro de busca e o usuário pode combinar termos de facetas diferentes para filtrar os resultados desejados.

Os componentes de representação anéis de sinônimo e arquivo de autoridade não foram identificados nas buscas das coleções do sistema Tainacan.

Diante do exposto, observa-se que os componentes de representação interagem com a camada estrutural do sistema de organização definindo os tipos de relações entre o conteúdo dos itens e dos grupos favorecendo, sobretudo, os usuários para a busca e navegação nos ambientes de informação digital.

7 CONSIDERAÇÕES FINAIS

Considerando a importância dos metadados e vocabulários controlados para a arquitetura da informação de um *site*, este estudo propôs-se a identificar os componentes de representação da informação adotados pelo sistema de organização do software Tainacan, de modo a oferecer aos usuários a informação relevante com mais clareza, rapidez e com satisfação ao navegar pelo ambiente digital.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

A análise foi realizada a partir dos esquemas e estruturas utilizados pelas coleções Biblioteca Modelo, da UFG e Retratos do Império, do MHN agregadas ao sistema de gestão de acervos digitais Tainacan.

Os resultados apresentam que o sistema de organização do Tainacan é caracterizado como esquema híbrido e estrutura principal hierárquica com o uso complementar das estruturas de modelo de banco de dados relacional e hipertexto. Caracteriza-se como esquema híbrido, pois a coleção da Biblioteca Modelo adota o esquema exato alfabético e cronológico e os esquemas ambíguos do tipo tópico (assunto), tarefa (funções de editar metadados) e público-alvo ao estabelecerem as categorias e subcategorias direcionadas aos leitores iniciantes, leitores com habilidade e leitores fluente. A coleção Retratos do Império adota o esquema exato alfabético na organização dos autores e cronológico para organização da data de produção das pinturas. Adota ainda, o esquema ambíguo orientado a tópicos (assunto, material e técnica) para definição de suas categorias e subcategorias.

A partir do modelo adotado neste estudo, a estrutura principal foi considerada hierárquica por possibilitar a construção de categorias e subcategorias com níveis e subníveis estabelecendo as hierarquias das coleções.

Para estabelecer a estrutura hierárquica, o Tainacan adota o vocabulário controlado denominado de Taxonomia, o qual permite estabelecer relacionamento hierárquico e o controle de ambiguidades e sinônimos.

O Tainacan adota como elemento de uso complementar a estrutura de modelo de banco de dados que permite relacionar categorias e subcategorias como filtro de busca. Na pesquisa avançada é permitido relacionar vários campos de metadados na expectativa de melhorar a precisão, evitando assim a recuperação de itens inúteis.

Ainda como estrutura complementar foi possível observar o uso de hipertexto na organização dos documentos, pois os campos de metadados destinado ao assunto são direcionados para outro documento dentro da mesma coleção. Nas coleções analisadas não foi identificado o direcionamento para páginas fora do Tainacan.

Os principais componentes de representação da informação identificados no Tainacan foram os metadados, esquema de classificação e taxonomia.

Nas coleções analisadas o padrão de metadados utilizado foi o Dublin Core, porém o Tainacan adota outros padrões reconhecidos para descrição de documentos.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

Para descrição dos documentos em campos de metadados observou que a Biblioteca Modelo da UFG adota as regras de catalogação estabelecidas pelo AACR2r, o que garante a padronização de entrada dos dados. No campo de assunto é apresentado apenas um termo e não adota linguagem controlada para indexação do assunto, o que dificulta a recuperação da informação exata. Na descrição dos documentos da coleção Retratos do Império não foi identificado regras para catalogação, contudo foi possível observar uniformidade no preenchimento dos campos; no campo termos de indexação observou-se o preenchimento de cinco a vinte assuntos.

O Tainacan segue os princípios da divisão para classificação das categorias e suas subcategorias, mas não observou um esquema de classificação decimal no preenchimento dos metadados.

Após a verificação dos componentes de representação presentes no Tainacan, recomenda-se o uso de listas para controle de autoridades e anéis de sinônimos para recuperação de termos equivalentes.

Para a Biblioteca Modelo da UFG recomenda-se inserir o código de classificação para as pessoas interessadas a recorrerem ao espaço físico da biblioteca.

Reforça-se a importância do estudo de um tesauro para as coleções desenvolvidas a partir do software Tainacan. Da mesma forma, recomenda-se aos gestores das coleções analisadas o uso de tesauro para o efetivo estabelecimento de relacionamentos equivalentes, hierárquicos e associativos entre os termos, além de controlar os problemas comuns da linguagem natural como ambiguidades, heterogeneidade, homografia e a polissemia.

Conclui-se que os ambientes digitais que adotam os componentes de representação em suas estruturas organizacionais transmitem aos usuários uma busca e navegação mais agradáveis e oferecem ao público a informação de forma compreensiva.

REFERÊNCIAS

ALVES, Rachel Cristina Vesu. **Metadados como elemento do processo de catalogação**. 2010. 132 f. Tese (Doutorado) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista Júlio de Mesquita Filho, 2010.

ALVES, Rachel Cristina Vesu; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. **Metadados no domínio bibliográfico**. Rio de Janeiro: Intertexto, 2013.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

ANSI/NISO. American National Standards Institute; National Information Standards Organization. **ANSI/NISO Z39.19**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda, Maryland, USA: NISO Press, 2005. 172 p. ISBN: 1-880124-65-3.

BRASCHER, Marisa; CAFÉ, Lígia. Organização da informação ou organização do conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9. 2008, São Paulo. **Anais eletrônicos ...** São Paulo. Disponível em: [http://cmappublic.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%C3%89\(2008\)-1835.pdf](http://cmappublic.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%C3%89(2008)-1835.pdf) . Acesso em: 14 set. 2018.

FUSCO, Elvis. **Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais**. São Paulo: Cultura Acadêmica, 2011.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Brinquet de Lemos, 2004.

MORVILLE, Peter; ROSENFELD, Louis. **Information architecture for the world wide web**. 3. ed. Cambridge: O'Reilly, 2006.

MUSEU HISTÓRICO NACIONAL (Brasil). Disponível em: <http://mhn.museus.gov.br/index.php/acervo/>. Acesso em: 05 ago 2019.

NOVELLINO, Maria Salet Ferreira. Instrumento e metodologias de representação da informação. **Inf.Inf.**, Londrina, v.1, n.2, p.37-45, jul./dez. 1996.

OLIVEIRA, Henry Poncio Cruz de. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. 202 f. Tese (Doutorado) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

OLIVEIRA, Henry Poncio Cruz de; VIDOTTI, Silvana Aparecida Borsetti Gregório; BENTES, Virgínia. **Arquitetura da informação pervasiva**. São Paulo: Cultura Acadêmica, 2015.

PIEDEDE, M. R. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1983.

REIS, Guilherme Almeida dos. **Centrando a arquitetura de informação no usuário**. São Paulo, 2007. Dissertação (Mestrado) - Escola de Comunicação e Artes, Universidade de São Paulo. 2007.

ROBREDO, Jaime. Sobre arquitetura da informação. **Revista Ibero-Americana de Ciência da Informação** (RICI), v. 1, n.1, p. 115-137, jul./dez. 2008.

ROSENFELD, Louis; MORVILLE, Peter; ARANGO, Jorge. **Information architecture**: for the web and beyond. 4. ed. Beijing. Boston: O'Reilly, 2015.

SILVA, Maria Amélia Teixeira da; PINHO NETO, Júlio Afonso Sá de; DIAS, Guilherme Ataíde. Arquitetura da informação para quem e para quem?: uma reflexão a partir da prática em

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

ambientes informacionais digitais. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da Informação, v. 18, n. 37, p. 283-302, mai./ago., 2013.

TAINACAN. Manual do usuário. Goiânia: UFG, 2015.

TAINACAN. Goiânia: UFG, 2019. Disponível em:
<https://www.gi.fic.ufg.br/biblioteconomia/colecao/bibliotecamodelo/>. Acesso em: 05 ago. 2019.

VIDOTTI, Silvana Aparecida Borsetti Gregório; CUSIN, César Augusto; CORRADI, Jiliane Adne Mesa. Acessibilidade digital sob o prisma da Arquitetura da Informação. In: GUIMARÃES, José Augusto Chaves; FUJITA, Mariângela Spotti Lopes. **Ensino e pesquisa em Biblioteconomia no Brasil**: a emergência de um novo olhar. São Paulo: Cultura Acadêmica, 2008.

WURMAN, Richard Saul. **Information Architects**. Zurich, Schweiz: Gingko Press, 1997.