



GT 8

Informação e Tecnologia

O GT 8 aborda Estudos e pesquisas teórico-práticos sobre e para o desenvolvimento de tecnologias de informação e comunicação que envolvam os processos de geração, representação, armazenamento, recuperação, disseminação, uso, gestão, segurança e preservação da informação em ambientes digitais.



SUMÁRIO

MODELOS CONCEITUAIS DE DADOS COMO PARTE DO PROCESSO DA CATALOGAÇÃO: PERSPECTIVA DE USO DOS FRBR NO DESENVOLVIMENTO DE CATÁLOGOS BIBLIOGRÁFICOS DIGITAIS <i>Elvis Fusco, Plácida Santos</i>	2448
INFORMAÇÃO E SOCIEDADE: POLÍTICAS DE PRESERVAÇÃO E DISPONIBILIZAÇÃO DE DADOS SOBRE A SEGURANÇA PÚBLICA EM MOÇAMBIQUE <i>Januário Albino Nhacuongue, Edberto Ferneda, Plácida Leopoldina Ventura Amorim da Costa Santos</i>	2466
ARQUITETURA DA INFORMAÇÃO COMO ARCABOUÇO PARA A INTEGRAÇÃO DA TECNOLOGIADA INFORMAÇÃO COM OS PROCESSOS ORGANIZACIONAIS E A INFORMAÇÃO ¹ <i>Márcio de Carvalho Victorino, Marisa Bräscher Basílio Medeiros, Rômulo Ferreira dos Santos</i>	2481
ANALISANDO AS QUESTÕES DE USABILIDADE E ACESSIBILIDADE DO PORTAL DE PERIÓDICOS DA CAPES <i>Sandra de Albuquerque Siebra, Jaciane Freire Santana, Denis Silva da Silveira</i>	2499
ARQUITETURA DA INFORMAÇÃO E ENGENHARIA SEMIÓTICA: UM ESTUDO DE CASO DA INTRANET DA UNIMED JOÃO PESSOA <i>Lilian Viana Teixeira Cananéa, Guilherme Ataíde Dias, Marckson Roberto Ferreira de Sousa, Maria Amélia Teixeira da Silva</i>	2517
<EM BUSCA DA COMPREENSÃO DA “BUSCA” NO CIBERESPAÇO> <i>Silvana Drumond Monteiro, Maria Júlia Carneiro Giraldes, Fernando Luiz Vechiato, Rogério Paulo Müller Fernandes, Nelma Camêlo de Araujo</i>	2536
ABORDAGENS PARA AVALIAR BIBLIOTECAS DIGITAIS <i>Izabel França Lima, Renato Rocha Souza, Guilherme Ataíde Dias</i>	2552
DADOS ABERTOS GOVERNAMENTAIS (OPEN GOVERNMENT DATA): INSTRUMENTO PARA EXERCÍCIO DE CIDADANIA PELA SOCIEDADE <i>Claudio Jose Silva Ribeiro, Reinaldo Figueiredo de Almeida</i>	2568
CULTURA DA WIKIPÉDIA COMO POLÍTICA DE CAPACITAÇÃO COMPARTILHADA EM MEIO DIGITAL DE SOCIALIZAÇÃO DO CONHECIMENTO <i>Maria José Vicentini Jorente</i>	2581
USO DE TECNOLOGIAS DA INFORMAÇÃO PARA ANÁLISE DE RELEVÂNCIA USANDO “CROWDSOURCING” <i>Marcos Luiz Mucheroni, José Fernando Modesto da Silva</i>	2598
AVALIAÇÃO DE USABILIDADE DA BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES: UM ESTUDO DE CASO <i>Fernanda Pereira, Gercina Ângela Borém de Oliveira Lima</i>	2610



ARQUITETURA DA INFORMAÇÃO: REPRESENTAÇÃO DA INFORMAÇÃO DE PRONTUÁRIO ELETRÔNICO DO PACIENTE <i>Virginia Bentes Pinto, Hamilton Rodrigues Tabosa,</i> <i>Silvana Ap. B. Gregorio Vidotti</i>	2629
ANÁLISE DE RISCO: UM MÉTODO PARA A PRESERVAÇÃO DIGITAL <i>Vildeane da Rocha Borba, Fanny do Couto Ribeiro, Marcos Galindo Lima,</i> <i>Májory Karoline Fernandes de Oliveira Miranda</i>	2643
SERVIÇOS DE RECOMENDAÇÃO EM BIBLIOTECAS DIGITAIS CIENTÍFICAS <i>Flavia Maria Bastos, Silvana Aparecida Borsetti Gregorio Vidotti</i>	2662
CLIO-I : PRIMANDO PELA USABILIDADE E ACESSIBILIDADE EM UM SISTEMA PARA GERENCIAMENTO E INTEROPERABILIDADE DE REPOSITÓRIOS DIGITAIS <i>Marcos Cardoso Jr., Sandra de Albuquerque Siebra, Marcos Galindo, Vildeane Borba,</i> <i>Majóry Miranda</i>	2676
O SISTEMISMO DE BUNGE: FUNDAMENTOS, ABORDAGEM METODOLÓGICA E APLICAÇÃO A SISTEMAS DE INFORMAÇÃO <i>Vinicius Medina Kern</i>	2693
O IMPACTO DO DESIGN EMOCIONAL NA RECUPERAÇÃO DA INFORMAÇÃO NO PORTAL DO INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT <i>Luciana dos Santos Nahuz</i>	2710
REPOSITÓRIO ELETRÔNICO INSTITUCIONAL: PRESERVANDO O CONHECIMENTO ATRAVÉS DE TECNOLOGIAS DA INTELIGÊNCIA <i>Rosilene Fernandes Machado, Guilherme Ataíde Dias</i>	2719
METADADOS PARA A REPRESENTAÇÃO DAS IMAGENS DIGITAIS <i>Ana Carolina Simionato</i>	2725
USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO MULTIMODAL <i>Emilio Evaristo de Sousa, Cláudio Gottschalg Duque</i>	2730
WEB SEMÂNTICA E SUAS CONTRIBUIÇÕES PARA A CATALOGAÇÃO: UM ESTUDO SOBRE ARQUITETURAS DE METADADOS E FRBR <i>Renata Eleuterio da Silva, Plácida Leopoldina Ventura Amorim da Costa Santos</i>	2738
MINERAÇÃO DE DADOS PARA APOIO À TOMADA DE DECISÃO DE DOCENTES NA AVALIAÇÃO DE ALUNOS DE GRADUAÇÃO A DISTÂNCIA <i>José Leonardo OLiveira Lima</i>	2744



MODELOS CONCEITUAIS DE DADOS COMO PARTE DO PROCESSO DA CATALOGAÇÃO: PERSPECTIVA DE USO DOS FRBR NO DESENVOLVIMENTO DE CATÁLOGOS BIBLIOGRÁFICOS DIGITAIS

Elvis Fusco, Plácida Santos

Resumo: O processo de catalogação ocupa-se dos registros bibliográficos, enquanto suporte de informação, servindo como base para a interoperabilidade entre ambientes informacionais, levando em conta objetos diversificados de informação e bases cooperativas e heterogêneas. Dentre as principais propostas da área de catalogação estão os FRBR - *Functional Requirements for Bibliographic Records*, que constituem novos conceitos nas regras de catalogação. Os FRBR mostram um caminho na reestruturação dos registros bibliográficos de maneira a refletir a estrutura conceitual de persistência e buscas de informação, levando em conta a diversidade de usuários, materiais, suporte físico e formatos. Neste contexto, o objetivo deste artigo é refletir e discutir, a partir de uma arquitetura conceitual de ambientes informacionais, baseada nos FRBR e na Modelagem Entidade-Relacionamento, o processo de catalogação no contexto do projeto de catálogos utilizando a metodologia computacional de Modelagem Conceitual de Dados, considerando a evolução dessa área no âmbito da Ciência da Informação em relação ao contexto da representação da informação com vistas ao uso e à interoperabilidade de todo e qualquer recurso informacional, que vise a preencher a lacuna entre o projeto conceitual de um domínio de aplicação e a definição dos esquemas de metadados das estruturas de registros bibliográficos. A modelagem de dados possibilita uma visão não linear dos elementos descritivos de um item bibliográfico e permite uma eficiente transmissão das mensagens contidas nas representações e a otimização do acesso e uso de ambientes catalográficos, por meio da intersecção entre os itens bibliográficos e as necessidades informacionais dos usuários. Nesse sentido, o processo de catalogação deve ser ampliado adicionando a ele a de definição das estruturas conceituais de representação dos itens bibliográficos e, para tanto, o profissional catalogador necessita buscar essas ferramentas que possibilitam um melhor desenvolvimento da área da catalogação.

Palavras-chave: Catalogação. FRBR. Modelagem de Dados.

1 INTRODUÇÃO

No cenário contemporâneo as informações digitais fazem parte do universo de bases informacionais e necessitam de estruturas de representação e mecanismos de acesso e entre elas podem-se citar publicações avulsas e periódicas, enciclopédias virtuais, bibliotecas digitais e repositórios institucionais e temáticos. As estruturas de representação e os mecanismos de acesso provocam uma evolução na gestão, na modelagem e no uso da informação em meio digital e nesse contexto



a catalogação, enquanto geradora do Tratamento Descritivo da Informação (TDI) e construtora de formas de representação dos recursos informacionais, passa a ser associada à usabilidade da informação digital, aos metadados, às ontologias no uso das taxonomias, à modelagem de dados e conteúdos, ao projeto de banco de dados e à recuperação e descoberta de recursos informacionais.

Esse cenário caracteriza-se pela indissociabilidade entre informação e tecnologia e, neste contexto, o conceito de biblioteca foi ampliado para uma organização provedora de informação diversificando-se, por exemplo, em bibliotecas digitais e repositórios institucionais, nesta tese, denominada unidade de informação.

A crescente introdução de padrões que procuram adequar a organização da informação às novas realidades, a publicação de regras, recomendações e estudos que apontam tendências e perspectivas da catalogação atual, constitui reflexo da evolução do conceito da catalogação que se tornou inegavelmente mais complexa e ampliada, com a rápida transformação dos formatos, padrões, funções, opções de exibição e busca, bem como, com algumas mudanças fundamentais, como as expectativas de controle bibliográfico (BAPTISTA, 2006).

São inúmeros e imensos os desafios no sentido de articular os avanços técnicos da Biblioteconomia tradicional com as novas possibilidades inseridas pela tecnologia. Exige-se cada vez mais um profissional hábil e capacitado para atuar em novos processos, bem como, se faz necessário o desenvolvimento de estudos mais aprofundados de novas metodologias de catalogação direcionadas ao tratamento da informação digital.

O processo de catalogação ocupa-se dos registros bibliográficos de documentos, enquanto suporte de informação, servindo como base para a interoperabilidade entre ambientes informacionais, levando em conta objetos diversificados de informação e bases cooperativas e heterogêneas. Segundo Lourenço (2005, p. 21)

A visão do código de catalogação, como fonte de consulta para a resolução de processos de representação, ressalta o papel deste instrumento como manancial de experiência prévia acumulada, preservando-se tipos e características do incontável número de itens que serviu para proporcionar paulatinamente a criação das regras hoje sistematizadas. Ressalta-se, entretanto que o conhecimento não é estático e novas necessidades sempre surgem impulsionando a atualização dos referidos instrumentos.

A catalogação continua sofrendo mudanças em sua natureza e processos, o que gera novos desafios e revisão das necessidades e conhecimento sobre essas mudanças. As novas tendências da catalogação, como por exemplo, os FRBR - *Functional Requirements for Bibliographic Records* (Requisitos Funcionais para Registros Bibliográficos) e FRAD - *Functional Requirements Authority Data* (Requisitos Funcionais para Dados de Autoridade), tomam autores clássicos e regras já estabelecidas, como os princípios de Cutter, base para a introdução de novos conceitos. Portanto, a estas tendências, todo conhecimento já adquirido nas estruturas bibliográficas deve ser mantido. Entretanto, deve ser expandido e novas tecnologias devem ser implementadas a partir dessa base conceitual já estabelecida e conhecida.



Alguns desafios são postos para o processo da catalogação, como a representação dos diversos tipos de materiais que surgem a partir das novas tecnologias e a interoperabilidade de bases informacionais heterogêneas. Essa complexidade deve ser abstraída do usuário, estabelecendo uma sólida teoria e compreensão das necessidades dos usuários, com a preservação de documentos digitais para acesso futuro e a definição de um nível de padronização que permita um nível global de cooperação.

Para tais desafios é necessário que os profissionais da Ciência da Informação discutam várias questões sobre a organização do conhecimento registrado e sua representação nas mais diferentes formas e práticas.

Os FRBR agem como um modelo de referência e, portanto, operam no nível conceitual, não levando a análise dos requisitos ao nível necessário para um modelo físico de dados. Esta característica é bem estabelecida conforme afirmação, “[...] demasiado abstratos e genéricos para ser um modelo de dados: nenhuma base de dados prática poderia realmente ser desenvolvida exclusivamente baseada naquelas entidades, atributos e relacionamentos que são definidos nos FRBR” (IFLA, 2006).

Isto leva à necessidade da reflexão e pesquisa para o desenvolvimento de padrões e modelos conceituais e lógicos de implementação utilizando esses requisitos.

As implementações dos FRBR se dão em um modelo de dados lógico e físico qualquer, valendo-se do seu modelo conceitual, que é baseado no modelo E-R que trabalha somente no âmbito conceitual.

Os aspectos atuais da catalogação levam a análises mais complexas no desenvolvimento de ambientes informacionais de mediação do conhecimento. As regras dos FRBR mostram um caminho na reestruturação dos registros bibliográficos de maneira a refletir a estrutura conceitual de persistência e buscas de informação, levando em conta a diversidade de usuários, materiais, suporte físico e formatos.

A utilização de uma arquitetura para ambientes informacionais tendo como base um modelo conceitual de requisitos bibliográficos pode corroborar a idéia de que o processo de catalogação sofreu uma evolução, passando da descrição para a representação de recursos informacionais sempre com vistas às expectativas e necessidades do usuário e a interoperabilidade de todo e qualquer recurso informacional. Considerando esta idéia, o papel do profissional da catalogação amplia-se de modo que este deve determinar não somente os elementos a serem representados, mas também participar da elaboração do modelo de dados a ser implementado para persistir as informações desses elementos.

O processo estabelecido a partir do ato de elaborar a representação de um objeto e descrever seu conteúdo, tornando-o recuperável com vistas ao uso encontra nos FRBR elementos que fortalecem o objetivo da elaboração de catálogos, aprimorando e introduzindo nesse processo uma nova perspectiva sobre a estrutura e as relações dos registros bibliográficos. A compreensão e adoção dos FRBR podem direcionar os responsáveis por códigos de catalogação, os fornecedores de sistemas de informação e os próprios catalogadores a repensarem seus produtos objetivando a atender as necessidades dos usuários. Todavia, é importante destacar que os FRBR não se constituem em um



novo código de catalogação, mas provoca impactos sobre sua revisão e construção de novos códigos e na definição de catálogos.

Além desses aspectos, a atividade específica da catalogação passa a fazer parte de um processo mais amplo e complexo com a introdução de novos materiais digitais e suas especificidades (por exemplo: objetos tridimensionais e documentos multimídia, estruturados e semi-estruturados), a crescente demanda por integração de ambientes informacionais e a inserção de teorias vindas de outras áreas como da Ciência da Computação.

Neste sentido, o objetivo desta pesquisa é refletir, a partir de uma arquitetura conceitual, lógica e de persistência de ambientes informacionais, baseada nos FRBR e na Modelagem Entidade-Relacionamento e estendido pelo uso dos conceitos da Orientação a Objetos, o processo de catalogação no contexto do projeto de catálogos utilizando a metodologia computacional de Modelagem Conceitual de Dados, considerando a evolução dessa área no âmbito da Ciência da Informação em relação ao contexto da representação da informação com vistas ao uso e à interoperabilidade de todo e qualquer recurso informacional, que vise a preencher a lacuna entre o projeto conceitual de um domínio de aplicação e a definição dos esquemas de metadados das estruturas de registros bibliográficos.

No intuito de criar uma infra-estrutura de implementação de ambientes informacionais automatizados na área da catalogação que permita que desenvolvedores e pesquisadores da Ciência da Informação possam utilizá-la no processo de análise, projeto e implementação de sistemas de informação de catálogos e unidades de informação, há a necessidade da utilização de uma recomendação que visa a nortear e a padronizar a estrutura dos registros bibliográficos.

2 O USO DA MODELAGEM CONCEITUAL DE DADOS NA CATALOGAÇÃO

A metodologia usada é derivada da ampliação do conceito da catalogação vista como um processo subdividido em duas fases:

1. O projeto do catálogo: fase em que o catalogador baseado em necessidades e requisitos informacionais define as estruturas dos objetos e elementos de representação e seus relacionamentos;
2. A descrição do objeto documentário: fase em que o catalogador realiza a entrada de dados em um registro baseado em um padrão de metadados e a definição dos pontos de acesso deste registro com base em normas de catalogação.

Com base na questão apresentada por Santos (1995, p.35), a seguir:

É preciso concentrar esforços para que a disciplina Representação Descritiva deixe de ser encarada como um recurso meramente técnico da Biblioteconomia, em que o treino da redação de fichas catalográficas serão suficientes para capacitar um profissional ao desempenho de sua função em um sistema de informação. Providências imediatas são necessárias para modernizar e expandir o preparo do bibliotecário no que se refere a catalogação. É importante que se pense na reformulação do conteúdo

e da metodologia da disciplina de Representação Descritiva, para ser assimilada de forma coerente pelo aluno. É preciso que ele compreenda o papel da catalogação e conheça sua relação e inter-relação no contexto informacional de um sistema. É necessário, enfim, se desmistificar a idéia de que o fazer da catalogação seja um produto absoluto e completo que se encerra nele mesmo.

Contempla-se a reflexão da extensão do processo de catalogação no que tange a etapa de projeto de um catálogo, no qual o profissional da informação detém as competências necessárias para, com seus conhecimentos prévios de representação tradicional, planejar e implementar modelos de dados, utilizando-se de métodos conceituais de modelagem.

Diante disso, apresenta-se o conceito de projeto de catálogo contemplando métodos e regras já estabelecidas na área da representação da informação sob os aspectos e requisitos dos FRBR e FRAD, somados a modelos oriundos da área da Ciência da Computação que estudam o tratamento da informação no âmbito conceitual, lógico e de persistência.

2.1 Bases Para o Projeto de um Catálogo

A catalogação tem como objetivos principais a descrição e a recuperação dos itens bibliográficos em ambiente centralizado ou interoperável. Assim, para atender esses objetivos, comumente são utilizadas regras de catalogação (e.g. AACR2) e padrões de metadados (e.g. MARC) como requisitos informacionais básicos para construção dos catálogos.

O profissional que projeta um catálogo tem como fontes básicas das necessidades informacionais do ambiente informacional e seu acervo:

- A Representação Descritiva;
- A Definição dos Pontos de Acesso;
- A Definição dos Dados de Localização.

Baseado nesses requisitos funcionais, o projetista do catálogo define, então, os elementos que farão parte dos metadados do catálogo (Fig. 01).

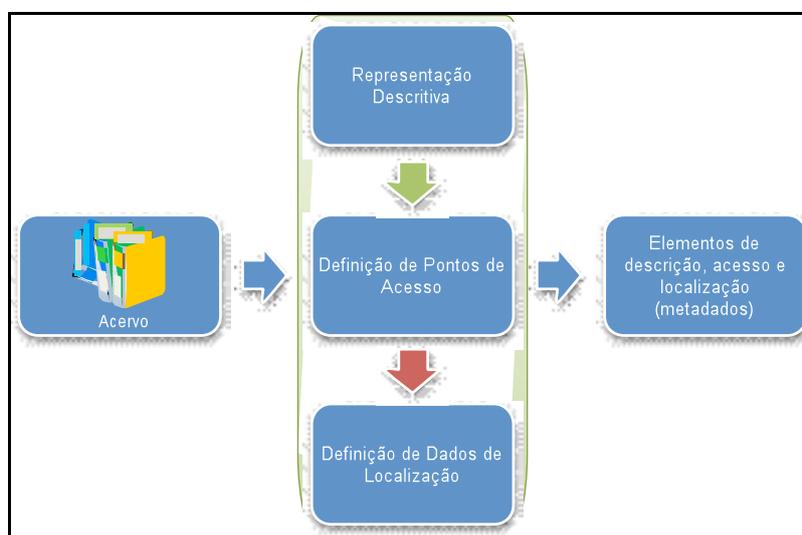


Figura 01 – Fontes das necessidades informacionais um catálogo



Independente de o profissional catalogador participar do processo de construção do banco de dados do ambiente digital, esses elementos de descrição, acesso e localização terão que ser representados neste banco de dados e, para isto, o profissional da computação utilizará esses elementos como base para a construção do modelo de dados.

Excluindo as etapas de modelagem lógica e física do banco de dados, na construção de um catálogo digital que utiliza como suporte um banco de dados relacional, a especificação conceitual (Modelagem Conceitual) da representação dos elementos de descrição, acesso e localização deve fazer parte do processo de construção do projeto do catálogo.

2.2 Aplicação da Modelagem Conceitual no Projeto de Catálogos

Normalmente o processo de catalogação não considera a representação conceitual dos elementos que farão parte dos metadados do catálogo, ficando a critério de um profissional da Ciência da Computação fazer o mapeamento dos elementos definidos pelo catalogador numa estrutura conceitual de entidade, atributos e relacionamentos entre os Itens bibliográficos.

É na Modelagem Conceitual que se desenvolvem modelos de representação das informações de um domínio como um acervo físico ou digital e esta idéia é compartilhada por outros autores dentro da área da Ciência da Informação.

Segundo Delsey (1997, p.1), “num nível teórico e prático, vários estudiosos especulam o desejo de (re)conceituar as estruturas utilizadas para registrar e armazenar dados bibliográficos com uma visão de exploração das tecnologias atualmente disponíveis”.

A Modelagem Conceitual pode auxiliar e melhorar o processo de desenvolvimento de ambientes informacionais digitais baseados no conceito de catalogação e a sua utilização nos FRBR corroborou essa idéia, como afirma Delsey (1997, p.3) “nos ajudar a reexaminar os princípios fundamentais que estão por trás do código de catalogação e a fixar direções para seu desenvolvimento futuro”.

Métodos de modelagem desempenham um papel importante como recurso metodológico para as áreas que têm fenômenos ligados à informação e conforme Burt & Kinnucan (1990), profissionais da Ciência da Informação podem encontrar nas técnicas de modelagem um mecanismo útil para capturar e comunicar seus conhecimentos sobre fontes de informação e sobre padrões de comportamento de quem busca informação.

Levando em consideração os processos da Modelagem Conceitual de Dados, a utilização somente de regras de catalogação como o AACR2 e/ou de um padrão de metadados como o MARC21, para modelar a estrutura conceitual e de persistência de um banco de dados de um sistema informatizado de catalogação, pode levar a algumas anomalias de informação como, por exemplo: redundância de dados; inconsistência de dados; anomalias de exclusão, inclusão e modificação.

Isso ocorre porque padrões de metadados como o formato MARC têm uma estrutura linear em suas estruturas de registros, ou seja, um modelo bidimensional de atributo-valor e determinam o formato de entrada e de intercâmbio de dados, sem determinar o conceito de domínio dos objetos representados com seus relacionamentos.

A forma de representação dos elementos de representação de Itens bibliográficos influencia a qualidade da informação nos processos de armazenamento, busca e recuperação dessas informações, e deve, portanto, fazer parte do processo de construção de catálogos digitais. A utilização de um padrão de metadados que tem como principal objetivo a entrada de dados e a interoperabilidade e não considera questões de qualidade de persistência da informação e, por isso, muitos ambientes digitais de informação que foram baseados somente nesses padrões apresentam problemas tanto nos processos de manipulação de dados quanto na recuperação das representações dos Itens bibliográficos.

Com objetivo de desenvolver um ambiente informacional que atenda as necessidades informacionais de usuários e unidades de informação e que também atenda os requisitos de qualidade da informação, é demonstrado na Fig. 02, um modelo de processo de catalogação contendo os requisitos e metodologias para desenvolvimento de um catálogo digital.



Figura 02 – Modelo de Processo de Catalogação

Nesta proposta os requisitos de informação que farão parte do escopo do ambiente informacional são formados pelo produto da representação descritiva, pontos de acesso, dados de localização, códigos e regras de catalogação e pelos padrões de metadados para manter compatibilidade com outros ambientes informacionais numa estrutura de interoperabilidade.

Esses requisitos gerarão um conjunto de elementos e propriedades de informação que podem ser definidos como a especificação dos requisitos funcionais e não funcionais de informação do catálogo.

Utilizando o conceito da Modelagem Conceitual por meio da metodologia do Modelo E-R, faz-se um mapeamento desses elementos de informação para a estrutura de entidades, atributos e relacionamentos pré-definidos dos FRBR, que já contém um modelo conceitual para registros bibliográficos sem chegar ao nível de implementação. Ou seja, nesse modelo de processo de catalogação proposto, o catalogador irá implementar um modelo conceitual de nível mais baixo de abstração, deixando toda a estrutura de informação pronta para o profissional da Computação somente mapear esta estrutura para um banco de dados. Essa idéia é corroborada por Delsey (1997, p. 2-3)



O maior ganho com a modelagem lógica da estrutura do AACR é que ajuda na mudança de enfoque no processo de catalogar as entidades ou objetos que procuramos representar nos nossos catálogos, passando das regras específicas para as suposições operacionais e princípios das regras e da estrutura formal do registro catalográfico para a estrutura lógica que está por trás dos dados no registro. A disciplina do próprio exercício de modelagem nos obrigaria a clarear nossos pensamentos a respeito dos conceitos que integram a lógica do código.

A Modelagem Conceitual pode trazer grandes contribuições às pesquisas na área de Ciência da Informação e tem uma grande importância na especificação de catálogos digitais, pois:

- Permite disponibilizar as informações de forma estruturada e eficiente;
- Evita a redundância de informações e aumenta a confiabilidade dos ambientes informacionais;
- Define um planejamento que deverá ser seguido pela equipe de catalogadores;
- Possibilita a reutilização de artefatos produzidos em outros catálogos e
- Facilita o desenvolvimento do ambiente digital e a manutenção do catálogo.

E nesta circunstância, a documentação gerada servirá de meio de comunicação entre a equipe.

3 CONSTRUÇÃO DO MODELO CONCEITUAL DE DADOS BASEADO NOS FRBR

Nesta seção, demonstra-se o processo de construção de um modelo conceitual de dados de um catálogo bibliográfico baseado nos requisitos dos FRBR e no Modelo Entidade-Relacionamento.

A construção deste modelo utiliza a notação gráfica E-R de Peter Chen (CHEN, 1976). Este modelo não esquematiza os atributos das entidades. A notação gráfica do software utilizado para diagramar as entidades e relacionamentos usa linhas duplas não tendo relação com entidades fracas da modelagem E-R.

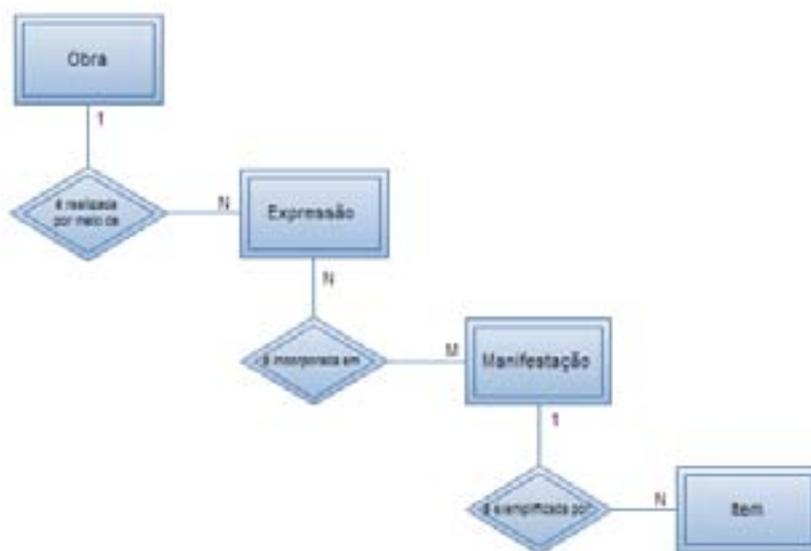


Figura 03 – Modelo Conceitual com as Entidades do Grupo 1



O modelo de dados mostrado na Fig. 03 tem como base as entidades do Grupo 1 dos FRBR e relacionamentos entre essas entidades. Bancos de dados de catálogos baseados nessas entidades mapeiam os elementos descritivos dos registros bibliográficos distribuídos em objetos informacionais (tabelas no Modelo Relacional e classes no Modelo Orientado a Objetos) gerados a partir dessas entidades e relacionamentos.

Na Fig. 04 tem-se o modelo de dados conceitual de um catálogo bibliográfico baseado nas entidades e relacionamentos dos Grupos 1, 2 e 3 dos FRBR e das entidades do FRAD incorporadas aos FRBR.

Este modelo contempla:

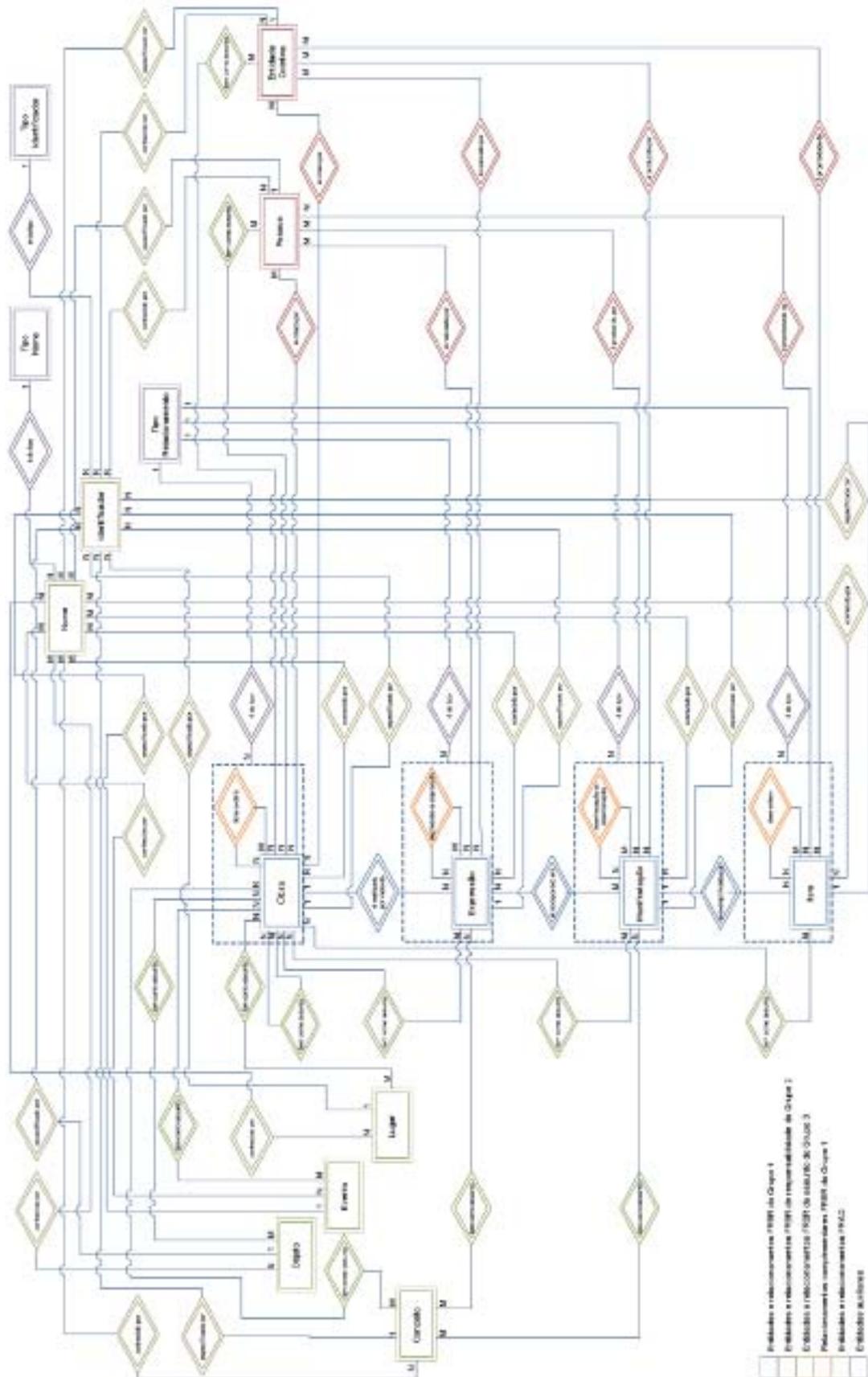
- Entidades do Grupo 1;
- Entidades do Grupo 2;
- Entidades do Grupo 3;
- Relacionamentos de responsabilidade entre as entidades do Grupo 1 e Grupo 2;
- Relacionamentos de assunto do Grupo 1 com as entidades do Grupo 2 e Grupo 3;
- Relacionamentos complementares do Grupo 1 e
- Entidades Nome e Identificador do modelo FRAD incorporadas aos FRBR.

Não foram especificados os atributos FRBR que são representados nas entidades especificadas.

A definição dos atributos cabe ao projetista (catalogador) diante dos requisitos funcionais levantados no processo de construção do catálogo. Nesse contexto apresentam-se duas técnicas em relação à especificação dos atributos no projeto do catálogo:

- I) Definição dos atributos no modelo conceitual, cabendo ao projetista (catalogador) do modelo a definição dos atributos que serão utilizados para representar a entidade dentro do cenário, esses atributos tem como objetivos cumprir as *user tasks*, além de atender a outros requisitos levantados no processo de construção do catálogo. Essa técnica é chamada aqui de “modelagem estática de atributos”.
- II) Definição de uma estrutura conceitual que permita ao catalogador, na parametrização do ambiente informacional para a construção do catálogo, definir os atributos que cada entidade bibliográfica irá conter, de acordo com as *user tasks* e os outros requisitos informacionais do cenário. Essa técnica é chamada aqui de “modelagem dinâmica de atributos”.

A Fig. 05 mostra o modelo para representar a técnica da modelagem dinâmica de atributos que permite ao usuário catalogador definir para cada entidade bibliográfica os atributos que serão utilizados no momento de catalogar o objeto informacional.



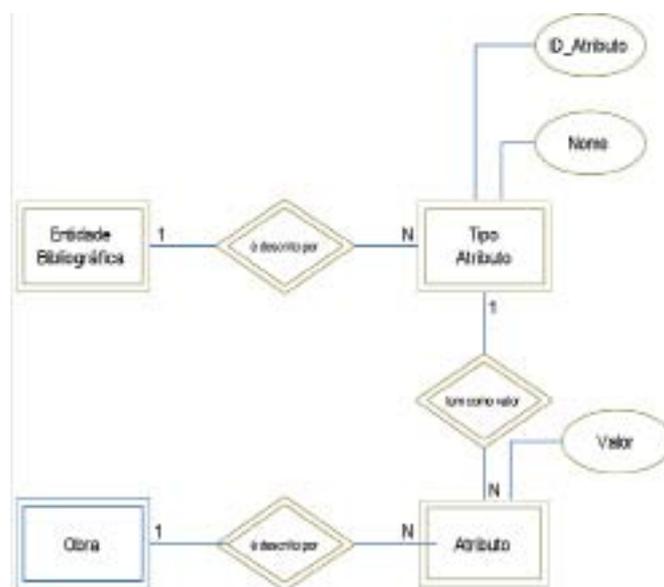


Figura 05 – Modelagem Dinâmica de Atributos

Neste caso, são definidos para cada instância da Entidade Bibliográfica, os atributos que serão utilizados nas interfaces de entradas de dados (entidade Tipo Atributo). Após essa definição, para cada instância de uma entidade bibliográfica, como Obra, podem ser criadas instâncias desses atributos pela entidade Atributo como mostrado na Fig. 06.

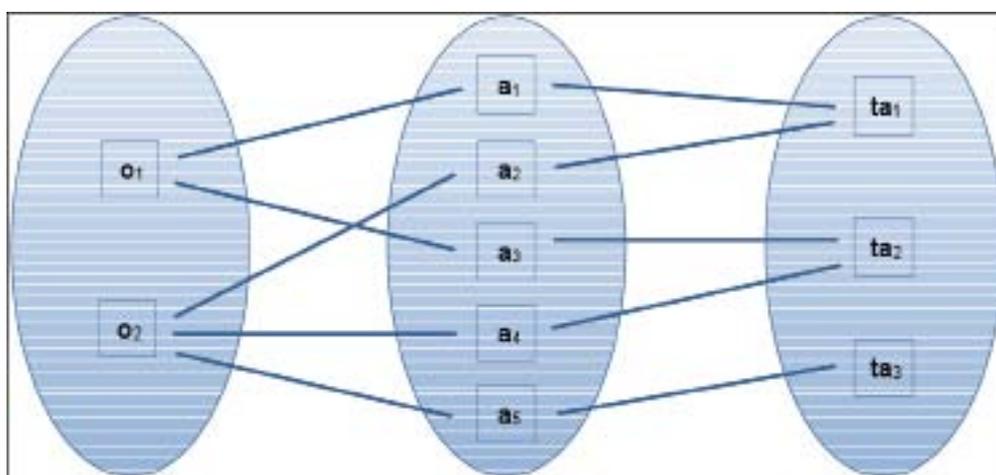


Figura 06 – Representação Gráfica da Modelagem Dinâmica de Atributos

Na Fig. 06 foram especificadas num cenário qualquer 3 instâncias para a entidade Tipo de Atributo referente à entidade Obra. Para a instância o1 da entidade Obra foram definidos para os tipos de atributos ta1 e ta2, os valores a1 e a3 respectivamente, para a instância o2 foram definidos para os atributos ta1, ta2, ta3, os valores a2, a4 e a5 respectivamente.

Nesta perspectiva de projeto, fica a critério do catalogador definir na implantação de um ambiente informacional de um catálogo especificar os atributos que serão utilizados para representar



cada entidade bibliográfica e isso deixa de ser papel do projetista catalogador.

Outro requisito de extrema importância em projetos de catálogos atuais é criar uma camada de interoperabilidade com outros sistemas computacionais baseados em padrões de metadados, neste contexto considerar que o modelo possa permitir uma integração com diferentes tipos de padrões de metadados traz uma característica de flexibilidade e de extensibilidade ao possibilitar à Unidade de Informação a liberdade de escolher integrar seus registros bibliográficos em diferentes formatos e padrões de metadados.

Para definição dos requisitos que demonstram um meio para existir no modelo, a integração com padrões de metadados, foram utilizados os estudos de Delsey (2006) que analisa e mapeia modelos FRBR com o padrão de metadados MARC 21. A partir desse estudo foram adicionadas no modelo as condições para que os elementos de dados do catálogo pudessem ser integrados com elementos de dados de padrões de metadados.

Para os FRBR, o registro bibliográfico é entendido no seu sentido mais amplo, cobrindo uma série de funções, não só elementos descritivos como também pontos de acesso, como assunto, por exemplo, elementos de organização e notas. Os atributos FRBR, portanto, nem sempre são similares aos elementos de metadados do MARC, podendo estar distribuídos em mais de um campo e subcampo (MORENO, 2006).

Nesse contexto, para criar a possibilidade de integração dos elementos bibliográficos do catálogo com elementos de dados de padrões de metadados é necessária a associação de entidades, relacionamentos e atributos estáticos do modelo, ou seja, que foram projetados fazendo parte do esquema do projeto, assim como, com os atributos dinâmicos que são estabelecidos pelo catalogador na implantação do ambiente informacional do catálogo.

Na Fig. 07, está a visão do modelo proposto para implementar a funcionalidade de integração com os padrões de metadados. Para tanto foi definida uma entidade para especificar os padrões de metadados que a Unidade de Informação irá trabalhar e para cada padrão são relacionados os elementos de metadados utilizados para integrar com o catálogo. Para integrar esses elementos às informações do esquema, foi criado um relacionamento entre as entidades Elemento Metadados e Entidade com cardinalidade “muitos-para-muitos” e nesse relacionamento são definidos os atributos “relacionamento”, “atributo”, “subcampo” e posição do elemento de metadados.

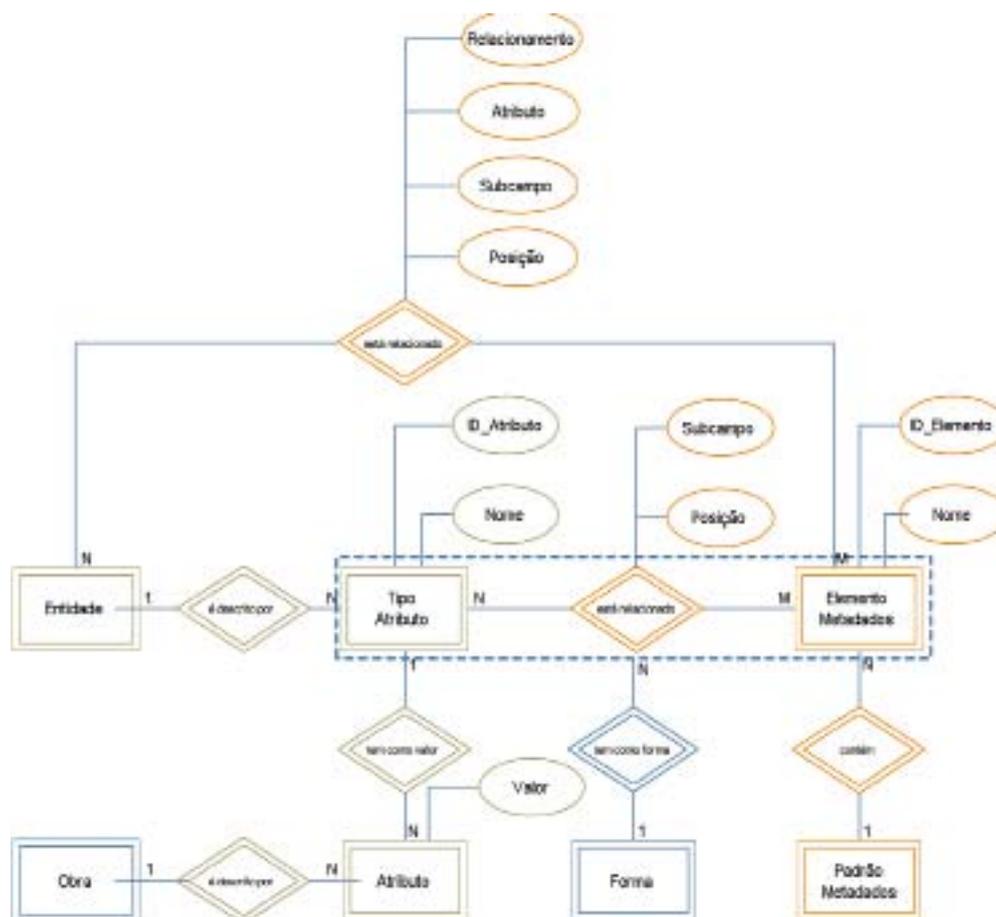


Figura 07 – Integração com padrões de metadados

Para os atributos dinâmicos, cada elemento de dados do padrão de metadados é relacionado com os tipos de atributos criados pelo catalogador para representar as entidades bibliográficas indicando, caso seja necessário, o subcampo e posição no registro bibliográfico.

De acordo com o Delsey (2006), a relação entre os atributos FRBR e o elementos de dados MARC se diferenciam, às vezes, pela forma do objeto, portanto, foi associada a entidade Forma ao relacionamento entre as entidades Tipo Atributo e Elemento Metadados.

Essa relação, vista na Fig. 07, entre os atributos dinâmicos e os elementos de metadados definidos pelo catalogador na implantação do catálogo deve ser feita também em todas as entidades, atributos e relacionamentos do modelo que tenham ligação com os elementos dos padrões de metadados.

Catálogos e sistemas computacionais baseados nesse modelo levam, na implantação do ambiente informacional, o catalogador a definir os atributos a serem utilizados para representar as entidades bibliográficas e a relacionar esses atributos com elementos de metadados dos padrões a serem utilizados.

A riqueza dos relacionamentos entre as entidades e dos atributos que representam essas entidades tem influência direta na capacidade que um ambiente informacional, baseado nesse modelo,



terá de possibilitar aos usuários do catálogo condições de navegar no universo bibliográfico a fim de conseguir atender suas expectativas informacionais.

Nesta pesquisa procurou-se representar as possibilidades de relacionamentos entre as entidades bibliográficas encontradas no relatório FRBR e a normalização dos atributos transformando-os como entidades associadas às entidades bibliográficas. A definição de atributos segue o nível básico de funcionalidades descrito nos FRBR para atender as user tasks.

A redução de um diagrama E-R em tabelas segue as seguintes regras:

- I) Toda entidade é convertida numa tabela ou relação, se E é um conjunto de entidades com atributos a_1, a_2, \dots, a_n , E será representada por uma tabela E com n colunas distintas, cada uma correspondendo a um atributo de E , cada linha da tabela corresponde a uma entidade (instância) do conjunto de entidades E .

Uma relação $r(R)$ é uma relação matemática de grau n nos domínios $\text{dom}(A_1), \text{dom}(A_2), \dots, \text{dom}(A_n)$, que é um subconjunto de um produto cartesiano dos domínios que definem R , sendo:

$\text{dom}(A_1)$: domínio do atributo a_1 ;

$\text{dom}(A_2)$: domínio do atributo a_2 ;

...

$\text{dom}(A_n)$: domínio do atributo a_n ;

Produto Cartesiano: $\text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n)$;

Relação R = subconjunto de $\text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n)$;

Grau de R = n (número de atributos do esquema da relação).

- II) Define-se um atributo identificador sintético para identificar unicamente (chave primária) as instâncias das tabelas.

Conceitualmente, entidades e relacionamentos são distintos, mas como no diferenciá-los num banco de dados? Para conjunto de entidades essa distinção é realizada por meio de uma superchave que é o conjunto de um ou mais atributos que permitem identificar uma entidade das demais.

Se K é uma superchave, qualquer conjunto de atributos que contenha K é uma superchave, no entanto, uma superchave pode conter atributos desnecessários. O que importa para o projeto são as superchaves menores possíveis, ou seja, aquelas em que nenhum subconjunto é superchave, nesse caso, são chamadas de chaves candidatas. Existindo mais de uma chave candidata na relação, o projetista escolhe dentre estas a chave que será utilizada como mecanismo principal para identificação de uma entidade como única, essa chave candidata escolhida é chamada de chave primária;



- III) Para os relacionamentos com cardinalidade “um-para-muitos”, coloca-se o atributo que forma a chave primária da relação “um” como atributo formando a chave estrangeira na relação “muitos”, criando, assim, um relacionamento lógico entre as relações.

Uma chave estrangeira é o mecanismo utilizado pelos bancos de dados para implementar relacionamentos entre tabelas e o conceito de integridade referencial. Dois conjuntos de atributo C e D compatíveis em que existe uma ordem entre os atributos de ambos os domínios tal que o primeiro atributo de C tenha o mesmo domínio do primeiro atributo de D , o mesmo valendo para os segundos atributos, assim por diante. Uma chave estrangeira, então, pode ser definida com um conjunto de atributos $D \subset R_j$ que não é chave em R_j , é compatível com outro conjunto de atributos $C \subset R_k$ que é a chave primária da relação R_k ;

- IV) Para os relacionamentos muitos-para-muitos, cria-se uma relação a partir desse relacionamento e os atributos que formam a chave primária de cada relação é levada à relação criada formando cada um deles uma chave estrangeira que relacionará a relação formada com as relações do relacionamento.

Se R é um conjunto de relacionamentos envolvendo os conjuntos de entidades E_1, E_2, \dots, E_n ; se (E_i) é o conjunto de atributos que forma a chave primária para o conjunto de entidades E_i ; se R possui atributos descritivos chamados $\{r_1, r_2, \dots, r_m\}$, R será representado por uma tabela com o seguinte conjunto de atributos:

$$\bigcup_{i=1}^n \text{chave primária } (E_i) \cup \{r_1, r_2, \dots, r_m\}$$

A partir desse modelo lógico, esquemas de bancos de dados (modelos físicos) podem ser construídos, gerando, assim, catálogos digitais baseados em bancos de dados relacionais com os requisitos dos FRBR.

CONCLUSÕES

A lacuna entre a estrutura conceitual dos FRBR e a implementação das representações lógicas e de persistência de um catálogo bibliográfico digital e os aspectos atuais da catalogação, que levam a análises mais complexas no desenvolvimento de estruturas de representação, enquanto geradora do Tratamento Descritivo da Informação que envolve a modelagem e a gestão dos recursos informacionais, foram os elementos motivadores e norteadores para o desenvolvimento desta pesquisa, a qual propõe a reflexão e a defesa da ampliação do conceito da catalogação e a extensão do papel do catalogador nessa ampliação do processo de construção de catálogos bibliográficos digitais.

No que concerne ao seu papel nos estudos atuais sobre o tema da representação no campo da Ciência da Informação e, da Catalogação em especial, para os propósitos desta pesquisa e a proposta



de um modelo de construção de projetos de catálogos bibliográficos, pode-se afirmar que os FRBR não devem ser considerados isoladamente. Toda e qualquer análise deve ser feita considerando a relação existente entre eles e as regras de catalogação, os formatos e padrões de metadados e os modelos conceituais de dados, assim como, a interoperabilidade nesses ambientes.

O Modelo E-R, utilizado pelos FRBR, permite que se realizem estudos na área de representação da informação e possibilita o aperfeiçoamento, do ponto de vista da persistência das informações, das estruturas dos esquemas dos registros, aumentando a riqueza das possibilidades de acesso aos documentos que satisfaçam as tarefas do usuário e aperfeiçoando, assim, a automação de catálogos em unidades de informação. Pode-se considerá-los como uma ferramenta importante para a modelagem de dados no sentido de desenvolver projetos de catálogos que conduzam às normas e padrões biblioteconômicos estabelecidos a satisfazer as necessidades atuais dos usuários no uso de catálogos digitais.

Em relação aos requisitos de representação de um registro bibliográfico, o relatório final do Grupo de Estudos FRBR da IFLA, mostra-se como um padrão para a construção de modelos referenciais de registros bibliográficos, mas sua estrutura de entidades, relacionamentos e cardinalidades de relacionamentos e de atributos não podem ser tomadas como exaustiva na modelagem de dados, pois eles são baseados no Modelo Entidade-Relacionamento, mas não contemplam todas as suas regras e normas na construção de projetos conceituais. Porém, a existência desse relatório mostra aos catalogadores um caminho para o desenvolvimento de projetos de catálogos e comprova que o conhecimento prévio do profissional especialista da área de Tratamento Descritivo e Temático da Informação utilizando um conjunto de metodologias de representação, torna o desenvolvimento de modelos de dados de catálogos mais aderente em relação às demandas informacionais dos usuários, trabalho este que anteriormente ficava a cargo de um profissional da área da Computação.

Para tanto, é necessário que o profissional catalogador estenda a sua atuação ampliando também o conceito de catalogação que passa da descrição para a representação de recursos informacionais sempre com vistas às expectativas e necessidades do usuário e a interoperabilidade de todo e qualquer recurso informacional e, conseqüentemente, agrega-se ao processo da catalogação a modelagem das estruturas bibliográficas no desenvolvimento de modelos conceituais que servirão de base para a construção de esquemas de banco de dados.

Constata-se que o Modelo Entidade-Relacionamento é uma metodologia que pode ser absorvida no processo do Tratamento Descritivo da Informação como ferramenta para melhorar a comunicação do catalogador com os profissionais da área da computação no desenvolvimento de sistemas informatizados de catálogos bibliográficos propondo para estes todo o modelo conceitual das estruturas de persistência dos registros bibliográficos.

Nesse contexto, regras e práticas de modelagem devem ser utilizadas pelo catalogador atuando como projetista do catálogo fazendo com que anomalias sejam retiradas do projeto e requisitos de qualidade da informação sejam homologadas no modelo.



No caso do modelo proposto vale ressaltar que o diagrama entidade-relacionamento resultante do processo de modelagem serve como base para a construção de projetos físicos (esquemas) de banco de dados de catálogos bibliográficos baseados nas estruturas dos FRBR, entretanto, a maior contribuição está no detalhamento do processo de construção do modelo demonstrando o uso dos requisitos presentes nos FRBR, comprovando que o catalogador de posse de seu conhecimento prévio de regras de catalogação tem condições de assumir a responsabilidade de projetar a estrutura conceitual de um catálogo bibliográfico.

A modelagem de dados possibilita uma visão não linear dos elementos descritivos de um item bibliográfico e permite uma eficiente transmissão das mensagens contidas nas representações e a otimização do acesso e uso de ambientes catalográficos, por meio da intersecção entre os Itens bibliográficos e as necessidades informacionais dos usuários. Nesse sentido, o processo de catalogação deve ser ampliado adicionando a ele a responsabilidade de definição das estruturas conceituais de representação dos Itens bibliográficos e, para tanto, o profissional catalogador necessita buscar essas ferramentas que possibilitam um melhor desenvolvimento da área da catalogação.

Essa ampliação do processo de catalogação permite a expansão do papel do catalogador que passa a desempenhar atividades que até então ficavam a cargo do profissional da Computação, como por exemplo, a modelagem conceitual do catálogo. Tendo em vista que o catalogador é o profissional especializado para identificar as necessidades em todos os níveis de usuários e que a estrutura do catálogo influencia o processo de busca da informação, esta pesquisa procurou, a partir da demonstração da construção de modelos conceituais de dados, como parte do processo de catalogação, baseada na perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais, ressaltar a importância do papel do catalogador no desenvolvimento de ambientes informacionais de catálogos bibliográficos digitais, bem como, apresentar um processo de catalogação mais eficaz, que vise facilitar o encontro, a identificação, a seleção e a obtenção de conteúdos informacionais.

Abstract: *The cataloguing process deals with bibliographic registers as data support, serving as a base to interoperability among informational environments and considering objects with different information and cooperative and homogeneous basis. Among the main proposes of the cataloguing area are the FRBR - Requirements for Bibliographic Records - which constitute a new concept in cataloguing rules. They present a manner of restructuring bibliographic registers so as to reflect upon the concept structure of persistence and information search, taking into account the diversity of users, material, physical support and formats. In this context, the present work is intended for reflecting and discussing, from a concept architecture of informational environments based on FRBR and Entity-Relationship Model, the cataloguing process in the context of catalogues project using computer methodology of Conceptual Data Modeling, thus considering the evolution of the field in the scope of Information Science in what concerns data representation towards use and interoperability of every and all information resource in order to fill the gap between concept project of an application domain and the definition of metadata schema of bibliographic registers structures. Data modeling offers a non-linear view of descriptive elements of a bibliographic item and allows an efficient spread of messages included in the representations, as well as the optimization of access and use of catalog environments by means of the intersection between bibliographic items and the*



information requirements of users. In this manner, the cataloguing process must be increased, adding to it the definition of concept structure of bibliographic items representation. In order to do so, the professional of the cataloguing area has to search for tools which allow a better development of the cataloguing activity.

Keywords: Cataloguing, FRBR, Data Modeling.

REFERÊNCIAS

BAPTISTA, D. M. A catalogação como atividade profissional especializada e objeto de ensino universitário. **Inf. Inf.**, Londrina, v. 11, n. 1, jan./jun. 2006.

BURT, P.; KINNUCAN, M. Information models and modelling techniques for information systems. **Annual Review of Information Science and Technology**, p.175-208, 1990.

CHEN, P. **The entity relationship model** – toward a unified view of data. **TODS**, 1,1 mar. 1976.

DELSEY, T. Modeling the logic of AACR. **International Conference on the Principles and Future Development of AACR**, Toronto, 23-25 Oct. 1997. *Proceedings...* Toronto: American Library Association: Library Association Publishing, 1998. p.1-16.

_____. **Functional analysis of the MARC21 bibliographic and holdings formats**. Washington: Library of Congress, 6 apr. 2006. Disponível em: < <http://www.loc.gov/marc/marc-functional-analysis/functional-analysis.html>>. Acesso em: 20 jun. 2010.

IFLA. FRBR Review Group. **Frequently Asked Questions about FRBR**. [version 9.0 latest, revision, oct. 2006]. Disponível em: <<http://www.ifla.org/VII/s13/wgfrbr/faq.htm>>. Acesso em: 02 nov. 2007.

LOURENÇO, C. de A. **Análise do padrão brasileiro de metadados de teses e dissertações segundo o modelo entidade-relacionamento**. Tese (Doutorado em Ciência da Informação). Universidade Federal de Minas Gerais. 2005.

MORENO, F. P. **Requisitos funcionais para registros bibliográficos - FRBR: um estudo no catálogo da rede bibliodata**. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2006.

SANTOS, P. L. V. A. C. Processos de representação descritiva e seu ensino nos cursos de graduação em biblioteconomia. Marília: **Cadernos da FFC**, 1995. v. 4, n. 1, p. 36-39.



INFORMAÇÃO E SOCIEDADE: POLÍTICAS DE PRESERVAÇÃO E DISPONIBILIZAÇÃO DE DADOS SOBRE A SEGURANÇA PÚBLICA EM MOÇAMBIQUE

*Januário Albino Nhacuongue, Edberto Ferneda,
Plácida Leopoldina Ventura Amorim da Costa Santos*

Resumo: o presente trabalho é síntese da dissertação do mestrado em Ciência da Informação, na qual se baseou em uma abordagem qualitativa para obter a descrição detalhada de fenômenos e comportamentos, através de dados com maior riqueza de detalhes e profundidade, colhidos por meio de vários métodos, em um processo do estudo de caso da Polícia da República de Moçambique (PRM), no que concerne a preservação e disponibilização de informações sobre a segurança pública para a sociedade. Neste contexto, objetivou-se analisar a criação do modelo de banco de dados para a gestão de informações em prol da segurança pública a partir dos conceitos da Ciência da Informação e descrever os elementos da arquitetura da informação para o ambiente da segurança pública. Basicamente, a fundamentação teórica esteve inculcada tanto nas percepções sobre os processos de criação de banco de dados orientados por Machado, Oliveira, Silberschatz e Sudarshan, como nos processos e metodologias da arquitetura da informação para ambientes digitais de Morville e Rosenfeld. Também se atentou à análise das políticas públicas de Meny e Thoenig e à Constituição e legislação que regula a segurança pública em Moçambique. O estudo culminou com a criação de um modelo de banco de dados que poderá ser usado pela organização para a preservação de dados produzidos no seu funcionamento e do ambiente informacional digital, por meio do qual se pode disponibilizar informações e prestar serviços de atendimento à sociedade, permitindo a sua participação direta na gestão e reforma de políticas sobre a segurança pública no país.

Palavras-chave: Banco de dados. Ambiente informacional digital. Segurança pública.

Abstract: *the present work is the synthesis of the thesis master's degree in Information Science, which was based on a qualitative approach to obtain a detailed description of phenomena and behavior, using data with greater detail and depth, collected through various methods, process in a case study of the Police of the Republic of Mozambique (PRM), concerning the preservation and availability of information about public safety to society. In this context, the objective was to analyze the creation of the model database to manage information in support of public safety from the concepts of information science and describe the elements of information architecture for the environment of public safety. Basically, the theoretical foundation was so ingrained perceptions about the process of creating database-driven Machado, Oliveira, Silberschatz and Sudarshan, and in processes and methodologies of information architecture for digital environments of Morville and Rosenfeld. It also looked to the analysis of public policies of Meny and Thoenig and the Constitution and laws governing public security in Mozambique. The study culminated in the creation of database model that can be used by the organization for the preservation of data produced in their operation and the digital information environment, through which it can provide information and care services to*



society, allowing their direct participation in management and policy reform on public security in the country.

Keywords: Database. Digital information environment. Public safety.

INTRODUÇÃO

As tecnologias de informação e comunicação que desde a origem da humanidade acompanharam o seu processo evolutivo, com a *Internet* ganham uma propulsão quantitativa e qualitativa de comunicação e organização entre diferentes atores sociais. Neste sentido, os órgãos administrativos são cada vez mais inseridos em uma das características do pós-modernismo sobre a velocidade ou instantaneidade na troca de informações e prestação de serviços públicos. O tema “informação e sociedade: políticas de preservação e disponibilização de dados sobre a segurança pública em Moçambique” se insere neste cenário, desdobrando-se como síntese da dissertação do mestrado em Ciência da Informação, concluída nos primórdios do ano de 2011. A partir da abordagem qualitativa, interpretam-se e compreendem-se os diferentes fenômenos, dados e ações relacionados à segurança pública, visando notabilizar o processo de preservação e, por conseguinte, disponibilização e recuperação no âmbito da informação e construção do conhecimento sobre a matéria, em Moçambique.

O estudo partiu de uma reflexão crítica sobre o defasamento da atual política de defesa e segurança em uso no país que culmina com a desvalorização de diversos aspectos que a atual era da sociedade da informação impera sobre qualquer ator social, mormente o foco sobre o digital e a ubiqüidade da informação, no que tange à satisfação das necessidades informacionais do usuário. Os serviços de informação para a sociedade devem pautar pela legitimidade inerente a cada estágio vivido por essa sociedade, para tal, os estudos de usuários que segundo Figueiredo (1994, p.7), são “investigações que se fazem para saber o que os indivíduos precisam em matéria de informação, ou então, para saber se as necessidades de informação estão sendo satisfeitas de maneira adequada”, são imprescindíveis para o processo de construção do conhecimento pela sociedade, baseada na reflexão crítica sobre os conteúdos apreendidos. Por isso na atualidade, disponibilizar a informação para a sociedade implica levar a cabo um conjunto de medidas que potencializem a democratização do acesso e o uso estratégico da informação, com recurso às novas tecnologias de informação e comunicação, em prol da massificação situacional aliada à vivência cotidiana na comunidade de inserção.

Um dos principais desafios para o avanço da democracia é o desenvolvimento contínuo das ferramentas de interação entre governo e sociedade (MAZONI, 2011). Neste sentido, ao lado dos governos eletrônicos surge o conceito de dados abertos governamentais numa injunção para a disponibilidade de dados para a sociedade, com a observância dos oito princípios prescritos pelo *OPENGOVDATA*, a saber: serem completos, primários, atuais, acessíveis, compreensíveis por máquinas, não discriminatórios, não proprietários e livres de licenças ou qualquer tipo de controle.

Para colmatar a situação descrita, o trabalho notabilizou a criação de um modelo de banco de dados para a preservação das informações produzidas no âmbito da segurança pública e o desenho de



um ambiente informacional digital, através do qual se possa recuperar os conteúdos armazenados para a sociedade. Deste modo, o estudo foi delimitado para a Polícia da República de Moçambique (PRM), um organismo público que entre outras funções, garante a ordem, segurança e tranqüilidades públicas, o respeito pelo Estado de Direito e a observância estrita dos direitos e liberdades fundamentais dos cidadãos¹. De realçar que no contexto geral, a defesa da ordem e segurança também é incumbida à sociedade em geral e a outros órgãos responsáveis pela aplicação da lei, como defesa do consumidor, saúde pública, defesa civil, gestão de calamidades, entre outros².

POLÍTICAS PÚBLICAS

As políticas públicas são ações governamentais que visam coordenar os meios disponíveis e as atividades dos entes privados para a satisfação do bem público. São determinados pelo sistema de governo, mas coordenadas com o envolvimento de diversos atores sociais, como públicos (políticos e burocratas), privados (individuais, coletivos e sindicatos), internacionais (Organizações não Governamentais) e sociedade civil (igrejas, mídia e líderes carismáticos), em observância dos princípios de legalidade, transparência e eficiência. A sua definição se sujeita a posteriores programas, planos e metas para o seu cumprimento.

Meny e Thoening (1992), consideram que o ciclo de uma política compreende as fases de identificação do problema, de formulação de soluções alternativas, de tomada de decisão, de implementação da decisão e de avaliação dos resultados. No caso concreto de Moçambique, a atual política de defesa e segurança carece da avaliação que possa conduzir à reforma ou ajuste à atual conjuntura política, econômica e social, pois se mostra defasada na medida em que foi aprovada pela Lei nº 17/97, de 1 de Outubro, para responder o momento que se vivia. Muitos aspectos demandados pelo advento da *Web* tanto no cenário do “*modus operandi*” criminal, como da resposta à sociedade se mostram descobertos pela referida política e por outros programas do governo.

Como se referenciou anteriormente, a garantia da segurança pública na atual era digital passa pela adoção de serviços de informação ou relações públicas que permeiam e encorajam a participação do cidadão. Ademais, as relações públicas segundo Bravo (1998), constituem parte da ciência social, na sua essência, um diálogo e uma arte, visando permitir a troca de opiniões, de sentimentos e de desejos entre pessoas situadas em níveis diferentes, sanando conseqüências negativas de incompreensão, ódio e desunião. Daí a importância de criação de infra-estruturas tecnológicas para o intercambio de dados, incentivando a participação da sociedade em todos os contextos sobre a segurança pública que podem alicerçar a tomada de decisões. Uma nova política ou programa sobre dados abertos governamentais só se torna efetivo quando haja um banco de dados para a preservação de informações, com normas explícitas para o intercâmbio de dados, dicionário de semântica ou vocabulário controlado de termos da área e mecanismos de visualização e processamento de dados em ferramentas de livre acesso para

1 Artigo 2º da Lei n.º 19/92, de 31 de Dezembro, conjugado com o artigo 254º da Constituição.

2 Artigo 12º da Lei n.º 17/97, de 1 de Outubro.



a sociedade. A participação da sociedade neste sentido garante a intervenção e enriquecimento das políticas do governo, permitindo o crescimento exponencial da área da segurança pública no país.

BANCO DE DADOS

O banco de dados segundo Oliveira (2004, p. 57), é “uma coleção organizada de dados e informações que pode atender às necessidades de muitos sistemas, com um mínimo de duplicação, e que estabelece relações naturais entre dados e informações”. Através do banco de dados, armazenam-se dados, criando um sentido inicial de informações que poderão ser utilizadas na sua natureza ou processadas pelos sistemas de informação em outras informações para satisfazer várias necessidades da organização. O banco de dados é uma extensão do *memex*, a memória auxiliar que segundo Vannevar Bush operaria por associação, visando à recuperação da informação, pois o *memex* é um dispositivo no qual qualquer indivíduo armazena todos os seus livros,

registros e comunicações, e que é mecanizada, para que possa ser consultado com extrema velocidade e flexibilidade. É um complemento ampliado íntimo de sua memória (BUSH, 1945, p.6, tradução nossa).

Contudo, conforme o autor chama atenção, para o aspecto da construção do conhecimento vale o pensamento criativo da sociedade e não o pensamento repetitivo da máquina, por isso, a Ciência da Informação desempenha um papel preponderante na criação de estruturas lógicas que auxiliam a sociedade na recuperação dos dados preservados, através de relacionamentos, cabeçalhos de assunto, listas controladas, listas de anéis de sinônimos, taxonomias, esquemas de classificação alfanuméricos, thesaurus, ontologias e folksonomias. Aliás, este papel interventivo da Ciência da Informação é enfatizado por Saracevic (1992) ao destacar o papel diferencial da área no cenário das relações de interdisciplinaridade, pela preocupação com aspectos sociais e humanos que transcendem a tecnologia.

Além da estrutura de armazenamento em si, um Sistema Gerenciador de Banco de Dados (SGBD), como conjunto formado por dados inter-relacionados e *softwares*, é usado na definição de estruturas para o armazenamento, mecanismos para a manipulação e segurança de dados (SILBERSCHATZ et al., 2006).

Um sistema gerenciador de banco de dados é caracterizado por compacidade, rapidez, integração e redundância, compartilhamento, segurança e padronização de dados, isto é, garante a disponibilidade de dados não repetitivos, seguros, que podem ser acessados rapidamente pelos usuários para as suas necessidades.

Para a criação do modelo de banco de dados para a preservação de informações sobre a segurança pública em Moçambique no que tange à criminalidade, o estudo baseou-se no funcionamento atual da Polícia da República de Moçambique, ilustrado pelo organograma da figura 1.

O Comando-Geral, conforme o organograma da figura 1, é composto pelo próprio Comando, pelos Conselhos da PRM e pelo Conselho de Ética e Disciplina; pelas Direções de Ordem e Segurança



Pública, de Investigação criminal, de Pessoal e Formação e de Logística e Finanças, bem como pelo Comando das Forças Especiais e de Reserva, pelos Departamentos de Estudos, Informação e Plano; de Informação Interna e de Relações Públicas, incluindo o Gabinete do Comandante e a Secretaria-Geral.

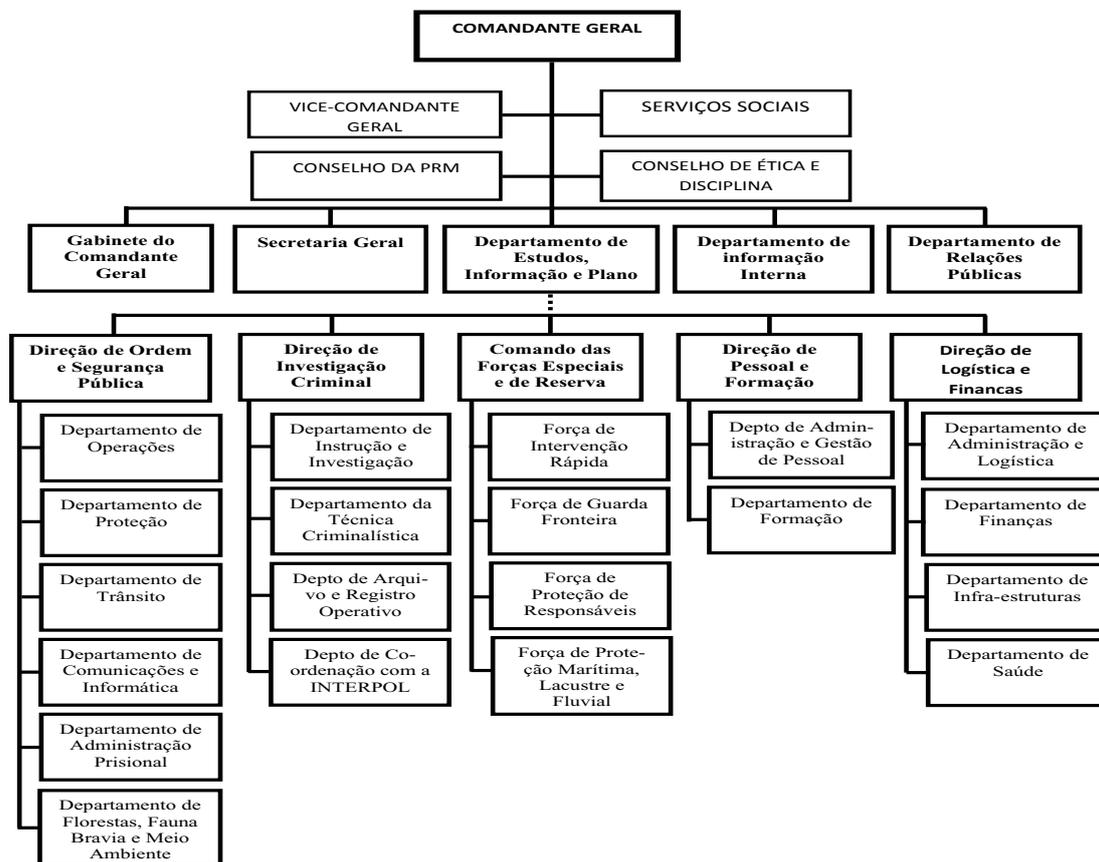


Figura 1: Organograma do Comando Geral da PRM.

Fonte: Decreto nº 27/99, de 24 de Maio.

A Direção de Ordem e Segurança Pública (DOSP), conforme o art. 13º do Decreto nº 27/99, de 24 de Maio, é responsável pela prevenção e combate ao crime, pelo funcionamento normal das instituições e circulação de pessoas e bens, pela patrulha, pela recepção de queixas e denúncias, entre outras ações. Na prossecução destes objetivos, coordena as atividades de diversos Departamentos, incluindo o de Trânsito que zela pela segurança rodoviária. A Direção de Investigação Criminal (DIC), conforme o art. 14º do mesmo Decreto tem a missão de investigar crimes e instruir os respectivos processos, sob a direção do Ministério Público e o Comando das Forças Especiais e de Reserva, por sua vez, tem a missão de intervir em situações especiais de manutenção da ordem pública.

Embora não estejam incluídos no organograma, os Comandos Provinciais constituem unidades da PRM com uma estrutura semelhante a do Comando Geral, representando o Comando Geral em cada Província. Os Comandos Distritais, por sua vez, representam a base da hierarquia do comando. Compõem-se de Secções que representam as Direções em outros níveis, de esquadras/delegacias, dos postos policiais e dos destacamentos das Forças Especiais e de Reserva.

MODELAGEM DE DADOS

Um modelo de dados é um conjunto de conceitos usados na descrição de estruturas lógicas e físicas do banco de dados, ou seja, define abstrações que facilitam ao usuário a compreensão e organização de todos os objetos de dados existentes em um contexto e necessários para a aplicação desse banco de dados. Silberschatz et al. (2006, p.5) consideram o modelo de dados como “coleção de ferramentas conceituais para descrever dados, relações de dados, semântica de dados e restrições de consistência” e acrescenta que existem quatro categorias de modelos de dados, a saber: modelo relacional, modelo de entidade – relacionamento, modelo de dados baseado em objeto e o modelo de dados semi-estruturados. No caso do estudo, baseou-se no modelo de entidade – relacionamento (MER) ilustrado no diagrama da figura 2, produzido no *software SmartDraw*.

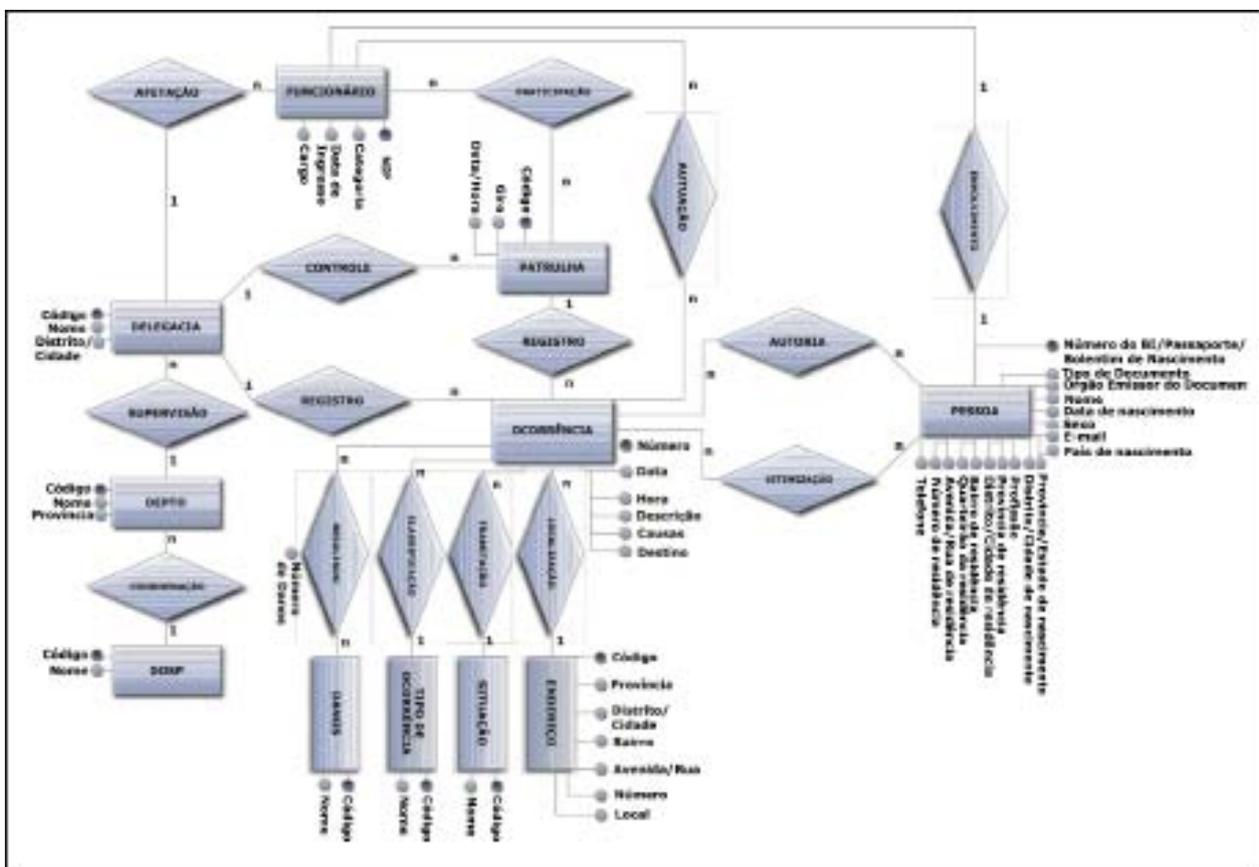


Figura 2: Diagrama entidade – relacionamento (DER).

Fonte: Elaborado pelo autor com base no contexto do funcionamento atual da PRM

O diagrama contém dez entidades relacionadas entre si, com os respectivos atributos necessários para a especificidade de informações a serem usadas pelos sistemas de informação ou para o auxílio à tomada de decisões. Os atributos de cada entidade são representados através de bolinhas, sendo as de maior tonalidade da cor referentes à chave primária de cada entidade, ou seja, o atributo único que não pode ser repetido nas tabelas de modo a garantir a integridade dos dados.

A entidade DOSP representa a Direção de Ordem e Segurança Pública no nível central, Provincial e Distrital e tem uma relação de coordenação 1 para n com os Departamentos (DEPTO), isto é, cada DOSP coordena vários Departamentos; enquanto que cada Departamento é coordenada apenas por única Direção. Na entidade DEPTO se encontram representados todos os Departamentos (Proteção, Trânsito e Proteção de Florestas e Fauna Bravia) e Destacamentos das Forças Especiais (Intervenção Rápida, Guarda Fronteira, Proteção de Responsáveis e Proteção Marítima, Lacustre e Fluvial) que intervêm diretamente nas ocorrências criminais. Igualmente, encontra-se representado o Departamento de Incêndios que também registra várias ocorrências no âmbito das suas competências.

Cada departamento, por sua vez, coordena as delegacias/esquadras da sua área de jurisdição, por isso o relacionamento também é 1 para n. Cada delegacia faz o registro das ocorrências e controla as patrulhas escaladas diariamente. As patrulhas e as delegacias são compostas por funcionários que registram as ocorrências, classificadas de acordo com o tipo de cada fato. As ocorrências produzem diversos tipos de danos humanos e materiais e a sua tramitação é controlada no âmbito da resposta policial. Qualquer pessoa pode ser vítima ou autor de alguma ocorrência, incluindo os próprios funcionários da organização que intervêm nas mesmas.

O Diagrama de Estrutura de Dados (DED) ilustrado na figura 3 mostra a articulação dos dados no banco, direcionada através de setas e cada entidade corresponde a uma tabela no nível físico do banco de dados. Os relacionamentos n para n implicam a transformação do relacionamento em uma entidade que igualmente será representada por uma tabela. Neste sentido, os relacionamentos afetação e participação, no modelo físico, são tabelas com chaves estrangeiras referentes às duas entidades com as quais se relacionam, respectivamente, totalizando dezesseis tabelas do banco de dados.

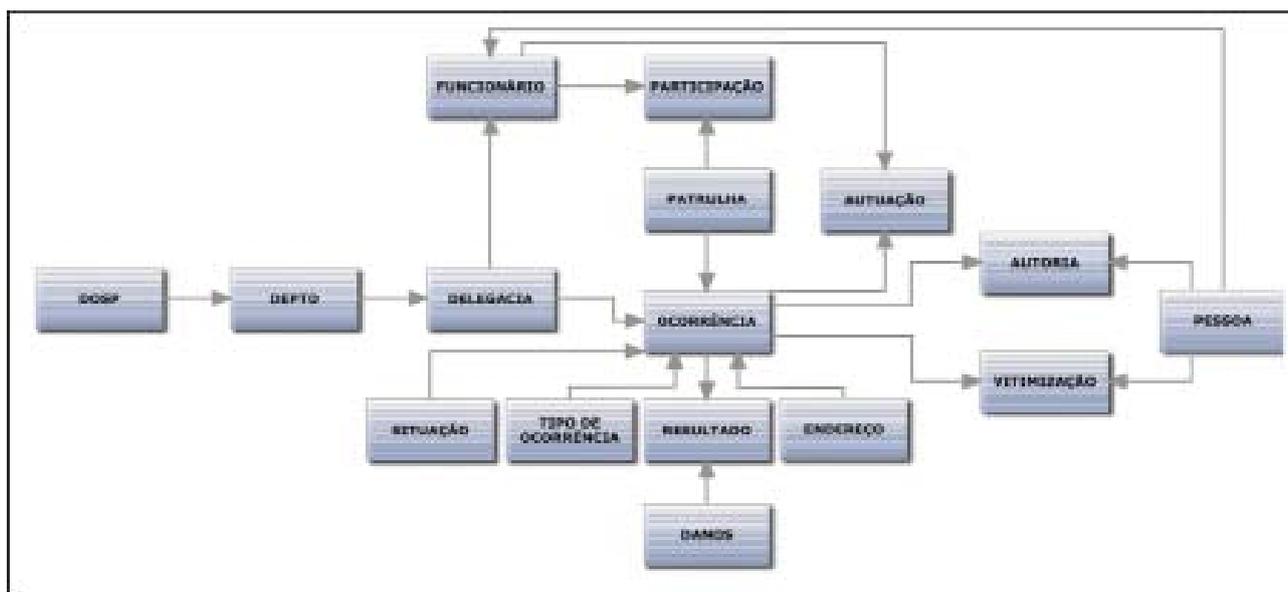


Figura 3: Diagrama de estrutura de dados (DED).

Fonte: Elaborado pelo autor com base no modelo entidade - relacionamento.



LINGUAGEM SQL

A *Structured Query Language (SQL)* ou linguagem estruturada de pesquisa teve sua gênese no início dos anos 70 nos laboratórios da IBM por forma a demonstrar as vantagens do modelo relacional de Edgar Frank Codd, publicado através do artigo *Relational Model of Data for Large Shared Data Banks* – modelo relacional de dados para grandes bancos de dados. O seu melhoramento na consulta e manipulação de dados levou o *American National Standard Institute (ANSI)* a torná-la padrão de linguagem em ambiente relacional, através do padrão SQL-86, seguida da SQL-99 e da versão mais recente, a SQL-2003 (MACHADO, 2004, p.315).

A linguagem SQL é usada para a definição, manipulação, controle, transação e consulta de dados.

A linguagem de definição de dados, segundo Silberschatz et al. (2006), garante a definição da estrutura e organização dos dados armazenados e suas relações, ou seja, é formada por um conjunto de comandos que permitem a definição e modificação dos esquemas das tabelas. A cláusula *not null* significa que o campo não deve ser nulo, em detrimento de ser a chave primária ou do valor das informações que irá conter. Deste modo, para a aplicação no sistema gerenciador de banco de dados *Interbase 6. Open Edition – 6.0.20*, usou-se a seguinte linguagem para criar as tabelas DOSP e DEPTO.

```
CREATE TABLE DOSP  
(cod_DOSP char (10) not null primary key,  
Nome varchar (80) not null)
```

```
CREATE TABLE DEPTO  
(cod_DEPTO char (12) not null,  
Nome varchar (60) not null,  
Provincia varchar (30) not null,  
Cod_DOSP char (10) not null REFERENCES DOSP (cod_DOSP) ON UPDATE CASCADE,  
Primary key (cod_DEPTO))
```

O mesmo processo foi observado em relação às restantes tabelas do banco de dados. A linguagem de manipulação de dados é um subconjunto de comandos para inserir, remover e modificar informações das tabelas. No caso do estudo, foram simuladas algumas informações, consubstanciando uma amostra da metodologia da operatividade policial. Por exemplo, para a manipulação das tabelas DOSP e DEPTO usaram-se as seguintes linguagens:

```
Insert into DOSP  
VALUES ('DOSP200', 'Comando da Cidade de Maputo');  
Insert into DOSP  
VALUES ('DOSP300', 'Comando da Província de Maputo');  
Insert into DOSP  
VALUES ('DOSP400', 'Comando da Província de Gaza');  
Insert into DOSP  
VALUES ('FER500', 'Comando da Província de Inhambane');
```



```
VALUES ('FER200', 'Comando das Forças Especiais e de Reserva da Cidade de Maputo');
Insert into DOSP
VALUES ('FER1200', 'Comando das Forças Especiais e de Reserva da Província de Niassa');
Insert into DOSP
VALUES ('SPB900', 'Serviço Provincial de Bombeiros de Tete')

Insert into DEPTO
VALUES ('DPP200', 'Depto de Proteção', 'Maputo - Cidade', 'DOSP200');
Insert into DEPTO
VALUES ('DPP300', 'Depto de Proteção', 'Maputo - Província', 'DOSP300');
Insert into DEPTO
VALUES ('DPT400', 'Depto de Trânsito', 'Gaza', 'DOSP400');
Insert into DEPTO
VALUES ('FIR200', 'Força de Intervenção Rápida', 'Maputo - Cidade', 'FER200');
Insert into DEPTO
VALUES ('FPMLF500', 'Força de P. Marítima, Lacustre e Fluvial', 'Inhambane', 'FER500');
Insert into DEPTO
VALUES ('FGF1200', 'Força de Guarda Fronteira', 'Niassa', 'FER1200');
Insert into DEPTO
VALUES ('DI900', 'Depto de Incêndios', 'Tete', 'SPB900')
```

Criadas as tabelas e inseridos os valores dos registros, o modelo do banco de dados apresenta-se conforme ilustra a figura 4:

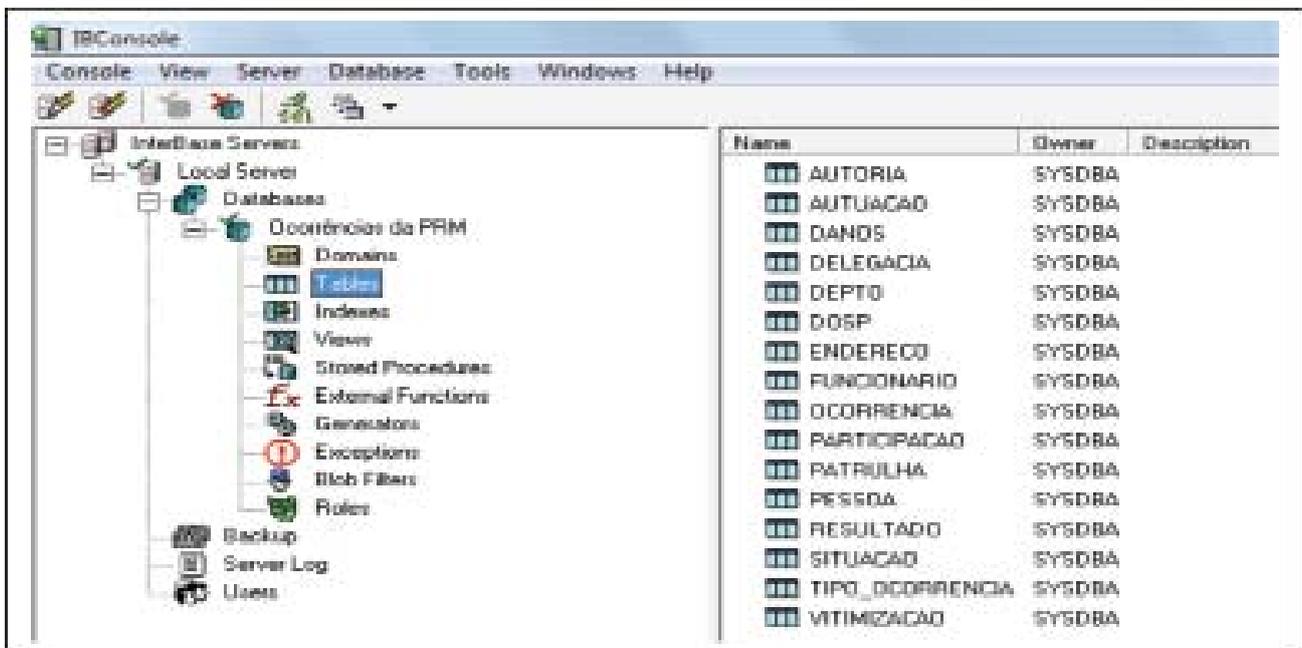


Figura 4: Modelo do banco de dados da PRM.

Fonte: Elaborado pelo autor com base no sistema Interbase 6. Open Edition – 6.0.20.



ARQUITETURA DA INFORMAÇÃO PARA AMBIENTES DIGITAIS

O estudo também visava o desenho de um ambiente informacional digital, através do qual a PRM pode disponibilizar as informações, tanto para os próprios funcionários, como para a sociedade em geral, de modo a contribuir com as suas reflexões críticas no processo de revitalização da área da segurança pública no país.

A expressão “arquitetura da informação”, segundo Lima-Marques e Macedo (2006), foi cunhada pela primeira vez em 1976 pelo arquiteto Richard Saul Wurman numa conferência organizada pelo Instituto Americano de Arquitetos, motivado pela necessidade de reunir, organizar e apresentar a quantidade e variedade de informações para maior e melhor acesso ao usuário. A arquitetura da informação é:

a combinação dos sistemas de organização, rotulagem, busca e navegação em *websites* e *intranets*; o desenho estrutural de ambientes informacionais compartilhados; a arte e ciência de estruturar produtos de informação e experiências que permitam usabilidade, além da facilidade e relevância para encontrar as informações (ROSENFELD; MORVILLE, 2006, p.4, tradução nossa).

A arquitetura da informação garante o *findability* ou a encontrabilidade que segundo Morville (2005, tradução nossa), se resume na qualidade de um objeto ser facilmente localizável ou um sistema ser navegável. Equivale dizer que não basta que as informações necessárias estejam armazenadas no banco de dados se o usuário não for capaz de localizá-las com facilidade e recuperá-las para a sua necessidade específica. Ademais, de acordo com Stewart (2008), o *findability* pode ter um impacto negativo na tomada de decisão, pois decisões críticas podem ser adiadas porque a informação encontrada, se houver, está incompleta ou conflitante ou pior ainda, más decisões podem ser decretadas quando nem teriam sido consideradas se a informação necessária fosse encontrada.

Os processos e metodologias de uma arquitetura da informação de acordo com Rosenfeld e Morville (2006, tradução nossa), seguem um programa que implica, em uma primeira instância, o desenho do projeto, através das fases da pesquisa, estratégia, *design* e implementação e, posteriormente, a administração.

ELEMENTOS DA ARQUITETURA DA INFORMAÇÃO

Segundo Rosenfeld e Morville (2006, tradução nossa), a arquitetura da informação é composta pelos sistemas de organização, rotulagem, navegação e busca. Os sistemas de organização são baseados em esquemas de organização da informação (alfabética, cronológica ou geográfica) para definir os tipos de relacionamento entre itens de conteúdos e grupos. Para garantir a acessibilidade, o conteúdo pode ser agrupado por tópicos, tarefas, público ou mesmo por nuvens de *tags* gerados pelo usuário durante a navegação. Também incluem estruturas de organização da informação hierárquica,



em hipertexto ou em *links*. Os sistemas de rotulagem são baseados no uso de etiquetas ou rótulos em forma de texto ou de ícones para transmitir um significado. As etiquetas auxiliam o entendimento do usuário sobre os esquemas de organização do ambiente e, por conseguinte, auxiliam-no na navegação, por isso, devem ser claros para serem facilmente compreendidos pelos mesmos. Os sistemas de navegação são baseados em elementos que fornecem o contexto e flexibilidade ao usuário para compreender a sua localização no ambiente e projetar o seu trajeto (*banner*, *menu*, conteúdo, mapa do *site*, índice e guias). Finalmente, os sistemas de busca permitem a localização e acesso aos conteúdos informacionais armazenados no ambiente.

Um sistema de busca é constituído pela interface da busca, através da qual o usuário formula a consulta; pela linguagem da consulta que inclui operadores booleanos *and*, *or* e *not* e operadores de proximidade, como ao lado e próximo; pelos otimizadores da consulta, como corretores ortográficos e dicionário de sinônimos; pelos recuperadores de algoritmos; pelo motor de busca propriamente dito; pelo conteúdo, normalmente armazenado em um banco de dados e pela resposta (ROSENFELD; MORVILLE, 2006, tradução nossa).

De salientar que devido ao crescimento de ambientes informacionais digitais, aliado à ambigüidade da linguagem que norteia os sistemas de classificação e do termo relevância no julgamento dos documentos recuperados, a Ciência da Informação vai proporcionar uma contribuição significativa, não só através de elementos invisíveis da arquitetura, como vocabulários controlados, tesouros e metadados, como também pela exploração de aspectos que norteiam o comportamento de busca dos usuários, para o aprimoramento da consulta através de termos variantes ou sinônimos. O processamento em linguagem natural enquanto conjunto de técnicas para a análise dos documentos em níveis lingüísticos vai permitir a normalização de variações lingüísticas e garantir a indexação automática de documentos produzidos pela organização.

A exploração das técnicas da representação da informação que se enquadram no núcleo epistemológico da Ciência da Informação enquanto mediação, propiciam uma riqueza de detalhe de informações descritivas sobre o objeto e o contexto, garantindo o *findability* dos recursos necessários à satisfação das necessidades de informação da sociedade.

No caso vertente deste estudo, os processos e metodologias não foram observados devido à exigüidade de fundos para o deslocamento e pesquisa na organização em estudo. O desenho do projeto do *website* baseou-se na experiência do autor, adquirida durante o vínculo laboral enleado com a instituição em estudo. Contudo, tratando-se de um modelo, a sua implementação será antecedida pelos processos de divulgação, discussão e enriquecimento do projeto, nos quais será evidenciada a metodologia proposta pelos autores citados.

O planejamento do modelo iniciou com o esboço do *layout* no *software Adobe fireworks CS3*, observando os elementos da arquitetura referenciados e a dimensão de 955x600 *pixels* para que a página fosse adequada aos monitores atuais com a resolução de tela de 1024 x 768 *pixels*. Também



se usou a *Hypertext Markup Language (HTML)* ou linguagem de marcação de texto, através de uso de *tags* para indicar as partes da página e produzir diferentes efeitos no navegador, como tabelas, linguagens de extensão (dia da semana, dia do mês, *login*, etc.), permitindo a disponibilização de conteúdo mais internacionalizado e acessível. Para garantir a interpretação pela semântica da página por qualquer dispositivo, como celular, usou-se a *eXtensible Hypertext Markup Language (XHTML)*, na versão 1.0 *transitional*, de modo a permitir o uso de etiquetas antigas. O desenvolvimento, a montagem, a diagramação, a manutenção, gerenciamento e publicação, foram feitos no *software Adobe Dreamweaver CS3*.

O modelo é composto por dezenove páginas (*home*, mapa do sítio, perguntas mais frequentes, quem somos, onde estamos, documentação, esquadra eletrônica, galeria de mídia, a PRM aconselha, a PRM comunica, formulários, *links* importantes, fale conosco, polícia de proteção, polícia de trânsito, polícia de investigação criminal, força de intervenção rápida, força de guarda fronteira e bombeiros). Todas as páginas contêm a mesma estrutura do cabeçalho, *menus* e rodapé, variando apenas o conteúdo, por isso para a sua arquitetura optou-se pelo uso de *template* (modelo).

DELEGACIAS VIRTUAIS COMO SERVIÇOS CÉLERES PARA A SOCIEDADE

As delegacias virtuais são serviços disponibilizados através da *Internet* que muitos países vêm adotando com vista a promover uma maior interação com o cidadão e flexibilizar respostas às diversas necessidades em prol da segurança pública. Neste sentido e de acordo com os *Websites* das delegacias virtuais do Estado do Rio de Janeiro e da Secretaria da Segurança Pública do Estado de São Paulo, através das mesmas, o cidadão pode apresentar queixas, informes, denúncias ou sugestões, registrar alguns tipos de ocorrências e comunicar o extravio de documentos, assim como fazer consultas sobre documentos recuperados, a partir de qualquer ponto de acesso à rede de *Internet*. Elas surgem como melhores alternativas à segurança pública, não só pela modernização das corporações policiais, como também pela redução da burocracia, informação à sociedade e aumento da resposta policial, requisitos fundamentais da reforma do setor público exigidos pelas políticas atuais da Administração Pública.

Uma das linhas de atuação das delegacias eletrônicas que o modelo do ambiente informacional ilustrado na figura 5 se propõe é a divulgação de informações para a autodefesa do cidadão, de endereços, telefones e horários de funcionamento de serviços de atendimento ao cidadão, de procedimentos em casos de delitos ou outro tipo de incidentes, bem como a apresentação de propostas de legislação para a punição de criminosos nos casos em que haja lacunas na lei, portanto, uma contribuição significativa na reformulação das políticas públicas.

Além das diversas informações que poderão ser obtidos do banco de dados, através do ambiente informacional, uma contribuição significativa do modelo para a sociedade consiste na criação do boletim eletrônico, por meio do qual o cidadão poderá fornecer os dados pessoais, dados e detalhes

da ocorrência, simplificando a burocracia e o tempo de espera nas delegacias físicas, para efeitos de queixas, denúncias e participações.



CONCLUSÕES

A criação do banco de dados e a arquitetura de um ambiente informacional digital vão impulsionar a sistematização e incremento do conhecimento na área da segurança pública em Moçambique, tanto pela preservação das informações produzidas na organização em diferentes coleções de documentos digitais em vários formatos, mídia e conteúdo, como pela viabilização da disponibilidade, localização e acesso às mesmas, através dos sistemas de organização, rotulagem, navegação e busca. Os esquemas e estruturas de organização que definem as características comuns das informações para o seu agrupamento lógico e os tipos de relações entre essas informações e seus grupos, respectivamente, vão desempenhar um papel importante na indexação e etiquetagem para melhor recuperação da informação do banco de dados. Ademais, o ambiente informacional digital, além de simplificar a burocracia nas delegacias mediante o provimento de serviços *on-line*, também vai garantir a interação contínua e reiterada entre a polícia e o cidadão, transmitindo confiança na ação policial. Indubitavelmente, o ambiente servirá de local aberto e apropriado para o exercício da democracia, de direito e de liberdade de expressão para a sociedade.

A Ciência da Informação vai reduzir a ambigüidade que caracteriza qualquer sistema de representação e recuperação da informação, agregando ao projeto uma possível ontologia aplicada à área de segurança pública e explorando as diversas dimensões que corporizam a representação do problema do usuário, de modo a garantir a eficiente e efetiva apropriação da informação pela sociedade.

Os sistemas de informação que permitem o processamento e análise do grande volume de dados armazenados por diferentes níveis de detalhes e objetivos organizacionais, por um lado, vão propiciar



a tomada de decisões céleres e, por outro, extrair comportamentos, associações e previsões para o desenvolvimento de algoritmos visando à produção do conhecimento na PRM e, por conseguinte, fundamentar o processo de reforma das políticas públicas em matéria de segurança. Neste contexto, relatórios específicos sobre ocorrências policiais por delito, tempo e espaço poderão ser acessados pela sociedade e por outras áreas de conhecimento para fundamentar pesquisas acadêmicas. Do mesmo modo, documentos e legislação sobre várias matérias policiais podem ser baixados para uso da sociedade.

REFERÊNCIAS

- BRAVO, O. **Relações públicas**: uma ciência, “um diálogo”, uma arte. Porto: Porto Editora, 1998.
- BUSH, V. As we may think. **Atlantic Monthly**, v.176, 1, 1945. Disponível em:
<<http://www.theatlantic.com/magazine/archive/1969/12/as-we-may-think/3881/1/>> Acesso em: 03 mar. 2010.
- DELEGACIA VIRTUAL DO ESTADO DO RIO DE JANEIRO. Disponível em:
<<http://www.delegaciavirtual.rj.gov.br/home.htm>> Acesso em: 22 fev. 2010.
- FIGUEIREDO, N. M. de. **Estudos de uso e usuários da informação**, Brasília: IBICIT, 1994.
- LIMA-MARQUES, M.; MACEDO, F. L. O. de. Arquitetura da informação: base para a gestão do conhecimento. **In**: TARAPANOFF, K. (Org.) Inteligência, informação e conhecimento em corporações. Brasília: IBICT, UNESCO, 2006. p.241-255.
- MACHADO, F. N. R. **Banco de Dados: projeto e implementação**. São Paulo: Érica, 2004.
- MAZONI, M. V. F. Apresentação. **In**: Dados abertos para a democracia na era digital, IV Congresso internacional software livre e comércio eletrônico – CONSEGI 2011, Brasília: Ministério das Relações Exteriores, 2011. Disponível em: <http://www.funag.gov.br/biblioteca/index.php?option=com_docman&task=search_result&Itemid=41>. Acesso em: 30 jul. 2011.
- MENY, I.; THOENIG, J. C. **Lãs políticas públicas**. Barcelona: Ed. Ariel SA, 1992.
- MOÇAMBIQUE. Constituição (2004). Disponível em: <<http://www.mozambique.mz/pdf/constituicao.pdf>> Acesso em: 11 dez. 2009.
- MOÇAMBIQUE. Decreto nº 27/99, de 24 de Maio. Aprova o Estatuto Orgânico, o Quadro de Pessoal, o Quadro de Funções de Comando, Direção e Chefia e os Organogramas da PRM. **Lex**: Coletânea de legislação penal e complementar por João Carlos Trindade: Centro de formação jurídica e judiciária, Maputo, 2005.
- MOÇAMBIQUE. Lei nº 17/97, de 1 de Outubro. Aprova a Política de Defesa e Segurança. **Boletim da República**, I Série, n.40, Maputo, 7 out. 1997. 3º Suplemento. Disponível em: <<http://www.pdgs.org.ar/Archivo/d000009a.htm>> Acesso em: 11 dez. 2009.
- MOÇAMBIQUE. Lei nº 19/92, de 31 de Dezembro. Cria a Polícia da República de Moçambique. **Lex**: Coletânea de legislação penal e complementar por João Carlos Trindade: Centro de formação jurídica e judiciária, Maputo, 2005.
- MORVILLE, P. **Ambient findability**. Sebastopol: O'Really, 2005.



MORVILLE, P.; ROSENFELD, L. **Information architecture for the world wide web**. 3.ed. Sebastopol: O'Really, 2006.

OLIVEIRA, D. P. R. **Sistemas de Informações gerenciais: estratégicas, táticas, operacionais**. 9.ed., São Paulo: Atlas, 2004.

PRINCIPLES OF OPEN GOVERNMENT DATA. Disponível em: <<http://www.opengovdata.org/home/8principles>>. Acesso em 11 jul. 2011.

SARACEVIC, T. Interdisciplinarity nature of Information Science. **Ciência da Informação**, Brasília, v.24, n.1, p.36-41, 1995.

SECRETARIA DE SEGURANÇA PÚBLICA DO ESTADO DE SÃO PAULO. Disponível em: <<http://www.ssp.sp.gov.br/Bo/>> Acesso em: 22 fev. 2010

SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. **Sistema de Banco de Dados**. 5. ed., Rio de Janeiro: Elsevier, 2006.

STEWART, D.L. **Building Enterprise Taxonomies**. Mokita Pess, 2008. p.1-22.



ARQUITETURA DA INFORMAÇÃO COMO ARCABOUÇO PARA A INTEGRAÇÃO DA TECNOLOGIADA INFORMAÇÃO COM OS PROCESSOS ORGANIZACIONAIS E A INFORMAÇÃO¹

Márcio de Carvalho Victorino, Marisa Bräscher Basílio Medeiros, Rômulo Ferreira dos Santos

Resumo: Este estudo teve por objetivo disponibilizar um processo de modelagem da informação fundamentado nos conceitos, métodos e técnicas de Organização da Informação preconizadas pela Ciência da Informação (CI). Foram utilizadas duas abordagens metodológicas para a realização da pesquisa: a Metodologia de Sistemas Flexíveis (*Soft Systems Methodology* – SSM), para o ciclo da pesquisa e a Arquitetura de Método Unificado (*Unified Method Architecture* – UMA), para a construção do processo de modelagem da informação. A validação da modelagem proposta ocorreu por meio de um estudo de caso realizado em uma organização pública brasileira. A principal contribuição deste trabalho consistiu na disponibilização de uma Arquitetura da Informação (AI), materializada por meio de um repositório informacional corporativo, composto por objetos informacionais, metadados e os sistemas de organização do conhecimento (SOC), mais especificamente, tesouros, taxonomias e ontologias. Todos estes artefatos são harmoniosamente conectados. A modelagem da informação proposta, além de materializar uma AI, documenta os processos organizacionais, suas decomposições, os componentes de software que os automatizam e as informações manipuladas por ambos. Assim, a documentação gerada permite relacionar a informação organizacional, com os componentes de software implementados e com os processos organizacionais concebidos durante a modelagem do negócio, proporcionando a integração da Tecnologia da Informação (TI) com processos organizacionais e com a informação manipulada pelos processos e pelos softwares.

Palavras-chave: Organização da Informação. Tecnologia da Informação. Processos Organizacionais.

1 INTRODUÇÃO

A informação é um componente intrínseco de quase tudo o que uma organização faz. Assim sendo, sem uma clara compreensão dos processos organizacionais e humanos pelos quais a informação se transforma em percepção, conhecimento e ação, as empresas não são capazes de perceber a importância de suas fontes de tecnologia da informação (CHOO, 2003).

Normalmente, as informações transitam pelas organizações sem que se tenha plena consciência do seu impacto, valor ou custo. Para o controle adequado das informações de uma organização, torna-se necessário buscar uma gestão efetiva da informação. (DAVENPORT, 1994).

Choo (2003), conforme apresentado na Figura 1, sugere que a gestão da informação, seja vista como a administração de uma rede de processos que adquirem, criam, organizam, distribuem e usam a informação.



Figura 1 - Modelo Processual de Gestão da Informação

Fonte: Choo (2003)

Nessa abordagem, a gestão da informação é analisada como um ciclo contínuo de seis processos correlatos:

- Identificação das necessidades de informação;
- Aquisição da informação;
- Organização e armazenamento da informação;
- Desenvolvimento de produtos e serviços de informação;
- Distribuição da informação;
- Uso da informação.

Portanto, um modelo processual de gestão da informação proporciona o uso eficiente da informação. As reações de uma organização interagem com as ações de outras organizações, gerando novos sinais e mensagens, aos quais se deve atentar e, dessa forma, manter ciclos do uso da informação (CHOO, 2003).

Na visão de Davenport (1998), os administradores precisam de uma perspectiva holística para o tratamento da informação nas organizações, para que possam assimilar alterações repentinas no mundo dos negócios e adaptarem-se às sempre mutantes realidades sociais. Essa abordagem foi denominada “Ecologia da Informação” e, além de exigir um modo holístico de pensar, fundamenta-se por meio de quatro atributos chave:

- Integração dos diversos tipos de informação;
- Reconhecimento de mudanças evolutivas;
- Ênfase na observação e na descrição;
- Ênfase no comportamento pessoal e informacional.

A ecologia da informação proposta por Davenport (1998), conforme apresentada na Figura 2, é composta por três ambientes: o informacional, o organizacional e o externo.



Figura 2 - Modelo Ecológico para a Gestão da Informação
Fonte: Davenport (1998)

O ambiente informacional é o núcleo da abordagem e abrange os componentes mais críticos, que são:

- Estratégia da informação: gira em torno da pergunta “o que queremos fazer com a informação nesta empresa?”;
- Política da informação: envolve o poder proporcionado pela informação e as responsabilidades da direção em seu gerenciamento e uso;
- Cultura e comportamento em relação à informação: esses dois fatores estão relacionados e são muito importantes na criação de um ambiente informacional bem-sucedido;
- Equipe da informação: pessoas ainda são os melhores “meios” para identificar, categorizar, filtrar, interpretar e integrar a informação;
- Processos de gestão da informação: esse componente mostra como o trabalho é feito e fornece uma descrição completa de como funciona cada parte do trabalho informacional;
- Arquitetura da informação: é um conceito que pode abranger muitos significados, no entanto, na perspectiva ecológica, significa um guia para estruturar e localizar a informação dentro de uma organização.

Já o ambiente organizacional, contém a posição global dos negócios, os investimentos em tecnologia e a distribuição física da informação.

Por fim, o ambiente externo representa a necessidade de monitoramento para a adequação da organização ao que está acontecendo fora dela. Esse ambiente é composto por mercados de negócios em geral, mercados tecnológicos e mercados da informação.

Nesse contexto, a pesquisa descrita não teve por objetivo implementar um modelo de gestão da informação específico e completo, no entanto, procurou-se disponibilizar um modelo de organização da informação compatível com os principais conceitos de gestão da informação, alinhado à visão holística segundo o modelo ecológico de Davenport (1998). Logo, o conceito de AI tornou-se o principal foco da pesquisa.



2 PROBLEMA DA PESQUISA

O armazenamento e o acesso à informação nos ambientes organizacionais foram significativamente facilitados com a evolução da Tecnologia da Informação e Comunicação (TIC) ocorrida nas últimas décadas. No entanto, conseguir reutilizar a informação existente nesses ambientes é um problema que se agrava, principalmente nas organizações orientadas a processos, que dependem de movimentos rápidos e eficientes da informação.

Nessas organizações, apesar das informações geradas por um sistema muitas vezes servirem de insumo para outros, esses diferentes sistemas, em geral, estão dispersos e não apresentam nenhum tipo de conexão entre si. Na maior parte das vezes, os sistemas de informação são desenvolvidos isoladamente, o que normalmente dificulta ou impede o reuso de informações pelos demais sistemas da organização.

Outro problema existente nas organizações consiste no fato de as metodologias de modelagem de processos e de engenharia de software, utilizadas para a modelagem e desenvolvimento de sistemas de informação computadorizados, não possuírem uma estratégia para documentar os aspectos semânticos da informação manipulada (VICTORINO; BRÄSCHER, 2009).

Essas metodologias vêm evoluindo consideravelmente nos últimos tempos, porém ainda apresentam deficiências no que diz respeito ao tratamento da informação, se comparadas com os métodos e as técnicas utilizados pela CI.

Nesse sentido, a questão da pesquisa é: Como organizar as informações organizacionais visando o seu compartilhamento entre sistemas de informação computadorizados?

3 ARQUITETURA DA INFORMAÇÃO

O termo ‘Arquitetura da Informação’ foi utilizado pela primeira vez pelo arquiteto Richard Saul Wurman, em 1976, que o definiu como a “ciência e a arte de criar instruções para espaços organizados”. O autor encara o problema de busca, organização e apresentação da informação como análogo aos problemas da arquitetura de construções que irão servir às necessidades de seus moradores (MACEDO, 2005).

Não há uma definição precisa sobre o que é ou o que constitui uma AI. Observa-se, dentre os vários pesquisadores que escrevem sobre o assunto, uma grande quantidade e diversidade de definições.

Brancheau e Wetherbe (1986) afirmam que AI consiste em um plano para modelagem dos requisitos informacionais de uma organização, que provê um modo de mapear as informações necessárias à própria organização, relativas aos processos do negócio e à documentação de seus inter-relacionamentos.

Rosenfeld e Morville (2002) propõem o modelo representado na Figura 3. Nesse modelo a AI é representada como a interseção de contexto, conteúdo e usuários. No espaço informacional de uma organização é necessário conhecer os objetivos do negócio da organização (contexto), estar consciente

aprimoramento das buscas realizadas pelos usuários com a delimitação do contexto (VICTORINO; BRÄSCHER, 2009).

Portanto, a AI provê uma visão integrada da informação que trafega entre os processos da organização.

O processo de Modelagem da Informação é composto por um conjunto de metodologias usadas de maneira integrada para gerar a AI descrita anteriormente. A Figura 5 apresenta os alicerces desse processo.



Figura 3 – Os Alicerces do Processo de Modelagem da Informação
Fonte: Victorino (2011)

A administração de dados gera os metadados; as metodologias para modelagem de processos registram os processos organizacionais; as metodologias para construção de tesouros geram tesouros; as metodologias para construção de taxonomias geram taxonomias; as metodologias para construção de ontologias geram ontologias e, finalmente, as metodologias para o desenvolvimento de sistemas de informação geram os componentes de software que são implementações dos processos organizacionais.

A integração das metodologias de modelagem ocorre de maneira iterativa, ou seja, a modelagem de processos, a modelagem dos sistemas de informação e a modelagem da informação sofrerão vários refinamentos até se atingir o produto final.

Essa abordagem integrada permite que a informação mapeada durante a modelagem de um processo seja manipulada por um componente de software relacionado ao processo.

O processo de modelagem da informação proporciona a integração da TI com os processos organizacionais e as informações, tendo em vista que a documentação gerada permitirá relacionar os componentes de software implementados com os processos concebidos durante a modelagem do negócio e as informações manipuladas por ambos.

4 OBJETIVO GERAL DA PESQUISA

Esta pesquisa tem por objetivo geral desenvolver um processo de modelagem da informação para ser empregado, em conjunto com as metodologias de modelagem de processos e as metodologias

da natureza e do volume de informações existentes e de sua taxa de crescimento (conteúdo), bem como, entender as necessidades e os processos de busca do público-alvo (usuários).

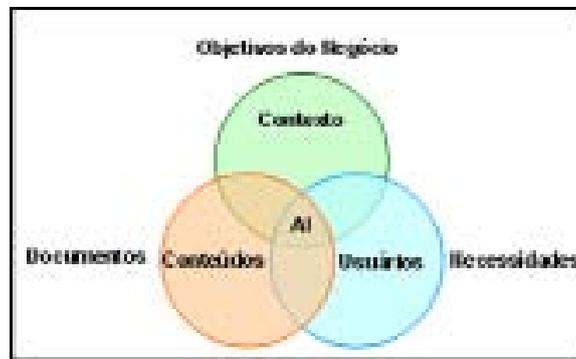


Figura 4 - Modelo de Arquitetura da Informação
Fonte: Rosenfeld e Morville (2002)

Do exposto, vê-se que o foco da AI está na representação semântica da informação, na organização de sua armazenagem e na otimização de sua recuperação. Em um ambiente de AI planejado, a organização da informação torna-se um elemento de vital importância para a garantia da qualidade na recuperação da informação.

A AI proposta nesta pesquisa é composta por usabilidade, metadados, tesouros, taxonomias e ontologias. A Figura 4 apresenta os alicerces desta AI.



Figura 5 – Os Alicerces de uma Arquitetura da Informação
Fonte: Victorino (2011)

Conforme ilustrado na Figura 4, para acessar a informação, o usuário interage com interfaces implementadas de acordo com critérios de usabilidade.

Assim sendo, os metadados descrevem o suporte e o conteúdo e servem como índices para a recuperação da informação. Os tesouros são utilizados para permitir ao usuário encontrar o termo que represente um determinado significado para o que procura.

As taxonomias navegacionais são utilizadas para permitir que os usuários leigos naveguem pelo conteúdo do repositório e, por esse motivo, são criadas levando em conta o comportamento do usuário. Já as taxonomias descritivas auxiliam os especialistas em suas buscas por informações. Por fim, as ontologias permitem o



de engenharia de software utilizadas para a modelagem e o desenvolvimento de sistemas de informação computadorizados.

5 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

A presente pesquisa classifica-se, quanto aos fins, como sendo qualitativa, metodológica e aplicada.

Duas metodologias foram utilizadas para a sua realização, a Metodologia de Sistemas Flexíveis (Soft Systems Methodology - SSM) e a Arquitetura de Método Unificado (Unified Method Architecture - UMA). A SSM foi utilizada para o ciclo da pesquisa propriamente dito, no entanto no estágio 4 desta metodologia (Construção de Modelos Conceituais), tornou-se necessária a adoção de outra metodologia capaz de permitir a especificação de um processo de Modelagem da Informação. Para essa parte da pesquisa, foi adotada a UMA.

5.1 Metodologia Utilizada para o Ciclo da Pesquisa

A SSM foi desenvolvida na década de 60 pela equipe de Peter Checkland, do Departamento de Sistemas e Administração de Informações da Universidade de Lancaster (MARTINELLI; VENTURA, 2005).

A Metodologia de Sistemas Flexíveis possui sete etapas distintas:

- Estágio 1: situação-problema não estruturada;
- Estágio 2: situação-problema estruturada;
- Estágio 3: definições fundamentais dos sistemas relevantes;
- Estágio 4: construção de modelos conceituais;
- Estágio 5: comparação dos modelos conceituais (4) com a realidade (2);
- Estágio 6: identificação das mudanças desejáveis e possíveis;
- Estágio 7: ações para melhorar a situação-problema.

Os estágios 1, 2, 5, 6 e 7 são atividades que envolvem as pessoas no mundo real e os estágios 3 e 4 destina-se ao pensamento sistêmico.



5.2 Metodologia Utilizada para a criação do Processo de Modelagem da Informação

Para o desenvolvimento do processo de modelagem da informação proposto, foi escolhido o metamodelo utilizado pelo RUP, o *Unified Method Architecture* (UMA) (SHUJA; KREBS, 2007). A UMA é baseada nas seguintes separações fundamentais:

- Separação de conteúdo de método versus a aplicação de conteúdo de método em processos;
- Definição de um mecanismo de extensibilidade opcional no método para gerenciamento em grande escala de repositórios de método e de processo;
- Separação de campos de descrição de método e orientação recomendados;
- Separação de elementos semânticos de notação em diagramas do processo.

A UMA separa o conteúdo do método e possui um núcleo reutilizável de sua aplicação em processos. O conteúdo do método descreve o que deve ser produzido, as habilidades necessárias requeridas e o detalhamento dele próprio, com explicações sobre como as metas de desenvolvimento específicas são atingidas, independentemente do posicionamento desses itens dentro de um ciclo de vida de desenvolvimento. Os processos obtêm esses elementos do método e os remetem para sequências semiordenadas, que são personalizadas para tipos específicos de projetos.

6 O PROCESSO DE MODELAGEM DA INFORMAÇÃO E SUA DOCUMENTAÇÃO

Modelagem da informação consiste em um conjunto de procedimentos, técnicas, ferramentas e documentos auxiliares que ajudam os profissionais de informação em seus esforços para representar o domínio observado e os objetos informacionais pertencentes a este domínio. A modelagem da informação contempla tanto a descrição física, características físicas do meio e do formato em que a informação está registrada, quanto à descrição do conteúdo informacional.

Para a construção do processo de modelagem da informação foi escolhido o metamodelo UMA (SHUJA; KREBS, 2007), composto por disciplinas e fases. As disciplinas caracterizam o conteúdo do método, ou seja, documentam o que deve ser feito e as fases caracterizam a execução propriamente dita das atividades que compõem o método.

A documentação das disciplinas é materializada por meio da construção de diagramas de atividades da UML (BOOCH *et al.*, 2005), sendo que as atividades são detalhadas em tarefas, papéis e produtos de trabalho. Uma das principais vantagens dessa abordagem é permitir ao usuário do processo de modelagem da informação configurar as disciplinas de acordo com as necessidades da organização alvo. Assim, o processo torna-se flexível tratando de maneira diferente organizações com objetivos ou porte diferentes.

As disciplinas do processo de modelagem proposto são:

- Requisitos informacionais;
- Análise da informação;
- Implementação;
- Validação;
- Disponibilização.

A disciplina requisitos informacionais descreve as atividades relacionadas ao levantamento das necessidades informacionais dos usuários. Sua finalidade é oferecer às pessoas que interagem com o domínio observado uma compreensão dos conceitos relevantes, estabelecer padrões de metadados e demarcar as fronteiras do domínio. Assim, esta disciplina proporciona uma compreensão dos conceitos de uma determinada área de atividade. A Figura 6 apresenta o fluxo de atividades desta disciplina.

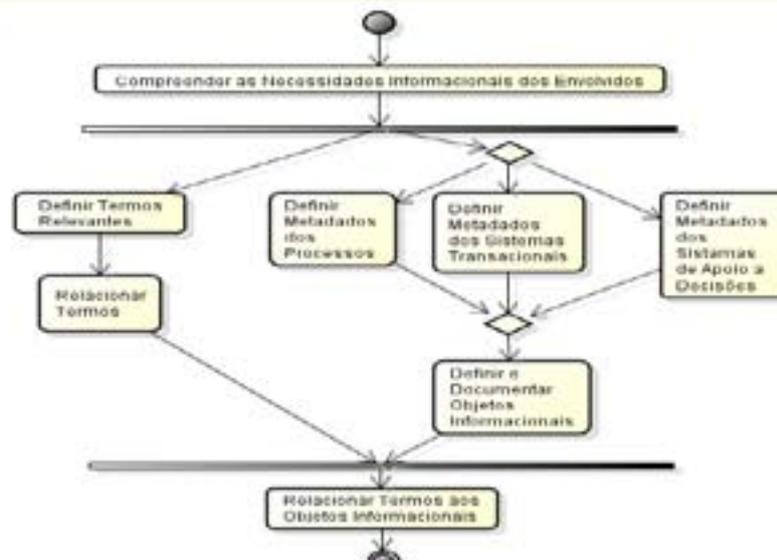


Figura 6 – Fluxo de Atividades da Disciplina Requisitos Informacionais
Fonte: Victorino (2011)

A disciplina análise da informação tem por finalidade transformar os requisitos informacionais em um conjunto de especificações de recursos para a representação da informação e do conhecimento. É durante esta disciplina que a AI será planejada e o Repositório Informacional Corporativo será projetado.

Segundo Café et al. (2003), Repositório Institucional “é a reunião de todos os repositórios temáticos hospedados em uma organização”. Lynch (2003) afirma que um repositório institucional é como um conjunto de serviços que uma universidade oferece para os membros da comunidade, para o gerenciamento e a disseminação de materiais digitais criados pela instituição e pelos membros da comunidade. Os autores relatam que os repositórios institucionais têm emergido nas universidades, mas estão espalhados dentro de outros tipos de organizações educacionais, desde colégios até institutos de pesquisa.

Do exposto, observa-se que existem objetivos em comum entre esses tipos de ambientes informacionais, como: armazenar, facilitar o acesso e disseminar informações.

Estendendo esse conceito para abranger, além dos objetivos educacionais, os objetivos organizacionais, este trabalho usa o termo “Repositório Informacional Corporativo” para referenciar o local onde as informações das organizações e respectivas descrições estão armazenadas.

A Figura 7 apresenta o fluxo de atividades da disciplina análise da informação.

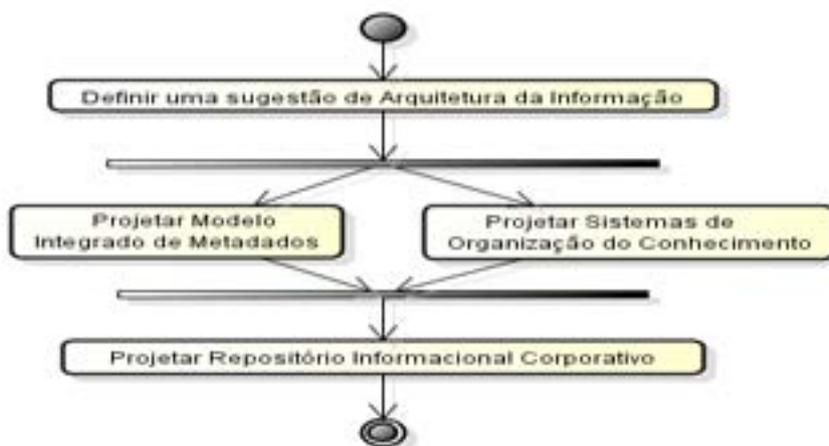


Figura 7 – Fluxo de Atividades da Disciplina Análise da Informação
Fonte: Victorino (2011)

A disciplina implementação define as estratégias para a materialização da AI planejada na disciplina Análise da Informação por meio da criação do Repositório Informacional Corporativo. A Figura 8 apresenta o fluxo de atividades desta disciplina.



Figura 8 – Fluxo de Atividades da Disciplina Implementação
Fonte: Victorino (2011)

A disciplina validação identifica os passos a serem seguidos para que os metadados e dos sistemas de organização do conhecimento sejam validados. A Figura 9 apresenta o fluxo de atividades desta disciplina.

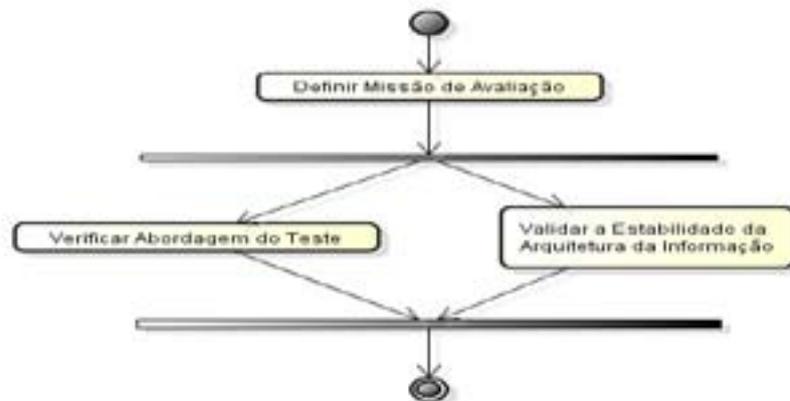


Figura 9 – Fluxo de Atividades da Disciplina Validação
Fonte: Victorino (2011)

A disciplina disponibilização descreve as atividades que garantem que o produto da modelagem da informação será disponibilizado a seus usuários finais. A Figura 10 apresenta o fluxo de atividades desta disciplina.

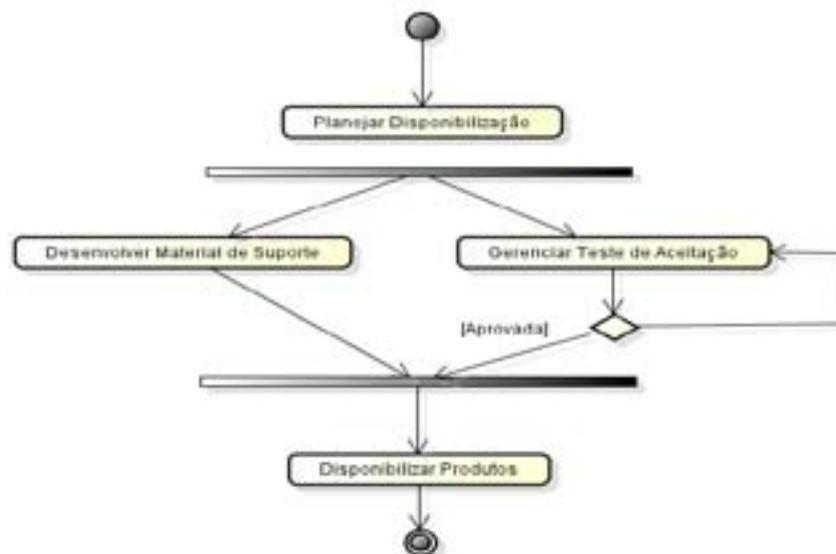


Figura 10 – Fluxo de Atividades da Disciplina Disponibilização
Fonte: Victorino (2011)

As fases do processo proposto são: iniciação, elaboração, construção e transição. A meta da fase de iniciação é identificar o escopo da modelagem da informação. Na fase de elaboração, a meta é validar a AI. Os SOC serão criados na fase de construção. Finalmente, na fase de transição, os produtos gerados nas fases anteriores serão disponibilizados integralmente para os usuários que passarão por treinamento quando for necessário.

1. OS PRODUTOS DE TRABALHO MAIS IMPORTANTES RESULTANTES DO PROCESSO DE MODELAGEM DA INFORMAÇÃO SÃO O REPOSITÓRIO INFORMACIONAL CORPORATIVO, METADADOS E OS SOC, MAIS ESPECIFICAMENTE, TESAUROS, TAXONOMIAS E ONTOLOGIAS.

A Figura 11 apresenta o “Repositório Informacional Corporativo”, para a materialização da AI apresentada anteriormente e sua conexão com os recursos utilizados na representação da informação e do conhecimento resultantes do processo de modelagem da informação. Esse repositório é composto por objetos informacionais gerados internamente; por objetos informacionais gerados externamente, porém copiados para o repositório; por referência a objetos informacionais externos; por metadados e por componentes dos SOC que devem ser persistidos.



Figura 11 - Repositório Informacional Corporativo
Fonte: Victorino (2011)

Os objetos informacionais internos são compostos pelos objetos informacionais gerados internamente e pelos criados externamente, porém, nesse último caso, eles passam por um tratamento para se tornarem aderentes ao modelo interno da organização. Já as referências aos objetos informacionais externos à organização têm o objetivo de facilitar a recuperação das informações externas de interesse da organização, sem comprometer sua capacidade de armazenagem, pois seria inviável, para grandes organizações, por exemplo, armazenarem, em seus repositórios informacionais, todo o conteúdo de seu interesse disponibilizado na web.

No repositório proposto observa-se, ainda, um repositório interno exclusivo para os metadados e os componentes dos SOC que devem ser persistidos.



A Figura 12 detalha como os SOC foram representados em tabelas relacionais. O conteúdo da coluna “Item” da tabela “Termo” foi retirado do Glossário das Forças Armadas (BRASIL, 2007).

A tabela “Termo” armazena o código do termo, sua grafia e sua conceitualização respectivamente nas colunas “Cod”, “Item” e “Conceito”.

As tabelas “Ontologia”, “Taxonomia Navegacional” e “Tesouro” armazenam os códigos dos termos que possuem associações e seus relacionamentos respectivamente nas colunas “TermoInicial”, “TermoFinal” e “Relacionamento”. Os relacionamentos são nomeados de acordo com o SOC representado, por exemplo, os tesouros terão relacionamentos “TA”, “TG” e etc. As ontologias têm os mais variados relacionamentos e as taxonomias navegacionais agrupam os termos de acordo com a percepção do usuário do domínio.

Termo		
Cod	Item	Conceito
1	Viatura	Veículo militar dotado de rodas, lagartas ou combinação de ambas.
2	Viatura de Combate	Viatura com ou sem blindagem, apropriada para ser empregada, por sua guarnição, em combate.
3	Guarnição	Grupo de homens que guarnece e opera uma instalação, equipamento, arma etc.

Ontologia		
TermoInicial	TermoFinal	Relacionamento
2	1	é um
1	3	Transporta

Taxonomia Navegacional		
TermoInicial	TermoFinal	Relacionamento
1	2	Agrupar

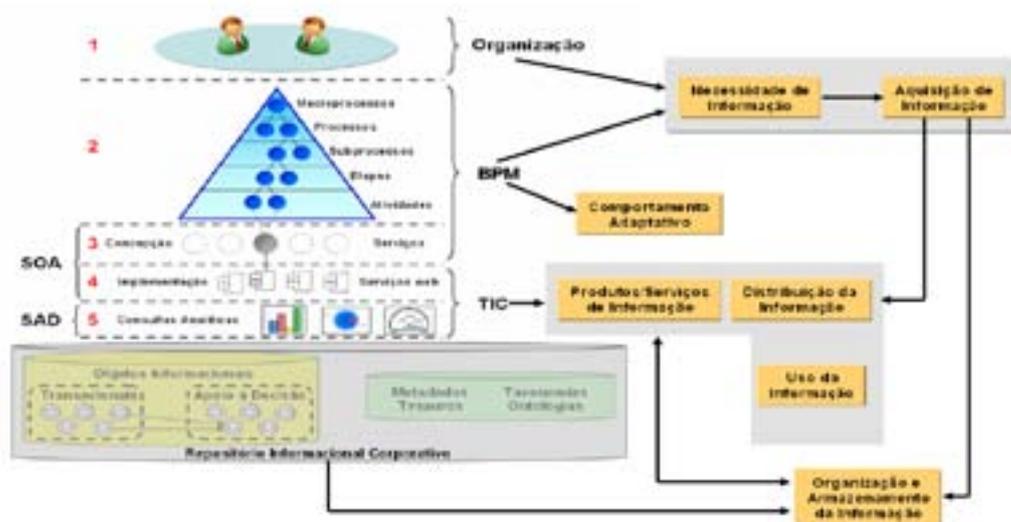
Tesouro		
TermoInicial	TermoFinal	Relacionamento
1	2	TG
1	3	TA

Figura 12 – Modelo de Persistência da Classe “Termo” e seus Auto relacionamentos
Fonte: Victorino (2011)

Com o modelo representado na Figura 12 é possível, a partir de um determinado termo, encontrar outros termos relacionado por meio da materialização das ontologias, taxonomias e tesouros.

7 O PROCESSO DE MODELAGEM DA INFORMAÇÃO E O FLUXO INFORMACIONAL NAS ORGANIZAÇÕES

A Figura 13 apresenta uma abstração de uma organização orientada a processos provida de AI (lado esquerdo da figura) juntamente com o conjunto de processos proposto por Choo (2003) para a gestão da informação (lado direito da figura). O objetivo da figura é apresentar como essa proposta pode colaborar com a gestão da informação nas organizações e o integrando da TI com os processos organizacionais e as informações manipuladas.



2.

Figura 13 – O Processo de Modelagem da Informação e o Fluxo Informacional nas Organizações

Fonte: Victorino (2011)

A camada 1 representa a organização no mundo natural, que pode ser entendida como um arranjo sistemático de duas ou mais pessoas que cumprem papéis formais e compartilham um propósito comum (Robbins, 2003).

Na camada 2, a mesma organização é vista, metaforicamente, como fluxo e transformação, isso porque é constituída por processos e suas decomposições (macroprocessos, processos, subprocessos, etapas, atividades). Uma das ideias principais em comparar uma organização a fluxo e transformação é justamente a compreensão de que a mudança é peça fundamental no contexto organizacional (MORGAN, 2002).

A camada 3 abrange a modelagem de serviços que são partes componentes de um processo organizacional. Essa camada consiste em uma estratégia para organizar as atividades em serviços dispostos em uma arquitetura, a fim de proporcionar flexibilidade e reuso. Esse é o ponto de partida para o planejamento da Arquitetura Orientada a Serviços (SOA).

Para lidar com a complexidade das organizações orientadas a processos, nas camadas 2 e 3 estão concentradas as atividades que envolvem a descoberta, o projeto e a entrega de processos de negócios adicionadas ao seu controle executivo, administrativo e supervisorio. Essa abordagem é chamada BPM (Business Process Management) (BALDAM et al., 2007).

Na camada 4 os serviços são implementados como componentes de software (mais comumente, serviços web), compondo os sistemas transacionais. A implementação e a forma de disponibilização desses serviços web materializam a SOA planejada na camada 3.

A camada 5 representa a implementação das consultas analíticas que atenderão aos usuários gerenciais, caracterizando os sistemas de apoio à decisão.



Assim, as camadas 4 e 5 representam a implementação dos produtos de tecnologia de informação e comunicação (TIC) para a automação da organização. Normalmente, o usuário final interage com essas camadas para obter as informações de seu interesse.

Finalmente, localizado na parte inferior da Figura 13, o repositório informacional corporativo armazena os objetos informacionais utilizados pelos sistemas transacionais e de apoio à decisão, além dos metadados relativos aos macroprocessos, processos, subprocessos, etapas, atividades, serviços, serviços web e consultas analíticas.

Outro componente do repositório informacional corporativo de fundamental importância é o repositório dos dados persistentes dos sistemas de organização do conhecimento (SOC), composto por tesouros, taxonomias e ontologias.

É importante ressaltar que, durante todo o processo de modelagem em uma organização, desde a sua concepção até a sua automação, o repositório informacional corporativo é carregado com objetos informacionais, metadados, informações sobre o domínio no qual a organização está inserida e seus inter-relacionamentos. Nesse repositório pode-se encontrar informações a respeito dos objetivos da organização, seus processos, serviços, componentes de software e conceitos do domínio de atuação, proporcionando o elo de ligação entre as várias camadas de abstração de uma organização.

No lado direito da Figura 13 estão representados os processos de gestão da informação preconizados por Choo (2003). Segundo o autor, a gestão efetiva da informação em uma organização inicia-se com os processos de Identificação das Necessidades de Informação e Aquisição da Informação. Esses dois processos são beneficiados pelo entendimento do propósito da organização (camada 1) e pelo mapeamento de seus processos utilizando técnicas de BPM (camadas 2 e 3). O processo de modelagem da informação proposto subsidia esses dois processos de gestão da informação por meio do registro dos metadados dos processos modelados, dos objetos informacionais manipulados e da contextualização de ambos.

As técnicas de BPM também proporcionam maior flexibilidade às organizações, subsidiando o **Comportamento Adaptativo**.

Os sistemas gerados com ferramentas de TIC constituem os principais produtos de informação consumidos pelos usuários finais, favorecendo os processos de **Produtos/Serviços de Informação, Distribuição da Informação e Uso da Informação**. O processo de modelagem da informação proposto documenta todos os objetos informacionais, os componentes de software e a semântica envolvidos, facilitando a busca por informação, realizada pelos usuários, para a solução de seus problemas.

O processo **Organização e Armazenagem da Informação**, subsidiará quase todos os demais, tem seus resultados materializados por meio do repositório informacional corporativo.

A modelagem da informação abordada também é compatível com a “Ecologia da Informação”, proposta por Davenport (1998). A proposta de modelagem contempla os quatro atributos chave, considerados fundamentais para uma abordagem ecológica, que são::



- Integração de diversos tipos de informação: o repositório informacional corporativo armazena os diversos tipos de objetos informacionais existente em uma organização;
- Reconhecimento de mudanças evolutivas: o repositório informacional corporativo pode ser configurado de acordo com as peculiaridades da organização e de seus sistemas de informações, possibilitando flexibilidade;
- Ênfase na descrição: o repositório informacional corporativo armazena, também, a descrição do conteúdo e do suporte dos objetos informacionais de uma organização;
- Ênfase no comportamento pessoal e informacional: os SOC, previstos na metodologia proposta, permitem contextualizar os objetos informacionais de uma organização, possibilitando o seu compartilhamento por sistemas distintos, bem como, facilitam a administração dos conteúdos informacionais e eliminam ou reduzem os significados múltiplos.

Um importante aspecto a ser ressaltado é que esta pesquisa apresenta uma proposta de estruturação de uma AI, recurso previsto por Davenport em seu modelo ecológico.

Observa-se que o processo de modelagem da informação apresentado, quando utilizado em uma organização orientada a processos, pode contribuir com os processos de gestão da informação propostos por Choo (2003) e proporcionar subsídios ao modelo ecológico concebido por Davenport (1998).

Do exposto, pode-se concluir que o mapeamento dos relacionamentos entre os processos organizacionais, os componentes de software e os objetos informacionais em um ambiente organizacional fornece o arcabouço necessário para a integração dos processos organizacionais com a TI e com as informações manipuladas. Uma maneira eficaz de se criar este arcabouço é por meio da utilização de uma abordagem sistemática para a modelagem da informação fundamentada nos conceitos, métodos e técnicas de Organização da Informação preconizados pela CI.

8 CONSIDERAÇÕES FINAIS E SUGESTÕES DE TRABALHOS FUTUROS

Esta pesquisa originou-se devido aos vários óbices encontrados em projetos de desenvolvimento de sistemas de informações computadorizados e à dificuldade de identificação, recuperação e compartilhamento da informação. Teve por objetivo disponibilizar um processo de modelagem da informação para ser empregado, em conjunto com as metodologias de modelagem de processos e as metodologias de engenharia de software utilizadas para o desenvolvimento de sistemas de informação transacionais e de apoio à decisão, nas organizações dispostas em torno de seus processos.

Verificou-se que a CI tem estudado o fluxo da informação nos mais variados ambientes, com critérios, princípios e métodos científicos. Esse fluxo, normalmente, é automatizado por meio do desenvolvimento de sistemas de informações computadorizados, utilizando metodologias de engenharia de software. No entanto, tais metodologias raramente utilizam os conceitos, métodos e técnicas da CI de maneira integrada e, geralmente, não levam em conta a influência das estruturas organizacionais.



O produto de trabalho mais importante resultante do processo de modelagem da informação é a AI, que é materializada por meio do repositório informacional corporativo, composto por objetos informacionais, metadados e os SOC, mais especificamente, tesouros, taxonomias e ontologias. Todos esses artefatos são harmoniosamente conectados.

No entanto, toda pesquisa apresenta limitações que podem ser de ordem teórica ou de ordem prática. Sendo assim, em continuidade ao trabalho da pesquisa sugere-se uma dissertação de mestrado que contemple a :

- Extensão da AI apresentada com a finalidade de orientar a organização da informação em diversas instituições públicas e privadas;
- Construção de um protótipo levando em consideração as metodologias propostas, para a partir do repositório informacional, identificar a relevância da informação sob o ponto de vista de seus usuários;
- Realização de estudos da interação de usuários com repositório informacional corporativo. O objetivo é compreender perfis e comportamentos com o propósito de projetar ambientes de acordo com suas necessidades informacionais e modos de interação.
- Comparação da satisfação do usuário da informação antes e após a implantação da AI proposta.

Abstract: *The aim of this study was to provide an information modeling process based on concepts, methods and techniques of Information Organization according to Information Science (IS). Two methodological approaches were used for this research: the Soft Systems Methodology – SSM – for research cycle and the Unified Method Architecture – UMA – for the building of information modeling process. The model validation occurred by means of case study of a Brazilian government organization. The main contribution of that first model was the Information Architecture (IA) that offered the corporate information repository, containing information objects, metadata and knowledge organization systems (KOS), more specifically thesauri, taxonomies and ontologies. All those artifacts are harmoniously connected. The information modeling proposed here, besides providing an IA, documents the organization processes, its decompositions, software components that make them automate and the information that is manipulated by both items. Therefore, the documentation that is generated allows us to relate organization information with software components that were implemented and also with the organization processes that were conceived during business modeling. It provides the integration between Information Technology (IT), organization processes and the information that is manipulated both by processes and softwares.*

Keywords: Information Organization. Information Technology. Organizational Processes.

REFERÊNCIAS

BALDAM, R. et al. **Gerenciamento de Processos de Negócios: BPM – Business Process Management.** São Paulo: Érica, 2007. 240 p.

BOOCH, G.; RUMBAUGH, J.; JACOBSON, I. **UML – Guia do Usuário.** 2. ed. Rio de Janeiro: Campus. 2005. 474 p.



- BRASIL. Ministério da Defesa. **MD 35-G-01: Glossário das Forças Armadas**. Brasília, DF, 2007.
- BRANCHEAU, J. C.; WETHERBE, J. C. Information Architectures: Methods and Practice. **Information Processing & Management**, v. 22, n. 6, p. 453-463, 1986.
- CAFÉ,L; MELO, B.A.; BARBOSA, E.M.F.; NUNES, E.M.A.N.; MÁRDERO ARELLANO, M. A. **Repositórios institucionais: nova estratégia para publicação científica na Rede**. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, XXVI, 2003, Belo Horizonte. Belo horizonte: INTERCOM – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2003.
- CHOO, C. W. A. **Organização do Conhecimento**. São Paulo: Senac, 2003. p. 403-421.
- DAVENPORT, T. H. **Reengenharia de processos: como inovar na empresa através da tecnologia da informação**. 5. ed. Rio de Janeiro: Campus, 1994.
- _____. **Ecologia da Informação**. 6. ed. São Paulo: Futura, 1998.
- LYNCH, C.A. **Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age**, 2003, ARL Bimonthly Rep, n. 226, p. 327-336. Disponível em <<http://www.arl.org/bm~doc/br226ir.pdf>>. Acesso em: 19 jul. 2010.
- MACEDO, F. L. O. **Arquitetura da informação: aspectos epistemológicos, científicos e práticos**. 2005. 190 p. Dissertação (Mestrado). Departamento de Ciência da Informação e Documentação da Universidade de Brasília. Brasília.
- MARTINELLI, D.P.; VENTURA, C.A.A. **Visão sistêmica e administração: conceitos, metodologias e aplicações**. São Paulo: Saraiva, 2005. 242 p.
- MORGAN, G. **Imagens da organização: edição executiva**. 2. ed. São Paulo: Atlas, 2002.
- INS, S. P. **Administração: Mudanças e Perspectivas**. São Paulo: Saraiva, 2003.
- ROSENFELD, L.; MORVILLE,. Information Architecture for the World Wide Web. 2. ed. Cambridge: O'Reilly, 2002. 461 p.
- SHUJA, A. K.; KREBS J. IBM Rational Unified Process Reference and Certification Guide. Boston: IBM Press Pearson plc, 2007. 307p.
- VICTORINO, M. C.** Organização da Informação para dar Suporte à Arquitetura Orientada a Serviços: Reuso da Informação nas Organizações. **2011. 276 p. Tese (Doutorado). Faculdade de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília - UnB. Brasília.**
- VICTORINO, M. C.; BRÄSHER, M. Organização da Informação e do Conhecimento, Engenharia de Software e Arquitetura Orientada a Serviços: uma Abordagem Holística para o Desenvolvimento de Sistemas de Informação Computadorizados. DataGramZero - Revista de Ciência da Informação - v.10 n.3 jun/09. Disponível em: <http://www.dgz.org.br/jun09/Art_03.htm>. Acesso em: 15 ago. 2011..



ANALISANDO AS QUESTÕES DE USABILIDADE E ACESSIBILIDADE DO PORTAL DE PERIÓDICOS DA CAPES

Sandra de Albuquerque Siebra, Jaciane Freire Santana, Denis Silva da Silveira

Resumo: O presente artigo analisa o Portal de Periódicos da CAPES sob a perspectiva da usabilidade e acessibilidade. Para teste da acessibilidade utilizou a ferramenta de avaliação *Hera*. Para averiguar a usabilidade recorreu-se à taxonomia de requisitos não-funcionais e testes com os usuários. Concluiu-se que o *site* apresenta os requisitos mínimos de acessibilidade e usabilidade, mas para tornar-se uma página plenamente utilitária e utilizada pela comunidade científica é necessário aprimorá-la.

Palavras-chave: Portal de Periódicos da Capes. Acessibilidade. Usabilidade.

1 INTRODUÇÃO

O Portal de Periódicos da CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior) está disponível para consulta há dez anos e tornou-se uma grande ferramenta para toda a comunidade científica brasileira e internacional que, através de suas bases de dados, acessam uma extensa gama de documentos científicos. Ao longo de sua existência, o portal cresceu exponencialmente em número de materiais dispostos, assim como em acessos pela comunidade acadêmica nacional. A biblioteca virtual do portal possui acervo com mais de 24.000 títulos com texto completo, 130 bases referenciais, bases dedicadas exclusivamente a patentes, estatísticas, entre outros (MEC, 2010) e a expectativa dos usuários é de que esse quantitativo venha a crescer ainda mais.

Neste cenário, aliar recursos de acessibilidade e usabilidade à interface do Portal de Periódicos da CAPES torna-se importante, na medida em que pode aproximar mais os usuários potenciais e reais para uso efetivo do mesmo. Isso porque quando estão realizando pesquisas, os usuários buscam respostas concretas e esperam interagir com o sistema de maneira rápida, simples e efetiva, sem necessitar de treinamentos ou da leitura de ajudas e manuais. Assim, este estudo teve por objetivo analisar a acessibilidade e usabilidade da interface do Portal de Periódicos da CAPES. Este portal foi escolhido por ser considerado estratégico para a pesquisa e desenvolvimento científico nacional. Esse tipo de avaliação torna-se interessante para a Ciência da Informação (CI), porque aos profissionais de tal área não basta apenas trabalhar com a disponibilização da informação – o que, por muito tempo acreditou-se que era suficiente para satisfazer a demanda social - mas a CI tem, também, a responsabilidade de opinar e interferir na maneira como a informação será apresentada, independente do formato (e.g.: digital ou analógica) e do meio de disponibilização (e.g.: *Web*). Assim, este estudo



se torna relevante, bem como os anteriormente realizados por Costa (2008) e Martinez *et al.* (2009) relacionados a usabilidade de versões anteriores da interface do Portal de Periódicos da CAPES.

O restante deste artigo está organizado como segue. A seção 2 trata do tema usabilidade e das regras de usabilidade que serão utilizadas para a realização do trabalho aqui proposto. A seção 3 trata do tema acessibilidade e de ferramentas que auxiliam a verificação de características de acessibilidade em sites Web. A seção 4 expõe a metodologia adotada para a realização deste trabalho. A seção 5 apresenta os resultados da realização do trabalho e as discussões a cerca dos mesmos e, finalmente, a seção 6 trás as considerações finais. Doravante neste trabalho o Portal de Periódicos da CAPES será chamado apenas de Portal.

2 USABILIDADE

O termo usabilidade está relacionado com a facilidade de aprendizado e uso da interface, bem como a satisfação do usuário em decorrência desse uso (NIELSEN; LORANGER, 2007). A norma ISO 9241 conceitua a usabilidade como a “*capacidade que um sistema interativo oferece a seu usuário, em determinado contexto de operação, para a realização de tarefas de maneira eficaz, eficiente e agradável*”. (ISO 9241, 1998 *apud* SANTOS, 2008). Para Nielsen e Loranger (2007, p.xvi) a usabilidade é um atributo de qualidade relacionado à facilidade de uso de algo. Mais especificamente, refere-se à rapidez com que os usuários podem aprender a usar alguma coisa, a eficiência deles ao usá-la, o quanto lembram daquilo, seu grau de propensão a erros e o quanto gostam de utilizá-la. Se as pessoas não puderem utilizar ou não utilizarem um recurso, ele pode muito bem não existir. Assim, especialmente no contexto da *Web*, a interface deve ser configurada de forma a não exigir sempre dos usuários um novo aprendizado para associar comandos e ações, tornando a interação e uso mais “intuitivos”. Isto é, processado inconscientemente, com um dispêndio menor de esforço cognitivo, facilitando assim a ação e minimizando a probabilidade de erros (SILVINO; ABRAHÃO, 2003).

Pensar na usabilidade e acessibilidade de determinado *site* é sinônimo de pensar no usuário, em qualquer etapa do projeto, seja no início, meio ou no fim. Quando não se pensa na usabilidade desde o princípio do projeto pode-se perder usuários, tempo e recursos para, posteriormente, corrigir os eventuais erros e inconsistências que emergirão. Silva alerta que:

a concepção de sistemas muitas vezes prioriza as exigências da informática antes de responder àquelas relacionadas ao usuário. Na maioria das vezes os profissionais de informática (projetistas) se empenham antes de tudo em definir as funções lógicas de um sistema sem de fato se preocuparem com as necessidades e habilidades físicas e cognitivas do usuário. (SILVA, 2007 *apud* IGNÁCIO; CARVALHO, 2008).

De acordo com Pressman, os requisitos não-funcionais de usabilidade da interface podem ser agrupados em duas categorias: os relacionados à entrada de dados e os relacionados à exibição da



informação. (PRESSMAN, 1992 *apud* FERREIRA; LEITE, 2003, p. 117). Os requisitos relacionados à exibição da informação são:

- **Consistência** - em um *site* na *Web*, ela é observada através da uniformidade dos comandos, funções e padrões regulares. Adicionalmente, a cor é um aspecto importante para garantir a consistência de uma página *Web*.
- **Feedback** - fornece uma resposta do sistema ao usuário, a toda e qualquer atividade desenvolvida por ele, seja uma confirmação de ação realizada, um alerta, uma sinalização de espera ou uma mensagem de erro.
- **Níveis de habilidade e comportamento humano** - cada indivíduo ao fazer uso de determinada página aplica seu referencial de conhecimentos no desenvolvimento de suas tarefas. Portanto, a página deve ser formatada de acordo com as diretrizes de acessibilidade, pois as mesmas garantem a acomodação de personalidades distintas dos mais experientes aos iniciantes e de forma independente de suas limitações. (FERREIRA; SILVEIRA; NUNES, 2009).
- **Percepção humana** - o cérebro humano reage a diversos estímulos, muitos deles externos. A observação do ambiente pode ser um estímulo, assim como a leitura de um texto, seja em formato analógico ou eletrônico. Por isso, deve haver a preocupação com a legibilidade da informação, que define características que podem facilitar ou dificultar a leitura da informação textual, por exemplo, brilho do caractere, espaçamento entre linha, comprimento da linha, entre outros (CYBIS, BETIOL, FAUST, 2010). Essa característica é importante para todos os grupos de usuário, mas, em especial, para as pessoas idosas e com dificuldades de visão. A leitura é uma atividade corriqueira e essencial em diversos contextos, e na *Web*, ela permanece com os mesmos padrões. O tamanho do texto, a fonte, a cor, os destaques (negrito, itálico, sublinhado) influenciam diretamente na percepção da informação pelo indivíduo.
- **Uso de metáforas** – de acordo com o Novo Dicionário Aurélio (2004) a metáfora é o emprego de palavra ou expressão em sentido figurado, que consiste na transferência de uma palavra para um âmbito semântico que não é o do objeto que ela designa, e que se fundamenta numa relação de semelhança subentendida entre o sentido próprio e o figurado. No desenvolvimento de interfaces é constante o uso de metáforas, por exemplo, para recortar um texto, geralmente utiliza-se um botão com o desenho de uma tesoura.
- **Minimização da carga de memória** – quando se faz uso de qualquer ferramenta, deve-se priorizar a disponibilização de recursos que remetam, de alguma forma, ao cotidiano dos indivíduos, a fim de minimizar a carga de memória dos mesmos.
- **Eficiência no diálogo, movimento e pensamentos** - Para projetar um *site* na *Web* é preciso pensar em todos os passos para desenvolvimentos das tarefas e para facilitar a leitura das informações. Para isso, as formas de diálogo da interface devem ser simples, objetivas e claras.



- **Classificação funcional dos comandos** – na exibição da página não é necessário o excesso informacional apenas para “encher” espaços vazios, deve-se priorizar o que é realmente importante. Também deve-se otimizar as opções de comando, para que não haja acessos dúbios de forma a confundir ou prejudicar a memorização do usuário.
- **Exibição exclusiva de informação relevante** – na exibição da página, não é necessário o excesso informacional apenas para “encher os espaços vazios”, deve-se priorizar o que realmente é importante. Também, deve-se otimizar as opções de comando, para que não haja acessos dúbios, de forma a confundir ou prejudicar a memorização do usuário.
- **Uso adequado de janelas** - as interfaces trabalham bastante com a utilização de janelas, no entanto é preciso cautela ao utilizar este recurso, para não sobrecarregar o usuário com diversas janelas simultaneamente.
- **Manipulação direta** – os objetos visuais (e.g. ícones) devem estar claramente visíveis durante as operações e ao término delas.
- **Uso de rótulos, abreviações e mensagens claras** - no desenvolvimento das páginas comumente é comum fazer uso de rótulos, abreviações e/ou mensagens indicativas. Ao fazer uso desses recursos deve-se procurar sempre a clareza e a objetividade. Textos do tipo “clique aqui” ou “siga” não são muito significativos neste contexto (FERREIRA; SILVEIRA; NUNES, 2009).
- **Projeto independente da resolução do monitor** – diversos fatores podem interferir na usabilidade de uma interface, um deles é a resolução do monitor (e.g. se o site foi desenvolvido para uma resolução específica de forma estática, os objetos visuais podem ficar em organizações duvidosas quando se usa uma resolução diferente da original). Assim, para que a resolução inadequada não interfira na qualidade gráfica do *site* é necessário definir a aparência dos componentes por meio de porcentagens do espaço disponível. Facilitando inclusive o acesso através de outras mídias, por exemplo, o aparelho celular.

Além destes, existem os requisitos relacionados à entrada de dados: **Mecanismo de ajuda, Prevenção de erros e Tratamento de erros** (FERREIRA; LEITE, 2003; FERREIRA; SILVEIRA; NUNES, 2009). Destes, neste trabalho, apenas foi avaliada a prevenção de erros que é a inibição de erros previsíveis do sistema. Em um bom projeto não se permite que o erro ocorra, por exemplo, se não há um objeto selecionado não há como copiar ou recortar algo no sistema, pois tais funções só podem ser executadas a partir de uma pré-seleção.

3 ACESSIBILIDADE

Comumente acesso e acessibilidade são tratados como sinônimos, no entanto, não são. Acesso é disponibilizar informação, porém a informação pode estar disponível, mas não acessível para



todos. Este é o ponto chave da acessibilidade: disponibilizar informação para todos os indivíduos independentes de suas limitações.

Segundo a Organização Mundial de Saúde (OMS), cerca de 314 milhões de pessoas apresenta algum grau de deficiência visual e 45 milhões são cegas. Mais de 80% desse contingente se encontra em países que estão em desenvolvimento (OMS, 2011). No Brasil, de acordo com o Instituto Brasileiro de Geografia e Estatística - IBGE o Censo 2000 revela que cerca de 14,5% da população possui algum tipo de deficiência¹ (IBGE, 2000). Desta forma, necessitam de alguma tecnologia assistiva (qualquer objeto, *software*, equipamento concebido para ajudar as pessoas deficientes ou com alguma limitação física de executar as atividades do cotidiano).

Ignácio e Carvalho (2008) apontam que acessibilidade refere-se à transposição de barreiras. Aplicada à Internet isto significa a eliminação de quaisquer obstáculos que impossibilitem ou dificultem o acesso das pessoas com deficiência, ao mesmo tempo em que pode beneficiar pessoas sem limitação. Isto é, corresponde à possibilidade de qualquer indivíduo, utilizando qualquer tipo de tecnologia de navegação (navegadores textuais, gráficos, especiais para cegos) ser capaz de acessar qualquer *site* e obter o entendimento completo da informação contida nele, além de ser possível total e completa habilidade de interação (DIAS, 2003; SONZA, 2008).

Para tratar as questões da acessibilidade na *Web* foram criadas diretrizes internacionais que, atualmente, estão em sua 2ª versão para a *Web 2.0: Web Content Accessibility Guidelines WCAG 2.0*. Tais diretrizes foram elaboradas em conjunto pelo consórcio *World Wide Web* (W3C, 2008), *Web Accessibility Initiative* (WAI), instituições e governos de diversos países. As diretrizes orientam a construção e o desenvolvimento de *sites* na *Web*. O consórcio desenvolve tecnologias interoperacionais (especificações, diretrizes, programas e instrumentos), a fim de possibilitar à comunidade *Web* desenvolver todas as suas potencialidades como fórum de informação, de comunicação, de comércio e de entendimento coletivo. A WAI, juntamente com organizações nacionais, objetivam tornar a *Web* acessível através de cinco domínios: tecnologia, diretrizes, instrumentos, educação e difusão, investigação e desenvolvimento (ALVES; ESTEVES, 2005).

O ajuste da usabilidade e acessibilidade permite o uso eficiente, produtivo, satisfatório, útil e seguro para o usuário do computador e destinatário da informação do conhecimento disponibilizado na rede. Ou seja, as condições de usabilidade e de acessibilidade podem assegurar que a informação disponibilizada na Internet possa ser compreendida e utilizada por todos, independente de limitações (IGNÁCIO; CARVALHO, 2008).

De acordo com Dias (2003) e Cybis, Betiol e Faust (2010) a acessibilidade é uma categoria de usabilidade. Um *software* ou *site* que não é acessível a uma determinada pessoa tão pouco pode ser considerado eficaz e eficiente. Assim como os fatores de usabilidade, a acessibilidade está intrinsecamente relacionada ao ambiente operacional de uso e às tarefas, necessidades e preferências dos usuários típicos.

¹ Maiores informações sobre tipos de deficiência podem ser encontradas em (QUEIROZ, 2011)



Para realizar a avaliação da acessibilidade em sites dispõem-se de algumas ferramentas automatizadas, tais como o *Hera* que foi desenvolvido por Carlos Benavídez, com a colaboração de Emmanuelle Gutiérrez y Restrepo e Charles McCathieNevile, especialmente para a Fundação Sidar. O *Hera* está disponível para a comunidade de desenvolvedores e projetistas assim como para o público em geral, gratuitamente, facilitando a revisão da acessibilidade das páginas *Web* (BENAVIDEZ, 2005).

O *Hera* é um avaliador de *software* que detecta um código HTML e faz uma análise do seu conteúdo, verificando se está ou não inserido no conjunto de recomendações das diretrizes de acessibilidade para o conteúdo da *Web* 1.0 – WCAG 1.0. O *Hera* efetua uma análise automática prévia da página, disponibilizando assim informações dos erros e inconsistências que merecem ajustes. As diretrizes são organizadas de acordo com três níveis de prioridade. A primeira prioridade trata das necessidades básicas que os desenvolvedores têm que satisfazer plenamente. A satisfação dessas prioridades é um requisito básico para que determinados grupos possam acessar documentos disponíveis na *Web*. Na segunda prioridade estão características que os desenvolvedores devem implementar para que não haja barreiras significativas ao acesso aos documentos eletrônicos. E, por fim, a prioridade de nível três que apresentam características desejáveis, que se efetivadas podem melhorar o acesso e a acessibilidade para quem possui algum tipo de deficiência.

4 METODOLOGIA

A análise da usabilidade foi direcionada a duas vertentes, sendo a primeira delas os requisitos de *software*, mais especificamente os requisitos não-funcionais de usabilidade que tratam da qualidade do sistema, descrevem as suas facilidades e estão relacionados, intrinsecamente, a aspectos negligenciados na engenharia de *software*: os fatores humanos. O desleixo com os fatores humanos constitui uma das principais razões da insatisfação dos usuários em relação ao produto (YEH, 1984; CHUNG, 1995 *apud* FERREIRA; LEITE, 2003, p. 117).

A taxonomia de Ferreira e Leite (2003) serviu como base para a análise da usabilidade do Portal. E adicionalmente foram realizados testes com os usuários. De que de acordo com Nielsen e Loranger (2007), na realização de testes com usuários, cerca de cinco avaliadores seriam suficientes, uma vez que, a partir do sexto avaliador, os erros começariam a se repetir. Logo, com um universo de cinco avaliadores é possível obter de 80% a 90% dos erros. Seguindo o critério de Nielsen, o teste da interface do Portal foi realizado com estudantes de graduação da Universidade Federal de Pernambuco - UFPE que, inicialmente, se prontificaram a participar do treinamento no uso do Portal de Periódicos da CAPES. Esse treinamento é realizado uma vez por mês em todas as bibliotecas do sistema da UFPE. O convite para participar do teste de usabilidade foi enviado para onze discentes inscritos no treinamento a ser realizado no dia 07/07/2011, no Centro Acadêmico de Vitória de Santo Antão e houve o retorno positivo de três deles.



Então, foi agendado o teste de usabilidade para o dia 22/06/2011. Durante o teste foi, inicialmente, solicitado que cada usuário realizasse quatro atividades no Portal: 1) buscar dois títulos de periódicos; 2) buscar duas bases de dados; 3) pesquisar a indicação das aquisições recentes inseridas na coleção e 4) realizar o cadastro de identificação pessoal no Portal. Essas atividades foram realizadas individualmente e foram acompanhadas por um observador que se valeu do método “pensando em voz alta”. Tal método consiste em uma observação direta de cada usuário utilizando o *site*. De acordo com Nielsen “*ouvir os ‘pensamentos’ de cada usuário permite entender as razões de suas ações e essas informações são muito importantes*”. (NIELSEN, 2007, p. 6). Ao final do teste, também foi passado um questionário para os participantes para eles pontuarem sua experiência de usabilidade no uso do Portal. O questionário aplicado foi baseado em um modelo utilizado pelo Núcleo de Formação Continuada de Profissionais da Educação – NUPE/ UNISSINOS (2011). Já para a análise da acessibilidade do Portal utilizada a ferramenta *Hera*.

5 RESULTADOS E DISCUSSÕES

Nesta seção serão apresentados os resultados das avaliações de usabilidade e acessibilidade, conforme a metodologia descrita na seção anterior.

5.1 Análise da usabilidade

A taxonomia de Ferreira e Leite (2003) serviu como base para a análise da usabilidade do portal e foram observadas as seguintes incongruências:

Consistência - No portal é possível visualizar alguns *links* externos, prática corrente no desenvolvimento das páginas da *Web*. Entre os *links* externos dispostos na página do portal encontra-se um que direciona para a Escola Britânica *Online* e este difere, substancialmente, do conteúdo do portal e não é de interesse direto do perfil dos usuários do mesmo, uma vez que ela é uma página de conteúdo para alunos de ensino fundamental e médio. Outro *link* permanente da página aponta para o BDTD – Banco Digital de Teses e Dissertações que ficou durante bastante tempo inativo e o link continuou durante todo o tempo a ser oferecido, direcionando a um link inativo (apontando uma página de erro). A escolha das cores foi bastante positiva no portal. Para testar o contraste da página imprimiu-se a página inicial em preto e branco e foi possível visualizar todas as informações. Tal fato mostra que a informação apresentada é passível de ser verificada independente da cor. No entanto, o mesmo não acontece em todas as seções. Na seção chamada “meu espaço” há uma informação que é quase invisível por conta de sua cor (vide Figura 1, texto circulado). O texto diz: “Caso não seja um usuário identificado”. Essa informação poderia ser lida mais facilmente se estivesse configurada como as demais informações similares, na cor verde.



Figura 1: Seção “meu espaço”. (Extraído de: CAPES, 2010).

Feedback - Na interface da página inicial do portal há bons exemplos de *feedback* tal como a indicação exata do usuário em relação a página inicial, isto é a seção em que ele se encontra e o caminho percorrido para tal fim. No entanto, esta sinalização só é apontada de forma superficial, uma vez que ela apenas indica o *menu* principal, mas as subseções não são identificadas; ao pousar o *mouse* sobre um *link*, o mesmo muda o formato de seta para “mãozinha” indicando assim que através da imagem ou texto é possível direcionar a navegação para outro local. Seria interessante haver um realce do ícone pela mudança de cor, destacando essa informação que pode ser tão útil ao usuário. Na seção “Meu espaço” há um formulário de preenchimento para os novos usuários e os campos obrigatórios estão identificados com um asterisco (*), o que é um padrão recorrente no desenvolvimento de páginas *Web*. No entanto, como praticamente todos os campos são obrigatórios e apenas um não mantém essa característica, seria mais explicativo e intuitivo que esta informação fosse apresentada inicialmente, de forma textual. Até porque, no tocante a acessibilidade, ao utilizar um símbolo para fornecer informações importantes, os usuários que dependem dos leitores de tela podem não conseguir obter tal informação (por exemplo, o leitor não leria o asterisco que marcaria os campos obrigatórios, mas poderia ler se a informação estivesse em formato textual no início do formulário).

Níveis de habilidade e comportamento humano - Na seção seguinte será apresentada a avaliação de acessibilidade que apresentará mais informações referentes a esse item. Mas, já pode ser dito que, entre outros fatores, é importante que o texto indique o idioma principal do documento e possibilite a permuta do idioma. Na página inicial do portal não se visualiza esta opção. No entanto, o código fonte do portal, sem as folhas de estilo, apresentam o “*lang tag*”, atributo que permitiria essa funcionalidade. A Figura 2 ilustra o mapa da página inicial do portal, com destaque para as caixas de idioma. O atributo “*lang tag*” é importante para a acessibilidade uma vez que ele aponta o idioma utilizado no documento, ou em partes dele, e os leitores de tela reconheceriam, de

forma automática, esse atributo e escolheriam o sintetizador de voz mais adequado para a leitura do trecho ou documento como um todo (FERREIRA; SILVEIRA; NUNES, 2009).



Figura 2: Código-fonte do mapa da página, sem as folhas de estilo. (Extraído de: *Hera* 2.1 Beta)

Percepção humana - A página do portal permite alterar o contraste da cor para ajudar pessoas com dificuldade de entendimento da informação ou se adequar ao gosto do usuário (vide Figura 3) A idéia seria válida se houvesse consistência no *layout* com a mudança da cor. No entanto, a página não mantém um padrão de contraste, fato que pode ser verificado na seção “Coleção”, da página inicial, onde a cor da fonte não foi alterada, permanecendo preta, mesmo a cor de fundo tendo sido alterada para uma cor escura, o que torna a leitura do texto difícil, assim como, na seção “Meu Espaço” (vide Figura 4).



Figura 3: *Layout* da página do Portal com uso da ferramenta de contraste. (Extraído de: CAPES, 2010)

Uso de Metáforas - Na página do portal não há uso de metáforas.

Minimização da carga de memória – Para diminuir a carga de memória e não ter sempre a obrigação de efetivar buscas, após cadastro no portal, a opção “meu espaço” permite o armazenamento dos resultados das buscas frutíferas para acesso posterior.



Figura 4: Layout da seção “Meu espaço” com uso da ferramenta de contraste. (Extraído de: CAPES, 2010).

Eficiência no diálogo, movimentos e pensamentos - A página do portal apresenta todas essas características, no entanto, há um *banner* com a informação “RBPG” apenas. Trata-se de uma sigla ainda não consagrada e incorporada pelo grande público, que significa: Revista Brasileira de Pós-Graduação. Para ser mais acessível e usual o ideal seria usar o nome completo da revista.

Exibição exclusiva da informação relevante - Essa premissa é em grande parte atendida na página do portal. No entanto, na seção “Meu espaço” (novo usuário) é apontado como vantagem o acesso a “Estatísticas”, “Treinamento e uso do Portal” e “Conteúdo personalizado”. Mas para se ter acesso a estatísticas basta acessar o *menu* “Institucional” → “Estatística de uso”, que direcionará para a mesma página. O mesmo acontece com os outros *links*: o “Conteúdo personalizado” tem o mesmo conteúdo de “Novas aquisições” e para o treinamento a mesma página pode ser acessada a partir de “Suporte” → “Treinamentos”.

Uso adequado de janelas - A página do portal utiliza este recurso para direcionar a páginas com domínios distintos de forma bastante satisfatória.

Os requisitos **Classificação funcional dos comandos, Manipulação direta; Uso de rótulos, abreviações e mensagens claras e Projeto independente da resolução do monitor** não foram avaliados dentro do escopo deste trabalho. Porém, além dos requisitos relacionados à exibição da informação foi analisado, adicionalmente, um requisito referente à entrada de dados: **a Prevenção de erros** - No preenchimento do formulário de novo usuário no *menu* “Meu espaço”, quando se preenche o campo relacionado ao vínculo institucional, é aberta uma nova caixa de diálogo para ser inserida a instituição a qual o usuário está vinculado e para confirmação do vínculo deve-se utilizar o botão “Incluir”. Porém, o uso desse botão “Incluir” é confuso e a sua localização espacial é próxima do botão “Enviar” para cadastrar os dados completos (vide Figura 5). Durante os testes realizados com os usuários percebeu-se insegurança por parte deles sobre qual botão clicar, pois eles acharam os botões confusos. Os usuários supunham que, ao clicar sobre o botão incluir, o cadastro também seria finalizado. Uma solução seria que, ao confirmar um vínculo institucional, uma caixa de diálogo fosse



aberta para incluir outro vínculo institucional (comum entre docentes) ou finalizar o cadastro.

Uma função importante que está ausente no portal e que sentiu-se falta durante os testes foi o cancelamento de tarefas em andamento. Por exemplo, durante uma busca, não é possível interromper a operação de forma direta, isto é, não há um ícone ou comando para efetivar o cancelamento.

Adicionalmente, a partir da análise dos testes de usabilidade percebeu-se que havia uma dificuldade por parte dos discentes em assimilar a metabusca, o seu significado e sua utilidade. Isso tornou-se ainda mais perceptível porque, em praticamente todas as atividades, os usuários recorreram a metabusca e, não obtendo sucesso no uso da mesma, buscavam outro caminho.

Nome da instituição	Vínculo	Nível
UNIVERSIDADE FEDERAL DE PERNAMBUCO	Docente	Mestrado

Figura 5: Layout do menu “Meu espaço” com a opção de vínculo institucional Extraído de: CAPES, 2010

5.2 Análise da Acessibilidade

A partir dos resultados obtidos com a utilização do *software Hera*, tratar-se-á alguns tópicos que mostram negligências no tocante à acessibilidade da página do portal. Vale ressaltar que as recomendações propostas pelo Consócio W3C estão em sua segunda versão que é focada na *Web 2.0*. No entanto, a ferramenta utilizada na presente pesquisa, assim como as demais disponíveis na *Web*, ainda estão baseadas na *Web 1.0*.

Erros de Prioridade 1

O avaliador *Hera* aponta para um erro grave: o não fornecimento de um equivalente textual para cada imagem disposta na página. Isso é importante porque quando se acessa o portal através de um leitor de tela o mesmo não conseguirá identificar do que se trata a imagem e poderá deixar de fornecer informação relevante ao usuário. No portal esse erro é grave porque além das imagens convencionais não terem equivalente textual, alguns títulos e rótulos do portal estão também no formato de imagem. Por exemplo, a Figura 6 ilustra algumas situações onde imagens são utilizadas em lugares onde um texto convencional poderia ser muito bem empregado, estendendo assim a acessibilidade a todos.

Outras imagens dispostas no portal são os *banners*. Estes apresentam o atributo “alt”, no entanto, a informação neste atributo é apenas de “Banner” (vide Figura 7), não especificando para qual *link* direciona, dificultando o acesso às pessoas com deficiência visual, pois as mesmas só saberão para onde o *link* do banner está levando, após clicar sobre o mesmo, o que pode ocasionar uma ida a um lugar indesejado. O conteúdo do atributo “alt” é lido pelos leitores de tela e informaria do que trata o *banner*, antes do mesmo ser clicado, apenas passado o mouse sobre o ele.



Figura 6: Textos no formato de Imagens dispostas no Portal de Periódicos da Capes. Extraído de: CAPES, 2010

Erros de prioridade 2

A página inicial do portal apresenta cinco *links* que são ativados mediante *scripts*². Quatro deles possibilitam uma forma de personalização que pode tornar a página mais adequada a determinadas necessidades, por exemplo, com a modificação do contraste da página ou do tamanho da fonte. Porém, o quinto *link* envolve a ativação de uma janela de ajuda ao usuário chamada “Ache sua área” (vide Figura 8). Esta janela tem como finalidade auxiliar o usuário a localizar uma área do conhecimento de seu interesse e, após essa localização, as buscas serão realizadas em um conjunto de bases pré-selecionadas relacionadas a área do conhecimento escolhida. A função “Ache sua área” é informativa e seu uso não é obrigatório. No entanto, por se tratar de uma função de ajuda deveria ser acessível a todos. Mas o uso de *script* inviabiliza sua universalidade, pois não existe um equivalente textual que possa ser lido pelos leitores de tela que mencione a não obrigatoriedade de uso dessa função ou que explicito o seu conteúdo.

Outros pontos a ressaltar são: de acordo com a avaliação do *Hera* a página do portal apresenta notações obsoletas do padrão W3C (2008). Por exemplo, a linguagem HTML é recomendada pelo consórcio W3C, porque não exige complementos (*plug-ins*) para a leitura de seu conteúdo, característica oposta dos arquivos PDF, que necessitam de ferramenta ou complemento

² Os *scripts* são utilizados para criar pequenos programas encarregados de realizar ações dentro do âmbito de uma página *Web*.



apropriado. No entanto, a linguagem HTML está na quinta versão e o portal apresenta 10 atributos considerados em desuso desde a versão 4.1, o que pode causar, em um futuro próximo, problemas na apresentação das informações.



Figura 7: Código-fonte dos links apresentados no Portal de Periódicos da Capes Extraído de: *Hera*, 2010

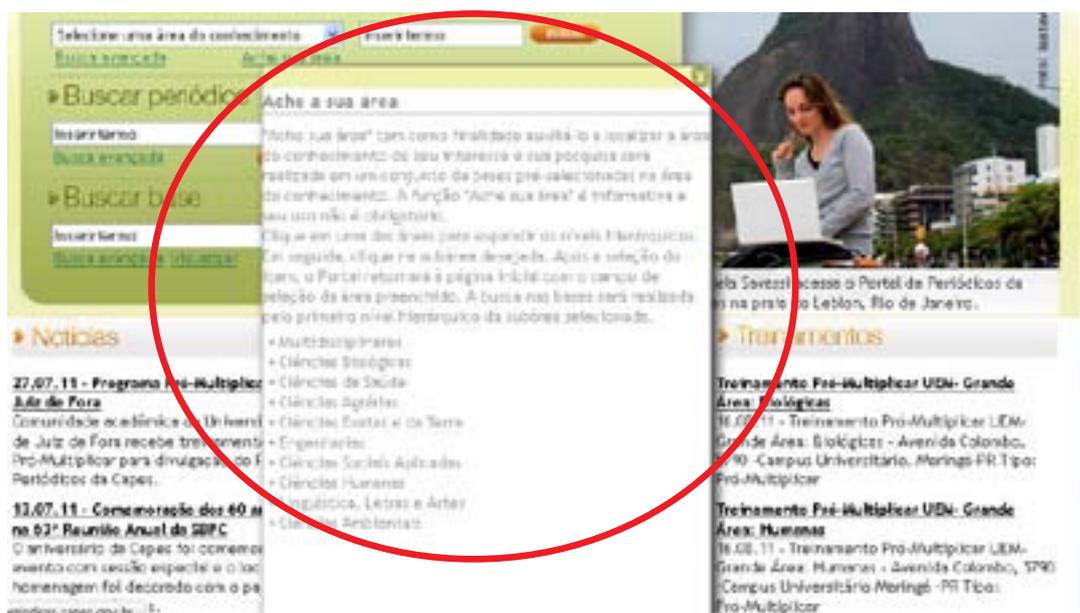


Figura 8: *Layout* da função “ache sua área”. Extraído de: CAPES, 2010



A diretriz 6.4 da W3C dispõe que em programas interpretáveis e *applets* deve-se assegurar que a resposta a acontecimentos seja independente do dispositivo de entrada, ou seja, a resposta deve poder ser acessada pelo *mouse* ou pelo teclado. O principal *menu do site* não está configurado para atender a esta diretriz. A Figura 9 ilustra o *menu* principal da página que é caracterizado como um *menu pull-down* na forma de um *applet* que é caracterizado pela discrição, uma vez que se apresenta apenas quando é acionado, por exemplo, quando o *mouse* está sobre o item. Este tipo de *menu* que o *site* recorre, economiza espaço na tela e é utilizado quando se deseja um *layout* mais limpo. O problema está no fato de que leitores de tela não conseguem determinar a existência do *submenu* quando este recurso é utilizado (Feliciano, 2010, p. 56). Por se tratar do *menu* principal é recomendado dispor seu acesso por mais de uma possibilidade.



Figura 9: *Menu pull-down*. (Extraído de: CAPES, 2010).

Erros de Prioridade 3

O avaliador *Hera* não identificou teclas de atalho de navegação no portal. Essas teclas são bastante utilizadas pelos indivíduos incapazes de visualizar a tela, uma vez que podem otimizar a navegação facilitando que se saltem as partes indesejadas ao usuário, proporcionando uma maior usabilidade.

Uma falha observada, independente da ferramenta de avaliação, foi que, ao ampliar a fonte da página para uma melhor visualização por parte de pessoas com baixo nível de visão, o *layout* muda bastante de forma inapropriada (vide Figura 10). Por exemplo, as letras O e P, por exemplo, da pesquisa alfabética ficam sem espaço tornando-se obscuras, escondidas atrás da imagem (vide Figura 10). Além disso, uma inconsistência foi observada: a coluna intitulada “Coleções” não aumenta o tamanho da fonte, como o restante da página, permanecendo o padrão anterior, sem aumento. Assim, uma opção disponibilidade para promover uma maior acessibilidade não satisfaz plenamente seus objetivos.



Figura 10: Página do Portal de Periódicos com a fonte ampliada. Extraído de: CAPES, 2010

6 CONSIDERAÇÕES FINAIS

A interface do Portal de Periódicos da CAPES é muito interessante, limpa visualmente, no entanto negligencia alguns aspectos básicos da usabilidade e diversos de acessibilidade.

A partir dos resultados obtidos nos testes e da análise de usabilidade feita pode-se inferir que a página do portal é assimilável. No entanto, a facilidade da busca pode ser observada apenas quando já se sabe o que pesquisar e, no universo da pesquisa, já há um título específico. Para pesquisar material sobre determinada área, sem o conhecimento explícito do que deve ser procurado, pode haver alguns entraves quanto à estratégia de busca como, por exemplo, o uso de termos inapropriados, dificultando a pesquisa. Outro ponto observado foi que na língua inglesa são obtidos mais resultados do que com a busca com o texto em português podendo configurar mais um empecilho para o usuário comum.

O portal está em funcionamento há mais de uma década e, em 2010, houve uma mudança significativa de sua interface, permitindo aos usuários terem acesso a mais recursos de busca e também houve uma melhoria na usabilidade. Porém, a acessibilidade permaneceu com falhas. De acordo com o avaliador *Hera*, a versão anterior do portal apresentava um erro de prioridade 1, seis erros de prioridade 2 e dois erros de prioridade 3. Já a versão atual apresenta um quantitativo maior de erros: três erros de prioridade 1, nove erros de prioridade 2 e dois erros de prioridade 3, ou seja, a versão atual, mesmo modificada e melhorada tornou-se menos acessível.



Assim sendo, em especial os problemas de acessibilidade precisam ser revistos e ajustados, a fim de tornar a navegação uma experiência rápida, simples e agradável para usuários comuns e com necessidades especiais. Isso poderá fazer do Portal de Periódicos da CAPES uma ferramenta completa, fácil de utilizar e referência entre toda a comunidade científica.

Abstract: *This article examines the CAPES Periodicals Portal, from the perspective of usability and accessibility. To test the accessibility, it was used the tool Hera. To ascertain the usability, it was appealed to the taxonomy of non-functional requirements and user testing. It is concluded that the site presents the minimum requirements of accessibility and usability, but to become a fully utilitarian page and used by the scientific community is needed to improve it.*

Key-words: Capes Periodicals Portal. Accessibility. Usability.

REFERÊNCIAS

ALVES, P. A.; ESTEVES, R. A. + Acesso: ferramenta para análise da acessibilidade em sítios web. In: CONGRESSO DE TECNOLOGIAS DO SOFTWARE LIVRE, 1., 2005, Coruña, Spain. **Anais...** Coruña, p.3-11. Disponível em: <http://bibliotecadigital.ipb.pt/bitstream/10198/1765/1/2005_%2baccesso%20corunha.pdf>. Acesso em: 10 jun. 2011.

Benavídez, Carlos. **Informação sobre o Hera**. 2005. Disponível em: <<http://www.sidar.org/hera/index.php.pt?ini=info>>. Acesso em: 01/07/2011.

CAPES. **Portal de Periódicos da Capes**. 2011. Disponível em: <<http://www.periodicos.capes.gov.br/>>. Acesso em 10 jun. 2011.

COSTA, Luciana F. da. **Usabilidade do Portal de Periódicos da Capes**. 2008. 236 f. Dissertação (Mestrado em Ciência da Informação) – Centro de Ciências Sociais Aplicadas, Universidade Federal da Paraíba, João Pessoa, 2008. Disponível em: <<http://dci2.ccsa.ufpb.br:8080/jspui/bitstream/123456789/99/3/Disserta%C3%A7%C3%A3o%20Luciana%20Costa.pdf>>. Acesso em: 10 jul. 2011.

CYBIS, Walter de Abreu; BETIOL, Adriana Holtz; FAUST, Richard. **Ergonomia e usabilidade: conhecimentos, métodos e aplicações**. 2.ed. rev. ampl. São Paulo: Novatec, 2010. 422 p.

DIAS, C. **Usabilidade na web: criando portais mais acessíveis**. Rio de Janeiro: Alta Books, 2003. 312p.

FELICIANO. Fábio Denilson de O. **Investigação de um modelo de avaliação da acessibilidade de portais na web**. 2010. 110 f. Dissertação (Mestrado Profissional em Ciência da Computação) – Centro de Informática, Universidade Federal de Pernambuco, Recife, 2010.



FERREIRA, Simone B. L.; LEITE, Julio C. S. do P. Avaliação da usabilidade em sistemas de informação: o caso do sistema submarino. **Revista de Administração Contemporânea – RAC**, São Paulo, v.7, n.2, p.115-137, abr./jun. 2003. Disponível em: <http://www.anpad.org.br/rac/vol_07/dwn/rac-v7-n2-sbf.pdf>. Acesso em: 10 jun. 2011.

FERREIRA, Simone B. L.; SILVEIRA, Denis S.; NUNES, R. R. Alinhando os requisitos de Usabilidade com as Diretrizes de Acessibilidade. **Chronos** (UNIRIO), Rio de Janeiro, v. 5, p. 33-48, 2009.

HERA: Folhas de Estilo para Revisão da Acessibilidade. **Avaliador de acessibilidade Hera**. 2011. Disponível em: <http://www.sidar.org/ex_hera/index.php.pt>. Acesso em: Jun. 2011.

IGNÁCIO, Edilson A.; CARVALHO, José O. F. de. Avaliação da acessibilidade de sites oficiais de pesquisa no Brasil por pessoas com deficiência. **Encontros Bibli**, Florianópolis, v. 13, n.26, 2008. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/7193/6641>>. Acesso em: 10 jun. 2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA ESTATÍSTICA. Censo demográfico do Brasil ano 2000. Disponível em: <<http://www.ibge.gov.br/home/presidencia/noticias/>>. Acesso em 20 jun. 2011.

MARTINEZ, M. L. et al. Estudo de usabilidade do portal de periódicos da CAPES: análise do perfil do usuário discente da UFPE. In: ENANCIB, 10., 2009, João Pessoa, PB. **Anais...** 2009. Disponível em: <<http://dei2.ccsa.ufpb.br:8080/jspui/bitstream/123456789/411/3/Corrigido%20GT%208%20Txt%20%20MARTINEZ,%20Maria%20Laura%20et%20al.%20Estudo%20de%20usabilidade.pdf>>. Acesso em: 10 jul. 2011.

MEC. **Portal de Periódicos**. 2010. (Livreto Institucional).

METÁFORA. In: FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário Aurélio da língua portuguesa**. 3.ed. ver. e atual. Curitiba: Positivo, 2004. p.1320.

NIELSEN, J. **Usability Engineering**. New York, NY: Academic Press, 1993.

_____. ; LORANGER, Hoa (Colab.). **Usabilidade na web**. Rio de Janeiro: Elsevier, 2007. p.ix-17.

OMS – Organização Mundial de Saúde. **Visual impairment and blindness**. 2011. Disponível em <<http://www.who.int/mediacentre/factsheets/fs282/en/>>. Acesso em 24 de Abril de 2011.

QUEIROZ, Marco. **Bengala Legal**. Disponível em <<http://www.bengalalegal.com/>> Acesso em 06 de Fevereiro de 2011.

SANTOS, Rodrigo C. dos. Revisão das métricas para avaliação de usabilidade de sistemas. In: GLOBAL BUSINESS AND TECHNOLOGY ASSOCIATION CONFERENCE, 2008, Madrid. **Anais...** Espanha: GBTA, 2008. Com o título original de: Systems Usability Evaluation Metrics Review.



Disponível em: <http://www.marcelomoraes.com.br/conteudo/marcelo/metricas_usabilidade.pdf>.
Acesso em: 10 jun. 2011.

SILVINO, Alexandre M. D.; ABRAHÃO, Júlia I. Navegabilidade e inclusão digital: usabilidade e competência. **RAE**, v.2, n.2, jul./dez. 2003. Disponível em: <<http://www.rae.com.br/eletronica/index.cfm?FuseAction=Artigo&ID=1808&Secao=CIENCIA&Volume=2&Numero=2&Ano=2003>>.
Acesso em: 15 jun. 2011.

SONZA, Andréa Poletto. **Ambientes virtuais acessíveis sob a perspectiva de usuários com limitação visual**. 2008. 298f. Tese (Doutorado em Informática na Educação) - Centro Interdisciplinar de Novas Tecnologias na Educação, Universidade Federal de Rio Grande do Sul, Porto Alegre, 2008. Disponível em: <http://www.redenet.edu.br/geral/siep_arquivos/Tese_Maio2008.pdf>. Acesso em: 10 jun. 2011.

UNISINOS. NUPE. **Questionário de avaliação da usabilidade**. Disponível em: <https://www.unisinos.br/nupe/_dados/Questionario.htm>. Acesso em: 10 jun. 2011.

W3C. **Recomendações de Acessibilidade para Conteúdo Web (WCAG) 2.0**. 2008. Tradução de Everaldo Bechara. Disponível em: <<http://www.ilearn.com.br/TR/WCAG20/>>. Acesso em: 10 jun. 2011.



ARQUITETURA DA INFORMAÇÃO E ENGENHARIA SEMIÓTICA: UM ESTUDO DE CASO DA INTRANET DA UNIMED JOÃO PESSOA

*Lilian Viana Teixeira Cananéa, Guilherme Ataíde Dias, Marckson Roberto Ferreira de Sousa,
Maria Amélia Teixeira da Silva*

Resumo: O artigo discorre sobre a aplicação de conceitos de Arquitetura da Informação e do Método de Avaliação de Comunicabilidade (MAC), utilizado pela Engenharia Semiótica, em um ambiente *Intranet*. A ascensão das Tecnologias Digitais de Informação e Comunicação (TDIC), que idealmente deveriam representar mais facilidade na geração de conhecimento, pode ocasionar dificuldades para a compreensão das informações, em virtude do pouco tempo que os usuários dispõem para interpretá-las. Com a finalidade de colaborar com estudos na área e propor possíveis soluções para uma melhor compreensão da informação disponibilizada, especialmente em ambientes institucionais, foi realizado um estudo considerando a Arquitetura da Informação e o MAC. Para validar a proposta, foi desenvolvida uma pesquisa sobre as necessidades de informações dos usuários em um ambiente de *Intranet* na Unimed João Pessoa, com a realização de uma coleta de dados que contou com a participação de 30 colaboradores, por meio de 5 grupos focais, caracterizando um estudo exploratório, com natureza quanti-qualitativa. Os resultados obtidos pelos métodos indicaram que cada um tem uma contribuição relevante e diferenciada para questões de Interação Humano-Computador (IHC). Considerando aspectos da Arquitetura da Informação, foram identificadas deficiências na organização e navegação da *Intranet*, enquanto que, através do Método de Avaliação de Comunicabilidade da Engenharia Semiótica, foi possível identificar as questões de comunicabilidade, relacionadas ao aprofundamento do conhecimento da mensagem de metacomunicação para o usuário. Os resultados do estudo são fortes indicativos de que muito se precisa avançar na concepção de melhores interfaces, evidenciando assim, a necessidade de maiores estudos das disciplinas em questão.

Palavras-chave: Arquitetura da Informação. Engenharia Semiótica. Método de Avaliação de Comunicabilidade. *Intranet*. Comunicação com o usuário.

1 INTRODUÇÃO

Atualmente as Tecnologias Digitais de Informação e Comunicação (TDIC) estão mudando não apenas as formas do entretenimento e do lazer, mas todas as esferas da sociedade. Esse pensamento é compartilhado por Santaella (2003), uma vez que acessar, em tempo real, informações sobre quase tudo que existe no mundo e estabelecer contato direto com as mais diversas fontes de informações, representa uma mudança de paradigma. Essa mudança pode ocorrer no trabalho (robótica e tecnologias para escritórios), nas atividades militares e de segurança pública (a guerra eletrônica), no consumo



(transferência de fundos eletrônicos), na comunicação e na educação (aprendizagem à distância), além de outros, influenciando o processo de mudança da cultura mundial.

Nesse contexto, a quantidade de informação produzida tem se elevado de forma significativa, conduzindo ao que Wurman (1991) denomina de “ansiedade de informação”, ocasionada principalmente pelo desafio de transformar informação em compreensão. Mais informação deveria representar maiores oportunidades para uma compreensão ampla do mundo, porém nem sempre a informação disponível atende as necessidades dos usuários.

Para tentar diminuir as barreiras entre o usuário e a informação, a Arquitetura da Informação (AI) tem buscado apresentar meios de estruturar o conteúdo disponível em ambientes informacionais digitais. De acordo com Morville e Rosenfeld (2006), a AI consiste na combinação entre esquemas de organização, rotulação, navegação, busca e representação de um sistema de informação, visando compreender e atender a três dimensões de variáveis para organizar as mesmas: usuários (suas necessidades, hábitos e comportamentos), conteúdo (volume, formato, estrutura – ou seja, o que será apresentado) e contexto de uso do sistema (objetivo do *site*, cultura e política da empresa, restrições tecnológicas, localização, entre outros).

Focando sua atividade na interação usuário-sistema, a AI possibilita várias soluções para os problemas do usuário, determinados pelas interpretações do projetista e suas decisões. Ao navegar pelas páginas, o usuário, muitas vezes, não percebe que as interpretações do projetista podem não ser ideais (ou até mesmo errôneas) e tratam o sistema como se ele fosse infalível ou como se ele (usuário) estivesse enganado quanto ao seu uso.

Aliado a isso, a Engenharia Semiótica (EngSem), que possibilita a compreensão da Interação Humano-Computador (IHC), privilegia os métodos qualitativos, e analisa os fenômenos envolvidos no projeto, uso e avaliação de um sistema interativo (DE SOUZA, 2005). Dessa forma, ao se analisar uma página na *Internet/Intranet* pela teoria da Engenharia Semiótica, juntamente com a pesquisa e conhecimentos sobre AI, torna-se possível sugerir representações que facilitem o entendimento e a navegação nos *sites*. A Engenharia Semiótica, que pode ser aplicada em diferentes fases do projeto, tem como foco passar a mensagem do que o sistema faz, permitindo que o usuário interprete de forma simples e direta o que o projetista expõe através da interface e a forma de comunicação do usuário-sistema, através de sons, símbolos, ícones, entre outros.

Na *Intranet* da Unimed JP, objeto de estudo desta pesquisa, correspondente a uma cooperativa médica formada por cerca de 1400 colaboradores, entre sua sede administrativa e seu hospital, eram percebidos problemas de comunicação com os usuários. Podiam ser percebidas dificuldades para realizar determinadas tarefas, perda de tempo para execução das tarefas mais usuais e dificuldades para o sequenciamento das mesmas, necessitando de um estudo para facilitar a sua utilização. Nesse sentido, a pesquisa teve como finalidade contribuir para a melhoria da interação dos usuários na *Intranet* da Unimed JP, facilitando, conseqüentemente, a disseminação das informações e aumentando a eficácia da comunicação interna da cooperativa médica. A necessidade do estudo originou-se da constatação da grande quantidade de usuários



e informações heterogêneas existentes e, conseqüentemente, da dificuldade de atender às necessidades informacionais dos usuários desse ambiente de comunicação empresarial.

Na pesquisa, foram considerados quatro sistemas da Arquitetura da Informação, correspondentes aos sistemas de organização, rotulação, navegação e busca, descritos por Morville e Rosenfeld (2006). Além desses sistemas, os autores apontam ainda para a existência de um quinto sistema correspondente à representação da informação (tesauros, metadados e vocabulários controlados), que não foi considerado em virtude destes elementos não estarem presentes no ambiente da *Intranet* da Unimed JP. Além disso, foi utilizado um dos métodos da Engenharia Semiótica, denominado Método de Avaliação de Comunicabilidade (MAC), cuja finalidade é avaliar, através da interface, a comunicação do projetista com o usuário, durante sua navegação pelas páginas.

Salgado (2007, p. 53) ressalta ainda que

Na abordagem da Engenharia Semiótica, por meio do MAC, o foco está na comunicação unilateral e integral, do projetista para o usuário, baseado em premissas sobre o usuário, seu contexto, seus gostos, preferências, capacidades e valores.

Nesse sentido, o MAC, da Engenharia Semiótica, combinada com os sistemas da AI, pode agregar seus conceitos de comunicabilidade a estrutura dos *sites*, tornando os mesmos ainda mais eficientes e ágeis para os usuários.

2 FUNDAMENTAÇÃO TEÓRICA

A facilidade com que se acessam informações pela *Internet* tem permitido obter detalhes sobre qualquer assunto em poucos instantes, durante vinte e quatro horas por dia, sete dias por semana. A informação faz parte do cotidiano de todas as pessoas, seja lendo um jornal, ouvindo uma música, assistindo a um filme, ou buscando detalhamento sobre determinado tema. A *Internet* possibilita o acesso às informações por todos os setores da sociedade, sendo a *Intranet* uma rede com as mesmas características da *Internet*, mas aplicada a ambientes restritos, tendo sido utilizada para a comunicação empresarial.

As técnicas e estratégias de comunicação utilizadas pelas empresas geralmente têm o objetivo principal de cumprir com sua função social e, assim, a construção de uma imagem organizacional, para que, a partir dela, a empresa seja legitimada pela sociedade. De acordo com o pensamento de Nassar (1995), a boa imagem institucional só pode ser conseguida através do relacionamento de todos os setores da empresa com seus públicos. Para tanto, torna-se imprescindível trabalhar com um planejamento estratégico de comunicação, em que seja possível desenvolver políticas de comunicação que permitam administrar a produção e os fluxos de informação dentro do ambiente organizacional. Desta forma, quando uma empresa desenvolve novos meios de comunicação, ela acaba abrindo novos canais, tanto ascendentes quanto descendentes, facilitando a interação entre as hierarquias.



Para que o processo de comunicação em ambientes informacionais digitais possa acontecer de forma satisfatória, se faz necessário o entendimento das informações dispostas nas páginas da *Internet/Intranet*, bem como a facilidade para navegar pelo espaço informacional visando à recuperação de informações consideradas relevantes. Para uma empresa, a organização das informações disponibilizadas na *Intranet*, além de proporcionar uma disseminação mais ágil, pode ser importante no processo motivacional, e na consequente eficácia da comunicação institucional.

Para auxiliar o processo de organização das informações, utilizando-se os conceitos da AI, torna-se necessário cumprir as etapas metodológicas de desenvolvimento das páginas visando facilitar a interação com o usuário. Tais etapas segundo Morville e Rosenfeld (2006) consistem em: Pesquisa, Concepção, Documentação, Implementação e Avaliação. No pensamento de Agner e Moraes (2003), esse processo deve envolver tanto a análise, quanto o projeto e a implementação de espaços informacionais. Nesse mesmo sentido, Reis (2007) comenta que no projeto das páginas, a AI é responsável por definir a estrutura, possibilitando a organização de uma espécie de esqueleto sobre o qual todas as demais partes irão se apoiar.

De acordo com Morville e Rosenfeld (2006), para manter uma relação harmoniosa entre a informação e o usuário, é preciso que as informações sejam “trabalhadas” ou “organizadas” por meio de quatro sistemas interdependentes, cada qual composto por regras próprias. São os sistemas de organização (determina como é apresentada a organização e categorização do conteúdo); de navegação (formas de se mover no espaço informacional); de rotulação (signos verbais e visuais para cada elemento informativo e de suporte à navegação); e de busca (perguntas que o usuário pode fazer e as respostas que irá obter).

Para Reis (2007), a tarefa de organizar as informações em ambientes digitais não apresenta simplicidade, pois estas apresentam características complicadas, como ambiguidade, heterogeneidade, diferenças de perspectiva, políticas internas, dentre outras.

Ainda de acordo com o pensamento do referido autor, por não haver metodologias e métodos padronizados na área de AI, grande parte dos desenvolvedores utilizam metodologias próprias desenvolvidas com base em suas experiências e estudos autodidatas. Entretanto, segundo Morville e Rosenfeld (2006), essa metodologia própria pode tornar o projeto subjetivo e inapropriado ao uso do público-alvo.

O fato é que sistemas de etiquetagem e organização são intensamente afetados pelas perspectivas dos seus criadores. Vemos isso no nível corporativo com sites organizados de acordo com divisões internas ou organogramas, com os agrupamentos, tais como marketing, vendas, atendimento ao cliente, recursos humanos e sistemas de informação. Como é que um cliente que visita este site sabe para onde ir para obter informações técnicas sobre um produto que acabou de comprar? Para projetar sistemas de organização utilizáveis, precisamos fugir dos nossos próprios modelos mentais de conteúdo rotulagem e organização. (MORVILLE; ROSENFELD, 2006, p. 27, tradução nossa).



A área de projeto oferece alguns métodos que podem ser muito utilizados na AI, principalmente na fase que envolve a concepção da interface. A área de Administração também pode fornecer contribuições por meio da aplicação de métodos na fase de concepção, gerenciamento e avaliação, onde de acordo com o pensamento de Camargo (2010, p. 57), pode auxiliar “principalmente no que se refere à obtenção de resultados mensuráveis (produtividade, custo-benefício e vantagem competitiva)”.

Outra área bastante discutida atualmente refere-se à usabilidade, que está associada aos estudos de IHC e à ergonomia. A usabilidade pode vir a complementar a AI, já que existe uma preocupação com a satisfação do usuário, visando o alcance de seus objetivos, considerando questões de eficiência, efetividade, flexibilidade, dentre outras. Sua aplicação deve ser realizada através de questionários aplicados aos usuários, baseados em modelos formais, base de conhecimento, *checklists*, ensaios de interação ou sistemas de monitoramento (CYBIS, 1999).

A partir desse contexto, é possível perceber que as áreas citadas não permitem traçar o perfil semiótico dos usuários, possibilitando prever possíveis problemas de comunicação entre o projetista e os referidos usuários. Para reduzir essa barreira, a Engenharia Semiótica, que corresponde a uma teoria explicativa de Interação Humano-Computador, busca facilitar o entendimento dos fenômenos envolvidos no projeto, uso e avaliação de um sistema interativo (DE SOUZA, 2005).

Ainda de acordo com De Souza (2005), a Engenharia Semiótica por ser uma área que possui metodologias e métodos bem definidos e trabalhados, pode oferecer uma base teórica e prática para a projeção de uma AI bem sucedida, pois está baseada em premissas sobre o usuário, seu contexto, seus gostos, preferências, capacidades e valores.

A Engenharia Semiótica está baseada na semiótica de Charles Peirce e Umberto Eco, que estuda os signos, ou seja, as representações das coisas do mundo que estão em nossa mente. Ela ajuda a entender como as pessoas interpretam mensagens, interagem como objetos, pensam e se emocionam (SANTAELLA, 2004).

A finalidade da Engenharia Semiótica é avaliar a qualidade de uso da interação, através da capacidade do projetista conseguir transmitir aos usuários, através da interface, o que, de fato pretendia com o *design* elaborado.

Para Prates (2007, p. 267),

[...] a análise fundamentada na teoria enriquece o entendimento que os designers têm do problema que estão tentando resolver com o sistema, e assim amplia as suas considerações em relação a possíveis soluções. Não é objetivo dessa teoria prever os resultados de uma ação ou levar o designer a buscar “a” solução correta para o problema em questão.

Na prática, segundo De Souza (2005), a Engenharia Semiótica possibilita perceber as aplicações computacionais interativas como artefatos de metacomunicação, em que o projetista é emissor da mensagem (interface) e o usuário, o receptor. Esta mensagem tem como objetivo comunicar ao usuário a resposta a duas perguntas fundamentais: “Qual a interpretação do *designer*/projetista sobre



o(s) problema(s) do usuário?”, e “Como o usuário pode interagir com a aplicação para resolver este(s) problema(s)?”

Para Salgado (2007, p. 43),

[...] o usuário concebe a resposta a estas perguntas à medida que interage com a aplicação. Assim, esta mensagem é unilateral, uma vez que o usuário recebe a mensagem concluída e não pode dar continuidade ao processo de comunicação naquele mesmo contexto de interação. Além disso, como esta mensagem (a interface) é ela mesma capaz de trocar mensagens com o usuário, ela é um artefato de comunicação sobre comunicação, ou meta-comunicação.

Na Engenharia Semiótica, existe o Método de Avaliação de Comunicabilidade (MAC), que traz para o mesmo contexto comunicativo os projetistas, usuários e sistemas. Através da observação dos usuários em um ambiente controlado, o método identifica rupturas de comunicação e associa “expressões naturais” a estas rupturas. A partir das rupturas identificadas o avaliador analisa os problemas de comunicação identificados e faz a reconstrução da metamensagem apontando os seus problemas. Para a realização do método, é preciso seguir cinco passos: preparação do teste, aplicação do teste, etiquetagem, interpretação e elaboração do perfil semiótico.

Um dos principais resultados possibilitados pelo MAC, de acordo com Salgado (2007), é de ampliar o conhecimento dos projetistas/*designers*, avaliadores e pesquisadores sobre como os usuários interpretam o artefato.

O MAC oferece aos avaliadores um caminho para identificação e análise dos pontos onde o designer pode ter falhado (ou vir a falhar) na comunicação da sua mensagem aos usuários. Oferece adicionalmente aos usuários que participam do processo de avaliação a oportunidade de manifestar o que não entenderam ou com que não concordaram em relação à interface. O MAC trabalha com diferentes testemunhos da interação que juntos, ilustram estratégias de interação usuário-sistema e também aspectos da metacomunicação designer-usuário. As rupturas de comunicabilidade ajudam os avaliadores a identificarem e explicarem pontos da interação problemáticos, assim como a informarem o processo de re-design através de recomendações e alternativas a serem exploradas. (SALGADO, 2007, p. 3)

Com o auxílio do MAC, utilizado pela Engenharia Semiótica, o projeto realizado seguindo seus princípios pode ser de grande utilidade para auxiliar ou, ainda, conduzir uma avaliação utilizando-se os sistemas da AI, além de gerar informações relevantes para a avaliação das necessidades de usuários.

3 ESTUDO DE CASO

Nesta seção são apresentados os resultados da aplicação do MAC juntamente com a AI, na *Intranet* da Unimed JP.

3.1 A INTRANET DA UNIMED JOÃO PESSOA

A *Intranet* da Unimed JP foi implantada há cerca de 13 anos, com o objetivo de divulgar as ações realizadas em todas as áreas da Cooperativa, bem como facilitar o entrosamento entre os colaboradores que trabalham na sede e no hospital.

Desenvolvida e atualizada por uma equipe interna do Departamento de Comunicação e *Marketing*, a referida *Intranet* pode ser considerada como um importante meio de comunicação interna da Unimed JP, uma vez que possui as funções de informar, integrar e oferecer serviços que facilitem o trabalho na Cooperativa.

Como veículo de informação, a *Intranet* conta com pautas que incluem assuntos relativos à rotina administrativa da Unimed JP; divulgação de cursos, treinamentos e palestras oferecidos pela Cooperativa; temas relacionados ao crescimento profissional, como novidades no mercado de trabalho; e assuntos na área de promoção da saúde dentre outros. As matérias são disponibilizadas em formato de texto, áudio e vídeo.

O acesso à página principal acontece por meio da inserção pelo colaborador de seu *login* e senha. Na figura 1 é mostrada a página inicial, após ter sido efetuado o *login*, da *Intranet* da Unimed JP.

Figura 1: Página inicial da *Intranet* da Unimed JP



Fonte: <http://intranet.unimedjp.com.br/>



Entre os serviços disponibilizados estão o contracheque, acesso ao *e-mail* corporativo, *Unimessenger* (programa interno de mensagens instantâneas), solicitação de suporte de informática, lista dos telefones da Cooperativa, formulários internos utilizados com frequência pelos colaboradores, cardápio das refeições servidas no refeitório do Hospital Unimed JP, lista de colégios e empresas parceiras que oferecem desconto para os colaboradores da Cooperativa, dentre outros.

3.1.1 Arquitetura da informação na *Intranet* Unimed JP: uma breve descrição

A *Intranet* Unimed JP está organizada (Sistema de Organização) de acordo com um esquema ambíguo orientado a tarefa, que organiza a informação por conjuntos de ações. Inclui seções bem definidas e mutuamente exclusivas (ou excludentes), que auxiliam na organização da informação, porém dificultam a navegação por meio de hipertextos. A heterogeneidade de conteúdos definida por Reis (2007, p.80) como a “mistura de diversos tipos de conteúdos (textos, imagens, vídeos, sons, etc.) em uma infinidade de formatos (html, pdf, ppt, swf, js, etc.)”, é bastante notória na *Intranet* Unimed JP, o que dificulta a elaboração de uma política única para organização e estruturação de seu conteúdo.

No Sistema de Navegação são utilizados, barra de navegação global (que possui sequência de *links* para áreas chaves do *site*) e elementos do Sistema de Navegação Embutido, com logomarca no canto superior esquerdo da página, conforme mostrado na Figura 1.

As páginas secundárias não são estruturadas com *Bread Crumbs* (uma espécie de guia de navegação, que pode facilitar a encontrabilidade nas páginas de um *site*) e não possui um espaço para pesquisa global (Sistema de Busca) por todo o *site*, apenas no espaço para matérias jornalísticas.

Já os rótulos (Sistema de Rotulação) são empregados nos títulos das páginas, nas opções do Sistema de Navegação e nos *links* contextualizados dentro da mesma, sendo compostos por palavras, imagens e sons.

3.2 PROCEDIMENTOS METODOLÓGICOS

Para contemplar à análise dos aspectos relativos aos usuários, o conteúdo e o contexto, foi realizado um estudo exploratório, com natureza quanti-qualitativa.

A pesquisa exploratória é indicada em situações nas quais as informações sobre o problema em estudo são restritas ou escassas, buscando aumentar o conhecimento sobre o tema pesquisado, clarificando conceitos e fornecendo subsídios para as etapas subsequentes de investigação. Ela é caracterizada pela flexibilidade e versatilidade com respeito aos métodos (CHURCHILL, 1999).

Quanto à abordagem foi uma pesquisa quanti-qualitativa, ao considerar aspectos tanto quantitativos quanto qualitativos, uma vez que ambos são complementares. A pesquisa quantitativa considera que tudo pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las, enquanto que a qualitativa aprofunda-se no mundo dos significados das ações e relações, não perceptível e não captável de forma matemática (MINAYO, 2004).

Os instrumentos aplicados foram o Método de Avaliação de Comunicabilidade (MAC), utilizado pela Engenharia Semiótica e um questionário semi-estruturado elaborado considerando



quatro sistemas da Arquitetura da Informação, correspondente aos sistemas de organização, rotulação, navegação e busca.

O MAC foi realizado por meio de seus 5 passos: preparação do teste, aplicação do teste (incluindo entrevistas pré e pós-teste), etiquetagem, interpretação e elaboração do perfil semiótico. Com a análise, pretendeu-se oferecer elementos para a reformulação da *Intranet* Unimed JP, objeto de estudo em questão. Já a AI foi utilizada na segunda etapa do MAC (pré e pós-teste), através de aplicação de questionário semi-estruturado.

3.3 APRESENTAÇÃO DOS RESULTADOS

Na sequência são apresentados os resultados obtidos pela aplicação dos passos do MAC.

3.3.1 Preparação do teste

Para participar da pesquisa, foram selecionados 30 colaboradores aleatoriamente, sendo 15 da Sede da Unimed JP e 15 do Hospital. O total de colaboradores foi distribuído em cinco grupos de 6 pessoas, tendo se mantido até o fim da pesquisa, por se tratar de um grupo focal.

A preparação do material para as etapas de avaliação envolveu a disponibilização de uma sala equipada com dois computadores, uma filmadora e um gravador de voz, tomando-se o cuidado de que as respostas de um respondente não viessem a influenciar nas respostas dos demais integrantes do grupo.

3.3.2 Pré-teste

Para avaliar a estrutura da *Intranet* da Unimed JP utilizou-se quatro sistemas da AI. A coleta e o registro das observações feitas geraram um conjunto de dados que se juntaram aos resultados alcançados pelos testes subseqüentes (específicos do MAC e teste pós-uso), aplicáveis ao estudo de caso específico.

Participaram da pesquisa 10 homens e 20 mulheres. Desse total 17 (56,7%) possuem formação superior, 3 possuem ensino fundamental (10%) e 10 ensino médio (33,3%).

Para todos os participantes (100%), navegar pela *Intranet* é uma tarefa considerada fácil, já que a “página é muito simples e familiar”, conforme afirmou um dos colaboradores.

Ao acessar a *Intranet*, 43% dos participantes disseram que utilizam mais o canal de ‘Serviços *on-line*’. Um dos funcionários afirmou que “o local possui serviços que gostamos de consultar e utilizar frequentemente, como o contracheque, extrato da Claro (telefonia móvel, do plano empresarial), suporte de informática e a rede social *Unimessenger*”. Os outros 57% afirmaram procurar mais informações no canal sobre ‘Colaboradores’ (27%), ‘Comunicação’ (20%) e ‘Serviços’ (10%) que, na página, é diferente de ‘Serviços *on-line*’. Embora o canal ‘Unimed’ possua toda a história da Cooperativa, nenhum usuário afirmou se interessar pelo mesmo. Um dos entrevistados relatou: “eu acho importante saber sobre a Unimed, mas a gente já leu tanto sobre isso que não tem mais por que ficar acessando”, disse um dos entrevistados. A afirmação sugere que, embora os colaboradores



não costumem acessar o canal ‘Unimed’, ele é importante e deve continuar fazendo parte do menu, principalmente por conter informações relevantes para os usuários externos a Cooperativa.

Quanto ao Sistema de Busca, apenas 12 usuários afirmaram utilizá-lo. Os outros 18 disseram que não sabem como utilizar o referido sistema. Um dos colaboradores afirmou que nunca viu o recurso na página ao dizer “Eu não sabia nem que tinha isso na *Intranet*”. Nesse contexto é importante ressaltar que o sistema de busca da *Intranet* Unimed JP prioriza apenas o histórico das matérias jornalísticas.

Apesar dos 30 colaboradores participantes considerarem as categorias do menu claras e objetivas e afirmarem que sabem exatamente para que servem os ícones, 60% acham que as informações não estão organizadas de forma adequada às necessidades dos usuários. Entre as sugestões de melhoria estão “dar mais importância aos vídeos e aos conteúdos voltados para os colaboradores e padronizar as cores dos *banners* disponibilizados na página”.

Nesse sentido, mesmo considerando a página fácil de navegar, quase 35% (10 pessoas) afirmaram ter dificuldades para encontrar as informações na página. Um dos participantes declarou “às vezes, não sei em que lugar posso clicar para achar uma informação que preciso naquele determinado momento”.

Este problema talvez explique o fato de 60% dos usuários (18 pessoas) afirmarem que se ‘perdem’ dentro da página, quando clicam nos *hiperlinks* disponibilizados. “Se eu saio da sala e volto, após um tempo fora, eu preciso voltar tudo para saber que caminho eu fiz”, declarou um dos colaboradores.

Quanto aos canais utilizados, os participantes foram unânimes em afirmar que gostam e consultam sempre os ‘aniversariantes do dia’, ‘lista de ramais’, ‘serviços *on-line*’, ‘classificados’ e o ‘RH responde’. A justificativa utilizada é que esses canais são muito voltados para facilitar a vida dos colaboradores.

Embora num primeiro momento os colaboradores tenham afirmado gostar do conteúdo disponibilizado na página, foi possível perceber algumas contradições no decorrer da pesquisa, conforme pode ser percebido pelos resultados das etapas subsequentes.

3.3.3 Etiquetagem de interação

Para esta etapa, foram utilizados um gravador de voz e uma câmera filmadora. O método foi necessário porque os sintomas de algumas etiquetas (correspondentes a expressões de comunicabilidade) numa aplicação deste tipo são diferentes, requerendo a inclusão de outras ferramentas para capturar a interação do usuário com o sistema.

Durante a etiquetagem do vídeo do grupo 2 (G2), por exemplo, foram identificadas algumas etiquetas “E Agora?” graças ao uso da etapa de etiquetagem da MAC; caso contrário seria muito difícil percebê-la, pois foram causadas por perda de orientação momentânea, com impacto pouco perceptível na execução das tarefas durante a navegação. No decorrer do teste, alguns integrantes do G2 disseram que muitos dos “E agora?” foram de surpresa e não negativos. Um exemplo de fala



foi “O que eu fiz!?! Hum, eu entrei algum lugar! Tô perdido!”, mas essa desorientação durou apenas alguns segundos, que sem a fala poderia ser interpretada como um tempo normal para planejamento de ações.

Ao visualizar diversos ícones na *Intranet*, um dos usuários emitiu a expressão “Ah, agora sim. Eu só não sei se é esse o objeto, mas pelo menos agora eu sei achar...”, indicando que tinha diversas opções de ícones para selecionar, como num menu de opções, mas não tinha certeza qual deles era o correto, como se dissesse “Cadê o objeto que estou procurando?”, num comportamento similar ao de um “Cadê?” típico.

Algumas rupturas não foram observadas ou identificadas nas falas nos vídeos, mas os participantes comentaram sobre algumas dificuldades na entrevista pós-uso, cuja análise pode ser realizada a partir dos dados mostrados na Tabela 1. Na referida Tabela é apresentado um sumário da distribuição das etiquetas por grupo, além dos totais gerais por etiqueta e grupo. Pode-se perceber que foram identificadas um total de 83 (oitenta e três) rupturas de comunicação, sendo as maiores ocorrências das etiquetas “Cadê?” (19 ocorrências) e “Onde estou?” (19 ocorrências).

Tabela 1: Distribuição das etiquetas por grupo

Etiqueta / Participante	G1	G2	G3	G4	G5	TOTAL
Cadê?	5	2	6	5	1	19
Pra mim está bom	0	1	0	0	0	1
Ué, o que houve?	3	1	1	2	1	8
Por que não funciona?	0	2	0	0	2	4
O que é isto?	3	4	4	2	5	18
Socorro!	2	0	0	0	0	2
Opa!	5	2	2	5	1	15
Onde estou?	4	3	5	3	4	19
E agora?	3	0	0	0	0	3
Total de rupturas	25	15	18	17	14	83

Fonte: Elaborada pelos autores a partir dos dados da pesquisa e adaptado de Salgado (2007)

3.3.4 Interpretação do perfil semiótico

De acordo com o disposto na Tabela 1, o grupo G1 foi o que apresentou o maior número de etiquetas, o que seria de certa forma esperado por ter participantes com pouco acesso à *Internet*. Seu alto índice de três ocorrências da etiqueta “E Agora?”, destacou-se uma vez que os demais não tiveram



nenhuma, indicando que o grupo não interpretou os signos da interface, correspondentes ao menu e ícones, da mesma forma que o projetista imaginou, indicando sua menor familiaridade com a página.

Embora duas das ocorrências do “E agora?” tenham sido momentâneas e causarem um impacto mínimo na execução da tarefa de navegação, uma delas foi bastante séria, onde um dos integrantes do G1 clicou em vários *links* da página, ocasionando assim diversos “E agora?” e “Epa!”, com muitas tentativas sem sucesso de retornar ao ambiente de navegação, mas eventualmente conseguindo, sem chegar ao ponto de desistir.

Ao navegar por diversos espaços da página, sem conseguir “se achar”, este mesmo grupo de usuários comprovou que não conseguiu perceber a intenção do projetista sobre o que fazer para sair da situação.

O maior número de ocorrências de “Ué, o que houve?” (três) deste grupo, também é um indicador de sua menor experiência com a *Intranet*. Uma dessas ocorrências também foi séria e provocou um novo ciclo de interação pela página, no qual novamente um dos integrantes do G1 teve de recorrer a tentativas e erros para perceber a intenção do projetista.

O grupo 5, apesar de ter tido um grande número de ocorrências de “O Que é isto?”, na realidade teve dúvidas com relação à categorização do menu da página, como, por exemplo, ‘Dicas de saúde’ como integrante do submenu ‘Colaborador’.

Ficou evidente que a interação da *Intranet* precisava oferecer algum outro recurso, como, por exemplo, *Bread Crumb*, para servir de orientação durante a navegação dos usuários da *Intranet*.

Além disso, ficou claro que alguns ícones usados na página como ‘Classificados’ e ‘Serviços’, disponibilizados na parte superior direita da página e no menu, respectivamente tinham influência no tempo de execução de uma tarefa.

Retornando à interpretação da etiquetagem, a etiqueta “Por que não funciona?” ocorreu para os grupos 2 e 5, e para o mesmo problema, correspondente a tentativa de acessar o *webmail*, que causou desistência ao abrir uma página com a seguinte informação “Há um problema no certificado de segurança do site”. Eles ficaram com receio de clicar na opção “Continuar neste site (não recomendado)” e atrair vírus para o computador de trabalho.

As ocorrências da etiqueta “Opa!” em todos os grupos apontaram em grande parte o mesmo problema, que foi o fato de clicarem na seção ‘Comunicação’ para solicitar a cobertura de eventos, enquanto o projetista, ao elaborar a página, esperava que o usuário procurasse na seção ‘Serviços’.

O engano ocorreu devido à dubiedade da categorização, já que uma solicitação de um evento é um serviço oferecido pela página, mas também é realizada pelo Departamento de Comunicação. O engano confundiu alguns participantes e causou outras etiquetas do tipo “E agora?” ou “Ué, o que houve?”, e em alguns casos apenas causando outros “Epa!”.

A etiqueta “Para mim está bom” foi utilizada por um dos integrantes do G2, após se enganar no item anterior e achar a solução do que era solicitado. Ao achar ‘Solicitação de eventos’ em ‘Serviços’, o usuário entendeu a intenção do projetista, concordando com o mesmo.



3.3.5 Elaboração do perfil semiótico

Com base no que foi percebido pelos 30 participantes durante as avaliações de comunicabilidade, foi produzida uma mensagem de metacomunicação projetista/usuário, respondendo às perguntas gerais, conforme apresentado por De Souza (2005): (1) Quem são os destinatários da metacomunicação do arquiteto da informação (qual sua concepção sobre os usuários)? (2) Quais de suas necessidades e desejos foram contemplados e por quê? (3) Quais formas de comunicação foram julgadas preferenciais e por quê? (4) Como funciona, o que realiza e não realiza a comunicação usuário-sistema e por quê (qual a lógica da página *web*)? (5) Qual a visão e razão do arquiteto da informação (qual o valor da página *web*)?

Nesta mensagem, foram destacados em negrito os principais pontos levantados nas etapas anteriores de pré-teste, análise e interpretação dos dados. A mensagem produzida foi:

Na minha interpretação, você é um usuário familiarizado com recursos de *Internet* e que utiliza as informações disponibilizadas na *Intranet Unimed JP* para ajudar em suas tarefas laborais diárias. Aprendi que você quer **navegar em uma página amigável, com ícones utilizados de forma clara** e em lugares que possam ser acessados rapidamente. Além disso, você necessita de um **Sistema de Busca, que não priorize apenas as matérias jornalísticas**, mas também outras informações contidas na página, para os casos em que você não conseguir acessá-las com facilidade. Para que não se “perca” na página, você também deseja um **Bread Crumb e Mapa do Site** uma categorização mais clara, evitando, ao máximo, a dubiedade. A página também deve dar **mais importância a vídeos e matérias jornalísticas** que visem ao bem-estar do funcionário e informações importantes sobre a cooperativa, mas as seções de ‘RH responde’ e outros serviços *on-line*, como Solicitação de Suporte de Informática devem ser reformuladas. **Quanto à cor, pode continuar do jeito que está** (verde e branco), com o fundo branco para evitar poluição visual.

De modo geral, os principais problemas identificados pela avaliação de comunicabilidade foram:

1. Ausência de um sistema de busca;
2. Ausência de um *Bread Crumb*;
3. Ausência de um Sistema de Navegação Remoto;
4. Estrutura e organização pouco claras e objetivas.

3.3.6 Pós-teste

Embora todos os 30 participantes da pesquisa tenham respondido, no pré-teste que a navegação é simples, nesta etapa apenas 6 afirmaram não sentir dificuldades para navegar na página, conduzindo a divergências entre o pré e o pós-teste. Além disso, para os 24 que sentiram alguma dificuldade ao navegar pela página, 5 apontaram desorganização das informações (categorização), 5 apontaram a



falta de *Bread Crumb* como fator dificultador, 9 pessoas criticaram o fato de não haver um Sistema de Busca, e 5 tiveram dificuldades em entender alguns ícones, perdendo-se no meio da navegação (rotulação), como já haviam afirmado no pré-teste.

Além de sentirem dificuldades na hora de navegar, a maioria dos participantes (18) também consideraram a página pouco interativa. Dos 12 participantes que consideraram a *Intranet* aberta à troca de informações com o usuário, 5 apontaram a ‘dica do funcionário’ como canal interativo mais eficaz, enquanto 7 consideraram a ‘frase do dia’ uma boa ideia de interação com os usuários. “A *Intranet* possibilita a participação da gente ao dar espaço para a gente sugerir frases e dicas que podem ajudar outros colaboradores”, afirmou um dos participantes.

Quando perguntados se a *Intranet* precisa passar por reformulações, 100% dos entrevistados afirmaram “sim”. Segundo eles, os vídeos devem ter mais destaque na página, o menu deve ser melhor categorizado, os *banners*, como já são coloridos, não precisam ganhar tanto destaque dentro da página. “Também é preciso dinamizar as respostas do Fale Conosco e do RH responde. Caso contrário, a página cairá em descrédito e passará a ser subutilizada”, alertou um dos participantes.

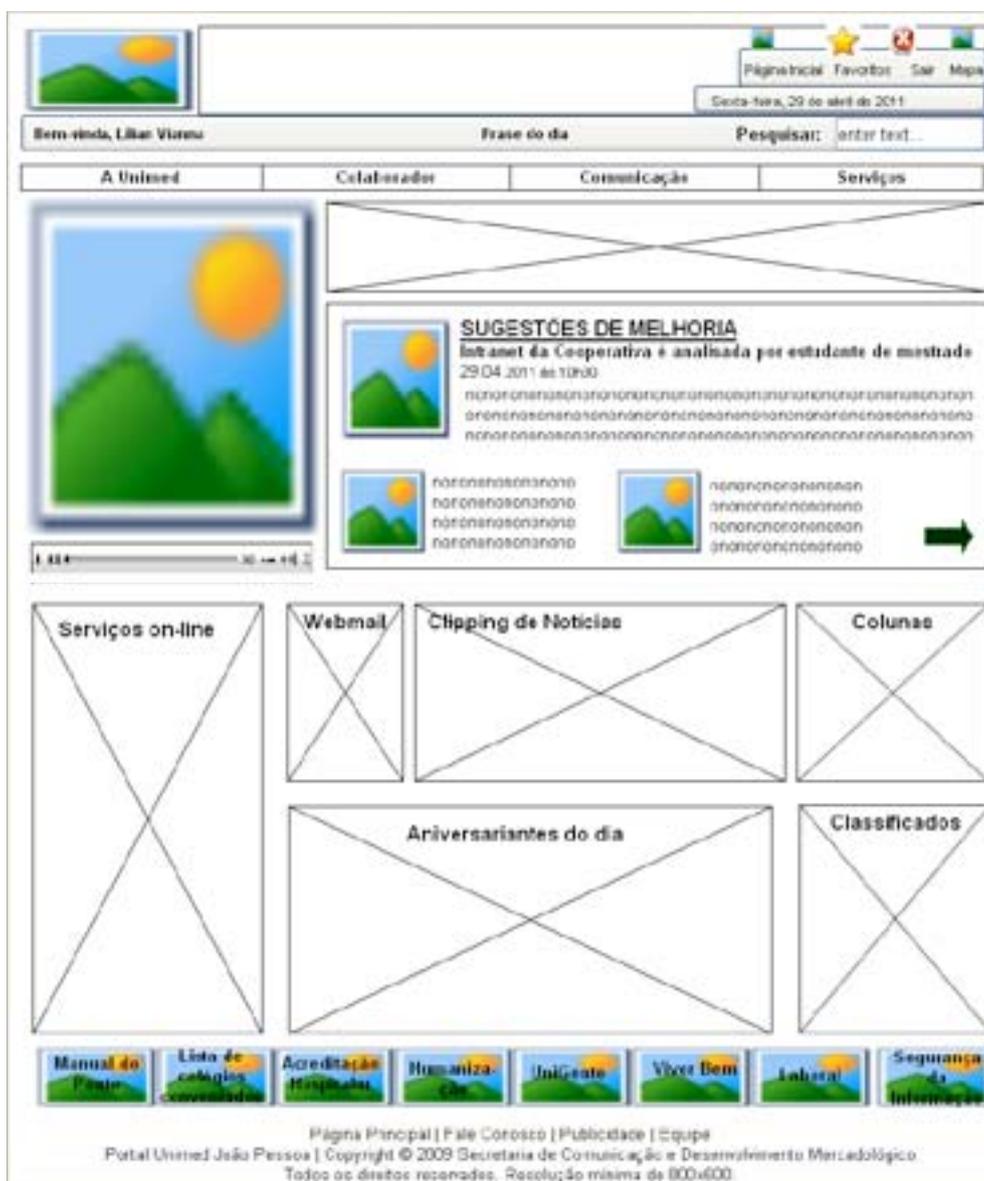
Quanto à estética visual, 27 entrevistados se mostraram satisfeitos com as cores utilizadas na página e acham que ela deve permanecer como está: *background* branco e detalhes verdes. Os outros 3 colaboradores afirmaram que a falta de padronização das cores utilizadas nos *banners*, juntamente com as fotos e o vídeo, tornam a página poluída, como afirmou um dos entrevistados: “Acho esses *banners* muito coloridos. Na minha opinião, todos deveriam ter a mesma cor, já que a imagem do vídeo e as fotos usadas na matéria já são muito coloridas”.

4 SUGESTÕES DE MELHORIA COM BASE NOS SISTEMAS DA ARQUITETURA DA INFORMAÇÃO

Os testes apontaram “ruídos” de comunicabilidade entre o projetista e os usuários. Baseando-se nesses resultados, buscou-se estruturar as mesmas informações da *Intranet*, com base nas necessidades apontadas, tendo como referência quatro sistemas da AI. Na Figura 2 são mostradas as sugestões de melhoria.



Figura 2: Sugestões de melhoria na página da *Intranet*



Fonte: Elaborada pelos autores com base em <http://intranet.unimedjp.com.br/>

É essencial que se procure manter ao máximo o aspecto da *Intranet*, como forma de causar o menor impacto possível, mas é imprescindível a atualização dos recursos existentes, além da inclusão de outros que venham a facilitar a interação com o usuário, a exemplo de um *Bread Crumb*, que pode vir a evitar que os usuários se percam pelo “caminho”, e a utilização eficaz de um sistema de busca já que no estágio atual, os usuários não sabem onde está exatamente a informação que precisam, sendo praticamente obrigados a navegar pelas páginas a busca do que desejam.

Em suma, as sugestões poderão servir para a melhoria da página, como forma de tornar mais eficiente a disseminação da informação. A partir dos dados coletados, foi possível perceber que não se faz necessário elaborar totalmente uma nova *Intranet*, basta apenas adequar de forma satisfatória as páginas às necessidades informacionais citadas pelos participantes.



Na sequência, são discutidas separadamente as alterações em cada sistema considerado na AI:

- SISTEMA DE ORGANIZAÇÃO

Algumas subcategorias foram consideradas ambíguas pelos usuários, se fazendo necessário mudanças na categorização do menu. A primeira mudança seria a organização das categorias em ordem alfabética, além da mudança da categorização de determinados itens, a exemplo de ‘Acreditação Hospitalar’ para ‘Atendimento de Qualidade’. Além disso, a inserção outras subcategorias, como ‘Humanização’. Na categoria ‘Comunicação’ ainda seria inserida a subseção ‘Galeria de Fotos’. Já em ‘Serviços’, os títulos das subseções ‘Recursos Humanos’ e ‘Comunicação e Marketing’ seriam substituídas por ‘Solicitações ao RH’ e ‘Solicitações ao Dep. Com’, dentre outras. Ao se considerar a preferência dos participantes da pesquisa, várias poderiam ser as seções alteradas. Além disso, os *banners* seriam colocados lado a lado no final da página. A ideia seria padronizar o tamanho e evitar a poluição da *Intranet* com excesso de imagens.

- SISTEMA DE NAVEGAÇÃO

Durante a pesquisa, os usuários realizaram sugestões de ações para melhorar o Sistema de Navegação, onde um dos problemas citados foi à falta de recursos que ajudassem a “se orientar” na página. Um dos exemplos apontados foi à ausência de um Mapa do Site, ao enfatizar um dos colaboradores que “às vezes, eu quero saber onde ficam guardados os cardápios do Hospital (Unimed JP), mas não sei onde encontrar; preciso ir procurando por cada espaço. Se tivesse um lugar com todas as partes da *Intranet*, ficaria mais fácil”. Outros aspectos referem-se também a falta de uma opção que possibilitasse mostrar a sequência de navegação. Nesse sentido, a sugestão seria a inclusão de um *Bread Crumb* (migalhas, em inglês), correspondente a um recurso que mostra o caminho das páginas visitadas em um *site*, procurando evitar dessa forma a desorientação do usuário.

- SISTEMA DE ROTULAÇÃO

O resultado da pesquisa também apontou para dificuldades sobre o entendimento referente à rotulação da página. Como forma de evitar a dubiedade na interpretação e uma possível perda de tempo do usuário durante a navegação, sugere-se a disponibilização de rótulos verbais e não verbais. Entre esses rótulos, a imagem de uma casa, com os dizeres ‘Página Inicial’, a imagem da estrela, juntamente com o ícone verbal ‘Favoritos’, o ‘Mapa do Site’, representado pela imagem de um fluxograma, além da opção ‘Sair’, juntamente com uma imagem representativa que sinalize o encerramento da sessão de utilização.

- SISTEMA DE BUSCA

Como a página só possui Sistema de Busca nas seções de notícias, os usuários que necessitam de informações em outras seções têm que navegar por toda a página, tentando “adivinhar” o local em que se encontram os dados que procuram. Para tentar resolver esse problema, sugere-se a existência de uma busca já na página inicial, com subcategorias remetendo a cada parte do *site*.



CONSIDERAÇÕES FINAIS

A *Intranet* Unimed JP, que há alguns anos não passava de um pequeno portal interno da empresa, com quase nenhuma funcionalidade para os negócios, tornou-se mais colaborativa, possibilitando o compartilhamento de arquivos das mais diversas formas.

Para ampliar os serviços disponibilizados, é imprescindível que o projeto seja centrado no usuário, que deve estar envolvido no processo desde o seu início, de forma a reduzir custos e evitar perda de tempo. Visando atender às necessidades dos usuários, pode-se utilizar a AI, que visa à organização dos conteúdos para serem facilmente encontráveis, em conjunto com conteúdos adequados alinhados aos objetivos da empresa. Nesse mesmo sentido, ao se analisar uma página na *Internet/Intranet* pela teoria da Engenharia Semiótica, juntamente com a pesquisa e conhecimentos sobre Arquitetura de Informação, podem tornar-se claras as sugestões, indicações e representações que facilitam o entendimento e a navegação nos *sites*.

Os resultados obtidos pelos métodos da AI e da EngSem, indicaram que cada um tem uma contribuição relevante e diferenciada para questões de IHC. Através da utilização da AI foi possível identificar deficiências na organização e navegação da *Intranet*. Entre as considerações, sugere-se que o Departamento de Comunicação da *Intranet* Unimed JP utilize novas formas de organizar e estruturar suas informações, de modo que atenda às necessidades de seus usuários e atenda à missão da Cooperativa, facilitando o acesso a aqueles que não estão muito familiarizados com a mesma.

Através do Método de Avaliação de Comunicabilidade da Engenharia Semiótica, foi possível identificar as questões de comunicabilidade, relacionadas ao aprofundamento do conhecimento da mensagem de metacomunicação do projetista para o usuário. Sua fonte de descobertas está focada principalmente na “intenção do projetista”. Através dos questionários pré e pós-uso, foi possível comparar a visão do usuário antes e depois de navegar na página, bem como identificar outras questões relacionadas à percepção do usuário com relação ao uso da *Intranet* Unimed JP. Esta etapa permitiu, ainda, que o usuário explicitasse a sua percepção da experiência, em vez de nos limitar apenas ao que o avaliador percebeu.

A partir de revisão da literatura realizada, até onde se tem conhecimento, esta foi à primeira vez que o Método de Avaliação de Comunicabilidade em conjunto com a Arquitetura da Informação foi aplicado em uma análise de *Intranet*. Os resultados alcançados podem servir de base para estudos futuros nessa área, além de que em trabalhos futuros, pode-se estender o cenário com a inclusão de outros métodos, a exemplo da usabilidade.

INFORMATION ARCHITECTURE AND SEMIOTICS ENGINEERING: A CASE STUDY AT UNIMED'S JOÃO PESSOA INTRANET

Abstract: This paper discusses the application of Information Architecture concepts and the Communicability Assessment Method, used by the Semiotics Engineering, in an Intranet environment. The rise of Digital Information and Communication Technologies (DICT), which ideally should



represent more facility in the knowledge creation, can cause difficulties for the information understand, given the short time that users have to interpret it. To collaborate with studies in the area and propose possible solutions for a better understanding of the information provided, especially in institutional settings, a study was conducted considering the Information Architecture and Semiotics Engineering. To validate the proposal, it was developed a survey on the information needs of Unimed João Pessoa Intranet users, with a data collection conducted with the participation of 30 employees through five focus groups. The study is characterized as being exploratory and quantitative and qualitative in its nature. The results of the methods indicated that each has a distinctive and significant contribution to Human-Computer Interaction (HCI) issues. Considering the aspects of Information Architecture was identified deficiencies in the organization and navigation of the Intranet, and through the Semiotic Engineering Communication Assessment Method was identified issues of communicability, related to deepening the understanding of meta message to the user. The study results provide strong indications that more progress is needed in better interface design, considering the need for further studies on the disciplines in question.

Keywords: Information Architecture. Semiotic Engineering. Assessment Communicability Method. Intranet. User communication.

REFERÊNCIAS

AGNER, Luiz; MORAES, Anamaria. **Navegação e arquitetura de informação na web: a perspectiva do usuário**. V.29, n.1, jan-abr. 2003. Disponível em: <<http://www.senac.br/INFORMATIVO/BTS/291/boltec291e.htm>>. Acesso em: 03 jun. 2009.

CAMARGO, Liriane Soares de Araújo de. **Metodologia de desenvolvimento de ambientes informacionais digitais a partir dos princípios da arquitetura da informação**. 2010. 287 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

CHURCHIL, Jr. G.A. **Marketing research: methodological foundations**. 7 ed. New York: Inter. Thomson Publishing, 1999.

CYBIS, W. A.; PIMENTA, M. S.; SILVEIRA, M. C.; GAMEZ, L. **Uma Abordagem Ergonômica para o Desenvolvimento de Sistemas Interativos**. 1999. Disponível em: <<http://www.helionet.varginha.com.br/files/Cybis.pdf>>. Acesso em: 02 fev. 2010.

DE SOUZA, C. **The Semiotic Engineering of Human-Computer Interaction**. Cambridge, MA. The MIT Press, USA 2005.

MINAYO, M. C. (Org). **Pesquisa Social: teoria, método e criatividade**. Petrópolis: Vozes, 2004.

MORVILLE, Peter; ROSENFELD, Louis. **Information architecture for the world wide web**. Sebastopol, CA: O'Reilly & Associates, 2006.

NASSAR, Paulo. **O que é Comunicação Empresarial**. São Paulo: Brasiliense, 1995.



PRATES, R. O. e BARBOSA, S.D.J. **Introdução à Teoria e Prática da Interação Humano-Computador fundamentada na Engenharia Semiótica.** In. T. Kowaltowski e K. Breitman (orgs.). Jornadas de Atualização em Informática, JAI 2007, p. 263-326.

REIS, Guilherme de Almeida dos. **Centrando a Arquitetura de Informação no usuário.** São Paulo, 2007. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação e Artes, Universidade de São Paulo, 2007.

SALGADO, Luciana. **CommEST: Uma Ferramenta de Apoio ao Método de Avaliação de Comunicabilidade.** Dissertação de Mestrado do Departamento de Informática da PUC-Rio. Rio de Janeiro, 2007.

SANTAELLA, Lucia. **Cultura e arte do pós-humano: da cultura das mídias à cibercultura.** São Paulo: Paulus, 2003.

_____. **Comunicação e Semiótica.** São Paulo: Hacker, 2004.

WURMAN, Richard Saul. **Ansiedade de informação: como transformar informação em compreensão.** São Paulo: Cultura Editores Associados, 1991.



<EM BUSCA DA COMPREENSÃO DA “BUSCA” NO CIBERESPAÇO¹>

*Silvana Drumond Monteiro, Maria Júlia Carneiro Giraldez, Fernando Luiz Vechiato, Rogério
Paulo Müller Fernandes, Nelma Camêlo de Araujo*

Resumo: A busca tem se tornado uma prática cada vez mais incorporada na sociedade contemporânea, não importando se ela implica na recuperação da informação e do conhecimento ou apenas em uma descoberta. O trabalho investiga a busca no ciberespaço operada pelos mecanismos, a partir de análise documental, em autores que pesquisam esse fenômeno social. Os enfoques contemplam a anatomia dos mecanismos, a anatomia da busca e a taxonomia da busca. Seja qual enfoque abordado pelo autor, a tecnologia, o usuário ou a linguagem, está relacionado a alguma das três dimensões da linguagem: a sintaxe, a semântica e a pragmática. Assim, essa nova gramática (sintaxe) de busca e seu contexto (semântica) englobam, cada vez mais, o usuário (pragmática) para que a busca faça sentido no ciberespaço.

Palavras-chave: Mecanismos de Busca. Busca. Ciberespaço. Sintaxe. Semântica. Pragmática.

1 INTRODUÇÃO

Os avanços sociais, culturais, políticos e tecnológicos têm modificado significativamente o modo do homem pensar, agir e se comunicar. Nesse cenário, permeado de aparatos tecnológicos, o fluxo veloz de signos, textos, imagens e sons configuram um ambiente hipercomplexo da cultura híbrida e cíbrida² das tecnologias, que impulsionam e incrementam a capacidade humana para a busca constante de informação e conhecimento.

A esse respeito, Santaella (2007, p. 183) assevera que “[...] o astronômico crescimento da Internet, seu imenso número de documentos e relações entre esses documentos nos obrigam a encontrar novas formas de orientação e busca.”

Quanto à busca, a plataforma *Web* é a ambiência contemporânea para o acesso à informação e ao conhecimento, dispersos no ciberespaço, e os mecanismos tecnológicos utilizados para a sua realização fazem parte desse contexto. Só o *Google* possuía, em 2007, mais de 1 (um) bilhão de buscas diárias (FRIEDMAN, 2007) e em 2010 sofreu mais de 500 mudanças em seu sistema de busca (GOOGLE, 2010).

A busca tem-se tornado um tema tão inquietante que Batelle (2006) ao conhecer o *Zeitgeist*

1 Parte dos resultados de pesquisa financiada pelo CNPq e pela Fundação Araucária (Paraná).

2 Termo formado pela união de cibercultura e híbrido, cunhado por Peter Anders, em 1999 (SANTAELLA, 2007, p. 132).



– “Espírito do Tempo” –, ferramenta criada em 2001 e encerrada em 2007, que media exatamente o que a sociedade buscava, referiu-se ao *Google* afirmando que ele “[...] está segurando o mundo pelos pensamentos.”

Nos últimos anos, a busca transformou-se num método universalmente compreendido de navegar pelo nosso universo de informações: assim como a interface do Windows definiu nossas interações com o computador pessoal, a busca define nossas interações com a internet. Ponha uma caixa de busca diante de quase qualquer pessoa e ela saberá exatamente o que fazer com esse instrumento. E é possível conhecer o agregado de todas essas buscas: ele constitui a base de dados de nossas intenções. (p. 4).

A busca como metáfora de interface do usuário e também como artefato cultural suscita ligações entre teorias e práticas e nesse contexto, essa base de dados de intenções e desejos desvela uma “[...] rica camada superficial sobre um sítio arqueológico de tecnologia que, ao longo do último meio século, criou o potencial para o surgimento de uma cultura inteiramente nova.” (BATELLE, 2006, p. 5).

É interessante observar que o autor considera a *Web* um acontecimento relativamente recente, não obstante, tem sua base sobre a Internet, a qual, por sua vez, está constituída sobre uma vasta rede de computadores de todos os tipos – de grande porte, minicomputadores, servidores, *PCs* e máquinas portáteis. Essa rede foi construída ao longo de quase três gerações; contudo, na última década, ela emergiu, especialmente a partir da ambiência simbólica do ciberespaço, em nossa consciência cultural.

A busca e seus motivos variam, assim como o enfoque teórico adotado pelos autores citados, de modo que tentaremos evidenciar algumas relações entre eles. Nesse sentido, apresentaremos um enfoque tecnológico, da **anatomia dos buscadores**, que acentua a interface de busca, com os aspectos sintáticos e semânticos dos resultados (MONTEIRO, 2009), passando pela **anatomia de busca** que procura evitar uma discussão eminentemente tecnológica para contemplar um mapa de experiência do leitor (MORVILLE; CALLENDER, 2010), até a **taxonomia da busca**, que intenta categorizar os motivos pelos quais as pessoas buscam informação no ciberespaço (BRODER, 2002).

Além dos enfoques de estudo dos autores supracitados, é preciso reconhecer que o homem é linguagem e as ações de busca no ciberespaço só acontecem a partir do uso da linguagem em seu contexto de atualização. Assim, a pragmática se impõe, posto que, considera a linguagem essencialmente como uma atividade de comunicação, de origem e de natureza social. (ARMENGAUD, 2006).

Em relação à linguagem e à tecnologia, Santaella (2007, p. 49) argumenta que a “[...] a técnica, hoje transmutada em tecnologia, remonta às origens da constituição do ser humano como ser simbólico, ser de linguagem.” Complementa também que, a tecnologia não nos é tão estranha



assim como parece ser, “[...] são prolongamentos de nossos corpos e da nossa mente.” formando uma “dobra”, conceito denso utilizado por Deleuze (1991) para explicar o sentido do signo.

Nessa linha de raciocínio, este trabalho apresenta discussões e estudos a respeito da busca no ciberespaço, no intuito, embora inicial, de estabelecer uma discussão teórica e pragmática do assunto. Isso dito, o texto é revisional, mas objetiva em sua argumentação tecer algumas conexões, frisa-se, preliminares, sobre o tema em tela.

2 A BUSCA: dimensões teóricas e pragmáticas

A busca faz parte de uma das categorias dos mecanismos, conforme pode ser visualizada no Quadro 1. A categoria, do ponto de vista ontológico, significa qualquer noção ou princípio que sirva de regra para a sua investigação (ABBAGNANO, 2003) e busca compreender as relações entre o todo e as partes. Já do ponto de vista lógico, significa os modos de predicar a substância, e implica ainda, em estudar as relações hierárquicas, das quais tentamos evitar, já que os objetos no ciberespaço estão configurados rizomaticamente.

A categorização dos mecanismos de busca, já consolidada na literatura (embora não reflita o mesmo com a terminologia empregada) é estudada pelo “Grupo de Pesquisa Informação e Conhecimento no Ciberespaço” <http://www.uel.br/grupo-pesquisa/ciberespaço/index_entrada.html> e contempla os seguintes aspectos dos mecanismos, a saber:

- anatomia; forma geral de organização ou indexação;
- ordenação dos resultados;
- apresentação dos resultado;
- paradigma semiótico (MONTEIRO, 2009).

A partir do esquema apresentado no Quadro 1, pode-se observar a primeira categoria, “anatomia dos buscadores”, que será destacada neste trabalho, com ênfase na “busca”.

QUADRO 1- Anatomia dos Mecanismos de Busca

CATEGORIA	CLASSIFICAÇÃO
ANATOMIA	<ul style="list-style-type: none"><i>Crawling</i> (varrer)<i>Indexing</i> (indexar ou gerar o índice a partir da base de dados)<i>Searching</i> (buscar através da interface de busca)

FONTE: MONTEIRO, 2009, p.72.

A anatomia é uma expressão já consagrada na literatura nacional e internacional, no entanto, a



designação da categoria pode aparecer também como arquitetura ou funcionamento dos mecanismos. Percebe-se, também que, pode haver uma discriminação maior nas etapas descritas no Quadro 1, como *crawler*, repositório, indexador, ordenador e apresentador (GOMES, 2008).

Na primeira etapa um programa denominado *crawler* “varre” o ciberespaço, periodicamente, com critérios específicos coletando todas as páginas possíveis que estão conectadas (*crawling*). A partir da base de dados gerada pelo *crawler* forma-se o índice (*indexing*) e por fim a interface de busca para o usuário (*searching*).

A criação do índice separa os dados em várias partes, fazendo listas de palavras associadas às respectivas *URL* (*Uniform Resource Locator*). Cada mecanismo de busca tem seu algoritmo, pelo qual se faz a análise de acordo com os métodos específicos de indexação que, segundo Batelle (2006), só o Google tem mais de 100 critérios definidos.

Na análise, o índice “etiqueta” seus dados por meio dos metadados e procede a indexação para fornecer os resultados de busca. Como destaca Cendón (2001), os mecanismos atuam de três distintas maneiras, a saber:

- a) incluem cada palavra do texto visível das páginas;
- b) extraem apenas a *URL* e as palavras que ocorrem com frequência, ou palavras e frases mais importantes contidas no título ou nos cabeçalhos e nas primeiras linhas;
- c) extraem os termos das *metatags*, etiquetas fornecidas para classificação do *site* por meio do campo “descrição” ou “palavras-chave”.

A próxima etapa é a interface de busca para o usuário, pois os dados após analisados, indexados e etiquetados:

[...] são jogados naquele que é chamado um índice de tempo de execução (*runtime index*) – uma base de dados pronta para oferecer resultados aos usuários. O índice de tempo de execução forma uma espécie de ponte entre a parte dos fundos (seu *crawl* e seu índice) e a parte da frente (seu servidor de perguntas e a interface com o usuário). (BATELLE, 2006, p. 18).

É justamente nesse processo de gerar o índice invertido que os algoritmos agem estabelecendo seus critérios de ordenação que aparecem na apresentação dos resultados.

Nota-se que, apesar de a grande parte da inteligência de um mecanismo estar contida na análise, do ponto de vista de Batelle (2006), o servidor ou a interface também tem recursos que garantem a potencialização da sintaxe de pesquisa, operando sua “semantização”, pois a busca tornou-se a metáfora de navegação:

No fim do dia, o santo graal de todas as ferramentas de busca é descobrir seu verdadeiro objetivo – o que você está procurando e em que contexto.



[...] Um exemplo de progresso nesta área está a identificação daquelas que são denominadas frases atômicas. Quando você digita uma pergunta de uma só palavra, por exemplo, ‘York’ você está procurando resultados para ‘Nova York’? (p.18).

Aquilo que Batelle (2006) se refere como “frase atômica”, Morville e Callender (2010) denominam *query*, isto é, pergunta sem a estrutura de frase, mas entendida ou respondida como texto. Mas, a “*query*” tem se tornado uma sintaxe, posto que quase 54,50% de todas as buscas usam duas ou três palavras e outros 15,98% usam apenas uma e 3,11% utilizam mais de seis (STATOWL, 2011). “Porém, de modo geral, a tendência é no sentido de acrescentar mais palavras às nossas perguntas à medida que navegamos por esta estranha nova gramática de palavra-chave.” (BATELLE, 2005, p. 22).

A partir do exposto, constatamos a dimensão e o mérito de abordagens teóricas e pragmáticas a respeito da busca no ciberespaço. Com o surgimento do ponto de vista pragmático, os estudos dos signos e da linguagem se distribuíram da seguinte maneira:

- a) a dimensão *sintática* estuda as relações dos signos entre si, das palavras na frase ou nas sequências de frases;
- b) a dimensão *semântica* trata da relação dos signos, palavras e frases com as coisas e com o estado de coisas; é o estudo conjunto do sentido, da referência e da verdade;
- c) a dimensão *pragmática* estuda a relação dos signos com os usuários dos signos, das frases com os falantes. (ARMENGAUD, 2006, p. 12).

Vale destacar que os signos entram em relações diferentes segundo a dimensão considerada. “Na dimensão sintática, os signos se implicam. Na dimensão semântica, os signos designam e denotam. Na dimensão pragmática, os exprimem (seus usuários).” (ARMENGAUD, 2006, p. 44). Dessa forma, o signo *im-plica*³ o sentido e o sentido *ex-plica* o signo, implicação e explicação, envolvimento e desenvolvimento são atributos do signo, pois “O próprio sentido se confunde com esse desenvolvimento do signo, como o signo se confunde com o enrolamento do sentido.” (DELEUZE, 2010, 84-5).

Em relação à busca, alguns mecanismos, no intento de mostrar alguns contextos de atualização de busca, mostram, por meio de *mashup*, uma lista de possíveis sentidos, no final da página de resultados, como no caso do *Google*.

Nessa direção, os recursos das interfaces têm incorporado contextos mais semânticos, denominados, por Morville e Callender (2010), *autocomplete* e *autosuggest*, conforme os exemplos das Figuras 1 e 2:

3

Plica, em latim, significa prega ou dobra, implicar quer dizer, dobrar, unir; já explicar é desdobrar.



FIGURA 1- Recurso de Autocompletar do Yahoo!

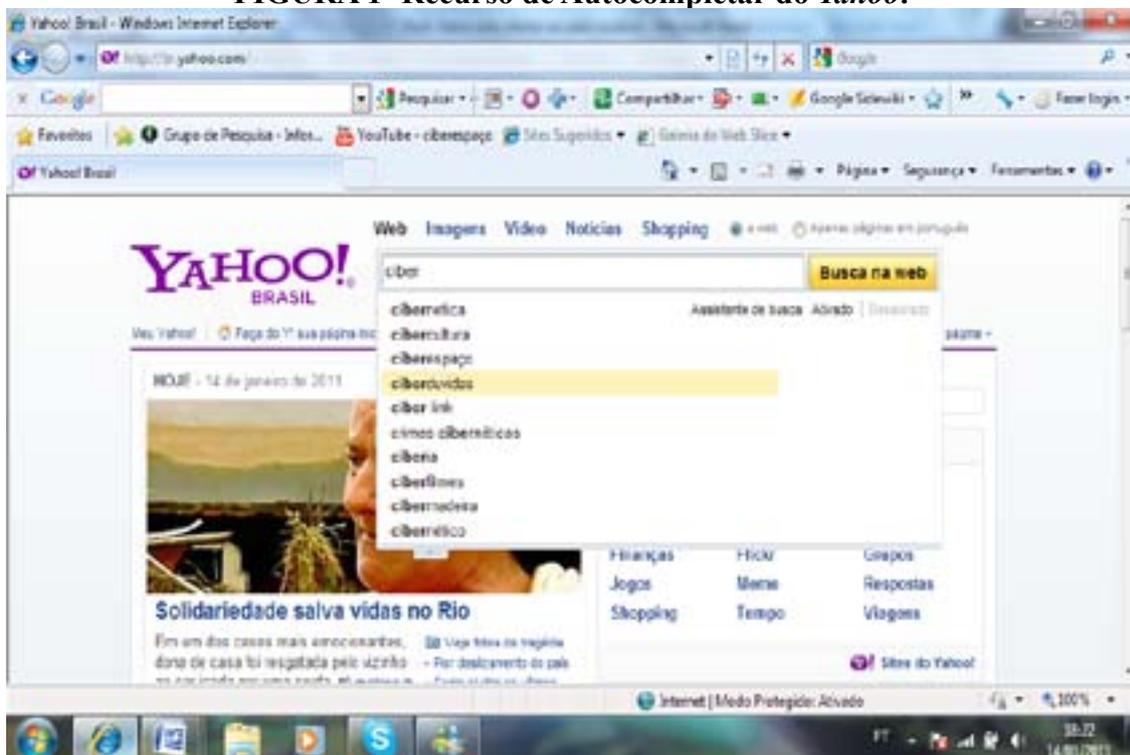
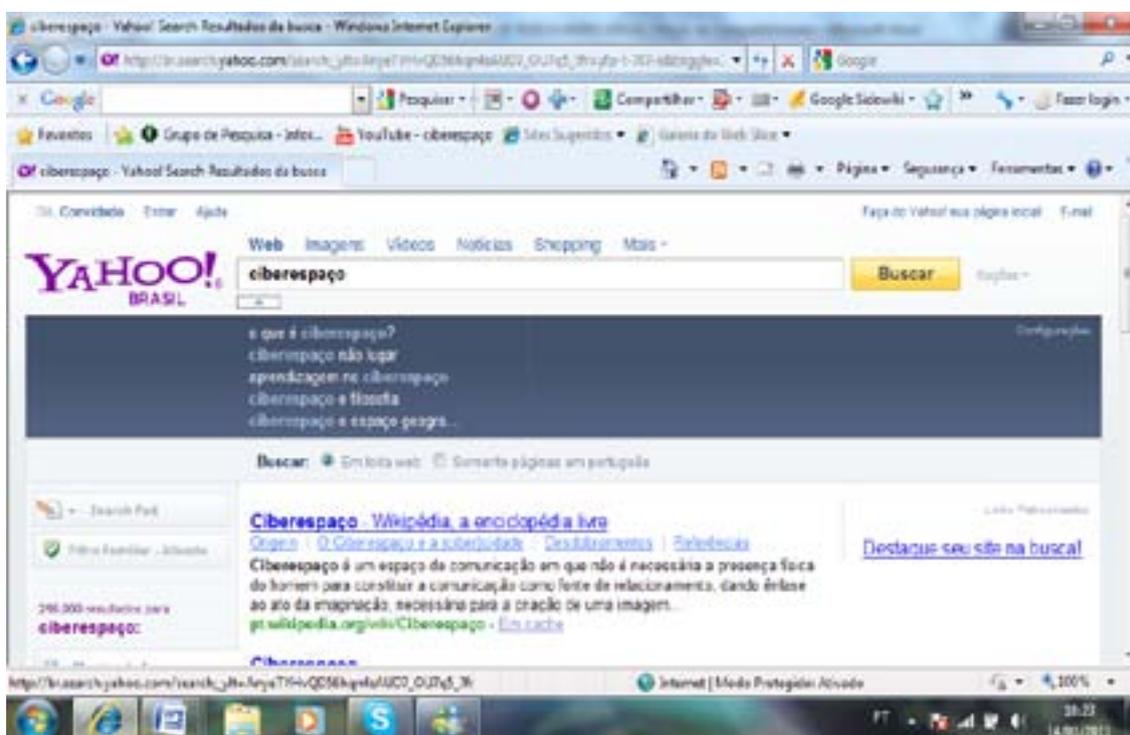
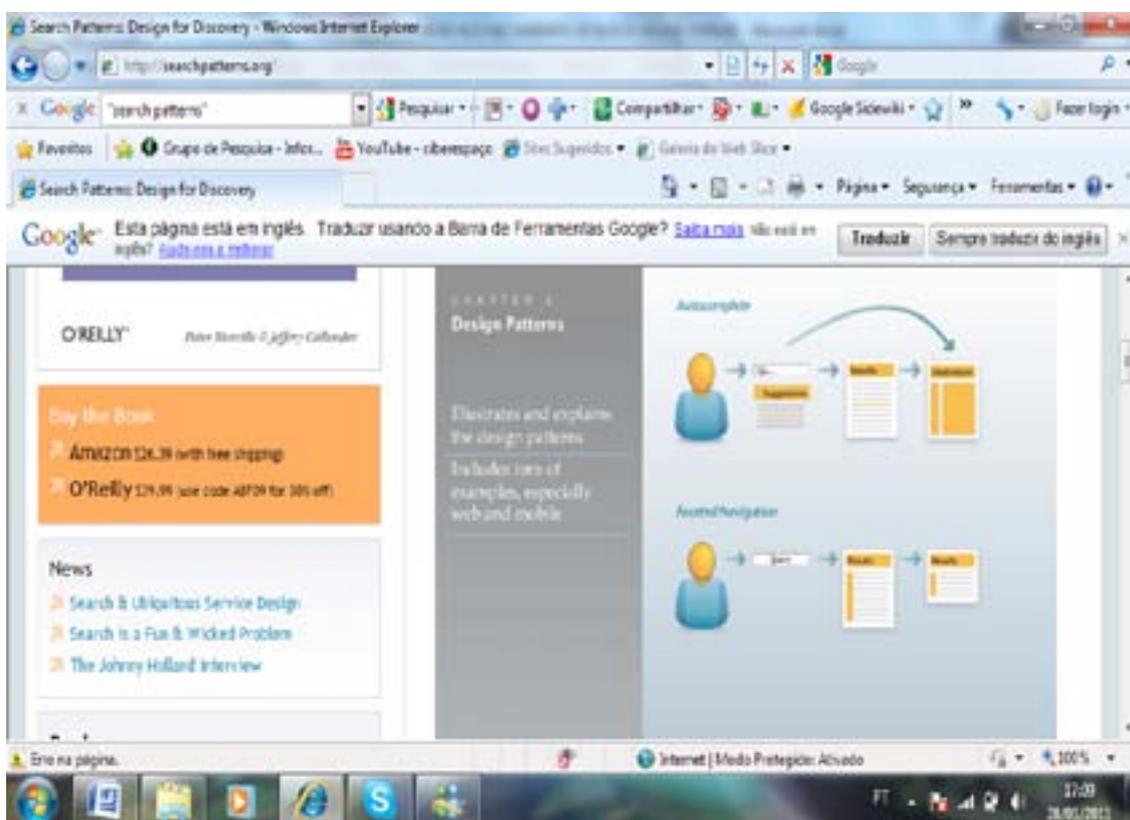


FIGURA 2- Recurso de Autosugestão do Yahoo!



Apesar de parecidos, não são a mesma coisa, pois o autocompletar evita a digitação completa da palavra de busca (*query*) projetando o histórico de outras pesquisas realizadas no ciberespaço, criando padrões de busca e alavancando algoritmos “mais inteligentes”, pois de acordo com Batelle (2006), os desenvolvedores utilizam esse *corpus* de dados para identificar conjuntos de propriedades gramaticais de linguagem; já os assuntos relacionados buscam atribuir contexto ao campo semântico de busca ao indicar as sugestões.

FIGURA 3- Ilustração da Busca por meio do Recurso Autocompletar



FONTE: MORVILLE; CALLENDER, 2010, *slideshare* p. 8. Disponível em: <<http://searchpatterns.org/>>.

Conforme a Figura 3, esses recursos de busca garantem, também, a rapidez, saltando etapas na consulta do índice de resultados, e mais, em plataforma móvel, é de extrema importância, por questões óbvias e podem ser aperfeiçoados com o comando de voz. Outro aspecto interessante, dessas ferramentas, nos buscadores visuais (não é o caso do *Yahoo!* em português) é a “desambiguação” da busca por intermédio da apresentação das imagens.

Esse termo significa tirar a ambiguidade do argumento de busca e conferir melhor precisão aos resultados retornados. Já na linguagem visual, com proeminência de representação no índice, como no caso de fotografias, que têm alto poder referencial e denotativo, a “desambiguação” é completa, chegando a 100%, pois atualiza algo singular, denota o objeto real, conforme o exemplo da Figura 4:



FIGURA 4- *Visual autocomplete*



FONTE: MORVILLE; CALLENDER, 2010, p. 84. Disponível em: <<http://www.flickr.com/photos/morville/4274337448/>>.

Essa é a grande discussão: os mecanismos de busca aperfeiçoarão seus resultados investindo em recursos nas interfaces de busca ou a *Web Semântica* se manifestará de forma mais abrangente no ciberespaço? Haverá uma hibridização entre as duas? Algumas dessas questões estão ainda sendo investigadas.

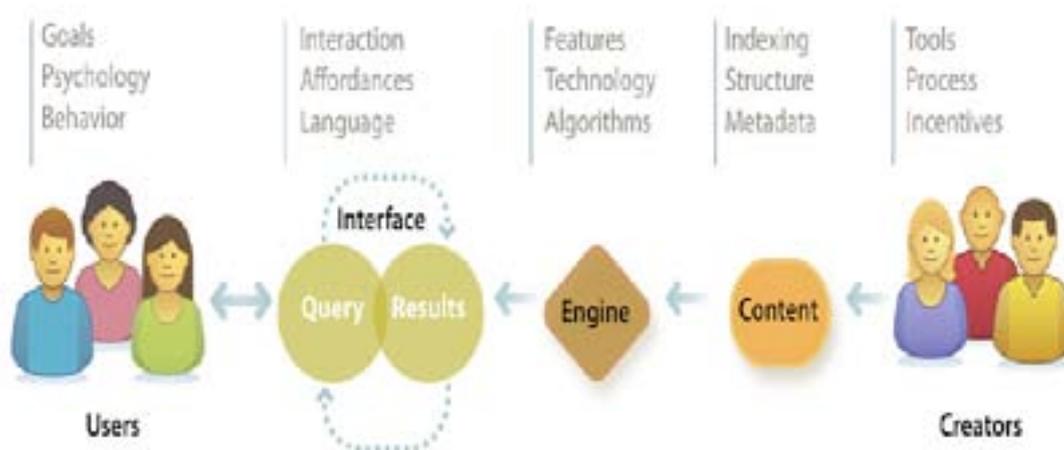
Morville e Callender (2010) afirmam que a categoria “anatomia”, nos buscadores, surge de um enfoque eminentemente tecnológico e propõem um mapa que desloca a atenção do *hardware* e do *software* para os elementos da experiência dos usuários, a saber:

- a) criadores;
- b) conteúdo;
- c) mecanismos;
- d) interface;
- e) usuários.

Afirmam que a anatomia é um poderoso instrumento para estudar partes, mais tradicionalmente, órgãos, já os mapas demonstram o relacionamento entre elas, ademais, o foco na forma e na estrutura, característico da anatomia, não deve distrair o foco na função e no contexto.

A partir desse enfoque, propõem outra anatomia de busca, englobando os seguintes aspectos nos elementos do mapa infracitado:

FIGURA 5 - *The Anatomy of Search*



FONTE: MORVILLE; CALLENDER, 2010, p. 25. Disponível em:
<http://farm5.static.flickr.com/4056/4274260470_9bc6be3466_b.jpg>.

Os autores abordam um contexto mais amplo que pode ser considerado pragmático, isto é, a relação do signo com o seu usuário, e seu conseqüente agenciamento. É justamente na dimensão pragmática que um signo define-se, na sua utilização, na combinação com outros signos e com os outros membros de um grupo social, e não esqueçamos que esse agenciamento é maquínico, portanto, ele acontece na interação com as TIC (Tecnologias da Informação e Comunicação).

De acordo com Morville e Callender (2010), é necessário considerar os objetivos do usuário e, para isso, os desenvolvedores precisam pensar na:

- precisão e revocação dos resultados;
- a busca com objetivo de apenas conhecer e encontrar um item (*lookup*), ou para descobrir, analisar e comparar um conhecimento (*learn*);
- o domínio que o usuário tem da questão de partida (se é especialista ou não, *domain expertise*);
- a plataforma em que efetua a sua pesquisa (*notebook, desktop, móbile*);
- seu comportamento e experiência.

Já na interface, Morville e Callender (2010) orientam os desenvolvedores a olhar para fora da caixa de busca. O que isso significa? Significa que algumas perguntas (ou respostas) deveriam fazer parte nesse quesito: *Who, What, Where, Why, When* e *How* (ou seja, 5-W e 1-H), que são na verdade, técnicas de redação jornalística.



As relações entre a parte dos fundos e da frente, a indexação e a busca, ou ainda, a organização e apresentação dos resultados devem ser pensadas holisticamente. Já Monteiro (2009) propõe pensar essas relações como dobras do signo, o avesso e direito de um mesmo processo, operado este por múltiplas sintaxes e semióticas virtuais em atualizações pragmáticas na busca.

Voltando ao conjunto de perguntas, sobre “quem” (*Who*) busca no ciberespaço, a resposta óbvia seria quase todo mundo. Não obstante, uma pesquisa realizada em 2004, nos EUA, mapeou os usuários, e à época descobriu que eles perfaziam um grupo de pessoas jovens, de alto nível de escolaridade e familiaridade no uso de TIC.

Responder a questão “o que” (*What*) as pessoas buscam no ciberespaço, implica em dizer que o homem está em busca da felicidade, satisfazer lacunas cognitivas, atualizar-se profissionalmente, de amor, amizade, lugares, do último lançamento, “Enfim [a busca] é movida pela inimaginável complexidade inerente à linguagem humana.” (BATELLE, 2006, p. 21).

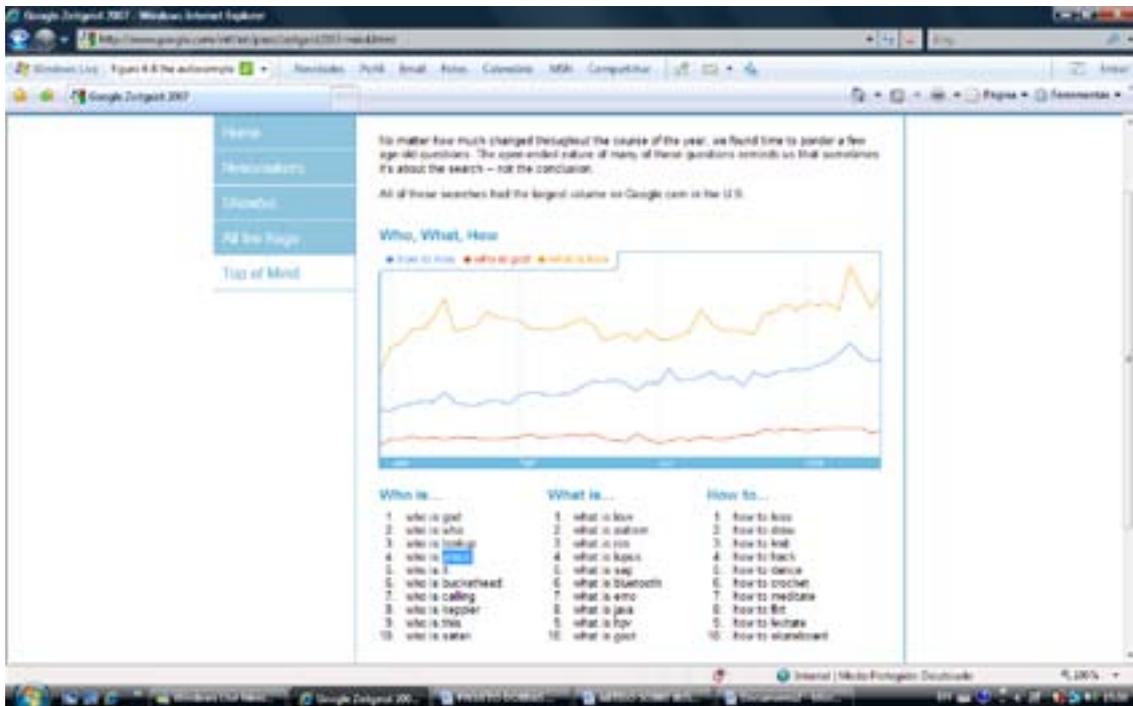
Em relação ao “onde” (*Where*) a literatura, de forma geral, constata que os mecanismos gerais (ou horizontais) são o lugar onde os usuários mais utilizam, estando à frente o *Google*, e é também neles que se pode responder a outra questão “por que” (*Why*), pelo menos de forma categorial, Broder (2002) intenta rastrear os motivos (próxima parte); mas por que buscamos? “Para recuperar aquilo cuja existência na *Web* conhecemos e para descobrir aquilo que supomos existir nela, seja uma aula de cerâmica ou um amigo perdido.” (BATELLE, 2006, p. 26).

Responder o “quando”, de acordo com Morville e Callender (2010), significa ponderar o tempo em que um usuário demora em uma interface de busca, já Batelle (2006) reporta-se ao período de tráfego na rede.

Mas, o *Zeitgeist* (mencionado) tem outras interessantes respostas em relação a algumas do conjunto de perguntas supracitadas, totalmente voltadas ao contexto da linguagem, ilustradas no *print* da Figura 6, mas que refletem exatamente o fluxo do pensamento humano *online*.



FIGURA 6 - *Who, What, How*



FONTE: GOOGLE Zeitgeist, 2007. Disponível em: <<http://www.google.com/intl/en/press/zeitgeist2007/mind.html>>.

Com efeito, por meio dos padrões de busca, o *Google* sabe “o quê” a sociedade procura e lembra, de alguma forma, a ordenação da linguagem em Deleuze (1998): as designações e seus preenchimentos por coisas (o que é aquilo? O que é isso?); as manifestações e suas efetivações por pessoas (quem é Fulano? Quem é Beltrano?).

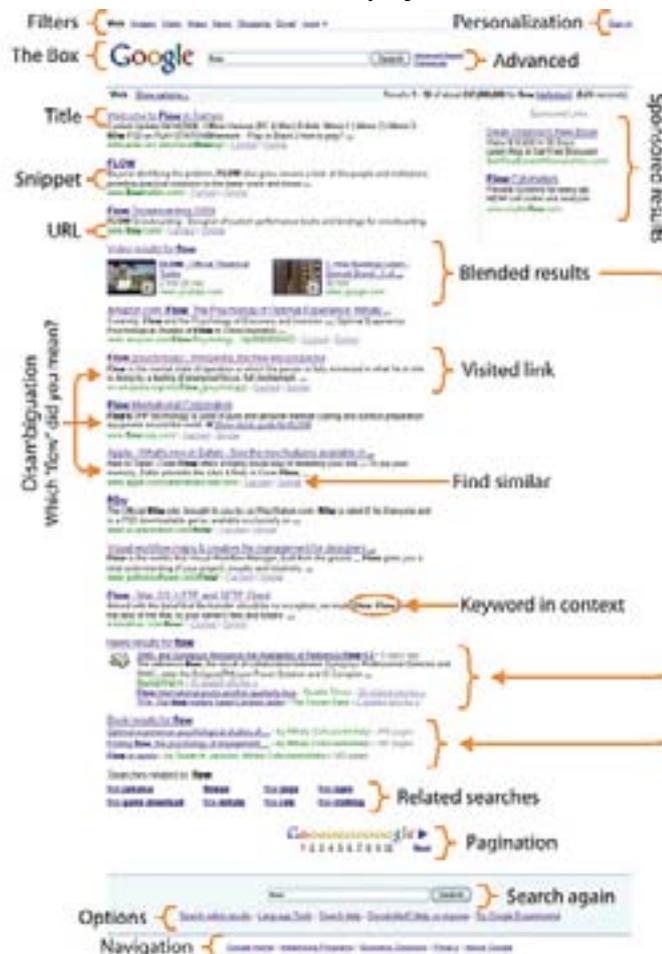
Batelle (2006, p. 20) destaca a pergunta “como” (*How*) fazendo uma interessante observação:

Finalmente, ao tratar do ‘como’ da busca, é importante dar uma olhada rápida nos métodos específicos que desenvolvemos como pesquisadores. A verdade é que somos incrivelmente preguiçosos. Digitamos umas poucas palavras e esperamos que a ferramenta de busca nos traga resultados perfeitos. Mais de 95% dos usuários nunca usam as funções avançadas de busca incluídas na maior parte das ferramentas [...].

Ainda Morville e Callender (2010) apresentam uma *SERP* (*Search Engines Results Page*), no circuito das questões, fazendo apenas uma menção: “[...] ela está entre o mais complexo e importante desafio do design.” (p. 31). O fato é que, quanto maior a base do buscador, maior a probabilidade dos resultados apresentarem-se relevantes. Não se deve esquecer, também, que atrás de cada grande interface de busca há um motor que produz resultados.



FIGURA 7- *The Anatomy of a SERP*



FONTE: MORVILLE; CALLENDER, 2010, p. 32. Disponível em:
< <http://www.flickr.com/photos/morville/4274260680/in/photostream/> >

No tópico *engine*, os autores referem-se à colaboração entre os engenheiros e os *designers* destacando a necessidade, destes últimos, o maior conhecimento sobre a tecnologia de busca, envolvendo os seguintes elementos: o mecanismo, o índice, os algoritmos e a interface. Abordam, ainda, o conteúdo, a indexação e a sua relação estreita com os metadados.

A busca é um fenômeno social, todavia Morville e Callender (2010) lembram que, “os criadores” não são só humanos, pois nesse agenciamento maquínico deve-se considerar também as tecnologias, os instrumentos e os processos de criação de sentido.

Broder (2002), que já foi diretor de tecnologia do *Altavista*, em seu artigo intitulado “A taxonomia da busca”, apresenta três motivos pelos quais as pessoas utilizam os buscadores, ou seja, os motivos navegacionais, informacionais e transacionais.

É interessante notar que a busca navegacional diz respeito ao uso dos mecanismos como portais de entrada no ciberespaço, aquilo que Batelle (2006, p. 25) define como “[...] o caminho mais curto para um site que sabemos que existe (a prática de digitar uma palavra conhecida para chegar a um site que se deseja visitar é chamada de pergunta-piloto).” O uso dos buscadores como portais, significa que:



Existem dois tipos de usuários que digitam a URL no sistema de busca ao invés de no campo de endereços do browser: aqueles suficientemente inexperientes para não compreender a diferença entre os dois e aqueles que são tão experientes que estão habituados a usar os buscadores como um portal para a internet. [...] Não importa se este comportamento é motivado por ignorância ou destreza, o resultado final é o mesmo: o buscador é o ponto focal da experiência online para todos os tipos de usuários da internet. (NIELSEN/ NETRATINGS *apud* FRAGOSO, 2007, p. 158).

A citação acima também responde a questão “Por que”, pois um dos usos dos mecanismos de busca é justamente esse, ou seja, ser redirecionado ao *site* sem a necessidade de “digitar”, no *browser*, toda a *URL* de um *site*.

Em relação aos motivos informacionais, os buscadores são utilizados como ferramentas para consulta de uma determinada informação ou conhecimento, a partir de termos previamente estabelecidos que possibilitem o acesso a quaisquer páginas e ambientes no ciberespaço.

Fazendo uma breve relação entre a busca navegacional e a busca informacional, é possível prever que esta última contribui para que haja certa continuidade no processo de busca em relação à primeira, tendo em vista que é realizada ou de forma muito estreita ou de forma muito ampla e o indivíduo ainda desconhece as possibilidades do ciberespaço relacionadas aos termos que utilizou. Consequentemente, ele fará novas tentativas até que consiga efetivamente encontrar o que procura (BRODER, 2002).

No que diz respeito à busca transacional, Broder (2002) entende que a intenção do indivíduo é realizar alguma atividade mediada pela *Web*. Nesse tipo de busca, há maior interação desse indivíduo devido ao fato que realizará uma determinada tarefa, como uma compra, o *download* de arquivos, a pesquisa em bases de dados, entre outros.

A partir dessa taxonomia para a busca, Broder (2002) compreende que os mecanismos se desenvolveram essencialmente em três gerações. Na primeira, havia principalmente uma página de dados (texto e formatação) e esses mecanismos estavam próximos ao modelo clássico de recuperação da informação, no qual não considera aspectos inerentes ao contexto de busca, ou seja, a Interação-Humano-Computador e os aspectos cognitivos, ou os aspectos semânticos da busca.

Na segunda geração, os mecanismos passaram a suportar buscas navegacionais e informacionais. Na atual terceira geração, segundo o autor, há tentativas de apresentação de dados de diversas fontes, com o objetivo de contextualizar o indivíduo no que diz respeito a sua busca e, ainda, compreender a necessidade por trás da busca, por meio de análise semântica, determinação do contexto entre outros aspectos.

Assim, os mecanismos objetivam articular e apoiar buscas navegacionais, informacionais e transacionais, sendo esta última o grande desafio, pois geralmente é realizada de maneira indireta e não contextualizada, o que indica o caminho para essa terceira geração atuando efetivamente



no ciberespaço e permitindo ao indivíduo a satisfação, também, de suas necessidades cognitivas (BRODER, 2002).

Seja qual for o motivo, busca não é mais um aplicativo isolado, uma ferramenta útil para achar algo na *Web*. “É nosso modo de navegar pelo único recurso infinito que move a cultura humana: o conhecimento.” (BATELLE, 2005, p. 241).

3 CONSIDERAÇÕES FINAIS

A “busca” tem-se consolidado na sociedade graças à prática de navegar no ciberespaço utilizando os mecanismos de busca como “oráculos pós-modernos”, apropriando-se da expressão de Soares (2008), e tanto faz se ela inclui a recuperação ou simplesmente proporciona a descoberta.

Os autores citados, neste trabalho, têm-se debruçado sobre este tema para estudar o fenômeno e todos estão contemplando, de alguma forma, as dimensões da linguagem, seja ela a sintática, semântica ou a pragmática, para entender essa nova gramática da *query*, seu contexto e sua utilização.

Vale salientar que a pragmática, no contexto do uso da linguagem nos mecanismos de busca, deve ser entendida no sentido deleuziano, ou seja, em que caso, onde, quando e como qualquer coisa surge, desenvolve-se e funciona? Assim, uma expressão evidencia práticas em funcionamento (ALMEIDA, 2003) cabendo-nos indagar sobre o funcionamento dos agenciamentos de linguagem nessas máquinas semióticas no ciberespaço.

Mas, uma pergunta feita por Batelle (2006) é inquietante. Existe a busca perfeita? Não obstante, a busca perfeita implica em memória ubíqua no ciberespaço. Há, ainda, mais outro elemento, que existe e insiste na dobra do signo no ciberespaço: a *Web* profunda, com suas camadas de invisibilidade e seus vários planos de existência que remetem ao conceito deleuziano de “mil platôs” (DELEUZE; GUATTARI, 1995, v.1).

Batelle (2006) acredita que a ubiquidade (a inserção e integração da informação e conhecimento no ciberespaço), a busca personalizada, a ascensão da *Web* Semântica, entre outros, poderão trazer a resposta perfeita, no que se pode duvidar facilmente, frente à complexidade humana.

Contudo, o aperfeiçoamento da sintaxe, a inclusão da semântica, a compreensão do contexto do usuário (pragmática) e o desenvolvimento dos algoritmos a partir da análise dos padrões de linguagem já ocorrem, e transformarão aquilo que chamamos de anatomia ou taxonomia em um ato cada vez **mais intuitivo**, na concretização daquilo que designamos como “busca” no ciberespaço. Haja vista as buscas realizadas em linguagem natural, no *Google*, como as frases que solicitam conversão de moeda, operações aritméticas, entre outras.

De acordo com pesquisas realizadas, devido às múltiplas sintaxes de indexação do conhecimento e da informação, operadas pelos mecanismos de busca, bem como pelas constantes atualizações de



seus algoritmos, não há somente um método válido, uma sintaxe geral que possa modelar os seus usos (MONTEIRO, 2006, 2009).

Assim, a busca é uma sintaxe em devir, de recursos variados, segundo as necessidades do usuário e os limites dos índices compilados pelos mecanismos, de tal modo que os operadores pragmáticos, envolvidos nesse processo, permitem construir um mapa de significados vigentes e atualizados, no momento da busca, no universo simbólico, virtual e movente que é o ciberespaço.

SEEKING THE COMPREHENSION OF SEARCH ON THE CYBERSPACE

Abstracts: *The search has become more and more embedded in contemporary society, whether it involves the retrieval information and knowledge or just a discovery. The Work investigates the search engines operated by the cyberspace, from documentary analysis, the authors in researching this social phenomenon. The approaches include the anatomy of search engines, search and its taxonomy. Whatever approach discussed by the author, technology, the user or the language is related to any of the three language dimensions syntax, semantics and pragmatics. Thus, this new grammar (syntax) search and its context (semantics) include, increasingly, the user (pragmatic), so the search makes sense on the cyberspace.*

Keywords: Search Engines. Search. Cybespace. Syntax. Semantics. Pragmatic.

REFERÊNCIAS

- ABBAGNANO, Nicola. **Dicionário de Filosofia**. São Paulo: Martins Fontes, 2003.
- ALMEIDA, Júlia. **Estudos deleuzianos da linguagem**. Campinas: Ed. UNICAMP, 2003.
- ARMEMGAUD, Françoise. **A pragmática**. São Paulo: Parábola Editorial, 2006.
- BATTELLE, John. **A busca**. Campinas: Campus; Rio de Janeiro: Elsevier, 2006.
- BRODER, Andrei. **A taxonomy of web search**. 2002. Disponível em: <<http://www.sigir.org/forum/F2002/broder.pdf>>. Acesso em: 2 maio 2010.
- FRAGOSO, Suely. Quem procura acha? O impacto dos buscadores sobre o modelo distributivo da *Word Wide Web*. **Eptic**: Revista de Economia Política de las Tecnologias da La Información y Comunicación, v.9, n.3, n.p., sept/dic. 2007. Disponível em: <<http://www.eptic.com.br>>. Acesso em: 15 mar. 2009.
- DELEUZE, Gilles. **A dobra: Leibniz e o barroco**. 5.ed. Campinas: Papirus, 1991.
- DELEUZE, Gilles. **Lógica do sentido**. São Paulo: Perspectiva: 1998.
- DELEUZE, Gilles; GUATTARI, Félix. **Mil platôs: capitalismo e esquizofrenia**. São Paulo: Ed. 34. 1995. v. 1.
- GOMES, Daniel. **Visibilidade.net**: estar presente na web não basta. Disponível em: <<http://visibilidade.net/tutorial/o-que-e-a-visibilidade.html>>. Acesso em: 02 maio 2008.



GOOGLE. Google Inc. Disponível em: <<http://www.google.com/>>. Acesso em: 08 fev. 2010.

GOOGLE Zeitgeist, 2007. Disponível em: <<http://www.google.com/intl/en/press/zeitgeist2007/mind.html>>. Acesso em 26 jul. 2011.

MONTEIRO, Silvana Drumond. O ciberespaço e os mecanismos de buscas: novas máquinas semióticas. **Ciência da Informação**, n. 35, v. 1, p. 31-38, jan./abr. 2006. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/663/577>>. Acesso em 08 ago. 2011.

MONTEIRO, Silvana Drumond. As múltiplas sintaxes dos mecanismos de busca no ciberespaço. **Informação & Informação**, v. 14, n. especial, p. 68-102, 2009. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/2027/3223>>. Acesso em: 8 fev. 2011.

MORVILLE, Peter; CALLENDER, Jeffery. **Search patterns: design for discovery**. Canadá: O'Reilly, 2010.

SANTAELLA, Lúcia. **Linguagens líquidas na era da mobilidade**. São Paulo: Paulus, 2007.

SOARES, Claudio. A anatomia da busca. **Pontolit**, 2008. Disponível em: <<http://blog.pontolit.com.br/autor/>>. Acesso em: 08 abr. 2009.

STATOWL. **Search engine keyword count usage**. Disponível em: <http://www.statowl.com/search_engine_keyword_count.php>. Acesso em: 08 ago. 2011.

YAHOO! Disponível em: <<http://www.yahoo.com.br>>. Acesso em: 08 fev. 2010.



ABORDAGENS PARA AVALIAR BIBLIOTECAS DIGITAIS

Izabel França Lima, Renato Rocha Souza, Guilherme Ataíde Dias

RESUMO: A pesquisa identifica abordagens usadas nas avaliações de bibliotecas digitais a partir da bibliografia de Cunha (2009) sobre a temática. A identificação das abordagens toma como base as levantadas por Saracevic (2005) em pesquisa sobre avaliação de bibliotecas digitais. Adotam-se os seguintes passos: identificar, na bibliografia, os artigos que tratam de avaliação de bibliotecas digitais; identificar os elementos determinantes das abordagens adotadas para avaliação de biblioteca digital. Metodologicamente, recorreu-se à abordagem qualitativa para tornar viável a proposta, adotando como categorias, da Análise de Conteúdo, as abordagens apresentadas por Saracevic (2005). Os resultados da pesquisa apontaram a abordagem, centrada no homem, como a que mais se apresenta nos estudos de avaliação de bibliotecas digitais, mas ressaltou-se que existem outras abordagens para avaliação, por isso devemos selecioná-las tendo em vista a sua adequação a diferentes públicos e objetivos de avaliação.

Palavras-chave: Biblioteca digital. Avaliação de bibliotecas digitais. Metodologias de avaliação de bibliotecas digitais.

1 INTRODUÇÃO

O estudo sobre bibliotecas está relacionado à própria trajetória do ser humano em sua busca constante pela informação para registrar, organizar, armazenar e usar o conhecimento, a fim de atender às suas demandas econômicas, políticas, sociais e culturais (PARENTE, 1999). É importante enfatizar que as técnicas de registro do conhecimento sempre fizeram parte das preocupações humanas. Elas se iniciaram com os povos primitivos, quando as primeiras tentativas de talhar a pedra exerciam um extraordinário poder de abstração. Com a evolução das sociedades, procurou-se aprimorar as formas de conhecer, passando por três fases: a oral (culturas tribais); a impressa ou visual (expansão e uso da cultura tipográfica) e a ciberespacial (predomínio das TIC) (BARRETO, 1998).

As bibliotecas digitais, como organização social, são constituídas por serviços e produtos diferenciados, têm a função de selecionar, organizar, indexar, disponibilizar, disseminar e democratizar a informação, sendo concebidas como redutoras das barreiras físicas e das distâncias, aspectos que sempre limitaram os serviços das bibliotecas físicas. Apesar das facilidades que as bibliotecas digitais oferecem, uma grande parte dos usuários demonstram ter dificuldades quanto ao acesso e ao uso, o que nos leva a suspeitar que os dispositivos são afetados por limitações de ordem técnica. Muitas vezes, ao interagirem com esses dispositivos, os usuários se deparam com empecilhos informacionais, técnicos e de acesso (GOMES, 1983).



A biblioteca digital se caracteriza por serviços e objetos de informação, disponíveis no meio digital, dotados de organização, estrutura e interface com vistas a suportar a interação dos usuários. Estudos sobre avaliação dessas bibliotecas são fundamentais para conhecer suas limitações e potencialidades, consideradas fatores relevantes no ambiente informacional, visto que por meio delas o usuário acessa o sistema com o qual interage.

Argumentamos que as bibliotecas digitais parecem necessitar de avaliação contínua, prática que possibilita ao gestor detectar problemas e buscar soluções para melhor uso desse dispositivo informacional.

Cunha (2009) e Saracevic (2004) afirmam que as bibliotecas digitais ainda são pouco avaliadas e apresentam reflexões acerca de como fazê-lo, bem como se é possível usar as mesmas metodologias aplicadas às bibliotecas tradicionais. Enquanto Saracevic (2004) relata que as discussões sobre bibliotecas digitais são abundantes, exceto quando tratam de avaliação; acrescenta que na literatura sobre sua avaliação podem ser encontrados dois tipos distintos de relatos: a) trabalhos que sugerem conceitos de avaliação, modelos, abordagens, metodologias ou discutem avaliação e b) trabalhos que relatam avaliação real, ou seja, aplicação de metodologias quantitativas (estatísticas) ou qualitativas (impressões).

Para Cunha (2009) essas indagações ainda não obtiveram respostas definitivas. Ao publicar a bibliografia que inclui publicações sobre a avaliação de bibliotecas digitais afirma que:

As bibliotecas tradicionais têm sido avaliadas desde os tempos remotos; por outro lado, as bibliotecas digitais, para justificarem os recursos financeiros, documentais e humanos nelas alocados, também necessitam serem avaliadas. Mas, como elas deverão ser avaliadas? Podem ser usadas as mesmas métricas empregadas nas bibliotecas tradicionais? (CUNHA, 2009)

Com a crescente demanda de informações, sobretudo em meio digital, o que está em jogo, na sociedade da informação e do conhecimento, é a necessidade de engendrar repertórios capazes de transcender o formato códex e a cultura impressa como únicas possibilidades de disponibilização, acesso e uso do conhecimento, para maximizar as de armazenagem.

A pesquisa objetiva identificar as abordagens usadas nas avaliações de bibliotecas digitais. Para isso, adotaram-se as seguintes etapas: a) Identificar na bibliografia os artigos que tratam de avaliação de bibliotecas digitais; b) Identificar os elementos característicos determinantes das abordagens adotadas para avaliação de biblioteca digital; c) Apresentar abordagem mais adotada para avaliação de biblioteca digital a partir dos modelos encontrados nessa literatura.

A consecução da pesquisa e seus resultados contribuem reunindo um conjunto de abordagens metodológicas sobre avaliação de bibliotecas digitais. A nosso ver reside aí sua importância inclusive para o desenvolvimento de novos estudos sobre tecnologias em ambientes digitais.



2 AVALIAÇÃO DE BIBLIOTECAS DIGITAIS

No domínio das bibliotecas digitais, tanto quanto em qualquer outro, os sistemas de avaliação são apresentados como solução para fatores intervenientes e questionamentos. Mesmo assim, a avaliação não é uma atividade ampla, nem mesmo em crescimento nas bibliotecas digitais. Na realidade, a avaliação é mais observada por sua ausência ou presença mínima na literatura sobre bibliotecas digitais, tanto na pesquisa quanto na prática; a avaliação parece ser uma exceção e não uma regra.

Segundo Saracevic (2004) as bibliotecas digitais são intrincadas, constituem mais do que sistemas tecnológicos e a avaliação de sistemas complexos é igualmente complexa. O autor afirma que não há interesse na avaliação e acrescenta “aqueles que fazem ou pesquisam bibliotecas digitais estão interessados em fazer, construir, implementar, abrir novos caminhos, operar [...] a avaliação é de pouco ou nenhum interesse, além de não haver tempo para isto” (SARACEVIC, 2004, p. 10).

Com o advento da World Wide Web, os avanços das tecnologias de computação, as pesquisas e projetos de criação e implantação de bibliotecas digitais cresceram muito desde os anos 2000. Com isso um número de pesquisadores e profissionais de diferentes áreas do conhecimento (por exemplo, ciência da informação; tecnologia de informação e ciência da computação) têm sido atraídos para esta área. Outro interesse recai sobre o ritmo e dimensão de projetos, entidades e atividades em torno das bibliotecas digitais.

Percebe-se isto por meio do crescimento de periódicos especializados, por exemplo, *D-Lib Magazine*, *International Journal of Digital Libraries*, *Journal of Digital Information*, de eventos específicos como *ACM Joint Conference on Digital Libraries*, *European Conference on Research and Advanced Technology for Digital Libraries*, *International Conference on Asian Digital Libraries*, Seminário Internacional de Bibliotecas Digitais e de organizações como *Digital Library Federation*, que visam especificamente este tema. No entanto, por trás deste cenário de rapidez existem, para alguns autores, pontos fracos que podem dificultar o progresso das bibliotecas digitais. Um deles consiste na falta de definição do conceito de biblioteca digital.

Há duas versões para alguns pontos conflitantes sobre biblioteca digital: uma na perspectiva do pesquisador que a vê como uma coleção de informação organizada e armazenada digitalmente (FOX et al, 1993; LESK, 1997); a outra pode ser encontrada, principalmente, a partir dos profissionais que atuam em bibliotecas, que compreendem a biblioteca digital como a extensão institucional das bibliotecas tradicionais em ambientes digitais (*Digital Library Federation*). Na opinião de Borgman (1999), essas visões prejudicam o avanço das bibliotecas digitais porque restringem as fronteiras entre as coleções digitais e as instituições.

A percepção sobre a necessidade de um acordo quanto à concepção de biblioteca digital motivou ao *D-Lib Working Group on Digital Library*¹, propor a seguinte definição, em reunião

¹ <http://www.dlib.org/metrics/public/>



realizada em 07 de janeiro de 1998, na Universidade de Stanford: “A biblioteca digital é uma coleção de serviços e objetos informacionais que facilitam para o usuário a utilização da informação, inclusive a apresentação e organização desses objetos, disponíveis direta ou indiretamente por meio de instrumentos eletrônicos/digitais”

A proposição do conceito no âmbito do grupo buscou o entendimento comum do que seja biblioteca digital, uma vez que visam a desenvolver um conjunto de métricas adequadas para avaliar e comparar a eficácia dessas bibliotecas.

Tammaro e Salarelli (2008) observam que os serviços das bibliotecas digitais são diferentes dos serviços das bibliotecas tradicionais; além disso, as possibilidades de medir o uso das publicações digitais diferem das usadas nas publicações impressas. Por conseguinte, os modelos de avaliação das bibliotecas tradicionais, em alguns aspectos, necessitam de adequações, enquanto outros não se aplicam quando tratamos de bibliotecas digitais.

Um dos problemas da avaliação de bibliotecas digitais encontra-se na coleta dos dados: falta exatidão nos dados sobre uso, faltam normas internacionais destinadas à mensuração das bibliotecas digitais; faltam diretrizes para enfrentar os problemas relativos às coleções digitais e à medição de seu uso (TAMMARO; SALARELLI, 2008).

Para Blandford et al. (2008) muitos estudos publicados sobre avaliação de bibliotecas digitais são relatos de avaliações de sistemas específicos, envolvendo testes com usuários ou avaliação de especialistas. Esses estudos de avaliação podem se basear em análise quantitativa, como os que envolvem o uso de *logs* de transação e em análise qualitativa, como entrevista ou observação. Tais estudos ilustram a diversidade de abordagens possíveis e a variedade de questões, ao avaliar bibliotecas digitais.

Zhang (2007) considera que a avaliação de bibliotecas digitais pode pedir abordagens e critérios também utilizados nas avaliações do sistema de recuperação de informação das bibliotecas tradicionais, mas é essencial desenvolver modelos de avaliação específicos para essa biblioteca. Até porque, segundo o autor, com um enorme consumo de recursos técnicos, financeiros e de pessoal empregado em cada projeto de implantação de uma biblioteca digital, este deve ser avaliado para garantir o resultado de seu desenvolvimento.

Sob esse ponto de vista, Marchionini (2000) considera que a essência da avaliação de bibliotecas digitais reside na recolha sistemática de dados e integração de pontos de vista diferentes, usando diferentes abordagens e de diferentes dimensões.

Saracevic, ao relatar pesquisa na Conferência LIDA sobre avaliação de bibliotecas digitais, identificou diferentes abordagens com denominações também estabelecidas e aplicadas por ele:

Abordagem centrada nos sistemas: amplamente utilizada. Envolve o estudo de algum aspecto de desempenho. [...]; **Abordagem centrada no homem:** também amplamente utilizada. Envolve estudos de comportamento em relação a dadas necessidades de informação, tais como busca de informação, navegação, pesquisa ou desempenho no cumprimento de tarefas atribuídas, seja pré-determinada ou observada em cenários naturais. [...]; **Abordagem**



centrada na usabilidade: Envolve avaliação de diferentes características, particularmente em relação a portais, pelos usuários. [...]; **Abordagem etnográfica:** Envolve estudos dos modos de vida, cultura e costumes em um ambiente de biblioteca digital. Também envolve o estudo do impacto de uma biblioteca digital em uma dada comunidade. [...]; **Abordagem antropológica:** Envolve o estudo de diferentes partes interessadas ou comunidades e suas culturas em relação à uma dada biblioteca digital. [...]; **Abordagem sociológica:** Envolve avaliação de ação situada ou comunidades de usuários no ambiente social de uma biblioteca digital. [...]; **Abordagem econômica:** Envolve estudo de custos, custo-benefício, valores e impactos econômicos (SARACEVIC, 2005, p. 5).

Para o autor, os “níveis de avaliação variaram de micro-avaliação de determinadas características até macro-avaliação do impacto de uma biblioteca digital em um campo” (SARACEVIC, 2005, p.5). Além das abordagens citadas, vários outros estudiosos destacam uma variedade de fatores intervenientes da biblioteca digital e da necessidade de incluir suas diversas perspectivas nas avaliações.

3 PROCEDIMENTOS METODOLÓGICOS

Acenando para o foco de interesse que gravita em torno da avaliação de *bibliotecas digitais*, o estudo é uma pesquisa bibliográfica, de caráter exploratório, analítico e comparativo. Metodologicamente, recorreu-se à abordagem qualitativa, para tornar viável a proposta, adotando um delineamento de pesquisa (coleta, interpretação e análise) baseada na análise de conteúdo, adotando como categorias de análise as abordagens apresentadas por Saracevic (2005), uma vez que as categorias podem ser definidas *a priori*. (BARDIN, 2010; FRANCO, 2007; VALENTIM, 2005).

No primeiro momento do levantamento bibliográfico recupera-se a produção científica sobre avaliação de biblioteca digital a partir da bibliografia exaustiva sobre biblioteca digital realizada por Cunha (2009). No momento seguinte, identificamos 17 artigos, no subitem avaliação de bibliotecas digitais da citada bibliografia. Entendemos que por tratar desse assunto deveriam tais artigos focar abordagens metodológicas sobre avaliação. Desse total, 10 artigos atendiam aos critérios estabelecidos nos objetivos dessa pesquisa. Esclarecemos que descartamos os textos que, mesmo indicados na bibliografia, tratavam de tema diverso da avaliação de bibliotecas digitais. A seguir, no Quadro 1, a relação dos artigos e os objetivos de cada um deles.



QUADRO 1 – Artigos citados na bibliografia consultada que abordam a questão de avaliação de bibliotecas digitais.

Nº	ARTIGO	OBJETIVO DO ESTUDO
	Blandford, A.; Adams, A.; Attfield, S.; Buchanan, G.; Gow, J.; Makri, S.; Rimmer, J.; Warwick, C. The PRET A Reporter framework: Evaluating digital libraries from the perspective of information work. <i>Information Processing and Management</i> , v. 44, n. 1, p. 4-21, 2008.	Conhecer como os sistemas de recuperação da informação, e em particular, as bibliotecas digitais podem ser avaliados para melhorar sua interação com os usuários.
	Choudhury, S.; Hobbs, B.; Lorie, M.; Flores, N. A framework for evaluating digital library services. <i>D-Lib Magazine</i> , v. 8, n. 7-8, July/August 2002	Analisar a aplicação de uma metodologia com base em um atributo multimodelo de indicativo econômico, utilizado para avaliar o amplo acesso a materiais impressos (CAPM)
	Fuhr, N. et al. Evaluation of digital libraries. <i>International Journal of Digital Libraries</i> , v. 8, n.1, 2007.	O objetivo é fornecer um conjunto de diretrizes flexíveis e adaptáveis para avaliação bibliotecas digitais.
	Hill, L. L.; Carver, L.; Larsgaard, M.; Dolin, R.; Smith, T. R.; Frew, J.; Era, M. A. Alexandria Digital Library : User evaluation studies and system design. <i>Journal of the American Society for Information Science</i> , v. 51, n. 3, p. 246-259, 2000.	Apresentar o desenvolvimento de uma geobiblioteca como plataforma de testes para avaliação pelos usuários
	Kilker, J.; Gay, G. The social construction of a digital library: a case study examining implications for evaluation. <i>Information Technology and Libraries</i> , v. 17, p. 60-70, June 1998.	Analisar a construção social da tecnologia (SCOT) como múltiplas perspectivas inerentes ao desenvolvimento e avaliação de bibliotecas digitais
	Marchionini, G.; Plaisant, C.; Komlodi, A. The people in digital libraries: multifaceted approaches to assessing needs and impact. 1999	Entender o impacto das bibliotecas digitais e avaliar como estas afetam o comportamento informacional humano
	Peterson, E. Evaluation of digital libraries using snowball sampling. <i>First Monday</i> v. 10, n. 5, May 2005.	Distinguir padrões de consumo e recolher declarações qualitativas a respeito do uso e a melhoria do site da Montana Sistema de Informação de Recursos Naturais (NRIs)
	Saracevic, T. Digital library evaluation: toward an evolution of concepts. <i>Library Trends</i> , v. 49, n. 2, p. 350-369, 2000.	Enumerar os desafios enfrentados pela avaliação de bibliotecas digitais e os requisitos fundamentais para avaliação
	Shen, Rao.; Vemuri, N. S.; Fan, W.; Fox, E. A. What is a successful digital library? In: <i>European Conference on Digital Libraries</i> , ECDL 2006	Desenvolver um modelo de biblioteca digital de sucesso a partir de uma perspectiva do usuário final, integrando diversas áreas de pesquisa.
110.	Yang, S. C. An interpretive and situated approach to an evaluation of Perseus digital libraries. <i>Journal of the American Society for Information Science and Technology</i> , v. 52, n. 14, p. 1210-1223, 2001.	Desenvolver teorias de interação humano-computador e aplicar em sistemas hipermídia de aprendizagem, buscando a compreensão do ensino-aprendizagem

Fonte: Dados da pesquisa, 2011.2.1 Dados da pesquisa



A partir dos procedimentos de recuperação da informação, nos artigos apresentados no Quadro 1, apresentamos uma síntese dos artigos nos quais identificamos as abordagens citadas por Saracevic (2005). Para cada um dos artigos há a referência e uma síntese desses textos que compõem o *corpus* da pesquisa, ordenados por ano de publicação. Em quadros, identificamos os termos usados para caracterizar o tipo de abordagem adotada.

3.2 Análise dos Dados

Como explicitado, os textos selecionados na pré-análise estão relacionados por ano de publicação, seguidos de uma síntese.

1998

KILKER, J.; GAY, G. *The social construction of a digital library: a case study examining implications for evaluation*. **Information Technology and Libraries**, v. 17, p. 60-70, June 1998.

Síntese: O artigo é um estudo de caso que focaliza múltiplas perspectivas inerentes ao desenvolvimento e avaliação de bibliotecas digitais. Utiliza a abordagem da construção social da tecnologia (SCOT) para descrever os diversos grupos sociais de usuários envolvidos com a avaliação da interface de uma biblioteca digital. Argumenta que as avaliações desta tecnologia devem ter em conta as interações entre as diferentes necessidades de tais grupos sociais. Característica de abordagem Sociológica.

1999

MARCHIONINI, G.; PLAISANT, C.; KOMLODI, A. *The people in digital libraries: multifaceted approaches to assessing needs and impact*. 1999

Síntese: A pesquisa adota uma abordagem multifacetada para avaliar as necessidades de informação e entender o impacto das bibliotecas digitais considerando as características e contextos das pessoas que será ou poderá ser usuárias dessas bibliotecas. A abordagem multifacetada para avaliação de bibliotecas digitais foi aplicada em três estudos de caso, com particular ênfase na avaliação das necessidades do usuário para desenvolver um protótipo de interface para a Biblioteca do Congresso como parte do Programa Nacional de Biblioteca Digital. Característica de abordagem etnográfica.

2000

HILL, L. L. et al. *Alexandria Digital Library: User evaluation studies and system design*. **Journal of the American Society for Information Science**, v. 51, n. 3, p. 246-259, 2000.

Síntese: Descreve a evolução do sistema e do design de interface da Alexandria Digital Library (ADL) e, em seguida, fornece uma visão geral das atividades de pesquisa e testes do projeto. Apresenta as metodologias de avaliação do usuário aplicadas e as descobertas resultantes



do trabalho com três grupos de usuários. O texto descreve a parte de avaliação do projeto. Característica de abordagem centrada no sistema.

SARACEVIC, T. *Digital library evaluation: toward an evolution of concepts*. **Library Trends**, v. 49, n. 2, p. 350-369, 2000.

Síntese: Apresenta conceitos para a avaliação de bibliotecas digitais que o autor reconhece como uma tarefa complexa e difícil. Enumera alguns dos desafios enfrentados pela avaliação dessas bibliotecas, bem como uma revisão dos esforços das pesquisas e práticas de avaliação além de requisitos fundamentais para avaliação das bibliotecas digitais. Não foram observadas quaisquer características das abordagens estudadas.

2001

YANG, S. C. *An interpretive and situated approach to an evaluation of Perseus digital libraries*. **Journal of the American Society for Information Science and Technology**, v. 52, n. 14, p. 1210-1223, 2001.

Síntese: Este artigo é baseado em estudos de alunos tentando usar um ambiente hipermídia como parte das atividades educacionais e reflete sobre a concepção e uso de sistemas hipermídia de aprendizagem. Além disso, o artigo é uma reflexão sobre a concepção e utilização do sistema Perseus. Característica de abordagem centrada no homem.

2002

CHOUDHURY, S. et al. *A framework for evaluating digital library services*. **D-Lib Magazine**, v. 8, n. 7-8, July/Aug. 2002.

Síntese: Discute sobre estudos de avaliação de biblioteca digital. O estudo centra-se em metodologias de avaliação que considera os resultados, impactos ou benefícios dos serviços de bibliotecas. Os autores aplicam o método do multiatributo, usualmente adotado em pesquisas de mercado, objetivando o *feedback* dos usuários de uma biblioteca digital sobre o custo provável a ser pago pela informação disponibilizada. Característica de abordagem econômica.

2005

PETERSON, E. *Evaluation of digital libraries using snowball sampling*. **First Monday**, v. 10, n. 5, May, 2005.

Síntese: A pesquisa é um estudo de caso, comparando resultados obtidos por dois métodos de entrevista com usuários de uma biblioteca digital. O estudo buscava conhecer a satisfação dos usuários com a coleção. Na primeira etapa os usuários foram entrevistados presencialmente e no segundo momento, os mesmos usuários foram consultados por e-mail. Os resultados constataram que o e-mail não é a maneira mais eficaz para conhecer



a satisfação do usuário de uma biblioteca digital. Características de abordagem centrada no homem.

2006

SHEN, R. et al. What is a successful digital library? In: EUROPEAN CONFERENCE ON DIGITAL LIBRARIES, 10, 2006. *Research and Advanced Technology for Digital Libraries*. Alicante, Spain: ECDL, Sep. 2006

Síntese: Apresenta um modelo de avaliação de sucesso de biblioteca digital a partir da perspectiva do usuário final, integrando diversas áreas de pesquisa. Identifica fatores que afetam a qualidade das informações e do sistema da biblioteca digital; aponta relações entre esses fatores e os indicadores de qualidade desenvolvidos por outros pesquisadores; desenvolve indicadores e medidas de qualidade para o catálogo e os serviços de uma biblioteca digital. Característica de abordagem centrada no homem.

2007

FUHR, N. et al. *Evaluation of digital libraries. International Journal of Digital Libraries*, v. 8, n.1, p. 21-38, 2007.

Síntese: Apresenta um estado da arte sobre avaliação de biblioteca digital. Em paralelo estabelece uma estrutura para avaliar sistemas de bibliotecas digitais de forma holística e faz recomendações com relação às ações estratégicas e necessárias a serem seguidas em avaliações desse tipo de biblioteca. Indica necessidade de avaliação.

2008

BLANDFORD, A. et al. *The PRET a rapporter framework: evaluating digital libraries from the perspective of information work. Information, Processing and Management*, v. 44, n. 1, p. 4-21, 2008.

Síntese: Relata aplicação de avaliação do sistema de recuperação da informação (IR) com foco no usuário. Apresenta um quadro para o planejamento de estudos de avaliação centrada no usuário que estabelecem sistemas no contexto da IR com aplicabilidade em avaliação de bibliotecas digitais. Característica de abordagem centrada no homem.

Para conhecer a abordagem adotada na avaliação de bibliotecas digitais presente nos artigos pesquisados, identificamos os termos apresentados que caracterizam a abordagem e usadas para definir as categorias de análise, como consta no Quadro 2 ordenados alfabeticamente. Os artigos que tratam de alunos, usuários ou assemelhados caracterizam o ser humano,



QUADRO 2 – Termos característicos das abordagens adotadas nos artigos

TERMOS CARACTERÍSTICOS DAS ABORDAGENS ADOTADAS NOS ARTIGOS	
Análise econômica do sistema	Inovação tecnológica
Avaliação de sistemas	Interação Humano-Computador
Busca de informação	Necessidade de informação
Comportamento de busca	Necessidades dos usuários
Contextos sociais	Navegação
Custo-benefício	Medição dos custos
Custos potenciais	Preço de serviço
Design de interface	Navegabilidade
Facilidade de uso	Novos sistemas de software
Grupos sociais	Satisfação do usuário
Impacto das bibliotecas digitais	Serviços de biblioteca
Indicadores de qualidade	Sistema de informação
Informações sobre seus usuários	Valor de uso

FONTE: Dados da pesquisa, junho/2011

Com o objetivo de conhecer qual abordagem prevalece na avaliação de biblioteca digital, optamos por trabalhar com o tema enquanto unidade de análise. Como as categorias de análises foram definidas *a priori*, com base nas abordagens apresentadas por Saracevic (2005) em pesquisa como as bibliotecas digitais são avaliadas, os termos encontrados nos artigos analisados foram agrupados por subcategorias de análise, nas categorias referidas.

No quadro a seguir, organizamos as subcategorias do quadro anterior em categorias que caracterizam as abordagens apresentados pelo autor, recorrendo às abordagens apresentadas por Saracevic (2005). Vale ressaltar que alguns termos são recursivos nos textos, inclusive os que adotam abordagens diferentes. Como resultado da identificação das abordagens adotadas, temos que a **abordagem centrada no homem** foi a mais frequente, aplicada em quatro dos textos pesquisados. As abordagens: **centrada nos sistemas**, **etnográfica**, **sociológica** e **econômica** foram identificadas cada uma em um texto e as abordagens **centrada na usabilidade** e **antropológica** não aparecem nos textos pesquisados. Finalmente, em dois textos que compõem o *corpus* da pesquisa não está clara



qual a abordagem usada.

QUADRO 3 – Categorias e sub categorias de análises presentes nos artigos analisados

Categorias de Análises	<i>Sub Categorias de Análises</i>
Abordagem centrada nos sistemas	avaliação de sistemas novos sistemas de software sistemas de informação
Abordagem centrada no homem	busca de informação comportamento de busca design de interface facilidade de uso indicadores de qualidade informações sobre seus usuários interação humano-computador navegabilidade navegação necessidade de informação necessidades do usuário satisfação do usuário serviços de biblioteca
Abordagem centrada na usabilidade	
Abordagem etnográfica	impacto das bibliotecas digitais contextos sociais
Abordagem antropológica	
Abordagem econômica	análise econômica do sistema custo-benefício custos potenciais medição dos custos preço de serviços valor de uso
Abordagem sociológica	contexto social grupos sociais

FONTE: Dados da pesquisa, junho/2011



Nos artigos analisados com base na pesquisa de Saracevic (2005) foram identificadas características que apontam para cinco abordagens, sendo que dois dos textos abordam a questão da avaliação de bibliotecas digitais teoricamente, deixando de apresentar termos que caracterizem a abordagem adotada. O resultado dessa fase apontou para a seguinte representação das abordagens: uma sociológica; uma econômica; uma etnográfica; uma centrada no sistema e quatro centradas no homem.

No quadro 3 observamos que a abordagem com maior incidência de termos que caracterizam subcategorias de análise foi a **centrada no homem**, identificada como já colocamos, em quatro textos. Na pesquisa de Saracevic (2005) a **abordagem centrada no sistema** foi amplamente utilizada e a **centrado no homem**, a segunda. Inferimos que a mudança se deve à questão temporal, visto que nos textos analisados de 2005 a 2008 surgem as pesquisas centradas no usuário, portanto com abordagem centrada no homem. O próprio autor justifica essa mudança quando afirma que “toda avaliação é também temporal. Algumas avaliações tiveram um componente temporal muito acentuado, particularmente em relação à avaliação de determinadas características tecnológicas” (SARACEVIC, 2005, p. 5).

A abordagem **centrada no homem** é utilizada em uma série de estudos que esclareceram o comportamento humano. As exigências, necessidades ou dificuldades encontradas na interação com o sistema, ficam evidentes nas subcategorias extraídas dos textos analisados como pode ser constatado no Quadro 3, ao apontar questões importantes para atender às necessidades do usuário, lembrando que ele é a razão de ser da biblioteca digital.

A abordagem **etnográfica** é apropriada para se obter um claro entendimento do papel e dos efeitos de uma prática ou construto em uma estrutura social ou organizacional mais ampla, por isso identificamos os termos como “grupos sociais” e “impactos sociais”. Já na abordagem **sociológica**, adequada para esclarecer as forças e os efeitos sociais, os termos localizados foram “grupos sociais” e “contexto social”, numa pesquisa que aborda o uso social das tecnologias de informação e comunicação. Quanto à abordagem **econômica**, é usada para representar fatores econômicos, como “análise de custo”, “valor de uso”, “preço de serviços”, adequado ao texto analisado que avaliou o custo aceito por usuários para acesso à informação.

As abordagens não encontradas nos textos pesquisados, no caso a **antropológica** envolvendo o estudo de diferentes comunidades e suas culturas em relação a uma dada biblioteca digital é um estudo muito específico, pouco encontrado na literatura. No entanto, a abordagem centrada na **usabilidade** é considerada uma ponte entre as abordagens centradas em sistemas e no homem. Muito utilizada em pesquisas que avaliam Web sites e portais, geralmente com aplicação de testes de usabilidades. O uso dessa abordagem na avaliação de bibliotecas digitais cresceu nos últimos anos, uma vez que objetiva detectar problemas e dificuldades de interação entre o usuário e o sistema. Talvez isso também contribua para o aumento das pesquisas com abordagem centrada no homem.



3.2.1 Compreendendo as subcategorias de análise

As subcategorias capturadas dos textos que compõem o *corpus* desta pesquisa são termos que identificam as abordagens apresentadas por Saracevic (2005), quando confrontados com a descrição apresentada pelo autor para cada tipo de abordagem. A seguir, analisaremos essas subcategorias tomando com base teórica o texto de Saracevic (2005) e o conteúdo dos textos analisados.

A **abordagem centrada no homem** agrupa o maior número de subcategorias por ser a mais identificada nos textos analisados, são elas: comportamento de busca; design de interface; interação humano-computado, navegação, satisfação do usuário, observadas no artigo de Blandford et al, (2008, p.1) e que podemos confirmar na afirmação do texto: “[...] apresentamos um quadro para o planejamento de estudos de avaliação centrado no usuário”. Já a subcategoria informações sobre seus usuários, claramente presente na frase “[...] muitos criadores de bibliotecas digitais querem obter mais informações sobre seus usuários do que os arquivos de log podem fornecer, por isso os estudos centrados no usuário estão começando a aparecer” do texto de Peterson (2005, p. 1), deixa evidente a abordagem usada na pesquisa. Na análise, localizamos as subcategorias facilidade de uso, indicadores de qualidade, navegação, necessidade de informação, satisfação do usuário. No trabalho de Shen (2006, p. 2) está patente a abordagem quando diz que “faz-se necessário entender as necessidades mutantes dos usuários ao interagirem com a biblioteca digital, e o comportamento de busca de informações dos usuários durante estas etapas”. No textos de Yang (2001) temos como subcategorias: interação humano-computador e design de interface, e a frase “Perseus Digital Library deve ser projetada para se adaptar facilmente às diferentes expectativas e preferências do aluno[usuário]”, confirmando a abordagem centrada no homem. Para ratificar a análise acima, Saracevic (2005, p.5) caracteriza essa abordagem como a que “envolve estudos de comportamento em relação a dadas necessidades de informação, tais como busca de informação, navegação, pesquisa ou desempenho no cumprimento de tarefas atribuídas” e acrescenta tratar-se de “estudos que esclarecem o comportamento humano, as exigências, necessidades, ou dificuldades encontradas”.

Na **abordagem centrada nos sistemas** temos as subcategorias: avaliação de sistema; novos sistemas de software e sistemas de informação, identificadas no texto analisado de Hill et al (2000, p. 1) que “descreve a evolução do sistema e do design de interface da Biblioteca Digital de Alexandria”, ao buscarmos o que é apresentado como características na abordagem de Saracevic (2005, p. 5) “[...]Envolve o estudo de algum aspecto de desempenho[...]ou algum design específico ou algum componente tecnológico”, constatamos que atende às características. Ressaltamos, ainda, que os objetivos da avaliação apresentada no texto também evidenciam o foco no sistema.

Quanto à **abordagem etnográfica**, identificamos como subcategorias os termos: impacto das bibliotecas digitais e impactos sociais, presentes no texto de Marchionini, Plaisant e Komlodi (1999), constatados quando os autores falam que dada à complexidade das necessidades de informação humana e da incerteza sobre os efeitos de novos sistemas, uma multiplicidade de dados é essencial para



orientar a elaboração e entender o impacto das bibliotecas digitais. Ponderam, ainda, que compreender os contextos sociais e culturais são esforços inescapáveis para isolar variáveis específicas para a avaliação. Esta afirmação corrobora Saracevic, (2005, p. 5) ao afirmar que “[...] envolve o estudo do impacto de uma biblioteca digital em uma dada comunidade”.

Na **abordagem econômica**, as subcategorias observadas são: análise econômica do sistema; custo-benefício; custos potenciais; medição dos custos; preço de serviços, consubstanciadas em Choudhury et al (2002, p. 6) ao afirmarem que a “[...] abordagem tem sido cada vez mais utilizada para a análise custo-benefício das obras públicas” e ainda, em “[...] uma avaliação completa de qualquer sistema ou projeto, custos e benefícios devem ser considerados”, e a teoria adotada nesse estudo menciona que a abordagem “Envolve estudo de custos, custo-benefício, valores e impactos econômicos (SARACEVIC, 2005, p. 5).

A **abordagem sociológica** apresenta as subcategorias: contexto social e grupos sociais conforme explicitam Kilker e Gay (1998, p. 1) ao informar que o “estudo de caso utiliza a construção social da tecnologia (SCOT) para o desenvolvimento e a avaliação de um protótipo de biblioteca digital [...] que busque atender necessidades de grupos sociais específicos”. Sobre esse ponto Saracevic (2005, p. 5) coloca que “Envolve avaliação de ação situada ou comunidades de usuários no ambiente social de uma biblioteca digital.”

4. CONSIDERAÇÕES FINAIS

A avaliação de bibliotecas digitais é ainda uma prática pouco adotada, inexistindo critérios gerais e padronizados para realizá-la. Para Saracevic (2005, p. 6) “vários esforços que se dedicam a desenvolver métricas de bibliotecas digitais não produziram, até agora, métricas generalizáveis e aceitáveis”. Observa-se, na literatura da área, uma vasta produção sobre bibliotecas digitais tratando de questões como projeto e arquitetura; metadados, ontologias, usuários, entre outras. Mas quando o assunto é avaliação existem trabalhos na literatura, especialmente internacional, que tratam quase sempre das abordagens, metodologias e conceitos, sintetizados no item 2.2.1 no qual se verifica os princípios que são adotados em cada uma dessas abordagens. Ainda segundo Saracevic, a produção de dados que pode corroborar efetivamente para a definição de métricas usadas como padrão para a avaliação de bibliotecas digitais ainda é escassa.

Quanto às abordagens para a avaliação, elas existem. Entretanto, devemos reconhecer a adequação de diferentes abordagens para diferentes públicos e para diferentes objetivos de avaliação. Claramente, cada abordagem tem pontos fortes e fracos, inexistindo uma abordagem “melhor” que as demais. Portanto, é ingênuo defender a predominância ou padronização de uso de determinada abordagem. O objetivo da avaliação deve servir como base para a seleção de uma ou mais abordagens adequadas.



APPROACHES TO EVALUATE DIGITAL LIBRARIES

ABSTRACT: *From the Cunha's bibliographical survey (2009) on the topic, this research aims to identify the approaches used in digital library evaluations. To accomplish our goal, the approaches identified by Saracevic (2005) in a research on digital library evaluation were taken as a basis. The following steps were adopted: identify, in the literature, the articles addressing digital library evaluation; identify determining elements of the approaches adopted in digital library evaluation. Methodologically, a qualitative approach was used, to make the proposal feasible, adopting the Content Analysis and Saracevic's approaches as categories. Results pointed the human-centered approach as the most frequently used in digital library evaluation studies, but it is worth mentioning that there are other evaluation approaches, so one should select them in view of the appropriateness of the approach to different audiences and evaluation goals.*

Keywords: Digital library. Digital library evaluation. Digital library evaluation methodologies.

REFERÊNCIAS

BARDIN, Laurence. **Análise de Conteúdo**: Lisboa: Edições 70, 2010.

BARRETO, Aldo de Albuquerque. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. **Ciência da Informação**, Brasília, DF, v.27, n.2, p.122-127, maio/ago. 1998.

BLANDFORD, A. et al. The PRET a rapporter framework: evaluating digital libraries from the perspective of information work. **Information, Processing and Management**, v. 44, n. 1, p. 4-21, 2008.

BORGMAN, C. L. What are digital libraries? competing visions. **Information Processing and Management**, v. 35, n. 3, p. 227-243, 1999.

CUNHA, Murilo Bastos. Bibliografia sobre o fluxo do documento na biblioteca digital. **DataGramZero - Revista de Ciência da Informação**. Rio de Janeiro, v. 10, n. 5, out. 2009. Disponível em: <http://www.dgz.org.br/out09/Art_01.htm>. Acesso em: 20 ago. 2010.

CUNHA, Murilo Bastos. Das bibliotecas convencionais às digitais: diferenças e convergências. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 13, n. 1, p. 2-17, jan./abr. 2008

CUNHA, Murilo Bastos. Desafios na construção de uma biblioteca digital. **Ciência da Informação**, Brasília, v. 28, n. 3, p. 257-268, set./dez. 1999.

FOX, Edward A. et al, Users, user interfaces, and objects: Envision, a digital library. **Journal of the American Society for Information Science**, v. 44, n. 8, p. 480-491, Sep. 1993.

GOMES, Sônia de Conti. **Bibliotecas e sociedade na primeira república**. São Paulo: Pioneira, 1983.

LESK, Michael. **Practical digital libraries: books, bytes, and bucks**. San Francisco: Morgan Kaufmann, 1997.



LÉVY, Pierre. **As tecnologias da inteligência: o futuro do pensamento na era da informática.** Tradução Carlos Irineu da Costa. Rio de Janeiro: Editora 34, 1993. (Coleção TRANS).

MARCHIONINI, Gary Evaluating digital libraries: a longitudinal and multifaceted view. **Library Trends**, v. 49, n. 2, p. 304-333, Fall, 2000. Disponível em: < <http://ils.unc.edu/~march/perseus/lib-trends-final.pdf>>. Acesso em: 20 jun. 2011.

PARENTE, André. **O virtual e o hipertextual.** Rio de Janeiro: Pajulin, 1999.

PEREIRA, Edmeire Cristina; RUTINA, Raquel. O Século XXI e o sonho da biblioteca universal: quase seis mil anos de evolução na produção, registro e socialização do conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 4, n. 1, p. 5 - 19, jan./jun.1999

PEREIRA, Maria de Nazaré Freitas. Bibliotecas virtuais: realidade, possibilidade ou alvo de sonho. **Ciência da Informação**, Brasília, v. 24, n. 1, 1995. Disponível em: <<http://www.ibict.br/cionline/viewarticle.php?id=583&layout=abstract>>. Acesso em: 20 maio 2010.

SARACEVIC, Tefko. Evaluation of digital libraries: An overview. In: DELOS Workshop on the Evaluation of Digital Libraries, 2004, Padova: University of Padua, 2004. Disponível em: <http://www.scils.rutgers.edu/~tefko/DL_evaluation_Delos.pdf>. Acesso em: 20 jun. 2009.

SARACEVIC, Tefko. How were digital libraries evaluated? In: CONFERENCE LIBRARIES IN THE DIGITAL AGE (LIDA). Dubrovnik, Croatia, 2005. Disponível em: <<http://comminfo.rutgers.edu/~tefko/articles.htm>> Acesso em: 10 out. 2010.

TAMMARO, Anna Maria; SALARELLI, Alberto. Medição e avaliação da biblioteca digital. In: _____. **A biblioteca digital.** Brasília, DF: Briquet Lemos, 2008. Cap. 13, p. 309-339.

VALENTIM, Marta Lúcia Pomim. Análise de conteúdo. In: VALENTIM, Marta Lúcia Pomim (Org.) **Métodos qualitativos de pesquisa em Ciência da Informação.** São Paulo: Polis, 2005. Cap. 6, p. 119-134.

ZHANG, Ying. **Developing a holistic model for digital library evaluation.** 2007, 248f. Dissertation (Doctor of Philosophy) - Graduate Program in Communication, Information, and Library Studies, The State University of New Jersey, New Brunswick, New Jersey, 2007.



DADOS ABERTOS GOVERNAMENTAIS (OPEN GOVERNMENT DATA): INSTRUMENTO PARA EXERCÍCIO DE CIDADANIA PELA SOCIEDADE

Claudio Jose Silva Ribeiro, Reinaldo Figueiredo de Almeida

1. INTRODUÇÃO

Atualmente, quando falamos de acesso à democracia e a políticas públicas, em geral, estamos nos referindo à transparência na gestão governamental e ao acesso às informações produzidas pelo governo, nos seus diversos níveis de estruturação (MEIJER, 2003). Reconhecidamente, o uso da internet para acesso a estas informações tem sido considerado como o meio mais efetivo de todos e a ele se convencionou chamar de governo eletrônico (ZHIYUAN, 2001).

Porém, a sociedade é cada vez mais exigente quanto à transparência governamental, por isso, o simples acesso às informações já não é o bastante, pois a capacidade de compartilhar, tratar e reutilizar estas informações é fator determinante para acompanhar as ações de governo.

Assim, alguns fatos vêm marcando de modo significativo as ações relacionadas a governo eletrônico. Dois deles, no entanto, causam reflexos significativos no desenvolvimento do tema em nosso país.

O primeiro fato é a publicação do decreto nº 6.932, de 11 de agosto de 2009, assinado pelo presidente da República, o qual “dispõe sobre a simplificação do atendimento público ao cidadão, ratifica a dispensa do reconhecimento de firma em documentos produzidos no Brasil, institui a ‘Carta de Serviços ao Cidadão’ e dá outras providências”. Para tanto, este decreto traz no seu art. 1º, incisos II e VI, que o Poder Executivo nas relações com cidadão deverá compartilhar informações e aplicar soluções tecnológicas para simplificar processos e procedimentos, propiciando melhores condições para tornar disponível a informação.

Além disto, o decreto prevê no seu art. 4º, que as bases de dados oficiais geridas pelo Poder Executivo, sejam colocadas à disposição de órgãos e entidades públicas interessadas no acesso às informações constantes nas mesmas.

O segundo fato determinante foi a publicação pelo presidente dos EUA, Barack Obama, do documento Memorandum on Transparency and Open Government (OBAMA, 2009). Este documento define as diretrizes do governo dos EUA para a sua gestão, no que tange ao acesso dos dados do governo americano, indo desde a questão da gestão interna da informação até a publicação dos dados do governo para o público em geral. Após a publicação deste documento, deu-se início a um conjunto de iniciativas nos EUA em torno do chamado Open Government Data, fortalecendo as ações



relacionadas ao tema e implicando na disponibilização de bases de dados oficiais para a comunidade em geral. Este fato também fomentou outras discussões relacionadas, principalmente sobre padrões tecnológicos referentes ao tema, tais como: RDF (Resource Description Framework), OWL (Web Ontology Language) e SKOS (Simple Knowledge Organization System); além de definições sobre o padrão de licença a ser atribuída aos dados publicados.

Assim, impulsionados por estes dois fatos, os executivos da área do governo federal brasileiro passaram a trabalhar com diretrizes para tornar as informações disponíveis preferencialmente em meio digital, de modo compartilhado e usando a internet para isto.

A idéia fundamental é tornar disponível com padrão aberto o maior número possível de bases de dados oficiais, cobrindo todos os aspectos da atividade governamental. Este esforço permitirá que além do acesso aos dados, o público em geral possa compartilhá-los e utilizá-los de modo automático, sem intervenção humana. Com isto, será possível gerar novos conhecimentos, novos serviços e produtos, viabilizando, inclusive, o aumento do papel fiscalizador da sociedade, por meio de validação dos dados oficiais fornecidos.

Para implementar esta proposta, uma das alternativas tecnológicas está baseada no uso da Web Semântica, com todas as tecnologias associadas à mesma, incluindo àquelas sugeridas como padrões a serem adotados para Dados Abertos pelo *Memorandum on Transparency and Open Government*. A Web Semântica tem como base o uso de URIs (*Universal-Uniform Resource Identifier*) para identificação e endereçamento de unidades de informação, além de RDF para representá-las. Assim, a estruturação do espaço Web não se fará apenas por meio de links e conexões entre documentos, mas através de uma nova abordagem, onde documentos e dados passam a ser ligados, fazendo uso de um conjunto de melhores práticas para publicar e tornar disponível o acesso aos mesmos. Isto introduz um conceito formulado por Tim Berners-Lee, o criador da Web, o de Dados Ligados ou *Linked Data* ou *Linked Open Data* (LOD), segundo o qual, deve ser uma prática expor e compartilhar, tanto os elementos de ligação de dados quanto às informações e aos conhecimentos.

Em síntese, Bizer, Heath e Berners-Lee (2011) observam que a noção de Dados Ligados (*Linked Data*) consiste em fazer uso da Web para criar ligações entre dados de diferentes fontes. Estas fontes podem ser tanto bancos de dados mantidos por outras organizações, quanto um conjunto de sistemas heterogêneos residentes em diferentes ambientes de computação. Tecnicamente, dados ligados se referem a dados publicados na Web que são estruturados de forma que sejam lidos por máquinas e com o seu significado definido explicitamente.

Bizer, Heath e Berners-Lee continuam e esclarecem que:

O projeto Linked Open Data (LOD) é a aplicação mais visível dos princípios de Linked Data, pois tem o objetivo de promover a noção de Web of Data por meio da identificação do conjunto de dados abertos que estão disponíveis sob uma versão livre de licenciamento, representados em formato RDF e publicados na Web (BIZER, HEATH, BERNERS-LEE, 2011, p.5)¹.

¹ Tradução dos autores para “The original and ongoing aim of the project is to bootstrap the Web of Data by identifying existing



A aplicação de *Linked Data* nos leva a um esforço para construção de um esquema de nomenclatura global, a fim de permitir a troca de dados. Em outras palavras, isto quer dizer que é necessário a elaboração de dicionários, com vocabulários que possam ser controlados e publicados, mas que possam ser expressos em modelos de alto valor, fazendo uso de ontologias e permitindo o incremento da semântica para estes dados ligados.

No Brasil, o desenvolvimento de projetos para suportar as ações de governo eletrônico e instrumentalizar a adoção da “Carta de Serviços ao Cidadão” têm exemplos em diferentes segmentos onde dados abertos estão agregando valor aos serviços. Ações ligadas ao Portal da Transparência (www.portaltransparencia.gov.br/) e ao sítio Comprasnet (www.comprasnet.gov.br/), são esforços brasileiros que fomentam o maior controle das ações do governo pela sociedade. Ademais, estes esforços podem ser estruturados em acordo com outras abordagens automatizadas, de forma a possibilitar uma maior participação popular propondo inovações em produtos e serviços, além de auxiliar na melhoria da eficiência dos serviços governamentais por meio da medição do impacto das políticas públicas implementadas (MANUAL, 2011).

Convalidando estes esforços, também é possível identificar em nível mundial, diversos projetos que materializam as ações e princípios citados anteriormente. Com o intuito de enriquecer este relato, seguem alguns exemplos sobre projetos para tornar visíveis gastos de governos: Tax Tree (www.mindtrek.org/2009/node/127), da Finlândia; e Where Does My Money Go (wheredoesmymoneygo.org – para onde vai o meu dinheiro?) da Grã-Bretanha, que mostram como os recursos dos impostos estão sendo gastos pelo governo. No Canadá, a abertura de dados permitiu que se economizassem 3.2 bilhões de dólares canadenses em fraudes fiscais de caridade. Outros sítios, como o Folketsting (folketsting.dk), da Dinamarca, acompanham as atividades parlamentares e o processo legislativo, permitindo o acompanhamento de projetos no âmbito do parlamento (MANUAL, 2011).

Também em nível mundial, no âmbito da melhoria de serviços e da participação popular, é possível citar projetos como o Find Toilet (www.findtoilet.dk) que mostra em um mapa todos os banheiros públicos da Dinamarca. Em uma primeira análise este projeto pode parecer de pouca utilidade, contudo, as informações disponibilizadas possibilitam que pessoas com diferentes problemas de saúde possam sair de casa com maior frequência. Na Holanda, o sítio Vervuilings Alarm (www.vervuilingsalarm.nl) avisa com uma mensagem se a qualidade do ar da vizinhança está perto de atingir um nível prejudicial. Há projetos como o Mapumental (www.mapumental.channel4.com), no Reino Unido, e o Mapnificent (www.mapnificent.net), na Alemanha, que permitem encontrar locais para morar de acordo com características indicadas, como duração do trajeto até o local de trabalho, além dos preços das casas e outros atrativos das regiões. Todos são todos exemplos de serviços que utilizam dados abertos. (MANUAL, 2011)

data sets that are available under open licenses, converting these to RDF according to the Linked Data principles, and publishing them on the Web.”



Portanto, diante deste novo contexto de disseminação de informações, este trabalho relata o projeto em desenvolvimento pela Dataprev (Empresa de Tecnologia e Informações da Previdência Social), que objetiva na sua fase inicial investigar e experimentar os princípios de LOD na disponibilização dos dados publicados pelo Anuário Estatístico da Previdência Social (AEPS²). Com isto, será possível não só o acesso aos dados contidos no Anuário, mas a extração dos mesmos de modo automático e para uso compartilhado, proporcionando à sociedade novas formas de interagir com estas informações providas pelo governo.

2. O ARCABOUÇO CONCEITUAL DO PROJETO

2.1 *Open Government Data* ou Dados Abertos Governamentais

Dados Abertos Governamentais são entendidos como o esforço para a publicação e disseminação das informações do setor público na Web, permitindo a reutilização e a integração destes dados. Além de propiciar a análise destes dados segundo vários pontos de vista, esta disseminação permite que a sociedade construa uma visão mais clara sobre o desempenho do governo frente às metas estabelecidas, bem como sobre o desenvolvimento de políticas públicas. Estas informações devem estar compartilhadas em formato bruto e aberto³, além de serem compreensíveis logicamente, pois estas características permitirão sua reutilização em aplicações informatizadas desenvolvidas pela sociedade (BERNES-LEE, 2009).

Segundo David Eaves, pesquisador canadense, a trilha para implementar *Open Government Data*, dentre outros fundamentos, deverá obedecer as três leis⁴ para promover o uso de Dados Abertos:

1. Se o dado não pode ser encontrado e indexado na Web, ele não existe;
2. Se o dado não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado;
3. Se algum dispositivo legal não permitir sua reaplicação, ele não é útil.

A estas leis, acrescentam-se os pressupostos sugeridos pela International Aid Transparency Initiative (IATI), contidos no Manual dos Dados Abertos⁵, e materializados em oito princípios⁶:

1. Dados precisam ser completos: todos os dados públicos estão disponíveis. Dado público é o dado que não está sujeito a limitações válidas de privacidade, segurança ou controle de acesso;

2 Esta publicação contém séries históricas importantes para a Previdência Social Brasileira e sociedade em geral, incluindo dados sobre renda e faixa etária, que podem, por exemplo, serem associados a informações de acidente de trabalho.

3 O formato bruto corresponde à forma primária em que o dado foi coletado pelo sistema informatizado disponível no governo, respeitando-se a segurança e privacidade de cada cidadão. O formato aberto diz respeito à independência de licenças comerciais e/ou softwares específicos para seu uso.

4 Traduzido de <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>

5 Este Manual foi lançado no 12º. FISL (Fórum Internacional de Software Livre). Disponível em <http://softwarelivre.org/fisl12> (MANUAL, 2011)

6 Traduzido de <http://www.opengovdata.org/home/8principles> e também disponível em (MANUAL, 2011)



2. Dados precisam ser primários: os dados são apresentados como os coletados na fonte, com o maior nível possível de granularidade e sem agregação ou modificação;
3. Dados precisam ser atuais: os dados são disponibilizados tão rapidamente quanto necessário à preservação do seu valor;
4. Dados precisam ser acessíveis: os dados são disponibilizados para o maior alcance possível de usuários e para o maior conjunto possível de finalidades;
5. Dados precisam estar em formato compreensível por máquinas: os dados são razoavelmente estruturados de modo a possibilitar processamento automatizado;
6. Dados precisam ser não discriminatórios: os dados são disponíveis para todos, sem exigência de requerimento ou cadastro;
7. Dados precisam estar disponíveis em formato não proprietário: os dados são disponíveis em formato sobre o qual nenhuma entidade detenha controle exclusivo;
8. Dados precisam estar livres de licenciamento: os dados não estão sujeitos a nenhuma restrição de direito autoral, patente, propriedade intelectual ou segredo industrial. Restrições sensatas relacionadas à privacidade, segurança e privilégios de acesso são permitidas.

Especificamente em relação a Dados Abertos Governamentais, junto aos oito princípios devem ser observados quatro aspectos fundamentais: disponibilização dos dados, publicação de informações sobre os dados disponíveis para facilitar o processo de intercâmbio, definição de plataformas tecnológicas capazes de dispor os dados de modo acessível (inclusive de modo automático) e licenças ou autorizações públicas a fim de permitir não só o acesso aos dados disponibilizados, mas o pleno uso dos seus conteúdos (AUER, 2007).

O primeiro aspecto a ser observado se refere ao processo de disponibilização das bases de dados oficiais por parte dos Governos, que além de observar os pressupostos listados acima, devem observar e garantir (AUER et. al., 2007; GRAY et. al., 2009):

- a) O sigilo e a inviolabilidade de dados pertencentes a pessoas físicas e jurídicas;
- b) A não rastreabilidade dos acessos praticados junto aos dados disponíveis;
- c) A promoção da previsibilidade legal e da segurança.

Um segundo aspecto fundamental é o que se refere à publicação de informações sobre os dados disponíveis para facilitar o processo de intercâmbio. Segundo Gray et. al. (2009), as bases de dados oficiais, uma vez tornadas acessíveis, deverão estar associadas a dicionários que detalhem no seu conteúdo informações referentes aos dados publicados, com semântica adequada ao público em geral.

Em “*An ontology for e-government public services*”, Vassilakis e Lepouras (2006), chamam a atenção para o terceiro aspecto fundamental que deve ser observado: as plataformas tecnológicas a serem empregadas. Visando atender a proposta de Dados Abertos Governamentais, estas plataformas deverão permitir que sejam atendidos os princípios citados acima, seguindo preferencialmente



padrões propostos por organismos internacionais multilaterais como o W3C e o IATI, objetivando assim a garantia do acesso às bases de dados publicadas.

O último aspecto observado nas discussões conduzidas sobre Dados Abertos, especialmente quando se aborda o conjunto de dados oriundos do âmbito governamental, está relacionado às questões de licença e autorização para acesso aos dados e aos seus conteúdos (AUER et. al., 2007; GRAY et. al., 2009). Neste aspecto crítico observa-se, inclusive, a diferença legal existente entre os conteúdos individuais de uma base de dados (privacidade no conteúdo de cada um dos campos de dados) e a coleção de conteúdos da base (privacidade global da base de dados) (DAVIES, 2010).

Como consequência deste último aspecto, é importante que os entes governamentais adotem políticas que eliminem direitos de uso e de autoria para as bases constantes dos programas de Dados Abertos Governamentais, adotando e institucionalizando o pelo uso de protocolos como o *Protocol for Implementing Open Access Data*. Este protocolo já vem sendo adotado pelos EUA e pela União Européia, institucionalizando o assim chamado *Public Domain* ou Política do Direito Público (GRAY et. al., 2009).

Portanto, o esforço de uso de Dados Abertos Governamentais deve ser acompanhado de uma política capaz de dispor dados oficiais de fácil acesso e plenamente compreensíveis a pessoas não técnicas, em formato legível para homens e máquinas, com previsibilidade legal, segurança e por fim devidamente atualizados (DAVIES, 2010).

2.2 Dados Abertos Governamentais no contexto brasileiro

Como observado na introdução deste relato, muitas ações estão sendo desenvolvidas na direção de trabalhar com dados abertos, tanto em nível mundial quanto no Brasil. Um dos exemplos mais contundentes sobre esta percepção foi a realização de evento recente em nosso país para tratar deste tema. O IV CONSEGI (Congresso Internacional – Software Livre e Governo Eletrônico⁷) teve como temática central o uso de Dados Abertos para a Democracia na Era Digital e trouxe diferentes pesquisadores para debater ações de implementação de dados abertos.

Além disto, o conjunto de experiências de outros países tem sido utilizado como referência para os esforços de implementação que estão sendo desenvolvidos pelo governo brasileiro. Nesta direção, os sítios do governo americano (www.data.gov), do Reino Unido (data.gov.uk), da Alemanha (www.offenedatten.de) e do Canadá (www.datadotgc.ca), apresentam diferentes visões em relação à estruturação e organização para o conjunto de dados que será alvo de publicação (DADOS, 2011).

Por outro lado, existem no âmbito do próprio governo brasileiro ações para melhorar a qualidade das informações e criar instrumentos para facilitar o intercâmbio de dados. As iniciativas em curso podem ser constatadas no relato de Franzosi et. al., que pode ser considerado como a gênese da problemática para o tratamento da informação nas bases de dados do Governo:

7 Disponível em www.consegi.gov.br



A imensa quantidade de código dos Sistemas Estruturantes Corporativos de Governo, somado à complexidade do negócio que apóiam e as inúmeras diferentes plataformas computacionais que os suportam remetem a diversas tentativas de evolução e integração desses sistemas. Esta dificuldade natural em implementar melhorias apontadas em ambientes complexos incentivou, ao longo de anos, a criação de um Cinturão de Dados, CD para apoio ao processo decisório de governo. [...] Os dados dos SGAs [Sistemas de Gestão Administrativa], muitas vezes extraídos e consolidados em planilhas, são acrescidos com outras informações declarativas, outros documentos, notícias, etc, constituindo uma base importante para tomada de decisão no nível estratégico do Governo. (DADOS, 2011, p. 25)

É dentro deste contexto que também está sendo desenvolvido o projeto DadosGov⁸. Este projeto objetiva criar um Catálogo de Informações Aberto para aprimorar a gestão pública e facilitar o acompanhamento pela sociedade. Desenvolvido pelo Comitê de Organização de Informações da Presidência da República (COI-PR), este catálogo foi construído a partir de alguns pressupostos para disponibilização de informações (DADOS, 2011).

- Apresentação das informações com organização em árvores temáticas e séries históricas;
- Uso de grupos de informação para facilitar a aquisição e atribuição da responsabilidade pela qualidade da estrutura de dados e seu conteúdo;
- Especificar e estruturar os requisitos do nível geral para os níveis mais específicos;
- Organizar o armazenamento de dados em formato padrão de catalogação.

Os trabalhos de modelagem deste catálogo também foram desenvolvidos com o apoio do Modelo de Dados Global, onde, por meio do mapeamento das Bases de Dados Eletrônicas dos sistemas do Governo Federal, buscou-se integrar a estrutura, a semântica e os processos envolvidos na atualização dos respectivos repositórios de dados.

Some-se a isto o conjunto de esforços para organização deste tema, que estão sendo desenvolvidos pela Secretaria de Logística em Tecnologia de Informação (SLTI) do Ministério do Planejamento. Em recente debate também realizado no IV CONSEGUI, a Sra. Miriam Chaves – da Secretaria de Logística e Tecnologia da Informação do Ministério do Planejamento – apresentou em linhas gerais a proposta para integração e uso da infraestrutura de LOD no governo (Figura 1).

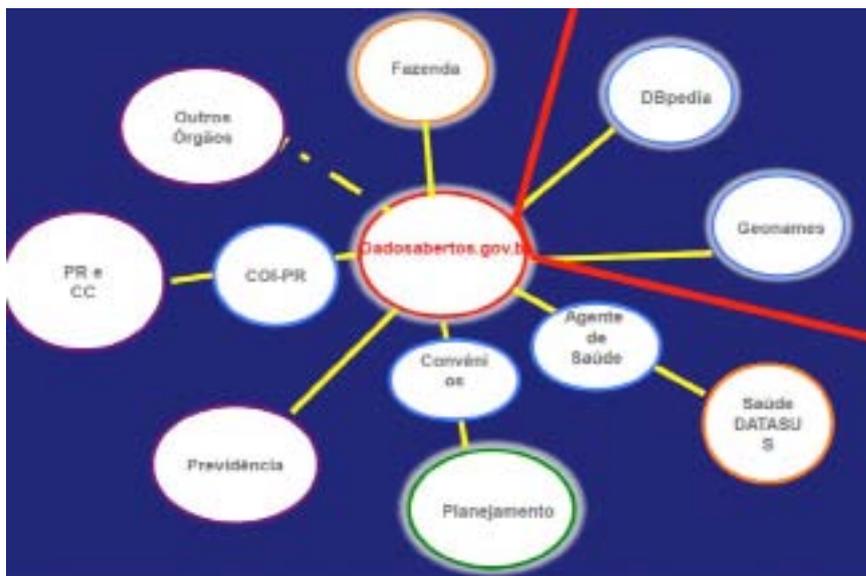


Figura 1: Estrutura LOD e Dados do governo (CHAVES, 2011)

Os esforços da SLTI culminam com a parceria entre o governo e a sociedade, representada por organismos de padronização, universidades e Organizações não governamentais, no sentido de desenvolver a INDA⁹ (Infraestrutura Nacional de Dados Abertos). Esta infraestrutura segue o modelo que foi praticado para geração da INDE¹⁰ (Infraestrutura Nacional de Dados Espaciais) e é composta por um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto no projeto e-Ping¹¹ (Projeto de Interoperabilidade no Governo).

É dentro deste cenário, repleto de alternativas e exemplos de estruturação segundo pressupostos para dados abertos, que se insere o projeto em desenvolvimento pela Dataprev.

3. VENCENDO O DESAFIO: ESTRUTURANDO E IMPLEMENTANDO O PROJETO DE DADOS ABERTOS NA PREVIDÊNCIA SOCIAL

Em sintonia com a evolução dos ambientes tecnológicos e para atender a estruturação proposta pela SLTI/Min. Planejamento, a Dataprev investiu na organização de um projeto para inserir a Previdência no contexto dos Dados Abertos.

Nesta direção, destaca-se, inicialmente, o papel da Dataprev, que administra grandes cadastros sociais do governo federal brasileiro. As bases de dados do sistema de Previdência Social pública, o Cadastro Nacional de Informações Sociais (que contém os registros referentes a todos os vínculos de contribuições previdenciárias e trabalhistas da população economicamente ativa brasileira) e o

9 Disponível em <http://wiki.gtinda.ibge.gov.br/MainPage.ashx>

10 Disponível em <http://www.inde.gov.br/>

11 Disponível em <http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padres-de-interoperabilidade>

Cadastro Nacional de Óbitos (que concentra o registro de todos os óbitos registrados em cartórios do Brasil), são exemplos do conjunto de dados que estão sob a guarda da Dataprev.

Portanto, para dar conta deste desafio, foi necessário compreender o domínio que seria alvo de publicação, compondo a fase de planejamento do trabalho (RIBEIRO, 2008). Esta fase foi encaminhada por meio de reuniões para estruturação do escopo, que foram conduzidas pelo Departamento de Gestão de Informações em parceria com a Presidência da Dataprev. Nestas reuniões foram encaminhados debates internos sobre o conjunto de necessidades de informação para a sociedade, principalmente as demandas provocadas pela comunidade de pesquisadores que já fazem solicitações de conjunto de dados para a Dataprev. Estes debates foram conduzidos no sentido de avaliação dos técnicos envolvidos sobre a relação das demandas, bem como a relevância destas para a comunidade.

O produto final destas reuniões foi o consenso sobre a publicação dos dados de acidente de trabalho presentes no AEPS. Apesar destes dados já serem publicados por meio de ferramenta que permite a extração de consultas pela internet¹², esta publicação não segue os princípios para uso de dados abertos que foram apresentados na seção anterior, dificultando o reuso destas informações pela sociedade em geral.

O AEPS juntamente com o Boletim Estatístico da Previdência Social (BEPS), que possui periodicidade mensal, contém dados referentes ao sistema previdenciário, os quais foram tratados, organizados e disponibilizados em séries agregadas antes de compor os dois instrumentos. Ambos, conforme citado anteriormente, estão disponíveis pela internet através do sítio da Previdência Social (<http://www.previdencia.gov.br/>), além dos exemplares publicados em outras mídias, como por exemplo, o uso de cadernos impressos.

Como forma de representar os grandes grupos de dados analisados, durante as reuniões de planejamento, foi apresentada o seguinte modelo de classes que elucida o conjunto de dados a serem publicados (Figura 2):

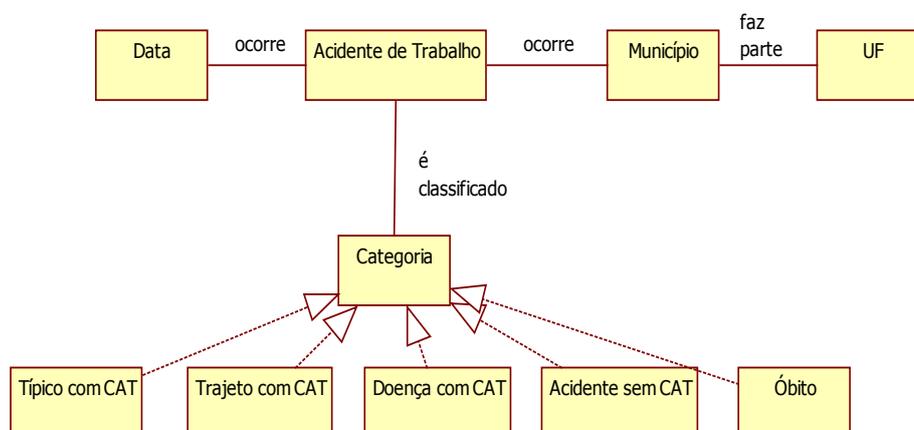


Figura 2: Modelo de Classes¹³ sobre evento de acidente de trabalho

¹² Disponível em <http://www3.dataprev.gov.br/infologo/>

¹³ Para facilitar a compreensão do modelo, a leitura pode ser traduzida por sentenças: Acidente de Trabalho ocorre em data; Acidente

Com o escopo para publicação definido, o debate sobre o desenvolvimento do trabalho foi encaminhado na direção de examinar as leis e princípios para a publicação de dados abertos. Neste sentido, foi necessário identificar um recorte de dados mais apropriado ao processo de publicação, pois o esforço para estruturação dos dados e de sua semântica precisava ser adequado aos prazos estabelecidos para a conclusão do projeto. Assim, foi possível chegar à representação a seguir como o modelo final que foi alvo de representação para publicação (Figura 3):

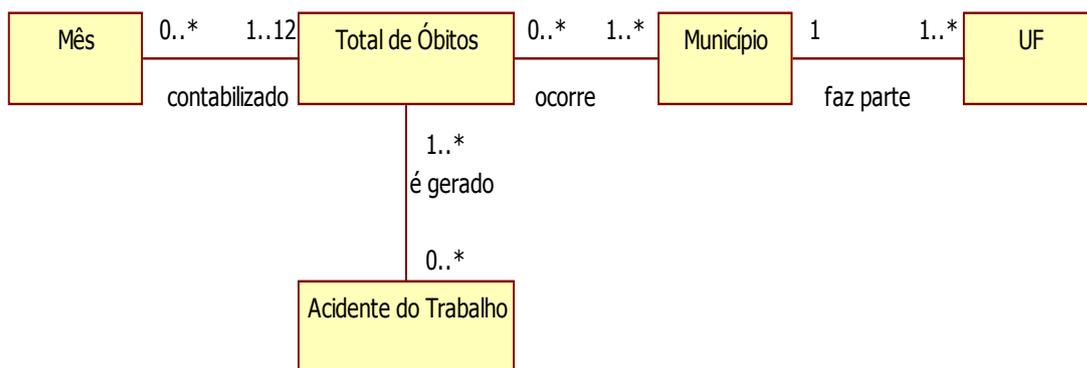


Figura 3: Modelo de Classes¹⁴ adotado com respectivas cardinalidades (ocorrências).

Com o entendimento do conjunto de dados, partiu-se para o processo de extração¹⁵ destes dados das bases operacionais do Anuário Estatístico, para em seguida dar início ao processo de transformação e carga para publicação.

Apesar da possibilidade de uso de ferramentas automatizadas para apoiar estas ações (tais como D2R Server, Triplify e Virtuoso RDF), estas etapas estão em curso e demandam um grande esforço, uma vez que a partir deste ponto serão geradas as representações em RDF, formato definido para ser usado pelos dados a serem disponibilizados, incluindo a representação semântica destes dados. Estas atividades são conhecidas como geração de triplas¹⁶ RDF ou ainda “triplificação”¹⁷ dos

de Trabalho ocorre em Município; Acidente de Trabalho é classificado em Categoria; São tipos de Categoria: Acidente Típico com CAT (Comunicação de Acidente de Trabalho), Acidente de Trajeto com CAT, Acidente que causa Doença com CAT, Acidente que causa Doença sem CAT e Acidente que causa Óbito. Neste modelo preliminar as cardinalidades (ocorrências) não foram identificadas.

14 De forma análoga à Figura 1 a leitura deste modelo também pode ser traduzida em sentenças: Total de Óbitos é contabilizado de um até 12 Meses e um Mês pode ter zero ou vários Óbitos; Total de Óbitos ocorre em nenhum Município ou em vários Municípios e um Município pode ter um ou vários Totais de Óbitos; Um Município faz parte de uma UF e uma UF pode ter um ou vários Municípios; Total de Óbitos é gerado por um Acidente de Trabalho ou por vários Acidentes e um Acidente pode ter nenhum Óbito ou então vários Óbitos totalizados.

15 Este processo faz uso de dinâmica para seleção de registros para publicação segundo uma janela temporal pré-determinada, com o uso de linguagem SQL (Linguagem de consultas para Bancos de Dados Relacionais - Structured Query Language).

16 Uma tripla RDF é composta por: [Sujeito ou Recurso, Predicados ou Propriedades, Valores (o conteúdo do objeto)] e para geração da parte semântica destas triplas é essencial o uso de ontologias, incluindo ontologias de fundamentação de forma a facilitar o intercâmbio de dados (GUIZZARD, 2005).

17 O fluxo para triplificação consiste em etapas para extração, alinhamento, diferenciação e complementação para carga (BIZER et. al., 2009). Outra referência para o assunto é o tutorial para triplificação, disponível em: <http://www.inf.puc-rio.br/~psalas/tutorial.html>



dados. Estas atividades estão previstas para serem concretizadas até o final de 2011.

Outra atividade de fundamental importância e que complementa este projeto, é a organização, revisão e disponibilização de todo o vocabulário relacionado ao conteúdo a ser publicado, sendo o mesmo associado a uma estrutura de metadados, a qual terá como componentes, além das definições, as informações sobre a composição dos dados (em caso de agregações) e suas respectivas origens, por exemplo.

4 CONCLUSÕES

A Web alterou radicalmente a forma de compartilhar conhecimentos, reduzindo as barreiras para a publicação e para o acesso aos documentos. Um dos maiores fatores de sucesso do ambiente Web é a liberdade que seus usuários têm em criar e disseminar informação, logo, o grande desafio é como trabalhar com este ambiente descentralizado e quase que anárquico, buscando um modelo de organização de informações.

Ligações entre documentos através de estruturas de hipertexto permitem aos usuários percorrer o espaço informacional onde os documentos estão publicados, por meio dos assim chamados navegadores Web, enquanto motores de busca associados com analisadores de ligações de documentos fazem com que os usuários localizem os documentos requeridos (JACOBS & WALSH, 2004 *apud* BERNES-LEE, 2009). Esta combinação foi a principal responsável pelo sucesso da Internet e pelo seu constante incremento (BERNES-LEE, 2009).

Porém, com a adoção de novas aplicações, tais como redes sociais e uso de estruturas semânticas para intercâmbio de dados entre ambientes automatizados, novos hábitos são estabelecidos na utilização da rede mundial de computadores. O compartilhamento de informações e dados, além da construção de ações colaborativas, são exemplos destas novas formas de interação que estão em prática na atualidade (AUER et. al., 2007).

Obama (2009) redefiniu o conceito de Governo Eletrônico ao definir que o mesmo, assim como toda a estrutura do novo governo norte-americano, deveria atender fundamentalmente a três princípios: transparência, participação e colaboração. Deste modo, o Governo Eletrônico norte-americano foi obrigado a se reelaborar, incorporando em definitivo o conceito de Dados Abertos, e mais além, o de Dados Abertos Ligados.

Alinhado a este momento é que a Dataprev, por meio da iniciativa para implementar um projeto baseado no conceito de Dados Abertos, insere o Governo Brasileiro no contexto mundial de iniciativas de igual natureza. Ademais, a empresa inicia um processo arrojado de desenvolvimento de competências tecnológicas estratégicas, além de prestar um enorme serviço ao cidadão comum, ao disponibilizar dados previdenciários num formato acessível e intercambiável.

Os próximos passos a serem trilhados no projeto aqui apresentado, serão a publicação em ambiente Web e a divulgação para a comunidade e sociedade em geral do AEPS em Dados Abertos, juntamente com todo o vocabulário associado ao mesmo. A publicação do mesmo terá sua divulgação,



além das estratégias de veiculação tradicionais, associada ao registro em mecanismos de buscas tradicionais da Web e estendida a outros mecanismos, especialmente orientados ao ambiente de dados abertos, tal como, o *Comprehensive Knowledge Archive Network* (CKAN).

O processo de aferição dos resultados alcançados passará pelo uso de redes sociais¹⁸, blogs¹⁹ e outros canais de comunicação, além de avaliações com membros da sociedade civil organizada (universidades, comunidades de software e representantes de organizações internacionais, tal como o escritório Brasil do consórcio W3C, além de Organizações não Governamentais – ONG's). Este é um dos primeiros esforços conduzidos pela Dataprev para incremento da participação da sociedade no desenvolvimento do nosso país. Outras ações precisarão ser desenvolvidas para tornar mais eficaz à publicação de dados abertos de governo.

A adoção de políticas que incrementem o acesso à informação é ponto primordial na construção de uma sociedade mais justa. Adicionado a isto, é possível afirmar que tanto o uso de novos métodos de trabalho quanto o apoio de instrumentos tecnológicos, são elementos essenciais para melhorar a disseminação e uso de sistemas de recuperação de informação. Assim, dotar a Web de facilidades para permitir o processamento automatizado das informações e conteúdo é um dos caminhos naturais para trabalharmos com a miríade de dados que estão presentes atualmente na Internet.

Por fim, cabe ressaltar que tanto o uso de ontologias de fundamentação (GUIZZARD, 2005) quanto o incremento da adoção de padrões de interoperabilidade no processo de desenvolvimento de novos sistemas, transformarão os sistemas computadorizados em verdadeiras fontes de informação para a sociedade, fazendo com que surja aquele que o pesquisador norte-americano Kingsley Idehen chamou de *citizen analyst* ou analista cidadão.

REFERÊNCIAS

1. AUER, S. et. al. DBpedia: A Nucleus for a Web of Open Data. 2007. **The Semantic Web, Lecture Notes in Computer Science**, 2007, Volume 4825/2007, 722-735. SpringerLink.com.
2. BERNES-LEE, T. **Putting Government Data online**. 2009. Disponível em <http://www.w3.org/DesignIssues/GovData.html> com download em 01 de julho de 2011.
3. BIZER, C, HEATH, T., BERNES-LEE, T. **Linked Data. The story so far**. Disponível em <http://eprints.ecs.soton.ac.uk/21285/1/bizer-heath-berners-lee-ijswis-linked-data.pdf>. com acesso em junho de 2011.
4. BIZER, C. et. al. DBpedia - A crystallization point for the Web of Data. **Web Semantics: Science, Services and Agents on the WorldWideWeb**, Vol 7, 2009, pp. 154–165
5. CHAVES, M. B. F. **Apresentação sobre ações de Governo para Dados Abertos**. Disponível em http://www.w3c.br/conferenciaegov/06_COI-PR_Mirian.pdf com acesso em julho de 2011.

18 Disponível em <http://twitter.com/#!/dataprev>

19 Disponível em <http://portal.dataprev.gov.br/category/blog/>



6. DADOS Abertos para a Democracia na Era Digital. Brasília: Fundação Alexandre Gusmão, 2011. 84p. Disponível em http://www.consegi.gov.br/sobre_consegi/livro-consegi-dados-abertos-para-a-democracia-na-era-digital?set_language=pt-br&cl=pt-br
7. DAVIES, T. Open data, democracy and public sector reform. **A look at open government data use from data.gov.uk**. 2010. Disponível em <http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf> com download em 20 de maio de 2011.
8. GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. Enschede: Netherlands. Telematica Institut Fundamental Research Series, No. 015 (TI/FRS/015). 2005. Disponível em <http://www.inf.ufes.br/~gguizzardi/OFSCM.pdf> com download em 13 de junho de 2011.
9. GRAY, J. et. al. **Unlocking the Potential of Aid Information**. 2009. Disponível em <http://www.unlockingaid.info/wp-content/uploads/2010/02/UnlockingAidInformation.pdf> com download em 13 de junho de 2011.
10. MANUAL dos Dados Abertos: Governo. 2011. Disponível em http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf com download em 01 de julho de 2011.
11. MEIJER, A. J. Transparent Government: Parliamentary and legal accountability in an information age, **Information Polity**, Vol. 8, Nrs. 1 & 2, 2003, pp. 67 – 78. Disponível em http://scholar.google.com/scholar?q=transparency+government&hl=pt-BR&as_sdt=0&as_vis=1&oi=scholar com download em 17 de setembro de 2011.
12. OBAMA, B. **Memorandum on Transparency and Open Government**. 2009. Disponível em (www.whitehouse.gov/the_press_office/Transparency_and_Open_Government/) com download em 23 de março de 2011.
13. RIBEIRO, C. J. S. **Diretrizes para o projeto de portais de informação: uma proposta interdisciplinar baseada na Análise de Domínio e Arquitetura da Informação**. 2008. 298. (Doutorado). PPGCI/UFF-IBICT. Rio de Janeiro
14. VASSILAKIS, C., LEPOURAS, G. **An ontology for e-government public services**. 2006. Disponível em <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.2496&rep=rep1&type=pdf> com download em 13 de junho de 2011.
15. ZHIYUAN, F. **E-Government in Digital Era: Conceptal, Practice and Development**. 2001. Disponível em http://scholar.google.com/scholar?hl=pt-BR&as_sdt=0,5&q=define+electronic+government com download em 17 de setembro de 2011.



CULTURA DA WIKIPÉDIA COMO POLÍTICA DE CAPACITAÇÃO COMPARTILHADA EM MEIO DIGITAL DE SOCIALIZAÇÃO DO CONHECIMENTO

Maria José Vicentini Jorente

Resumo: Apresenta-se laboratório em desenvolvimento no âmbito da disciplina História da Cultura, ministrada nos cursos de Arquivologia e Biblioteconomia da Unesp (Marília), que se vincula à linha de pesquisa Informação e Tecnologia: alunos do segundo ano da graduação participam na melhoria de 27 verbetes em língua portuguesa na Wikipédia. Busca-se capacitar para a leitura e a escrita científica no ambiente digital, habilitar para a identificação e recuperação de informações e para interpretação e compreensão de aspectos formais e conteudísticos da informação e sua reorganização. Compreende assim atividades de busca, seleção, remix e republicação de textos, imagens, áudio e vídeos na convergência das diversas fontes de informação hipertextuais, apoiadas por tutores com domínio estratégico dos ambientes digitais. Aderiu-se para isto ao projeto internacional da Fundação Wikimedia de Embaixadores de Campi universitários. Visa também induzir comportamentos de compartilhamento e colaboração em rede com finalidade de criar hábitos necessários para empoderamento informacional no Brasil. Como metodologia, se otimiza o trabalho de indivíduos já formados na cultura wiki, e cria-se na disciplina programas de compartilhamento de informações com viés mais especializado, dando ao ambiente digital maior credibilidade. O ambiente tem sua própria sintaxe que ajuda no aprendizado das habilidades complementares à leitura e à escrita e se oferece como repositório aberto em que a informação poderá ser reutilizada. Trata-se, assim, de estratégia de autonomia e suficiência que considera conhecimentos intersemióticos na edição, visualização e compreensão de informação na web social. Propõe-se segunda etapa de verificação da confiabilidade do ambiente após o trabalho de consolidação e divulgação da melhoria dos verbetes.

Palavras-chave: Informação e Tecnologia. Cultura Digital e Compartilhamento. Wikipedia.

1 INTRODUÇÃO

A cultura contemporânea se define pela prevalência de novas poéticas mediadas por interfaces tecnológicas que provocam mudanças radicais na natureza das mensagens convergidas nestas poéticas.

As possibilidades de amplo acesso ao conhecimento - definido como novas formatações e/ou como processos de reconhecimento e recriação internos aos indivíduos - devem, como consequência, propiciar novas formas de inteligência e de consciência. Novos modelos mentais de formatação e de construção pensamental edificadas por novas fusões ou integrações conceituais surgem auto-organizados a partir das hipertextualidades e intersemioticidades que se implicam nestes processos generacionais. Tais modelos mentais se sobrepõem em camadas que articulam enciclopedicamente as ofertas de informação apresentadas com aparências diversas nos contextos também distintos das vivências e das trocas em meio digital.



Ofertas são dispostas aos sujeitos que interagem com o conhecimento objetivado de maneira fractal ofertando, por esta característica, múltiplas formas de visualização e de *design* que as diferentes abordagens evocam como caminhos.

Esses novos relacionamentos entre os indivíduos e o conhecimento criam no sistema da cultura contemporânea rupturas que de forma ideal provocarão mudanças conceituais na sociedade. Criam também novas tradições e crenças plenamente realizadas por uma nova geração de indivíduos nativos digitais que com elas interagem por meio de novos métodos de aproximação, de aquisição, de organização, de produção de nova arquitetura e de novo *design* na utilização e na reutilização das informações.

As postulações antecedentes são de uma contemporaneidade que, no entanto, não é homogênea na realização das possibilidades acima descritas, pois a cultura tampouco se oferece como um tecido uniforme. O próprio termo cultura oferece complexidades que devem ser exploradas para sua contextualização.

Na década de 1970, momento referido como de passagem para a discutida pós-modernidade, Hanna Arendt, em **“A crise da cultura” (1974)**, ainda remete o conceito cultura à sua origem romana “colere” que significa cultivar, habitar, tomar conta, criar e preservar (ARENDR 1972, p.265). A autora aponta no texto que entre os gregos, por outro lado, prevalecia o conceito de cultura como fabricação. Fabricação envolve *techne*, artificios técnicos. A combinação das duas conceitualizações determinava, para Arendt, o conteúdo e o significado de cultura provindos do mundo clássico e que se manteve na modernidade (ARENDR, 1972, p. 266).

Abraham Moles em **“Sociodinâmica da Cultura” (1974)** reflete, por outro lado, que modernamente, em 1793, século XVIII, a palavra cultura já estava presente em dicionários alemães como um conceito que traduzia apenas conjuntos de vestígios artificiais deixados no meio ambiente pelo homem.

Estes vestígios, por via estatística, caoticamente estimulam o cérebro, embutidos em linguagens multidimensionais e impressão gradual de sentidos a conceitos inicialmente vazios (MOLES, 1974, p.13), mas preenchidos de significados no percurso de sua utilização. A variação de suportes e de linguagens, e os fluxos contínuos desta variação é que mudam, para Moles, a própria estrutura do pensamento:

Os elementos do mobiliário cerebral do homem da rua são, de preferência, os cartazes do metrô, o que ouviu no rádio ou na televisão na véspera, o último filme que assistiu, o jornal que lê ao encaminhar-se para o trabalho, as conversas dos colegas de escritório e os bate-papos; o que aprendeu na escola e um nevoeiro vago de noções passadas. Seus conceitos de encruzilhada, as idéias integradoras de sua percepção de fatos e de coisas, impõem-se a ele por uma via estatística muito diferente da via da educação racional, cartesiana, com elevado grau de coerência e em cujas virtudes ele continua a acreditar. (MOLES, 1974, p.13).

Via estatística é, segundo o autor, a somatória da informação circulante na vida cotidiana, nos dados recolhidos em função das necessidades imediatas. Somente após o recolhimento de certo volume de informações é que se obtêm determinadas estruturas pensamentais.



Bastante relevante em Moles é a sua redefinição do termo cultura como “cultura mosaico”, pois tal redefinição já aponta para as mudanças impressas ao conceito na pós-modernidade. O autor compara cultura a um “[...] tecido fibroso composto por fragmentos de conhecimentos desordenados ligados por relação de proximidade, de época de aquisição, de assonância, de associação de idéias.” (1974, p.19). Enfatiza que a cultura não é mais o fato de uma educação (ele acrescenta a universitária) racional, mas de fluxos de conhecimentos transmitidos por meios diversos.

Outro conceito importante em Moles é o da impregnação e imersão da e na esfera das mensagens, contrapostas a uma dissonância entre vida e educação. O autor fala em canais de cultura, referindo-se a sua sedimentação em determinados sistemas sociais, pois o compartilhamento dos bens culturais está naturalmente ligado aos meios físicos de transferência da informação.

Daí já se depreende que a cultura contemporânea, representada por novos ambientes transformados com bases na comunicação da informação, na aprendizagem e no conhecimento, apresenta possibilidades distintas de interação com o fluxo das informações circulantes que tem papel significativo na construção coletiva do conhecimento e das relações humanas envolvidas neste momento de transição histórica.

O conceito de informação é, em tal transição, relacionado com um paradigma social, porém, apoiado nas tecnologias correntes, que propõe como parâmetro essencial a consideração de contextos específicos: vivências que levam em conta valores e crenças dos seus ambientes de inserção.

Tais crenças filtram e dão forma à percepção, contextualizando a informação por meio daqueles princípios ou leis, padrões que regulamentam e estruturam pensamentos, linguagens e ações inseridos em um recorte de local e de tempo.

2 O CONTEXTO CULTURAL BRASILEIRO NO INÍCIO DO SÉCULO XXI

Levando em conta a definição de cultura mosaico como parâmetro que se firmou na pós-modernidade, a realidade da cultura contemporânea em países emergentes como o Brasil tem sido marcada por tentativas de melhorar a capacidade de ação e de interação dos sujeitos atuantes na cultura para alcançar objetivos e interesses que dão sentido à existência humana nas relações cotidianas contextualizadas, mas, especialmente, que também coloquem tais países em melhores condições de competitividade no panorama dos mercados globalizados. Que possibilitem, portanto, a um número crescente de pessoas acessar e interagir com os estoques de informação ou de objetivação do conhecimento.

Ao se pensar a Ciência da Informação como área do conhecimento que se consolida e se legitima no contexto da cultura contemporânea assim configurada no Brasil, entende-se também que determinadas práticas necessitam ser implantadas em todos os setores que podem contar com as Tecnologias de Informação e Comunicação (TIC) neste sentido de melhor fluidez de novos conteúdos e formatações. A geração e o uso da informação digital tornam-se dessa maneira, na contemporaneidade, preocupações inerentes ao campo da Ciência da Informação.



Paradoxalmente, no contexto brasileiro, repleto de ambivalências e com relevantes áreas de escassez de ambientes tradicionais ou digitais de disponibilização de informação, a Ciência da Informação também enfrenta, como todas as ciências na contemporaneidade, o problema da explosão da informação.

Nesse sentido, é pertinente que a preocupação do domínio se volte para um melhor relacionamento com os saberes da cultura contemporânea: por razões óbvias, a via digital, com o avanço e domínio das TIC é um destes necessários saberes e pode contribuir para uma melhoria de vida do cidadão brasileiro, em todas as atividades permeadas pela utilização da informação e do conhecimento.

Fatores técnicos e tecnológicos aliados às transformações estruturais no conhecimento proporcionado aos indivíduos e à sociedade afetam os processos de aquisição desses conhecimentos e reforçam as características da própria rede em que se produzem. Não há, entretanto, garantias de aquisição automática de conhecimento simplesmente por se estar inserido em uma rede de compartilhamento, mesmo que o conectar-se, possa indubitavelmente encurtar muitos caminhos e facilitar a concretização dos processos de transferência de informação.

Com esta consciência, o compartilhamento sistematizado dos conhecimentos por meio de intercâmbio social e cultural em processos de redes de socialização da cultura, resultantes da codificação dos conhecimentos individuais, dos grupos e das organizações, pode ser acelerado e melhorado.

A justificativa para o levantamento da hipótese de tais melhorias é que nesses processos grupais facilita-se a visualização das informações com níveis variáveis de complexidade, dependendo das características dos públicos a que se destinam.

Resultante da elencada problemática, a busca que se denota dos discursos no domínio de uma Ciência da Informação integrada aos novos parâmetros da contemporaneidade é o de propostas de práticas sociais transformadoras, conscientes e indutoras da necessária capacitação e de uma melhor compreensão dos desafios que se propõem contemporaneamente: com a ampliação das trocas informacionais na segunda metade do século XX e as novas condições que a Web Social patrocinou em inícios do século XXI os sujeitos que interagem com as TIC, de forma geral, e os especialistas em diversos campos do saber, em particular, vêem-se frente a possibilidades e desafios cotidianos propostos por recursos ainda não explorados.

No que diz respeito às TIC, formas induzidas de aprendizado direcionadas para acelerar os processos naturais de compreensão e conscientização no seu uso, com novas estratégias e metodologias, podem ser introduzidas e praticadas com vistas a reequilibrar, nos diversos setores da vida diária, as situações em que se verificam a incorporação acelerada de uma pleiade de transformações tecnológicas.

Nesse contexto, o computador é uma metamídia híbrida (ANDERS, 1997, p.5), possibilitadora da produção, do armazenamento e da distribuição da informação provinda de outras hipermídias



híbridas, por sua capacidade de traduzí-las, absorvê-las, descrevê-las e de ser meio para recuperação, construindo uma linguagem tecnológica universal.

Nos ambientes criados e mediados via computador, o hipertexto representa uma gramática que se constrói sob a lógica de nós estruturados em rede. Por esta lógica, coloca em ação interativa uma diversidade crescente de formas e de códigos de modelação da informação de maneira similar àquela utilizada pela estrutura cerebral humana que de forma natural e orgânica organiza o conhecimento no córtex. O aprendizado das convergências dessa lógica se realiza de maneira simulatória, aperceptiva, mas participativa. Convergências aqui se cristalizam em diferentes formas de construção e disseminam informação fundida nas diversas estruturas de codificações ou linguagens, transformando-as em nós da rede rizomática do conhecimento.

2.1 Avaliação internacional e posição do Brasil nos exames de letramento e de ciências

A categoria cognitiva dos que interagem com as TIC tem sido, neste sentido, alvo de pesquisas e avaliações relacionadas às capacidades de um aprendizado que visa à autonomia e independência no seu uso e quanto às suas complexidades, propiciando visão crítica dos meios e estímulo dos capitais cultural, social e intelectual do contemporâneo.

Quanto ao desempenho dos jovens em relação ao conhecimento considerado relevante para o desenvolvimento, os países membros da Organização para Cooperação e Desenvolvimento Econômico (OCDE) possuem um sistema de testes de rendimento escolar, aplicado no Brasil pelo Instituto de Estudos e Pesquisas Educacionais (Inep/MEC). O INEP administra todos os procedimentos da pesquisa do Programa Internacional de Avaliação de Alunos (PISA), no momento que jovens aleatoriamente escolhidos passam para o ensino médio.

O objetivo da avaliação e determinar até que ponto os jovens que se encontram nesse estágio final da chamada educação obrigatória adquiriram “[...] habilidades essenciais para a participação efetiva na sociedade.” (INEP, 2010), com a principal finalidade de produzir indicadores sobre a efetividade dos sistemas educacionais a cada três anos.

Em 2000, o foco dessas provas foi a Leitura e letramento; em 2003, Matemática e em 2006, Ciências. A página do INEP informa que:

Alguns elementos avaliados pelo PISA, como o domínio de conhecimentos científicos básicos, fazem parte do currículo das escolas, porém o PISA pretende ir além desse conhecimento escolar, examinando a capacidade dos alunos de analisar, raciocinar e refletir ativamente sobre seus conhecimentos e experiências, enfocando competências que serão relevantes para suas vidas futuras.

As afirmações acima são enfatizadas no mesmo texto como questões propostas: “Até que ponto os jovens adultos estão preparados para enfrentar os desafios do futuro? Eles são capazes de analisar, raciocinar e comunicar suas idéias efetivamente? Têm capacidade para continuar aprendendo pela vida toda?” (INEP, 2010).



O enfoque das avaliações é assim enfaticamente ligado pelo PISA às competências “necessárias à vida moderna” (INEP, 2010), pensado dentro das fronteiras do modelo de aprendizagem dinâmica, com vistas à adaptação em um “mundo em constante transformação” (INEP, 2010). O exame PISA de letramento em leitura pretende desta maneira, avaliar:

- a forma material da leitura pela utilização de textos que incluem passagens em prosa e outros tipos de documentos como listas, formulários, gráficos e diagramas, pois o modelo seguido pressupõe que os indivíduos encontrarão diferentes formas de escrita pela vida e não somente as encontradas na escola;

- a habilidade de identificação e de recuperação de informações, de desenvolver compreensão dos textos e de interpretá-los, refletindo sobre os seus conteúdo e forma, e de construir argumentações para, a partir da experiência de leitura, defender um ponto de vista;

- a habilidade de identificação do uso para o qual o texto foi construído: romances, cartas e/ou biografias são escritos para uso “pessoal”; enquanto documentos oficiais ou pronunciamentos são para uso “público” e um manual ou relatório, para uso “operacional”.

O Letramento em Ciências, por outro lado, avalia o conhecimento de conceitos científicos para a compreensão do mundo natural:

[...] a capacidade de reconhecer questões científicas, fazer uso de evidências, tirar conclusões com bases científicas e comunicar essas conclusões. São utilizados conceitos científicos que serão relevantes para serem usados pelos alunos tanto no presente quanto no futuro próximo” (INEP, 2010).

A avaliação acontece também em três dimensões:

- a dos conceitos científicos, necessários para compreender fenômenos do mundo natural e as mudanças decorrentes de atividades humanas;

- a dos processos científicos, centrados na capacidade de adquirir, interpretar e agir com base em evidências: reconhecimento de questões científicas, identificação de evidências, elaboração de conclusões, comunicação dessas conclusões, demonstração da compreensão de conceitos científicos.

- a das situações científicas, selecionadas principalmente da vida cotidiana das pessoas.

Em 2006, no relatório da avaliação sobre o *letramento em leitura* do PISA, o Brasil ficou em último lugar entre 32 países industrializados. Em Ciências (2006) o resultado do PISA não foi muito melhor do que o de letramento em leitura, de 2000; até porque é de se esperar que desempenhos deficitários de leitura também influam sobre a compreensão dos conteúdos dos textos científicos.

O fato é que em ambas as situações os resultados colocam o país em níveis muito baixos de proficiência: no Brasil apenas 0,5% dos estudantes alcançou o Nível 5 de proficiência e nenhum chegou ao Nível 6; é no Nível 2 que os estudantes começam a demonstrar as competências científicas que lhes permitirão participar ativamente em situações da vida que tenham relação com Ciência e Tecnologia.



Em 2009 os resultados da quarta participação do país no PISA foram melhores, amplamente divulgados pelos meios de comunicação, e o Brasil foi considerado um dos países cujo desempenho mais cresceu desde 2000 - embora ainda esteja entre as últimas posições no *ranking* internacional, uma colocação em 53º lugar entre os 65 países participantes do exame. À frente estão o Uruguai, o Chile e o México.

Desta informação pode-se destacar a significação de letramento em leitura, problema só aparentemente restrito às preocupações de educadores e da área da educação, antecedendo mesmo as questões do letramento científico, pois este último está objetivado principalmente em formato textual das informações.

Verifica-se a partir desses dados que a incapacidade de decifrar um texto escrito não se deve somente à pobreza, mas provavelmente a um erro sistêmico no ensino da leitura que leva a uma compreensão superficial dos textos e, conseqüentemente, a falhas no processamento cognitivo das informações ali constantes pela não assimilação dos seus códigos sintáticos e ligações semânticas.

Torna-se, portanto, premissa essencial para o Brasil o domínio pleno da linguagem escrita - que será ainda sobreposta pelas intersemiotidades entre as linguagens e codificações que abordam os sujeitos diariamente nos seus diversos campos de atuação.

Desnecessário seria dizer que, se prévio ao desenvolvimento e incorporação das TIC no âmbito da aquisição de conhecimento já eram observadas deficiências, conforme os relatórios acima citados, tais deficiências se potencializam ainda mais com essas TIC, que irão impor a necessidade do seu aprendizado de uma forma genérica, entre usuários e, principalmente, entre mediadores da informação.

2.2 Inclusão info-digital nas universidades brasileiras e a CI

Se determinadas competências são exigidas do cidadão na Sociedade da Informação, no Brasil as escolas desde o ensino fundamental e as Universidades, em particular, têm papel fundamental no encaminhamento das políticas de capacitação e de formação, desenvolvendo o seu próprio potencial de estudo, pesquisa e produção de conhecimento por meio da ampliação dos conceitos de acesso à informação, ao conhecimento e à produção do saber, para melhor aproveitamento das TIC.

Aos cursos de graduação e formação de profissionais na área de informação cabe o reconhecimento desse contexto em que estão inseridos, da realidade social do país, da região, das comunidades que a compõem e dos meios disponíveis para o exercício da disseminação de informação e criação de conhecimento no sentido de melhorar as condições de vida das populações que representam. Nesse sentido, a formação do profissional da informação, bibliotecário ou arquivista, futuro mediador das interações dos sujeitos que buscam significações nos signos para gerar novos conhecimentos, bem como a sua capacitação para a socialização do saber, deve obrigatoriamente relacionar-se primeiro, embora não somente, aos estoques de informação construída e à sua transferência à sociedade por meio de comunicação, nos específicos contextos sociais em que o profissional atuará.



Como aponta Aldo Barreto em *Os agregados da informação; - Memórias, esquecimento e estoques de informação* (2000, p.1), os estoques de informação representam itens de informação organizados, ou não, segundo critério técnico, do que o autor denomina instrumentos de gestão da informação cujo conteúdo interessa a comunidade de receptores.

Por essas razões, o cenário atual de interatividade entre comunidades globais se evidencia com o uso das TIC atrelado às metas da Sociedade da Informação no Brasil: as redes abertas e descentralizadas de informação e comunicação poderão contribuir para uma revolução na informação de maneira realmente ampla, compreendendo ações inerentes à diversidade cultural de seus usuários com a participação mais pró-ativa dos profissionais da informação.

Deve-se pensar em uma discriminação positiva entre os brasileiros, pois sem um aprendizado que leve em conta as diferenças nos diversos contextos do país se homogeneizarão diferenças; principalmente quando essas diferenças não puderem ser explicitadas por aqueles não familiarizados com a utilização dos meios. Vale lembrar o alerta de Barreto (2000, p.3) de que,

Grandes estoques crescentes de informação, que se acumulam em um tempo sem limites, degeneram a vivência cotidiana em que o conhecimento se realiza no indivíduo. A sintonia do sujeito consciente se dispersa em um mundo de informações irrelevantes, imprecisas e ultrapassadas e com uma distribuição inadequada (BARRETO, 2000, p.3).

As interfaces homem-máquina oferecem problemas aos sujeitos que com elas interagem em grande parte do contexto nacional atualmente e esse é um problema a ser tratado conjuntamente a um atrasado letramento anterior, procurando resolver conjuntamente ambos os desafios; se a informação necessita ser comunicada para reelaborar-se em conhecimento, o seu apoio ou suporte representacional necessita ser apreendido como processo tecnológico que é - escrito, imagético, sonoro ou multimidiático.

Os suportes de informação determinam formas de acumulação ou estocagem e também de abordagem para sua “leitura” e compreensão. Envolve-se nesse processo de determinação o universo da informação - considerada em sua materialidade e que propõe formas concretas de leitura - e o do leitor da informação - que vai determinar o verdadeiro conteúdo da informação pelas suas possibilidades e limites de leitura. Esses universos devem convergir para o equilíbrio na real apropriação dos materiais informativos.

Os processos digitais informacionais contemporâneos - que compreendem novas formas de relacionamentos entre diferentes subsistemas de informação - criam e ampliam de maneira transdisciplinar novas possibilidades de reflexão e de exploração do conhecimento neles veiculado.

O momento histórico de um país emergente, das dimensões e diversidades como as do Brasil, impõe relevantes mudanças de princípios e de paradigmas na organização da sua cultura - uma transição entre ciclos; contudo, o mundo assim constituído convida à realização das potencialidades humanas em latência, à espera dos estímulos culturais que propõem sujeitos da informação em ação.



O conceito de aprendizado, diante desse painel se modifica, nega a primazia do território escolar visto da maneira tradicional da informação e do conhecimento emitidos de um para muitos, valoriza a iniciativa individual e o auto-didatismo tanto nas buscas individuais quanto nas demandas sociais. Aprendizado implica atualmente também no compartilhamento da informação e do conhecimento que tem transformado a própria face do planeta.

O compartilhamento é uma viável forma de superação dos problemas identificados em realidades como a brasileira, pois no Brasil verifica-se, paralelamente às dificuldades detectadas pelo PISA, o surgimento de grupos significativos de sujeitos que convivem com as novas formas de aprendizado mediadas pelas TIC: tais grupos, que foram denominados internacionalmente de geração Y e de geração Z, a dos nativos digitais, na situação de diversidade mencionada no Brasil se sobrepõem, amalgamando-se.

Geração Y foi um conceito sociológico de definição do *coorte* dos nascidos entre finais de 1970 a 1990 em países com as características bem definidas como pós-industriais e que determinaram a chamada pós-modernidade; já a de 1990 até os dias atuais foi conceituada como geração Z. As definições de passagens de uma geração à outra, entretanto, são aproximadas mesmo nesses países. Suas características são extensivas a uma padronização global na verdade impossível, pois a extensividade planifica uma heterogeneidade de situações, dificulta consensos na definição de suas fronteiras temporais e conflitua também a própria conceituação, tanto no que diz respeito ao seu surgimento quanto à sua amplitude terminológica: enfim, o que se delineia no momento sob as denominações de geração Y e/ou Z é a inclusão aproximada daqueles atualmente com idades entre 16 a 27 anos.

Levando em consideração que no ambiente digital se processam conjuntamente protocolos tecnológicos e convenções sociais, é fato que os primeiros determinam como os computadores interagem. As convenções sociais, por outro lado - com a ampliação dos hábitos de *linkagem* e a crescente disponibilização de acesso e de regras de engajamento nos *sites* - tratam de como as pessoas gostam e/ou estão possibilitadas a interagir:

[...] estamos somente arranhando a superfície do que poderia ser alcançado com uma investigação científica mais profunda de seu design, operacionalidade e impacto sobre a sociedade. A *Web* permanece uma plataforma universal: independentemente de qualquer instrumento de hardware específico de uma plataforma de software, de linguagem, cultura ou falta de habilidades. (BERNES-LEE, 2007, tradução nossa).

A rede mundial de computadores, como conjunto de ambientes digitais de aprendizado e entretenimento, enfatiza aspectos de cooperação e de empoderamento como participação, conversação, auto-arquivamento, código visualizável e contextualizável, permitindo a constante agregação de dados, pois a sua característica de configuração rizomática tem grande potencial como veículo para a expressão da diversidade cultural, para a prestação de serviços *on-line* e para o desenvolvimento de práticas educacionais inclusivas.



O fator essencial na identificação das mudanças culturais tem sido atribuído ao fato dos sujeitos envolvidos no processo conviverem desde muito jovens, ou mesmo desde o nascimento, com as possibilidades geradas pelas mídias de massa, como a tevê, pela interatividade dos jogos eletrônicos, pela portabilidade e capacidade crescente de armazenamento de música nos dispositivos musicais (MPs), pelos celulares e as tecnologias de mobilidade, pela Internet e nela a Web, por seus aplicativos e codificações que resultam em representações convergidas por diversas linguagens e, finalmente, por mais convergências e interoperabilidade entre os objetos tecnológicos, as TIC e as tecnologias da inteligência.

O importante é destacar que os jovens descritos anteriormente conviveram sistematicamente, em seus lares, nas escolas, ou em outros ambientes, com a web gráfica (GUI), que surge a partir do programa Mosaico, em 1993. A modalidade gráfica é responsável pela grande expansão da rede mundial de computadores devido à sua clareza e à facilitação no uso das interfaces. A denominação “gráfica” vem da disposição de imagens junto aos textos, no lugar de disponibilizá-las em janelas separadas como nas tecnologias anteriores. O Mosaico foi descontinuado em 1997, mas os seus sucessores, o Internet Explorer, o Netscape, o Mozilla, o Chrome, ou outros produtos gratuitos da Google mantêm as características GUI (Graphic User Interface) que proporcionam, por sua vez, a experiência interativa de que atualmente pode-se fazer parte quando se acessa informação na rede.

Nos 1990 o Mosaico representou uma enorme revolução tecnológica porque propiciava uma visão convergente das codificações textual e imagética, e assim criava especial interesse entre a comunidade leiga de usuários. Como consequência, o uso disseminado deste tipo de aplicativos GUI - blogs, wikis, etc - provocou o contemporâneo oceano fluído de informações e patrocinou elementos de fixação da internet e da web que representaram as relevantes mudanças definidoras de uma nova conformação da web, primeiro como web gráfica, depois como web social e de compartilhamento.

A aquisição de informação auto gerida por provocações do meio digital determinou que os jovens da atualidade atuassem plenamente no mundo profissional e criassem empreendimentos como os das plataformas de blogs, as das wikis, entre elas a da Wikipédia (e de suas versões transnacionais), a do Facebook e de similares: inovações que transformaram os ambientes convencionais de entretenimento, de aprendizagem e de trabalho, com suas atitudes diferenciadas pelas formas de construção do conhecimento que desenvolveram.

Destaca-se, como exemplo do poder destas mídias que, segundo dados do ano de 2010, a Wikipédia tornou-se a maior enciclopédia do mundo, criada e mantida por mais de 100 mil colaboradores voluntários de todas as partes do planeta. A enciclopédia digital tem acessos mensais de mais de 388 milhões de pessoas que procuram um número dinâmico, portanto, de atualização diária, de mais de 16 milhões de artigos em mais de 260 línguas. Todo o conteúdo é livre para uso e modificação.



Como resultado, a auto-gestão na busca de conhecimento que estas mídias trazem embutidas em si já transformou seu entorno de atuação multi-geracional, construiu novos valores e visivelmente redesenhou a cultura como um todo.

Multi-tarefas e mais rápidos no acesso, identificação, utilização e reutilização das informações, os jovens da contemporaneidade têm sua atenção simultaneamente focada em vários assuntos do ambiente digital, do uso intensivo das TIC e de seus aplicativos

Considerando, assim, as mudanças nos princípios de aquisição de conhecimento e de habilidade, surgem algumas situações que devem ser atendidas por pesquisas, trabalhos e avaliações na contemporaneidade enfocando competências, como as do PISA.

Deve-se considerar também que tipo de qualidades são realmente avaliadas pelo PISA no mundo e nos países emergentes como o Brasil: o tipo de avaliação que o PISA propõe deve estar de acordo com novos tipos de conhecimento e de habilidades que estão se desenvolvendo nas novas gerações? Ou se deve estabelecer um padrão independente dos meios de tramitação das informações e das condições socio-econômicas e políticas dos países em que o exame é aplicado? Como as novas formatações da informação impactam os interagentes do processo informacional, comunicacional e de aquisição de conhecimento? Que mudanças se impõem sobre os aspectos criativos, estéticos e culturais nos processos de construção do conhecimento? Que formas de deshierarquização ou de nova hierarquização, de legitimidade, de papel dos pares e de confiabilidade da informação depositada de muitos para muitos estão surgindo? Qual o papel da informação nesta conexão entre pessoas e qual o papel da construção de novas comunidades com bases em interesses informacionais e de aquisição de conhecimento? Finalmente, e especificamente para o domínio científico que lhe cabe, como estas mudanças interferem nos próprios paradigmas anteriormente traçados pela Ciência da Informação?

3 LABORATÓRIO DE COMPARTILHAMENTO DE CONHECIMENTO EM AMBIENTE DIGITAL WIKIPÉDIA¹

Embora tais questionamentos sejam todos de extrema importância, este trabalho não pretende cobrir um número tão extenso de aspectos, porém detem-se sobre a apresentação da experiência embrionária que está se desenvolvendo no departamento de Ciência da Informação da Unesp de Marília. Foca-se na pesquisa de como a CI e suas estratégias de aproximação e facilitação do acesso e do compartilhamento do conhecimento pode contribuir, criando em um dos ambientes das TIC - a Wikipédia em português - meios de produção e de construção de conhecimento compartilhado.

Tem-se por princípio que atividades de cunho laboratorial, por simulatórias que são, poderão fornecer material não só para a análise, como para o desenvolvimento de novas competências exigidas na contemporaneidade, partindo do uso estratégico das mídias digitais na produção e na apropriação do conhecimento. Isto se dá por meio de simulação das atividades de busca, de seleção a partir do

¹ Palavra grafada com acento quando o termo representa o ambiente em língua portuguesa.



critério de qualidade e relevância, do *remix* de informação e republicação de textos, de imagens, de áudio e de vídeos, do controle dos elementos do ambiente digital e da convergência das diversas fontes de informação hipertextuais.

As ações anteriormente citadas seguem princípios da organização dos participantes de acordo com as tendências culturais dessa nova geração, porém com o apoio de tutores, cujo domínio estratégico dos ambientes digitais possa servir como fio de Ariadne² que conduza a metodologias de criação coletiva de conhecimento. Aderiu-se para tanto ao projeto que internacionalmente visa formar nas universidades compartilhadores de informação nos ambientes wiki. Este projeto está em expansão fora do Brasil, contando com embaixadores em *campi* de importantes escolas americanas e européias que se juntaram ao projeto da Fundação Wikimedia.

O Programa de Embaixadores de Campus da Wikipedia consiste de otimizar o trabalho de indivíduos já formados na cultura wiki, entre professores e alunos das universidades do mundo, para mediar voluntariamente o aprendizado do compartilhamento de informações nas Wikipédias em diversas línguas com um viés mais especializado, dando ao ambiente digital maior credibilidade.

Os embaixadores de *campi* são voluntários (graduandos, pós-graduandos, professores, funcionários, ou simplesmente pessoas ligadas por relações diversas de proximidade) que ligados a professores em determinadas disciplinas dão suporte, treinamento e assistência presencial aos estudantes das universidades em: editar a Wikipedia; trabalhar lado a lado para incorporar e implementar assuntos tratados em sala de aula aos verbetes da Wikipedia; recrutar novos colaboradores para a edição e tratamento da informação no ambiente e, principalmente, participar de atividades de apoio ao conhecimento livre. Os embaixadores também devem de maneira ideal organizar eventos nos *campi* para encorajar pessoas a editar a Wikipedia.

O programa fora do Brasil é focado essencialmente no compartilhamento. Adicionado a esta perspectiva, no início do ano de 2011, iniciou-se nos cursos de Arquivologia e Biblioteconomia do Departamento de Ciência da Informação da Unesp de Marília um laboratório de utilização da Wikipédia em português como ambiente de compartilhamento de informação e conhecimento no âmbito da disciplina História da Cultura.

No Brasil, a inovação de uso proposta, porém, consiste em simultaneamente preparar neste laboratório os alunos na escrita científica, capacitando-os nessa estrutura de formatação da informação condutora de conteúdos específicos com vistas a minimizar na universidade alguns dos problemas de letramento identificados pelo PISA nessa geração que já não terá oportunidades de solucioná-los no ensino fundamental e médio. Por outro lado, o ambiente digital escolhido para receber a experiência da escrita, a Wikipédia em português, ainda denota falta de consistência e de confiabilidade em relação aos verbetes comparados aos que são editados em outras línguas, como no inglês. A edição de melhorias nos verbetes escolhidos trará benefícios à comunidade usuária, portanto.

Se no Brasil não há muitos voluntários que se disponham a colaborar para a melhoria dessas

2 Metáfora para método que permita seguir vestígios de caminhos anteriormente traçados.



condições, o ambiente da Wikipedia mostrou-se ideal para o laboratório por várias razões: a primeira justificativa está na formatação explicitada pelas regras do ambiente, pois tais regras de escrita e edição dos textos científicos enciclopédicos estão muito bem dispostas pela própria organização do ambiente e o controle de qualidade do que se deposita como informação é feito de maneira ampla por revisores dedicados, zelosos e distribuídos em rede, corrigindo as postagens segundo critérios claros.

O ambiente digital hipertextual da Wikipedia torna-se, dessa maneira, uma via de duas mãos no que diz respeito ao compartilhamento: se de um lado pode ajudar no aprendizado das habilidades complementares da leitura e da escrita, de outro se oferece como um repositório aberto em que a informação depositada poderá ser reutilizada por pessoas falantes de língua portuguesa onde quer que estejam.

Esse ambiente foi escolhido porque quem busca informação na Wikipedia pode fazê-lo simplesmente utilizando o *site* como tradicionalmente se utilizava uma enciclopédia convencional. Porém, pode também criar uma conta de usuário e participar, compreendendo primeiramente a interface do usuário e relacionando-se com diferentes formas abertas de contribuição e de compartilhamento. Tal usuário deixa o papel passivo de quem apenas recebe os benefícios do artigo pronto, ou aparentemente pronto, pois os artigos estão em permanente transformação em busca de melhorias. Pode dispor-se a entender os processos e as características de construção de artigos com qualidade crescente.

Em números, o sistema Wikipedia se acresce de cerca de 1200 artigos diariamente, segundo dados da própria Fundação e, embora tenham em sua grande maioria pouca visibilidade, os voluntários enviam por volta de 4 milhões de contribuições mensais, compartilhando conhecimento, mas também outros insumos, como o do tempo necessário para disponibilizar e objetivar esses conhecimentos. Como a informação está em fluxo dinâmico, não há uma perspectiva de finalização das atividades sistêmicas de compartilhamento e de validação do conhecimento continuamente objetivado.

Os usuários cadastrados como contribuidores passam a ter uma identidade na Wikipedia, simples e importante instrumento de interação; assim têm acesso a mecanismos de criação de novas páginas; têm acesso também a recursos como o de páginas vigiadas que propiciam o controle da edição de artigos e a movimentação daqueles e naqueles já existentes. Há mecanismos para o acréscimo de imagens, o que demanda dos sujeitos conhecimento dos procedimentos reguladores deste outro conjunto de codificações e das leis que determinam o uso de imagens na atualidade. Tais atividades são gerenciadas pelos próprios participantes que ganham posições na hierarquia da Wikipédia à medida de suas participações.

A cultura da Wikipedia se expandiu em subsistemas em que contribuições podem ser especializadas e feitas em diversos níveis, além da autoria pura e simples: no nível sintático, um subnível é denominado Wikignomo, em que contribuições de pequenas correções gramaticais podem ser adicionadas; outro é o da Formatação, em que *experts* na linguagem de formatação textual utilizada na wiki aprimoram a apresentação das informações. Ainda no nível sintático, há as figuras do Ilustrador, especializado na inserção de imagens e do Revisor, que repassa os textos.



No nível das políticas do meio como o da mediação de assuntos controversos também existem figuras de contribuidores como a do Mediador que permeia as discussões e a do Mantenedor, uma espécie de vigilante de questões político-filosóficas.

A complexidade da cultura wikipedista reproduz as figuras de especialização do mundo exterior a ela e desenvolve muitas habilidades úteis aos futuros profissionais da informação. Por essa razão, principalmente, a Wikipédia em língua portuguesa e seus regradados conjuntos de procedimentos foram adotados como base para o laboratório de aprendizado da escrita científica com objetivos de melhoria das competências para essa aquisição de conhecimentos específicos a que a maioria dos alunos dos referidos cursos da Unesp, provindos do ensino público fundamental e médio, não teve acesso.

A escolha visa à atuação futura destes que atualmente são alunos dos cursos, mas que serão multiplicadores de conhecimento na mediação eficaz da informação, quando formados como profissionais da informação. Trata-se de uma estratégia de empoderamento que se refere à busca de autonomia e suficiência no processo de geração, tratamento, uso, re-uso e preservação de informação, considerando conhecimentos intersemióticos quanto aos instrumentais de edição, visualização e compreensão de informação gráfica na web social.

A utilização de conceituações interdisciplinares e das relações entre ciência, tecnologia, arte, criação e sociedade e o aprendizado da resolução criativa de problemas reforçam o embasamento adequado para a avaliação e aplicação do hipermídia como meio para um aprendizado pró-ativo e para o desenvolvimento das técnicas, metodologias e estratégias de forma coletiva.

As referências teóricas foram relacionadas com os temas da História da Cultura para adequar-se ao projeto da Fundação Wikimedia que visa uma expansão estratégica da enciclopédia digital ganhando concomitantemente qualidade e profundidade. A proposta do trabalho laboratorial está também de acordo com o projeto da docente da disciplina, que determina as formas de ações para o triênio 2010 – 2013. O laboratório foi concebido para cobrir áreas distintas de capacitação, tanto no *design* da informação digital, quanto no domínio da linguagem, da literácia e da administração de ambientes informacionais digitais.

Cerca de sessenta alunos do segundo ano, provenientes de cursos de Biblioteconomia e Arquivologia estão participando do projeto de melhoria de 27 possíveis pontos de acesso em língua portuguesa na Wikipédia. Os verbetes relacionados com a disciplina de graduação História da Cultura são: cultura cibernética, informação compartilhada, comunicação ciberespacial, comunicação de massa, comunicação mediada por computador, convergência tecnológica, cultura, cultura digital, cultura popular, cultura visual, *design* de informação, hipertexto, cultural história, história cultural brasileira, imperialismo cultural, inclusão social, indústria cultural, infografia, linguagem visual, meios de comunicação social, mídia, sociedade da informação, tecnologia da informação, tecnologias da informação e comunicação, web 2.0.

O campus da UNESP tornou-se o primeiro no Brasil a ter um embaixador da Wikipédia, aluno



quarto-anista do curso de Biblioteconomia e pesquisador FAPESP, cuja pesquisa trata de ambientes wiki. Seu trabalho se dá na linha de pesquisa Informação e Tecnologia e se desenvolve na temática de pesquisa Web Social e Cultura Digital: cooperação, colaboração e compartilhamento.

4 CONSIDERAÇÕES

Ainda em andamento, o trabalho laboratorial no ambiente da Wikipédia, pela sua estrutura heterogênea, mas harmonizada, anteriormente descrita, busca capacitar os participantes do projeto para as formas materiais de leitura (listas, formulários, gráficos, diagramas, etc) que os indivíduos leitores e produtores de informação encontram pela vida, como propõe o PISA.

Nesse sentido, o projeto procura habilitar os participantes para a identificação e recuperação de informações, para a compreensão e interpretação de aspectos formais e conteudísticos dos textos lidos em fontes diversas, porém confiáveis, bem como sua posterior reorganização no ambiente da Wikipédia como melhorias nos verbetes.

Embora as contribuições devam ser feitas individualmente, com os usuários registrados para avaliação sistemática dos resultados, a organização dos participantes em grupos de discussão reforça a necessidade dos indivíduos de desenvolver capacidades de construção de argumentações e defesa de pontos de vista para a melhoria dos verbetes escolhidos.

Finalmente, ao ler e reescrever inúmeras vezes os verbetes ligados à disciplina História da Cultura que deverão melhorar, o letramento em ciências é reforçado, capacitando os sujeitos ao reconhecimento das questões científicas, de materiais, métodos e proposições, implícitos nos textos estudados visando a sua passagem para o ambiente da Wikipedia.

Após esta primeira etapa, como mais uma atividade de preparação para a escrita científica em ambiente digital, propõe-se uma segunda etapa de verificação da confiabilidade do ambiente após o trabalho de consolidação e divulgação de informações sobre Cultura na Wikipédia em língua portuguesa. A ideia de validação é focada na verificação de produção de informação acadêmica de qualidade, no compartilhamento de informações entre os alunos e instituições de ensino superior, bem como entre falantes da língua portuguesa e, finalmente, entre os usuários da Wikipedia em geral. Também visa verificar o comportamento do compartilhamento e na colaboração dos estudantes como voluntários na tentativa de criar hábitos necessários para o empoderamento informacional no Brasil.

Todo o processo considera a contemporaneidade como um espaço/tempo de atuação multidisciplinar em que múltiplas tarefas relativas à informação são desenvolvidas; volta-se para a necessidade da formação de bibliotecários e arquivistas que atuarão para as gerações futuras no terceiro paradigma de Ciência da Informação; volta-se, em última instância, para a formação de profissionais que possam ter instrumentos cognitivos para reconhecer os elementos específicos de seu nicho profissional relacionado à informação, à análise documental, à descrição de recursos, e aos núcleos de processos de armazenamento de recursos informacionais.



A hipótese antes levantada de que novas abordagens do conhecimento veiculado compartilhadamente no ambiente digital potencializará indivíduos pensadores, pesquisadores, criadores e disseminadores de conteúdos informacionais e potencializará também a sua incorporação no conhecimento coletivo, evidentemente leva em conta o favorecimento do desenvolvimento social e a contribuição dos profissionais da informação na superação de deficiências como as demonstradas pelo PISA.

A Wikipédia em português se configura como um meio em que interagem homem e máquina. A busca de soluções dos problemas dos sujeitos em tal processo interativo passa pela necessidade de perseverantes esforços do tipo laboratorial que atualmente se empreende na UNESP de Marília, no sentido da criação de novos hábitos que influam simultaneamente no letramento propriamente dito, nas relações com o mundo digital e nas especificidades de pesquisa e de comunicação científica, procurando resolver conjuntamente os desafios que cada campo propõe.

Wikipedia Culture as shared capacitating politics in digital media of knowledge socialization

ABSTRACT: Presents a developing laboratory within History of Culture, subject taught in the Library and Archive Studies at UNESP (Marília). The project is based upon the foundations of research line Information and Technology: students from the second under graduation year participate in the improving of 27 entries in Portuguese language Wikipedia. The aim is to capacitate for scientific reading and writing in digital media, habilitate for the information identification and recuperation and for the interpretation and understanding of formal and contents aspects and its reorganization. It includes activities of search, selection, remix and republishing of texts, images, audio and videos in the convergence of diverse hypertext information sources, supported by tutors with strategic abilities in digital environments. In this sense it was adhered to the international Wikipedia Foundation University Campus Ambassadors project. It's also aimed to induce sharing and collaboration behaviors with the purpose of creating necessary habits for the informational empowerment in Brazil. As methodology is to optimize the work of individuals already trained in wiki culture and to create in within the subject information sharing programs with a more specialized bias, giving greater credibility to the digital environment. The environment syntaxes helps in the learning of the complementary skills of reading and writing and offers itself as an open repository from which information can be reused. It is, thus, an empowerment strategy in the search of autonomy and self reliance considering intersemiotic knowledge in the edition, visualization and understanding of information in the social web. A second step of a verifying research on the environment's credibility after the consolidation and dissemination of the entries improvement work is proposed.

Keywords: Information and Technology. Digital Culture and Sharing. Wikipedia.



REFERÊNCIAS

- ANDERS, P. **Cybrids**: integrating cognitive and physical space in architecture. Disponível em: http://www.isea-webarchive.org/mmbase/attachments/30193/ISEA_97_8.pdf. Acesso: 12 ago. 2011.
- ARENDT, H. **Entre o passado e o futuro**. São Paulo: Ed Perspectiva, 1972.
- BARRETO, A. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. In: **Ciência da Informação**, Brasília, v.27, 1998. 1998. Disponível em: http://www.scielo.br/scielo.php?pid=S0100-9651998000200003&script=sci_arttext&tl=. Acesso em: 11 ago. 2011.
- _____. Os agregados da informação: memórias, esquecimento e estoques de informação. In: **DataGramZero - Revista de Ciência da Informação**, Rio de Janeiro, v.1, n.3, 2000. Disponível em: http://www.dgz.org.br/jun00/Art_01.htm. Acesso em: 20 jul. 2011.
- BERNES-LEE, T. Testimony of sir Timothy Berners-Lee. In: **Digital future of the United States: the future of the world wide web – part I**. Disponível em: <http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web>. Acesso em: 15 ago. 2011.
- INEP. **Instituto Nacional de Ensino e Pesquisa**. 2011. Disponível em: <http://www.inep.gov.br/internacional/pisa/>. Acesso em 01 ago. 2011.
- JORENTE, M.J.V. Digital inclusion initiatives in Brazil: improving education and information seeking behavior through government-academic partnerships. In: **Bulletin of the American Society for Information Science and Technology**, v. 34, p. 30-34, 2008. Disponível em: http://www.asis.org/Bulletin/Feb-08/FebMar08_Jorente.html. Acesso: 02 ago.2011
- _____. **Tecnologias, mídias, criação e hipertextualidade na transformação de informação em conhecimento interativo**. 2009, 260 f. Tese (Doutorado em Ciência da Informação) – Marília, Universidade Estadual Paulista, 2009. Disponível em: <http://acervodigital.unesp.br/handle/123456789/37004>. Acesso: 02 ago. 2011.
- INGWERSEN, P. Conceptions of Information Science. In: VAKKARI, P.; CRONIN, B. (Ed.). **Conceptions of library and Information Science**: historical, empirical and theoretical perspectives. London: Taylor Graham, 1992. p. 299-312.
- MOLES, A. **Sociodinâmica da Cultura**. São Paulo: Ed. Perspectiva, 1975.
- SANTOS, P.L.V.A.C.; BOTENTTUIT, A.M ; GROSSI, A.; JORENTE, M. J. Redes de informação e perspectiva de inclusão digital: um olhar para a educação. In: JORNADA PEDAGÓGICA, 12., 2008, Marília, **Anais eletrônicos...** Marília: UNESP, 2008. p.1-15.



USO DE TECNOLOGIAS DA INFORMAÇÃO PARA ANÁLISE DE RELEVÂNCIA USANDO “CROWDSOURCING”

Marcos Luiz Mucheroni, José Fernando Modesto da Silva

Resumo: Este trabalho analisa os conceitos fundamentais de recuperação da informação e propõe uma ferramenta que usa o poder das multidões que a conexão eletrônica em rede propiciou e cujo fenômeno inicialmente apenas perceptível para o código livre agora penetrou em outras áreas como a administração, o ativismo político social, o marketing. Este recrutamento on-line ficou conhecido pelo nome de *crowdsourcing*. Uma área sensível a estas mudanças é a Ciência da Informação, onde a relevância de documentos é importante. Anteriormente era feita somente por especialista porém agora se propõe ao uso do *crowdsourcing*. Neste sentido, o objetivo do texto é comentar tais conceitos e apresentar uma ferramenta para avaliação de relevância denominada TERCO. A ferramenta proposta retorna as avaliações de relevância sobre textos, artigos e publicações em um arquivo ao usuário com o contexto XML, permitindo que também ele se torne um disseminador. Pode-se dispor desta ferramenta em torno de qualquer comunidade discursiva que queira se tornar colaboradora na análise e no processo de relevância de documentos.

Palavras-chave: Tecnologias da Informação, Relevância, *Crowdsourcing*.

1. INTRODUÇÃO

As tecnologias tornaram o conhecimento, de mero apêndice da produção geral da sociedade, em uma força de trabalho central e imprescindível e que o processo de conexão global promoveu, tornando-a uma fonte de desenvolvimento.

Este paradigma, no aspecto da informação, detém uma característica social importante que permite a reorganização da força de trabalho, integrando-a de modo a formar um todo com o trabalho manual (que era antes dividido em diversas unidades estanques), originando um capitalismo cognitivo (BOUTANG, 2007).

Surgem a partir disto novas práticas interativas, que estimulam a criação de colaborações, propiciando um fluxo e aumentando a integração entre trabalhos (LE COADIC, 2003), gerando uma inteligência coletiva (LEVY, 1999) e instrumentos capazes de tomar decisões em conjunto.

Estes aspectos possibilitam a produção e utilização intensiva de informação, tornada um requisito essencial em qualquer empresa ou instituição (pública ou privada), incluindo as “multidões” como força cognitiva.

O desenvolvimento das tecnologias da informação coloca o conhecimento como fonte de valor e de poder e provoca profundas alterações na organização do trabalho (TOFFLER, 1990), com



a passagem do modelo conhecido como taylorista-fordista para o modelo chamado de especialização flexível (MACHADO, 1993).

Há agora uma ênfase na capacidade de mudar de especialidade, chamada em termos educacionais de transversalidade. Assim, este conceito de flexibilidade implica em mudança do processo educativo, das relações sociais entre gêneros e idades, e dos sistemas de valores (SCHAFF, 1995), e uma revalorização de todas as pessoas, que fazem parte desta multidão.

A sociedade atual, agora chamada de: sociedade pós-industrial (LYOTARD, 2003), sociedade informática (SCHAFF, 1995), sociedade do conhecimento (TOFFLER, 1990), sociedade “tecnizada” (MACHADO, 1993) ou sociedade em rede (CASTELLS, 2003), apresenta uma nova forma social onde o poder das multidões (*crowdsourcing*) (HOWE, 2009), e a entrada do conhecimento (força produtiva), valorizam o trabalho imaterial (HARDT, 2005), pelo uso e aplicação da subjetivação do trabalho (GORZ, 2005), no assim chamado: capitalismo cognitivo (COCCO, 2003).

O papel que desempenha a internet e o desenvolvimento de conteúdos através da web (MARTINEZ USERO, 2007), ainda que polêmico em torno de aspectos de relevância e validação destes conteúdos e o fato de historicamente ocupado pela Recuperação da Informação, tornam inegável que esta mudança altere a força de produção, não apenas do conhecimento, mas de todo o trabalho produtivo que, de alguma forma, está vinculado a ele. Esta é a ligação entre o aspecto da discussão em torno do *crowdsourcing* e sua relação com a importante área da Recuperação da Informação (RI).

Há uma nova relação onde todos podem ser atores e ativos, conforme conceito de Castells (2003, p.8):

“A Internet é um meio de comunicação que permite, pela primeira vez, a comunicação de muitos com muitos, num momento escolhido, em escala global. Assim como a difusão da máquina impressora no Ocidente criou o que McLuhan chamou de a ‘Galáxia Gutemberg’, ingressamos agora num novo mundo de comunicação: a Galáxia da Internet”.

Isto não seria possível sem movimentos na web que tornam a disseminação da informação possível e aberta - além das necessidades de indexação da documentação (MARTINEZ UBERO, 2006).

A organização do conhecimento de forma a incorporá-lo no produto faz emergir o conceito de “trabalho imaterial”, porém a avalanche de informações nem sempre é organizada segundo critérios de relevância e certificação em contextos (HERNANDEZ e outros, 2007).

Quem financia estes projetos inicialmente foram pessoas e organizações públicas interessadas neste tipo de desenvolvimento. Em 2006, Michael Sullivan criou o termo *crowdfunding* para coletar fundos através das redes eletrônicas; há iniciativas para financiar blogs e jornalismo (CROWDFUNDING, 2009), mecanismos de micro empréstimos para pequenos negócios, como por exemplo, para o terceiro-mundo, como oferecidos pela organização Kiva (KIVA, 2011).

A preocupação na qual se concentram não só pesquisadores, mas projetistas de sistemas, cientistas da informação e também algumas empresas é como organizar todo este “produto” gerado desta sociedade em redes, e que garanta alguns fatores essenciais dentro do princípio da relevância. Nesse sentido, os testes de Cranfield para indexadores em linguagens naturais (CLEVERDON, 1997) fizeram uma proposta levando em consideração os seguintes aspectos:

- Selecionar os mais importantes conceitos, sendo registrados na linguagem natural dos documentos;
- Listar os termos simples de cada um dos conceitos; e
- Combinar as formas diferentes de um mesmo conceito (usando chaves lógicas).

Propostas de ambientes multimodais, incluindo multimídias, multi-linguagens em formas livres de contexto podem ser formas de recuperação de informação (INGWERSEN; JÄVELIN, 2005).

Nestes ambientes está implícito um contexto onde os objetos de informação são interoperáveis tendo entidade física em alguma mídia em sistemas de RI.

A figura 1 ilustra estes tipos de ambientes de informação:

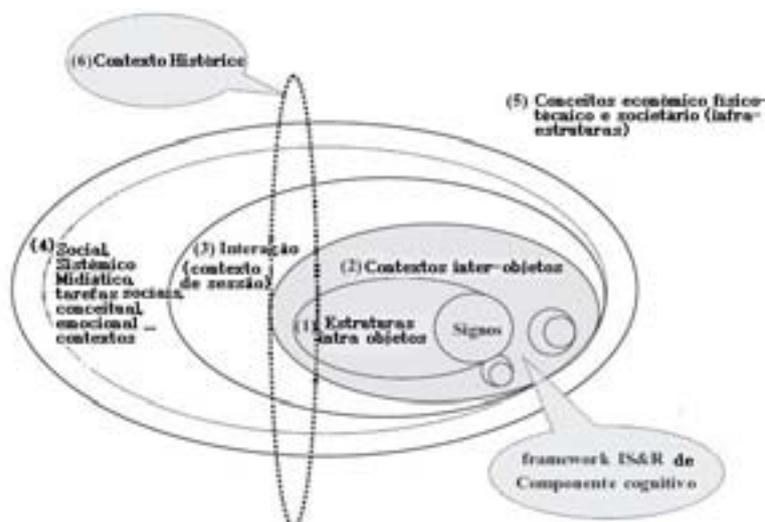


Figura 1 - Recuperação de informação em diversos níveis
(Fonte: Ingwersen e Järvelin, 2005).

Na figura 1, as estruturas intra-objetos estão relacionadas aos signos então nelas definidos, como os tipos feitos no contexto inter-objetos: (1) podendo haver outros ambientes com estruturas e signos diferentes produzindo um contexto inter-objetos; (2) o conjunto do contexto inter-objetos é o que produz um “Framework de Componente Cognitivo” definido por Ingwersen & Järvelin (2005), e que cria a possibilidade de uma sessão contexto (3).

Este conjunto ambiente e frameworks produzidos tecnicamente estão dentro de uma primeira camada contexto “social”, das “tarefas sociais”, das “mídias” utilizadas, dos conceitos adquiridos e do ambiente emocional (4). Esse, por sua vez, está dentro do ambiente “societário” que permite esta



Ambientes on-line que envolvam a descrição semântica de conteúdos - por isto esta Web é chamada Web Semântica - devem envolver a metalinguagem de descrição de conteúdos XML (eXtende Markup Language) e uma sintaxe que estruture os dados desta linguagem em um contexto Inter-objetos de documentos que estejam no acesso on-line, embora a RDF (Resource Definition Format) seja a mais conhecida e utilizada, a linguagem XSD (XML Schema Definition) já é um padrão recomendado desde 2001 pelo W3C.

Neste aspecto, o objetivo do texto é comentar os conceitos de relevância e *crowdsourcing* como subsídio ao desenvolvimento de sistema que ofereça à área da Ciência da Informação um modelo de avaliação de textos, artigos e publicações. É enfatizado um modelo para realização de testes através da ferramenta denominada: TERCO (TERC On-line), disponível on-line (www.wlinfo.com.br/TERCO). Fundamenta-se a modelagem da ferramenta em base teórica e metodológica descrita a seguir.

2. AS ESTRUTURAS INTER-OBJETOS

A escrita certamente não é inata. Entretanto, a escrita pode ser internalizada e tornar-se um hábito da mente. O que é natural parece mais intimamente e obviamente humano. Por essa razão não queremos insistir no fato de que a escrita é uma tecnologia; queremos que a habilidade da escrita seja natural. Nós gostamos de nossas ferramentas e máquinas, mas também gostamos da ideia de que podemos existir sem elas. Isolar nossa tecnologia nos dá a sensação de autonomia e nos permite reafirmar a diferença entre o natural e o meramente artificial. Desde o tempo de Rousseau, muita da hostilidade dirigida contra a moderna tecnologia foi enraizada na crença de que, conforme nossa tecnologia se torna mais complexa, nós nos tornamos criaturas de nossas próprias criações. (BOLTER, 1991, p. 36-37, tradução nossa).

A organização de busca e consultas fundamentadas na Ciência da Informação agora devem suportar estruturas Inter-objetos. No cenário em que muitos documentos estão on-line e uma grande parte em acesso aberto, especialmente os documentos acadêmicos e científicos, tais estruturas assumem importância.

Assim, além de estudos de relevância que poderiam ser feitos a partir de uma grande força tarefa (recrutada entre alunos, professores e outros voluntários que queiram exercer voluntariamente este tipo de trabalho), também os documentos e estruturas devem suportar uma nova relação entre usuário/produtores de conhecimento e os objetos.

Várias definições de relevância podem ser encontradas no campo da Ciência da Informação, destacando duas historicamente conhecidas:

- Uma medida da informação veiculada por um documento de consulta (GOFFMAN, 1964);
- Um critério usado para quantificar o fenômeno que envolve indivíduos (usuários) julgando relações, utilidade, importância, o grau de: correspondência, proximidade, adequação, pertinência, valor ou serem munidos de documentos ou representações de uma exigência de informação, necessidade, causa, declaração, descrição da pesquisa, tratamento, etc. (REES, 1966).

infraestrutura físico-técnica (5). Infraestrutura que está dentro de um contexto histórico (6). Importante notar que esta etapa interage diretamente com os níveis do “Framework de Componente Cognitivo”.

Neste contexto inter-objetos, na camada denominada “interação”, é possível inserir novos conceitos de critérios de relevância semi-automáticos a partir da linguagem XML, gerando mensagens automáticas de contextos, ao mesmo tempo em que sugerem testes de consultas aos leitores e autores; a estrutura de documentação OAI (*Open Archives Initiative*) auxilia neste caso (MUCHERONI; CABRAL, 2009).

A recuperação da informação, conceito clássico em Ciência da Informação, incorpora o conceito de relevância e de busca de dados, que se modificou profundamente a partir do desenvolvimento da Web (BAEZA-YETES; RIBEIRO-NETO, 1999); entretanto, há uma necessidade de se recuperar conceitos tais como: relevância e validade de documentos on-line na Web.

3. A RELEVÂNCIA E O USO DO CROWDSOURCING

A partir do trabalho de Sperber e Wilson (1986), os critérios de relevância passam a ser detalhados com base no estudo da comunicação humana firmada na ideia fundamental de que as informações comunicadas possuem uma garantia de relevância e onde aparecem as ideias de contextos, graus de compreensão e sua relação com a cognição.

O conceito de relevância é uma temática que compartilha com muitas características comuns com o campo dos Sistemas de Recuperação da Informação (SRI), em especial a multidisciplinaridade. As discussões envolvem basicamente a falta de consenso sobre a natureza e significado da relevância (não existe um único tipo de relevância) e a forma de operacionalizar em desenho e avaliação dos sistemas de informação (SALVADOR OLIVAN, 2008).

A figura 2 ilustra de forma sistematizada estes tipos:

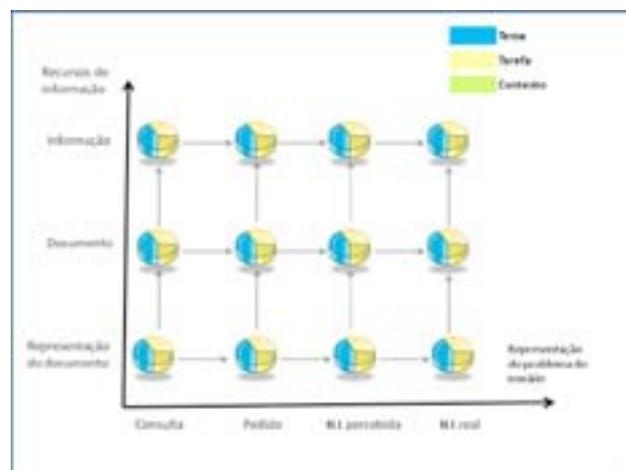


Figura 2. Tipos de relevância: recursos e representação pelo usuário.
(Fonte: SALVADOR OLIVAN, 2008)



- 1) Formada por recursos de informação: informação, documento, e representação do documento;
- 2) Formada pela representação do problema do usuário: necessidade de informação real, necessidade de informação percebida, necessidade de informação expressada em linguagem natural, estratégia de busca na linguagem do SRI;
- 3) Formada por três componentes em cada uma das entidades descritas nas duas primeiras dimensões: tema (a área temática de interesse do usuário), tarefa (atividade realizada pelo usuário com os documentos recuperados), e contexto (envolve tudo relacionado com o tema e tarefa, que influem em como se realiza a busca e a avaliação dos resultados); e,
- 4) Representada pelo intervalo de tempo que transcorre desde o surgimento da necessidade da informação até sua solução.

As possíveis relações entre as diferentes entidades contidas nas quatro dimensões apresentadas mostram os diferentes tipos de relevância existentes. O processo torna-se mais complexo quando se considera quem avalia a relevância: usuário, sistema, especialista no tema ou intermediário. Quem avalia pode restringir os tipos de relevância medidos ou utilizados. Assim, por exemplo, só o usuário pode determinar a relevância de um documento para suas necessidades de informação real ou percebida.

Segundo Saracevic (2007), os tipos de relevância podem ser resumidos, de acordo com área e motivações em:

- Relevância do sistema (ou algorítmico): relação entre a consulta e os objetos da informação (documentos) de um sistema recuperado ou não recuperado, por um determinado procedimento ou algoritmo.
- Relevância temática: relação entre o tema expressado em uma consulta e o tema coberto pelos documentos recuperados.
- Relevância cognitiva ou pertinência: relação entre o estado de conhecimento e a necessidade de informação cognitiva de um usuário e os documentos recuperados.
- Relevância situacional ou utilidade: relação entre a situação, tarefa ou problema e os documentos recuperados.
- Relevância afetiva ou motivacional: relação entre as intenções, objetivos e motivações de um usuário e os documentos recuperados.

Sumariamente, para os objetivos deste texto, relevância significa estabelecer um critério de validação para ambientes de recuperação da informação.

Anteriormente, este processo era uma tarefa difícil e cara por implicar no pagamento de pessoas dispostas a realizarem alguns testes. O procedimento era feito por um conjunto de voluntários



recrutados entre graduandos, profissionais aposentados, e bibliotecários com disponibilidade de lerem cuidadosamente documentos e verificar a relevância através de alguns testes.

O processo era custoso (pessoas deviam ser pagas) e limitado (encontrar pessoas com tempo disponível), havendo, também, a necessidade de conhecimento para realizar os testes, o que requeria um grupo de especialistas interessados em manter o processo de relevância da informação.

Assim, construídos os aspectos de relevância como um dos fundamentos de Recuperação da Informação, e estabelecida a possibilidade do uso da “multidão”, analisa-se agora sua evolução.

Com o início do TREC em 1992, foi possível pensar em uma avaliação parcial de documentos em textos completos. Não é possível imaginar que, toda vez que o documento é lido, ele seja analisado, então foi pensada uma abordagem onde a cada N documentos lidos uma avaliação é feita, o que se denominou TREC (*Experiment and Evaluation in Information Retrieval*) (VOORHEES, 2007).

Entretanto, era processo caro, devido ao pagamento de avaliadores. Alguns trabalhos atuais discutem a eficiência destes métodos em ambientes contextuais, conforme apresentado na figura 2 (SARACEVIC, 2007).

No estudo de Saracevic há discussão sobre a forma como as pessoas se comportam acerca de relevância. Demonstra que a maioria não tem critérios contextuais, o que não é uma surpresa. Até porque pessoas com o tempo livre não é o mesmo que habilitadas.

Outros estudos sobre a questão têm explorado o comportamento do usuário como um sinal de avaliação. O simples clique do usuário não pode ser utilizado para avaliar a qualidade de uma pesquisa, uma vez que a falta de um clique pode indicar tanto uma satisfação com a informação, como não consegue dizer sobre sua importância (JOACHIMS; RADLINKSKI, 2007), e sua evolução com a web (ROSE, 2004) e em redes sociais (OLIVEIRA; IRVING, 2008).

Assim, deve-se pensar em uma terceira abordagem usando a Rede, conforme proposta por Alonso, Rose e Stewart (2008), numa técnica chamada TERC (*Technique for Evaluating Relevance by Crowdsourcing*), que adapta um experimento para avaliar a inteligência coletiva na rede (BARR; CABRERA, 2006) e sendo utilizado pela Amazon em um procedimento chamado MTurk (AMAZON, 2009).

Howe (2006), cunhou o termo *crowdsourcing* em artigo para revista Wired, onde descreve tarefas que são terceirizadas para um grande grupo de pessoas ao invés de executados por um funcionário (HOWE, 2009). Desta forma, *crowdsourcing* pode ser visto como convite aberto às pessoas para resolver um problema ou realizar uma tarefa que até envolva um valor monetário em troca de tal serviço.

O *crowdsourcing* proposto tem estilo “Web 2.0” por prever que, a partir do aumento da participação interativa de um grande número de pessoas on-line, cada pessoa poderá optar pela adoção da ferramenta através da mensagem XML recebida. Uma vez cadastrada no sistema, poderá também sugerir questões que possam ser avaliadas quanto à relevância.



4. Ferramenta de Relevância – Proposta

A ferramenta proposta neste trabalho é uma extensão para o TERC, que é um processo de disseminação do teste pela Web. É denominada TERCO (TERC On-line), e pode efetivamente explorar a potencialidade do conceito de multidão, na Web.

A ferramenta comentada está disponível on-line (www.wlinfo.com.br/TERCO). A finalidade é possibilitar que o usuário gere o seu próprio teste, customizando-o em seu repositório pessoal de documentos. Com base na pergunta que ele queira formular e utilizando o formato proposto pelo TERC, gera-se um código XML com a questão, sendo enviada ao e-mail do usuário como resposta.

O formato proposto para o TERCO é semelhante ao do TERC, com o campo DTD CDATA [[questão]] variável, de acordo com a proposta do usuário.

Os campos do XML para o TERCO, adaptado de TERC são:

```
<Questao>
  <QuestaoIdentificador>Questao1</QuestaoIdentificador>
  <DisplayNome>Questao 1:</DisplayNome>
  <NecessarioE>true</NecessarioE>
  <QuestaoConteudo>
    <ConteudoFormatadoPeloUsuario><![CDATA[
O seguinte texto é relevante para capitalismo cognitivo ?]]>
  </ConteudoFormatadoPeloUsuario>
</QuestaoConteudo>
  <RespostaSpecification>
    <SelecaoResposta>
      <SugestaoEstilo>radiobutton</SugestaoEstilo>
      <Selecoes>
        <Selecao>
          <SelecaoIdentificador>ir</SelecaoIdentificador>
          <Text>Irrelevante</Text>
        </Selecao>
        <Selecao>
          <SelecaoIdentificador>mr</SelecaoIdentificador>
          <Text>Pouco relevante</Text>
        </Selecao>
        <Selecao>
          <SelecaoIdentificador>fr</SelecaoIdentificador>
          <Text>Muito relevante</Text>
        </Selecao>
        <Selecao>
          <SelecaoIdentificador>hr</SelecaoIdentificador>
          <Text>Altamente relevante</Text>
        </Selecao>
      </Selecoes>
    </SelecaoResposta>
  </RespostaSpecification>
</Questao>
```

A questão gera, além da mensagem em XML, um código HTML, que pode ser usado pelo usuário para disponibilizar em seu repositório. Já, no exemplo abaixo, é apresentada na figura 3 a interface do TERCO:



Figura 3. O ambiente TERC On-line (TERCO).

O ambiente de desenvolvimento foi construído em PHP (Hypertext Pre-Processor), que é uma linguagem de script adequada para o desenvolvimento Web, sendo *open-source*. O aplicativo, em breve, terá todo o código disponibilizado. Na consolidação de desenvolvimento foi realizado um teste inicial com 32 alunos de graduação do curso de Biblioteconomia da Universidade de São Paulo, que resultou promissor para desenvolvimento do ambiente.

A tabela a seguir tem apenas o efeito dos testes iniciais. O ambiente de validação depende de um ambiente mais amplo que possa inserir o *crowdsourcing*.

Tabela 1. Teste de relevância feito para um conceito de “capitalismo cognitivo”.

Conceito	Irrelevante	Pouco	Muito	Altamente
Relevância			X	
No. Usuários	2	5	20	5

O ambiente deverá incorporar ferramentas para permitir a colaboração de diversos grupos independentes na realização de seus próprios testes de relevância.

A validação do ambiente depende de um amplo repositório multimodal de documentos que está em fase de construção, sendo proposta de um projeto correlato.

4 CONCLUSÕES PRELIMINARES

A tradição da Ciência da Informação na Recuperação da Informação é significativa, como prova a força de eventos como o TREC.

A preocupação de pesquisadores, projetistas de sistemas, e cientistas da informação com a



organização da informação gerada pela sociedade em rede, leva ao desenvolvimento de conceitos baseados em princípios de relevância; princípios estes que subsidiam a proposta deste texto de estabelecer um critério de validação para recuperação da informação. Neste sentido, é destacado o conceito do *crowdsourcing* para descrever tarefas que podem envolver grupos de pessoas na sua execução.

A proposta descrita apresenta a ferramenta denominada TERC, destinada à realização de testes de relevância no ambiente da web. Baseada no conceito da “multidão”, opera com a participação das pessoas para resolver um problema ou realizar uma tarefa.

Desta maneira, a ferramenta proposta, além de realizar testes de relevância, permite ao usuário incluí-la em algum ambiente ou, ao menos, gerar seu próprio teste on-line.

O recurso tem o estilo “Web 2.0” de interatividade por envolver usuários on-line e a facilidade de adoção. Espera-se que o contato e uso da ferramenta, bem como a discussão sobre a mesma, estimule seu próprio aperfeiçoamento.

Abstract: *This paper examines the fundamental concepts of information retrieval and proposes a tool that uses the power of crowds, the connection provided electronic networking, and whose only noticeable phenomenon initially for free now entered the code in other areas such as administration, activism sociopolitical, marketing, the online recruitment, it became known as crowdsourcing. An area sensitive to these changes is the Information Science, where the relevance of documents is important, but previously was only done by a specialist, now proposes the use of crowdsourcing. In this sense, the goal is the text comment on such concepts and present a tool for assessment of relevance called TERC. The proposed tool returns the evaluations of relevance to texts, articles and publications in a file the user with the XML context, also allowing him to become a disseminator. You can have this tool around any discourse community that wants to become a collaborator in the process of analysis and relevance document's.*

Keywords: Information Technology, Relevance, Crowdsourcing.

REFERÊNCIAS

ALONSO, O.; ROSE, D. E.; STEWART, B. Crowdsourcing for Relevance Evaluation. **ACM SIGIR Forum**, vol.. 42, n. 2, 2008.

AMAZON Mechanical Turk: Artificial Artificial Inteligence. Disponível em: <http://www.mturk.com>. Acesso em: 30-06-09.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. USA: Addison-Wesley, 1999.

BARR, J.; CABRERA, L. F. AI Gets a Brain. **ACM Queue**, vol. 4, n.4, p. 24-29, 2006.



BOUTANG, Y. **Le capitalismo Cognitif: la nouvelle Grande Transformation**. Paris: Amsterdam, 2007.

BOUTANG, Y. M. Marx no século XXI: uma triste história de adeus ao socialismo ou uma outra coisa? **Periferia**, vol. 1, n. 1, 2009. Disponível em: http://www.febf.uerj.br/periferia/V1N1/yann_mboutang.pdf. Acesso em: 27-07-2009

CASTELLS, M. **A Galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade**. Rio de Janeiro: Jorge Zahar Ed, 2003.

CLEVERDON, C. The Cranfield Tests on Index Language Devices. Em: JONES, K. S.; WILLETT, P. **Readings in Information Retrieval**. San Francisco: Morgan Kaufmann, 1997. p. 47- 59.

COCCO, G.; SILVA, G.; GALVÃO, A. P. **Capitalismo cognitivo**. Rio de Janeiro: DP&A, 2003.

CROWDFUNDING, Journalism Idioplataform, Disponível em: <http://idioplatform.com/2009/05/crowdfunding-journalism/>, acesso: 2009.

GOFFMAN, W. On relevance as a measure. **Information Storage & Retrieval**, n.2, p.201-203, 1964.

GORZ, A. **Conhecimento, Valor e Capital**. Rio de Janeiro: Anna Blume, 2005.

HARDT, M.; NEGRI, A. **Multidão: guerra e democracia na era do império**. Rio de Janeiro: Record, 2005.

HERNANDEZ, N. H.; MOTHE, J.; CHRISMENT, C.; EGRET, D. Modeling context through domain ontologies. **Information Retrieval**, vol. 10, p. 143-172, 2007.

HOWE, J. The Rise of Crowdsourcing. **Wired**, vol. 14, n.6, 2006.

HOWE, J. **O poder das multidões**. Rio de Janeiro: Elsevier, 2009.

INGWERSEN, P.; JÄRVELIN, K. **The Turn: Integration of Information Seeking and Retrieval in Context**. New York: Springer-Verlag, 2005.

JOACHIMS, T.; RADLINKSKI, F. Search Engines that Learn from Implicit Feedback. **IEEE Computer**, vol. 40, n. 8, 2007.

KIVA. Empower people around the world, Disp. em: <http://www.kiva.org/>, Acesso: 2011.

LE COADIC, Y. **A Ciência da Informação**. 2. ed. Brasília: Briquet de Lemos, 2003.

LEVY, P. **Cibercultura**. São Paulo: Ed. 34, 1999.

LYOTARD, J. **A condição pós-moderna**. Lisboa: Gradiva, 2003.

MACHADO, L. Sociedade industrial X sociedade tecnizada. **Universidade e Sociedade**, ano 3, n. 5, p. 32-37, 1993.

MARTINEZ USERO, J. Á.; LARA NAVARRA, P. **La organización del conocimiento em Internet**. Barcelona: UOC, 2006.

MARTINEZ USERO, J. Á.; LARA NAVARRA, P. **La producción de contenidos web**. Barcelona: UOC, 2007.



MUCHERONI, M.L.; CABRAL, W. A. S. Authorship and management in OER using OAI. **In.** 1st. International Workshop on “D4PL - Designing for Participatory Learning”, 2009. Suécia: Kövde, 2009.

OLIVEIRA, E.; IRVING, M. Redes virtuais: da discussão teórica às potencialidades contemporâneas para a consolidação de redes sociais. **Revista Textos de la CiberSociedad**, n.13, 2008. Disponível em: <http://www.cibersociedad.net>. Acesso em: 27-07-09

REES, A. M. The relevance of relevance to the testing and evaluation of document retrieval systems. **Aslib Proceedings**, vol.18, n.11, p. 316-324, 1966.

ROSE, D. E. Why Is Web Search So Hard... to Evaluate? **Journal of Web Engineering**, vol. 3, n. 3/4, p. 171-181, 2004.

SALVADOR OLIVAN, J. A. **Recuperación de la información**. Buenos Aires: Alfagrama, 2008.

SARACEVIC, T. Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part III: Behavior and Effects on Relevance. **Journal of the American Society for Information Science and Technology**, vol. 58, n.13, p.212-214, 2007.

SCHAFF, A. **A sociedade informática**. São Paulo: Brasiliense, 1995.

SPERBER, D.; WILSON, D. **Relevance: Communication and Cognition**. Cambridge, MA: Harvard University Press, 1986.

TOFFLER, A. **Powershift**. Rio de Janeiro: Record, 1990.

VOORHEES, E. TREC: Continuing Information Retrieval's Tradition of Experimentation. **Comm. of the ACM**, vol. 50, no. 1, 2007.



AVALIAÇÃO DE USABILIDADE DA BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES: UM ESTUDO DE CASO

Fernanda Pereira

Gercina Ângela Borém de Oliveira Lima

Resumo: As bibliotecas digitais têm por objetivo disponibilizar informações para seus usuários com maior rapidez e facilidade de busca e acesso. Nessas bibliotecas, a interação dos usuários com o acervo realiza-se por meio de suas interfaces de ligação entre usuário e sistema. Esse trabalho propõe um estudo de caso sobre a usabilidade da Biblioteca Digital Brasileira de Teses e Dissertações, mantida pelo Instituto Brasileiro de Ciência e Tecnologia, BDTD-IBICT, (1) evidenciando a contribuição dos princípios de usabilidade para melhoria das interfaces de bibliotecas digitais; (2) mapeando os problemas de usabilidade da interface escolhida para o estudo; e (3) propondo soluções para os problemas encontrados. Discute-se como as avaliações de usabilidade podem ser aplicadas na avaliação de bibliotecas digitais com vistas a melhorar a interação humano-computador. A interface foi avaliada por meio de um método analítico (avaliação heurística) e um método empírico (teste com usuários). Na avaliação heurística, os dados foram coletados por meio de três formulários próprios para cada uma das três etapas da avaliação. O uso desses documentos orientou os avaliadores e evitou perdas de informações durante a avaliação. Para a avaliação empírica, foram utilizados questionários e anotações das observações feitas durante a avaliação. Nessa pesquisa qualitativa os dados colhidos foram analisados indutivamente. A fundamentação teórico-metodológica abordou as bibliotecas digitais, sua usabilidade e interfaces. Os resultados apontam para uma lista de problemas relacionados à usabilidade que interferem diretamente na interação do usuário com a BDTD-IBICT.

Palavras-chave: Biblioteca digital. Usabilidade. Avaliação heurística. Teste com usuário.

1 INTRODUÇÃO

A criação de um acervo digital sem uma política de indexação e pesquisa sobre seu usuário potencial pode acarretar em violação de alguns princípios de usabilidade e insatisfação do usuário. A falta de qualidade no serviço de entrada de dados no sistema acarreta problemas na recuperação de documentos e interfere diretamente na satisfação de uso do sistema pelo usuário.

A usabilidade analisa a qualidade de uso destas interfaces através de técnicas e métodos que buscam identificar problemas que afetam a satisfação do usuário durante a interação. Esses métodos podem ser aplicados durante a criação ou remodelação das mesmas, visando melhorá-las tornando-as mais agradáveis ao usuário. Através de avaliações de usabilidade pode-se verificar a facilidade de uso da interface. Uma interface é considerada fácil de usar quando atende as recomendações de



usabilidade e é indutiva para seus usuários proporcionando satisfação durante a navegação.

Este estudo se propõe a fazer duas avaliações de usabilidade, a avaliação heurística proposta por Nielsen e uma avaliação empírica, para avaliar a usabilidade da Biblioteca Digital Brasileira de Teses e Dissertações mantida pelo Instituto Brasileiro de Ciência e Tecnologia, BDTD/IBICT. A escolha pelo estudo dessa interface se deve ao fato dela integrar, em um único portal, os sistemas de informação de teses e dissertações existentes nas instituições de ensino e pesquisa brasileiras, além de estimular a publicação desses documentos em meio eletrônico.

2 REVISAO DE LITERATURA

Nesta breve revisão de literatura serão apontados trabalhos que abordam a importância da avaliação de usabilidade para as interfaces de bibliotecas digitais.

Os estudos sobre as Bibliotecas Digitais foram iniciados na década de 1980, e, atualmente, o assunto conta com ampla literatura, tanto nacional quanto internacional. Apesar de não contar com um conceito preciso e respaldado por toda a área de CI, vários autores buscaram definir as bibliotecas digitais, conforme se pode observar na literatura. Nonato *et al.* (2008, p. 126) comentam a falta de consenso entre os autores para a definição do termo, ressaltando como ponto de concordância das definições o tipo de acervo dessas bibliotecas que é disponibilizado em formato digital independentemente da forma como foi criado.

Os requisitos de usabilidade, quando aplicados na construção de *sites* de bibliotecas digitais, possibilitam inúmeros benefícios aos seus usuários. No entanto, poucos trabalhos na área de Ciência da Informação buscaram explorar as contribuições que os estudos de usabilidade podem trazer ao desenvolvimento deste tipo de biblioteca.

Blandford e Buchanan (2003) discutem como trazer os usuários para a biblioteca digital e proporcionar a eles experiências satisfatórias e produtivas. Os autores defendem a idéia de que as bibliotecas digitais são instrumentos poderosos de veículo da informação e que só terão um impacto proporcional ao seu investimento quando houver aceitação por uma comunidade ampla de potenciais utilizadores. Os autores apontam a falta de consenso sobre quais seriam os principais critérios para avaliar a usabilidade das bibliotecas digitais. Segundo eles, alguns pontos específicos não podem deixar de ser observados nas interfaces de bibliotecas digitais, são eles: familiarização do usuário com a estrutura da biblioteca (tipo de conteúdo, mecanismos de busca) e consistência de estilo de interação. É ressaltada ainda a importância de um profissional bibliotecário para apoiar os usuários dessas bibliotecas e garantir sua eficácia.

Jeng (2005) em seu artigo “*what is usability in the context of the digital library and how can it be measured?*” analisou como a usabilidade foi definida no contexto da biblioteca digital e ainda propôs um modelo de avaliação para verificar a eficácia, eficiência, satisfação e capacidade de aprendizado. O autor ressalta o rápido desenvolvimento desse tipo de biblioteca embora os meios



de avaliação não tenham se aprimorado no mesmo ritmo. Em seu trabalho, o autor considerou tanto a quantificação dos elementos de desempenho (tempo, taxa de precisão, medidas para completar tarefas) quanto os critérios subjetivos como a satisfação.

Fuhr et al (2007) acredita que a avaliação das bibliotecas digitais é importante para sua aceitação pelos usuários. O autor reconhece que geralmente as bibliotecas digitais são instrumentos de apoio para atividades em contextos específicos e que precisam ser avaliadas para determinar o quanto são eficazes. A principal razão de uso desse tipo de biblioteca se concentra em torno da informação e do acesso. Segundo o autor a pesquisa em usabilidade em bibliotecas digitais apontou para o que denomina síndrome: “*librarians know better*”, ele argumenta que os bibliotecários conhecem as necessidades que um sistema deve cobrir, enquanto os desenvolvedores estão ignorando as necessidades dos usuários primários. Conclui enfatizando que a avaliação das BDs é crucial para sua evolução e aceitação desses sistemas.

A preocupação com as interfaces das bibliotecas digitais existe, uma vez percebido que, se após dar um tratamento de qualidade as informações, elas forem inseridas em uma interface com baixa usabilidade os usuários não conseguirão recuperar e poderão até mesmo abandonar o sistema. Como citado acima, vários pesquisadores têm defendido a ideia de que é preciso aumentar cada vez mais a usabilidade das interfaces humano computador com o propósito de se permitir que os usuários consigam recuperar as informações que buscam de maneira eficaz, eficiente e confortável.

Kafure e Cunha (2006) enfatizam a importância de as ferramentas tecnológicas (FT) proporcionarem usabilidade aos usuários, para que possam utilizá-las confortavelmente para atingir seus objetivos. Os autores expõem ainda a importância de os projetistas conhecerem os usuários para definir os aspectos do sistema e definir, além das informações que estarão presentes, o modo como elas serão apresentadas na interface da ferramenta tecnológica.

Segundo Kafure (2004), a concepção de Interfaces Humano-Computador para sistemas de informação em bibliotecas progrediu bastante, mas ainda é grande a discrepância entre o modelo mental do usuário e a representação computacional.

Alguns trabalhos já publicados relacionados diretamente à avaliação de usabilidade em bibliotecas digitais são: Marcelino (2008), cuja pesquisa avaliou o *site* da biblioteca do Instituto Nacional de Pesquisas Espaciais (INPE) a partir do usuário e da usabilidade. Dias (2008) recomenda testes de usabilidade para se obter um referencial das necessidades de usuários de bibliotecas virtuais e digitais. Nascimento (2006) estudou cenários, tendências de *websites* de bibliotecas universitárias e conceitos de usabilidade para, então, avaliar a usabilidade da Biblioteca Central da Universidade de Brasília, usando guia de avaliação heurística, entre outras técnicas, e verificou problemas de usabilidade quanto à falta de linearidade na navegação, duplicidade de *links* e uma estrutura que impede o usuário de executar suas tarefas em tempo hábil. Bohmerwald (2005) destaca a importância de os profissionais em Ciência da Informação se dedicarem aos estudos de usabilidade, para agregar conhecimentos dessa área, sejam sobre necessidade, uso ou recuperação de informação. Nascimento



e Amaral (2010, p. 99) apresentam em seu livro um estudo de caso sobre o *website* de uma biblioteca universitária, os autores concluem dizendo que os gestores e desenvolvedores de *websites* de bibliotecas devem fazer da usabilidade uma aliada que os auxilia a conceber conteúdos e estruturas de navegação adequadas que permitam desenvolver serviços para um campus universitário compostos de variados perfis. Observa-se que os trabalhos voltados para a avaliação de bibliotecas digitais utilizam na maioria das vezes testes com usuários e fazem apontamentos para estudos de usabilidade, (Fernandes, Insfran e Abrahão 2011).

3 BIBLIOTECAS DIGITAIS DE TESES E DISSERTAÇÕES

Apesar de não contar com um conceito preciso e respaldado por toda a área de CI, vários autores buscaram definir as bibliotecas digitais, conforme se pode observar na literatura. Nonato *et al.* (2008, p. 126) comentam a falta de consenso entre os autores para a definição do termo, ressaltando como ponto de concordância das definições o tipo de acervo dessas bibliotecas que é disponibilizado em formato digital independentemente da forma como foi criado. Essa dependência tecnológica e a ligação de tarefas através de interfaces exigem que aspectos como usabilidade e arquitetura da informação sejam considerados para que se alcancem objetivos tais como melhorar a produtividade dos usuários.

As BDTDs das instituições brasileiras estão em diferentes estágios de desenvolvimento. No entanto, as instituições mantêm o objetivo comum de facilitar o acesso e impactar suas pesquisas disponibilizando o conhecimento gerado. Aquelas que oferecem acesso a toda a produção de teses e dissertações, através de documentos completamente digitalizados em formato *pdf*, partes dos textos, referências e resumos autorizados pelos autores, favorecem seus usuários e facilitam a recuperação.

O IBICT mantém, em parceria com as Instituições brasileiras de Ensino Superior (IES), a BDTD que objetiva integrar, em um único portal, os sistemas de informação de teses e dissertações. A partir de uma base centralizada de metadados, as IES podem compartilhar sua produção, alimentando o repositório com seus metadados. Para participar é necessário que as bibliotecas participantes façam adoção do MTD-BR que é compatível com o padrão Dublin Core e o padrão ETD-MS da NDLTD, assim como implementação do protocolo OAI-PMH para exposição dos metadados. Esse protocolo funciona segundo padrões de interoperabilidade aliado a dois atores principais, são eles: o provedor de dados que administra o depósito de publicação, expondo metadados para a coleta automática e o provedor de serviços que fornece serviços de informação com base nos metadados coletados junto aos provedores de dados.

Como forma de incentivar a parceria entre essas bibliotecas, o IBICT fornece um *software* para as instituições montarem suas BDTDs e, para aquelas que preferem outros *softwares*, o IBICT pede que utilizem algum que mantém padrões de interoperabilidade, para que consigam fazer a varredura das bases de dados. Dessa forma, as instituições de ensino e pesquisa atuam como provedores de dados, enquanto o IBICT atua como um agregador disponibilizando-os para coleta por outros provedores.



Nesse processo, as instituições participantes criam suas bibliotecas digitais, enquanto a Biblioteca Digital de Brasileira de Teses e Dissertações faz a varredura dessas bases, indicando em seu portal o endereço eletrônico para os documentos. Cabe, no entanto, a preservação dos documentos fornecidos às instituições participantes.

Essa iniciativa de implantação de BDTDs nas instituições de ensino, apesar de trazer diversos benefícios aos pesquisadores, como o de facilitar a pesquisa, permitindo que ela seja feita em um único portal, ainda não foi acatada por todas as instituições de ensino superior do país. Mas a tendência é que, vendo os benefícios trazidos, as instituições venham a participar desse compartilhamento.

4 USABILIDADE EM BIBLIOTECAS DIGITAIS

A definição mais abrangente e aceita sobre o termo usabilidade é a dada pela ISO 9241-11 (1998), que o define como: “capacidade de um produto ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto de uso”.

Na Ciência da informação, a usabilidade aparece nos estudos que abordam a necessidade de informação do usuário, nos estudos do comportamento de busca e uso de informação, e na avaliação de sistemas, considerando o usuário como componente essencial no processo de interação. A Ciência da Informação vê a informação como algo construído pelos seres humanos através de processos cognitivos. Procura compreender como as pessoas constroem sentido, focando então no usuário, nas suas situações particulares de uso da informação e no que ocorre antes e depois de suas interações com o sistema.

Para que uma biblioteca digital proporcione ao usuário um uso eficiente de seus recursos, ela precisa ter uma boa interface, e para avaliar a qualidade da interface, um dos critérios utilizados é em relação à usabilidade. Essas bibliotecas devem optar por interfaces que melhorem a produtividade dos usuários.

A qualidade da interface de uma biblioteca digital é de suma importância para seu uso eficiente. De acordo com Bohmerwald (2005, p. 95) os critérios de usabilidade fornecem parâmetros para medir a eficiência da interface e revelam como se dá a interação entre usuário e sistema. Apesar de já ser reconhecida a importância da qualidade das interfaces de bibliotecas digitais, Diez (2009, p. 9) alerta para o número pequeno de estudos que investigam diretamente a influência do desenho dessas interfaces e a forma como se dá a interação com os usuários. Para prosseguir em seu desenvolvimento as BDs precisam contar com pesquisas que contemplem toda sua complexidade, variedade de usos e usuários, e que permitam evoluir de modo que assegurem sua aceitação por parte da comunidade a que se dirigem.

Em um sistema digital, o tratamento da informação dos documentos e a usabilidade precisam estar presentes, pois é inútil que as etapas do tratamento de descrição física, análise de assunto, classificação e representação, sejam efetuadas com qualidade, se elas não refletirem em uma interface



que atenda aos critérios de usabilidade. É preciso prover aos usuários possibilidades como diferentes tipos de buscas e páginas atraentes, que contenham informações objetivas e diretas, que correspondam ao seu modelo mental, evitando sobrecarga cognitiva.

A interface deve funcionar como um mediador entre a visualização e a representação de redes de informações, permitindo que o usuário siga as ligações entre os conceitos que estão sendo explorados.

5 METODOLOGIA

O universo de estudo escolhido para essa pesquisa foi a Biblioteca Digital Brasileira de Teses e Dissertações mantida pelo IBICT, portanto terá como amostra a interface dessa biblioteca. As avaliações foram feitas buscando identificar problemas de usabilidade da interface da BDTD mantida pelo IBICT. Os métodos escolhidos foram um método analítico e um empírico de avaliação de usabilidade, são eles: a Avaliação Heurística e o Teste com Usuário. Desta forma, a interface terá uma avaliação sob o ponto de vista de avaliadores e outro dos usuários.

Esta pesquisa é de natureza aplicada, pois objetiva gerar conhecimento para aplicação prática dirigida a um problema específico. Quanto aos objetivos, trata-se de uma pesquisa exploratória, pois busca familiarizar-se com o problema, torná-lo explícito e apontar possíveis soluções. Quanto à forma de abordagem do problema, propõe-se uma pesquisa qualitativa. Os dados serão analisados indutivamente, centrando-se na compreensão e explicação das heurísticas e demais recomendações de usabilidade citadas no referencial teórico. Ao final teremos uma avaliação sob o ponto de vista do avaliador e outra do usuário.

5.1 Desenvolvimento e etapas metodológicas da Avaliação Heurística

A avaliação foi estruturada tendo como base os seguintes documentos: DAUSW, APUSW E RAUSW, descritos abaixo. Esses documentos foram desenvolvidos pelo professor Clarindo Pádua (2010) do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais (UFMG), para a realização de avaliações heurísticas. DAUSW, Descrição de Avaliação de Usabilidade do *Software*, auxilia a estruturação da avaliação, através desse documento os avaliadores poderão se orientar durante a avaliação individual; APUSW, Análise de Problemas de Usabilidade do *Software*, identifica o problema encontrado, é utilizado durante a avaliação individual, seu uso facilita o registro dos problemas observados; RAUSW, Relatório de Avaliação de Usabilidade do *Software*, descreve as considerações finais da avaliação. Esse relatório objetiva reunir em um documento final os objetivos e resultados da avaliação. Esses documentos foram adaptados para atender ao objetivo dessa pesquisa. O uso desses relatórios estruturados facilita o registro das informações durante a avaliação e evita a perda de informações importantes ao final do processo.

Antecedendo a essas etapas seguiu-se a sugestão de Nielsen (2003), de que se utilizem entre



três e cinco avaliadores, para este tipo de avaliação foi criado um grupo de três avaliadores, constituído pelos seguintes componentes: um especialista em avaliações de usabilidade (designer, mestrando em ciência da informação); um analista com conhecimento sobre o domínio e tarefas do sistema a ser avaliado (bibliotecário, mestrando em ciência da informação); e, por último, um analista de sistemas (doutorando em ciência da informação), todos avaliadores com interesse em temáticas relacionadas com usabilidade e arquitetura da informação.

A avaliação heurística foi desenvolvida em três etapas, descritas a seguir. Na primeira etapa os avaliadores tomaram conhecimento do objetivo da avaliação, do objetivo da BDTD, do perfil dos usuários a serem atendidos pela interface e das heurísticas para avaliação. Esses esclarecimentos serviram para orientar as avaliações. Para esta avaliação não foi excluída nenhuma das heurísticas propostas por Nielsen, pois elas são gerais e observou-se que todas deveriam ser consideradas. Sendo elas: (1) Visibilidade e reconhecimento do estado ou contexto atual do sistema; (2) Compatibilidade com o mundo real; (3) Controle e liberdade do usuário; (4) Consistência e padrões; (5) Prevenção de erros; (6) Reconhecimento ao invés de memorização; (7) Flexibilidade e eficiência; (8) Projeto estético minimalista; (9) Diagnosticar e corrigir erros; (10) Ajuda e documentação.

Na segunda etapa, cada avaliador realizou a avaliação da interface individualmente, evitando que a influência dos demais avaliadores em suas observações. Nessa avaliação individual, os avaliadores tiveram uma visão geral do conteúdo e da arquitetura da BDTD. Eles puderam navegar livremente pela biblioteca, anotando as ocorrências de problemas de usabilidade, justificando cada problema, baseando-se nas heurísticas de usabilidade, além de registrar a localização e o grau de severidade de cada problema.

O grau de severidade dos problemas encontrados foi baseado nas recomendações de Nielsen e Loranger (2007). Os autores classificam o grau de severidade de um problema segundo a combinação de três fatores: (1) Frequência com que o problema ocorre: se é comum ou raramente experimentado. Se apenas um número pequeno de usuários for prejudicado, então esse será considerado um problema cosmético; (2) Impacto que esse problema terá sobre o usuário: de difícil ou fácil superação. Esse tipo de problema pode irritar os usuários, fazê-los perder horas de trabalho e até mesmo sair do sistema; e, por último, (3) Persistência do problema: verificar se ele aparecerá somente uma vez, e se os usuários aprenderão sobre o problema, conseguindo superá-lo ou se eles continuarão repetindo o erro. Alguns problemas de usabilidade não são persistentes, pois à medida que as pessoas o identificam, elas conseguem solucioná-lo, superando-o com facilidade no futuro. No entanto, algumas interfaces são tão confusas que causam desorientação frequente nos usuários, merecendo uma classificação de gravidade mais alta.

A combinação desses fatores resultou em uma escala para pontuar a gravidade do problema, que segundo Nielsen (2005), é a seguinte: 0 = corresponde a um problema cosmético: embora deva ser trabalhado, não é uma prioridade; 1 = corresponde a um problema pequeno, de baixa prioridade; 2 = corresponde a um grande obstáculo ao uso do sistema: deve ter prioridade alta para ser resolvido; 3



= corresponde a um problema grave, que impede o uso do sistema e deve ser imperativo consertá-lo.

Foi solicitada aos avaliadores a realização das seguintes tarefas específicas, que são consideradas importantes para a interface analisada: Fazer uma busca simples; Cadastrar usuário; Fazer uma busca avançada; Consultar o glossário; Montar uma estante virtual; Criar acesso direto às Teses e Dissertações. O resultado da análise individual, realizada na segunda etapa, foi uma lista com todos os problemas de usabilidade encontrados pelos avaliadores.

A terceira e última etapa metodológica teve como subsídio o resultado das avaliações dos três avaliadores. Nela ocorreu a consolidação dos resultados da avaliação segundo consenso dos avaliadores e foram feitos apontamentos para melhorar a usabilidade da interface. Foram eliminadas as considerações repetidas e se buscou o consenso sobre as considerações contraditórias. Essa fase se desenvolveu considerando-se os problemas apontados na avaliação, as heurísticas avaliadas e as recomendações sobre usabilidade com base no referencial teórico.

Ao fim dessa etapa, obteve-se uma lista de todos os problemas identificados pelo grupo de avaliação, com a identificação do grau de severidade. Nessa lista foram explicitados os problemas críticos da interface e aqueles puramente cosméticos, que, embora influenciem sutilmente o funcionamento do sistema, têm solução também, recomendada.

5.2 Desenvolvimento e etapas metodológicas para o Teste Empírico de Usabilidade

Para a preparação da avaliação foi tomado como base a descrição do método de avaliação empírica de usabilidade proposto por Prates e Barbosa (2003). A coleta de dados foi feita através de anotação em tempo real, uso de gravação de vídeo e questionário. Para o teste desta pesquisa foi utilizado o laboratório para testes de usabilidade do Instituto de Ciências Exatas (ICEX-UFMG) que segue todas as recomendações para esse tipo de teste.

Segundo Prates e Barbosa (2003) a preparação dos testes em laboratório passam pelas seguintes etapas:

(1) determinação do objetivo da avaliação; (2) seleção das tarefas; (3) seleção dos usuários participantes; (4) geração de material para o teste; (5) execução de teste piloto; (6) execução dos testes em laboratório; (7) análise dos dados coletados.

Para determinar o objetivo da avaliação o avaliador deve estabelecer quais os critérios prioritários. Neste trabalho foi avaliado se a interface oferece facilidades de uso para atender a satisfação de seus usuários.

De acordo com os objetivos da avaliação, o avaliador deve selecionar as tarefas a serem executadas durante o teste, que poderão fornecer indicadores sobre os objetivos a serem atendidos. Neste caso, deve-se tomar cuidado com o tempo exigido para cada tarefa do teste, pois as tarefas devem ser descritas na ordem em que aparecerão para os usuários. Não é recomendado que o tempo de execução de cada tarefa exceda a vinte minutos, e o tempo total do teste ultrapasse a uma hora,



para que os participantes não se sintam cansados. As tarefas foram entregues aos participantes uma a uma, assim que resolvidas era entregue a seguinte sucessivamente até o cumprimento de todas.

A seleção dos usuários participantes deve ser feita de acordo com o perfil daqueles que utilizarão o sistema de fato. Em alguns casos o sistema se destina a usuários com perfis variados, é então recomendado que os testes sejam feitos com usuários que atendam a cada perfil. Para o experimento proposto nessa pesquisa foram selecionados 9 participantes. Foram escolhidos 3 participantes dentro das três grandes áreas do conhecimento, ciências exatas, humanas e biológicas que estivessem cursando ou que já concluíram pós-graduação. Não foram separados perfis de usuários novatos e experientes, por se tratar de pesquisadores, alunos de pós-graduação, pressupondo que todos já tivessem conhecimento do domínio avaliado.

O material necessário para execução do teste inclui: *script* de apresentação e explicação do processo de teste aos usuários, formulário de consentimento do usuário, texto com a descrição da tarefa para garantir que todos os participantes tenham as mesmas informações e um questionário para coleta de opinião do usuário sobre o sistema. Conforme anexo. Após assegurado que a interface continha suporte necessário para a realização das tarefas previstas, foi entregue aos usuários uma descrição das cinco tarefas que eles deveriam executar. As tarefas foram entregues uma a uma, à medida que uma tarefa era executada durante o teste era entregue a próxima.

O teste piloto foi feito com os mesmos avaliadores que fizeram a avaliação heurística, para assegurar que o experimento seria possível para ser aplicado aos participantes do teste empírico, antecipando algum problema que pudesse ocorrer durante a avaliação. Como não foi constatado nenhum problema, não foi preciso fazer alterações.

A execução dos testes em laboratório exige um ambiente adequado para os testes e atenção para as questões éticas envolvidas neste tipo de teste. É necessário dar condições aos participantes para que eles se sintam à vontade e ajam de forma natural nesse ambiente controlado.

Foi explicado ao participante seu papel no teste, deixando claro que não era a sua capacidade que estava sendo testada e sim a do aplicativo. Além disto, deve-se também dizer quais são seus direitos enquanto participante.

Alguns pontos foram esclarecidos para garantir aos usuários o entendimento sobre o teste:

- Os objetivos do estudo e exatamente como deveria ser a colaboração de cada participante.
- As expectativas do anonimato, enfatizando que os dados particulares identificados durante o teste não são divulgados.
- Possibilidade de interrupção do teste, caso o participante desejasse.
- Garantia de que se algum relato do usuário fosse divulgado, seria com seu consentimento e seu anonimato será preservado.

Os direitos do participante estão baseados na Portaria nº 196/96¹ que é um dispositivo legal

1 <http://www.sbh.com.br/pdf/etica/PesqSeresHumanos.pdf>



que trata de pesquisas realizadas com seres humanos. Apesar de concebida na área de saúde, regulamenta os testes que envolvem seres humanos em outras áreas por não existir uma lei específica para área. Segundo essa portaria, o usuário deve ser informado sobre seus direitos, condições de realização do teste, como gravações das telas, áudio, vídeo, observações. Para garantir tal procedimento foi elaborado um termo de consentimento que garante aos usuários os seus direitos. Esse termo permite ao usuário acrescentar novas condições ao acordo e deve ser assinado por ele e pelo avaliador. Ao assinar esse termo o usuário se declara ciente de seus direitos e autoriza o uso de todos dados gerados.

Após a apresentação do laboratório para teste e assinatura do termo de consentimento entregou-se a lista de tarefas para os usuários. A partir deste momento os participantes foram observados e começou-se as anotações para sua posterior análise de dados. Cabe ao avaliador nesse momento fazer as devidas anotações e responder apenas perguntas que não se relacionem com aquilo que será observado para não comprometer o objetivo do teste.

Nesta fase do teste, o avaliador pode fazer perguntas que julgue importante aproveitando ao máximo a contribuição do participante. Ao perceber alguma dificuldade o avaliador pôde interferir com perguntas, tais como: O que você está tentando fazer? O que esperava ao clicar nesse botão? Precisa de alguma dica? Tem alguma dúvida? Cabe ainda ao avaliador explicar ao participante sobre como fazer alguma tarefa abortada por ele.

Após o teste o pesquisador responsável pôde colher a opinião pessoal do participante, sugestões, tirar dúvidas sobre as ações observadas pelo usuário através de questionário.

A análise dos dados coletados serviu para gerar o relatório do teste. Além dos dados, ainda têm-se as anotações das observações de cada avaliador e os questionários de satisfação preenchidos por cada participante. O relatório descreve o teste feito, os problemas encontrados e propostas de solução.

O pesquisador foi o responsável pela observação dos participantes e dos avaliadores e pela coleta de dados para posterior interpretação e identificação de problemas da interface do aplicativo. Os documentos (roteiro do avaliador para execução da avaliação, roteiro de tarefas para o avaliador, termo de consentimento, lista de tarefas dos participantes, questionário) utilizados durante a avaliação serviram de apoio aos avaliadores e também para garantir que os participantes tivessem as mesmas informações.

Ao fim dessas etapas foi possível identificar problemas específicos de usabilidade da interface, verificar se a interface é usável, bem compreendida e se os usuários se sentem satisfeitos ao utilizá-la.

6 ANÁLISE DOS RESULTADOS

Em seguida, é apresentado o resultado das avaliações propostas neste trabalho: a avaliação heurística de usabilidade (analítica) e a avaliação com usuários (empírica).



6.1 Avaliação Heurística de usabilidade da BDTD

Após a avaliação e as considerações de cada um dos avaliadores, procedeu-se a consolidação dos resultados. Ao final, os avaliadores enumeraram 15 problemas, os quais, segundo eles, poderiam comprometer a qualidade da usabilidade da interface. A fim de ilustrar os comentários acima, o quadro abaixo apresenta uma síntese do resultado da avaliação, demonstrando os problemas identificados, a heurística não atendida, a estimativa de seriedade e a sugestão de solução.

PROBLEMA IDENTIFICADO / HEURÍSTICA NÃO ATENDIDA	SUGESTÃO DE SOLUÇÃO
1 - Tamanho da fonte pequeno, / H8	Aumentar tamanho da fonte
2 - Rodapé confuso, ultrapassando os limites da tela. / H6	Definir um estilo para que o usuário possa identificar a existência de um <i>menu</i> com subitens atrás daquele texto.
3 - Opção cadastro encontra-se pouco visível. / H10	A opção cadastro deveria aparecer na tela de início e estar disponível em todas as páginas.
4 - Quanto aos critérios de busca, não fica claro se os critérios utilizados são conjunção ou disjunção. / H7	Fornecer flexibilidade ao usuário, de modo que possa usar conjunção ou disjunção na aplicação dos critérios na busca avançada. Ampliar os parâmetros de busca.
5 - Ausência de instrução sobre o uso dos operadores. / H10	Colocar uma dica (<i>hint</i>) sobre como utilizar os operadores de busca.
6 - Dificuldade para modificar uma busca. / H7	Criar opção de voltar à página de busca, exibindo os últimos parâmetros inseridos e permitindo a alteração destes parâmetros para uma nova busca.
7 - Falta de estilo padrão para texto, <i>links</i> e comandos, o que pode induzir o usuário ao erro / H4	É necessário que o <i>designer</i> estabeleça uma padronização visual para cada um dos elementos de interação (<i>links</i> , botões, <i>menus</i> , etc.).
8 - O sistema poderia ser mais flexível, oferecendo, por exemplo, a opção imprimir o resultado das pesquisas, nas telas de resultados. / H7	Poderia ser implementado um comando de impressão, para que, com maior rapidez, o usuário possa imprimir a relação de documentos encontrados.
9 - O sistema não oferece dicas para o preenchimento dos campos nas telas de procura. / H7	Colocar uma dica (<i>hint</i>) sobre como preencher os campos que admitem uma única forma de preenchimento.
10 - Mensagens com linguagem técnica. / H2	As mensagens de erro devem apresentar uma linguagem clara para o usuário, sem uso de códigos.
11 - Ausência de caixa de busca nas telas. / H1	Acrescentar caixa de busca.



12 – Na tela instituições parceiras, a barra de rolagem é extensa. / H1	Acrescentar caixa de busca; melhorar a forma de apresentação das instituições.
13 - Duplicidade de informações. / H8	A caixa “destaques” deve mostrar um conteúdo relacionado ao contexto do <i>menu</i> global, ou seja, suas informações devem estar distribuídas segundo a opção escolhida no menu global.
14 - Os <i>breadcrumbs</i> não informam o caminho exato percorrido pelo usuário para chegar à determinada página. / H1	Deve ser feito um estudo bem elaborado sobre ‘Arquitetura da Informação’ do site, a fim de corrigir esses problemas. O usuário deve saber em que ponto da navegação ele se encontra.
15 - Não há ajuda. A biblioteca possui algumas dicas de como guardar o site em “favoritos”, do navegador, mas não existe qualquer orientação ou dica sobre como utilizar o sistema./ H10	As páginas devem fornecer ajuda e documentação. Seria interessante que, em todas as páginas, houvesse uma caixa de ajuda.

A associação de uma heurística para cada problema foi feita segundo consenso dos avaliadores, no entanto, alguns problemas podem estar associados a outras heurísticas, tais como: Problema 3 (opção de cadastro pouco visível) poderia também estar associado a H6; Problema 4 (falta de clareza nos critérios de busca) associado também a H5; Problema 5 (falta de instrução sobre os operadores) associado também a H5; Problema 7 (falta de estilo dos *links* e comandos) também associado a H5; Problema 9 (dicas de preenchimento dos campos nas telas de busca) associado também a H5; Problema 10 (linguagem técnica) também associado a H9; Problema 11 (ausência de caixa de busca) associado a H7.

Após a análise dos problemas, pode-se concluir que as heurísticas mais desrespeitadas foram H7 (flexibilidade e eficiência de uso) e H10 (ajuda e documentação), sendo apontados quatro problemas referentes à flexibilidade e à eficiência de uso, além de outros três, relacionados à ajuda e à documentação. Sobre a H1 (visibilidade e reconhecimento do sistema) e H8 (projeto estético minimalista), foram identificados dois problemas em cada. Já em H2 (compatibilidade com o mundo real), H4 (consistência e padrões), H6 (reconhecimento do sistema) e H9 (diagnosticar e corrigir erros), detectou-se um problema em cada. Por fim, não foi atribuído diretamente nenhum problema à H3 (controle e liberdade do usuário) e à H5 (prevenção de erros).

Quanto à gravidade dos problemas encontrados, cinco deles foram considerados como pequenos problemas de usabilidade, que precisam ser resolvidos com baixa prioridade. Outros três foram denominados grandes problemas de usabilidade, que devem ser analisados com alta prioridade. Finalmente, sete dos problemas encontrados foram considerados catástrofes de usabilidade, que precisam ser solucionados imediatamente.

Além dos problemas enumerados e citados acima, a discussão dos avaliadores identificou alguns *bugs*, ou seja, problemas de implementação que comprometem a eficiência do sistema, não



sendo, entretanto, de usabilidade, propriamente. Apesar de a avaliação proposta se direcionar apenas à qualidade de uso da interface, esses problemas (codificação de caracteres e formulário de *login*, citando alguns exemplos) despertaram a atenção dos avaliadores, pois podem influenciar, diretamente, a qualidade da interface, prejudicando, assim, toda a experiência de uso.

Com relação à codificação de caracteres, há um problema com seu registro no banco de dados. Haja vista que algumas bases de dados não suportam caracteres latinos como “ç”, “~” e outros. Por exemplo, a palavra “Avaliação” foi guardada no banco como “Avaliação” e “AvaliaÃÃo”. Então, se for feita uma pesquisa por cada um desses termos, os resultados obtidos serão diferentes.

É interessante verificar que a pesquisa realizada com a utilização do termo “avaliação” retorna 22518 documentos. Já a pesquisa feita por meio do termo “avaliãã”, oferece 1865 documentos, conforme pode ser observado na figura abaixo. Nesse sentido, pode-se perceber que a revocação das buscas é bastante reduzida em função da negligência acerca da codificação de caracteres. Com efeito argumentativo, nota-se que a diferença entre o número de documentos recuperados é de 20653 documentos.

Outro problema encontrado refere-se ao formulário de “*login*”. Mesmo após a ativação da conta criada, não é possível fazer o *login*. Esse erro é difícil de qualificar, uma vez que se trata de um problema de implementação do banco de dados, que, apesar de não ser um problema com a interface em si, atrapalha bastante o usuário.

Cabe registrar, por fim, que o método avaliação heurística, embora não substitua os testes com usuários, permite uma avaliação global da interface e é eficaz para identificação dos problemas de usabilidade de um sistema. Constitui um método básico da engenharia de usabilidade relativamente fácil de ser aplicado, seu uso propicia aprendizado aos avaliadores e, principalmente, melhorias para a interface em análise. Convém destacar, ainda, que, apesar dos problemas de usabilidade do sistema, permite aos avaliadores identificar *bugs* do sistema.

6.2 Avaliação Empírica de usabilidade da BDTD

Os resultados do teste empírico serão apresentados dentro de cada uma das tarefas solicitadas aos participantes.

- **Tarefa 1**

Você pensa em começar uma pesquisa sobre: compositores eruditos brasileiros. Para isso, deseja saber se já existe algum documento sobre esse tema.

Tarefa: Verificar se existe alguma tese ou dissertação sobre esse tema. Caso encontre algum documento, abra-o para leitura.

Os principais problemas apontados pelos participantes durante o teste foram: (a) Tamanho da fonte dos textos pequena; (b) Alguns ícones que apontavam para o documento remetiam a outro *site*,



e não diretamente para o documento. Isso faz com que o usuário tenha de aprender a trabalhar na nova interface apresentada; (c) Ícones pouco representativos; (d) O texto dos *links* são pouco indutivos e não possuem estilo apropriado; (e) Na tela de resultados, aparece um dado indicando uma porcentagem à esquerda. Entretanto, não há informação sobre seu significado; (f) Excesso de ícones e comandos em cada resultado recuperado.

Os *links* para os documentos conduzem às bibliotecas digitais de suas instituições de origem. Esse fato causou surpresa em todos os participantes, que estranharam o aparecimento de uma nova tela com leiaute diferente daquele do *site* que estavam navegando. A título de explicação, esse projeto de avaliação não cobre cada uma dessas interfaces, limitando-se a avaliar a interface da Biblioteca Digital Brasileira de Teses e Dissertações mantida pelo IBICT. Além dos problemas evidenciados, foram percebidos outros que impedem a interação, como a falta de compatibilidade com alguns navegadores, tal qual observado com o navegador *Mozilla Firefox*, que é um dos mais utilizados.

Para a segunda tarefa apresentada, tinha-se a expectativa de que o participante utilizasse o recurso “procura avançada”. A tarefa solicitava ao participante que fizesse uma pesquisa envolvendo aspectos descritivos e temáticos. Para cumpri-la, o participante deveria fazer uma busca, filtrando por assunto (país e data). Dentre os participantes, somente três conseguiram cumprir a tarefa. De maneira a contribuir para o entendimento, abaixo, segue a maneira como a tarefa foi apresentada.

• Tarefa 2

Você está fazendo um trabalho sobre: ensino à distancia no Brasil. Para isto, precisa fazer uma revisão de literatura que busque teses e dissertações que tenham tratado esse tema entre 2005 e 2010. **Tarefa:** Buscar os documentos que tratam esse tema entre os anos de 2005 a 2010. Quantos documentos você conseguiu recuperar?

Os principais problemas apontados durante o teste foram: (a) O *link* para a tela de início conduz à página de resultados da pesquisa anterior, e não para a tela inicial do *site*; (b) Falta padronização na entrada de dados; os resultados para buscas com os termos “ensino à distância” e “ead” são diferentes; (c) Dúvidas quanto aos campos “país”, “grau” e “idioma”, os quais estavam dispostos no formato *combo-box* ou *menu-list*. Nesses campos, os usuários não estavam certos se, ao ignorar um dos campos, o sistema retornaria todos os itens da base de dados ou nenhum.

Na terceira tarefa, era esperado que os participantes conseguissem encontrar o formulário de cadastro e, feito isso, realizassem seus cadastros. Todos os participantes conseguiram cumprir essa tarefa, apesar de algumas dificuldades. A tarefa foi apresentada da seguinte maneira:

• Tarefa 3

Você encontrou vários documentos relevantes para sua pesquisa na Biblioteca Digital Brasileira de Teses e Dissertações. Agora, gostaria de ter acesso a outros recursos dessa biblioteca, como serviço de alerta e de estante digital. Para isso, você precisa efetuar seu cadastro na Biblioteca. **Tarefa:** Fazer seu cadastro

Os principais problemas identificados no desenvolvimento dessa tarefa foram: (a) Dificuldade



para localizar a tela de cadastro; (b) Mensagem confusa na tela de cadastro; (c) O estilo dos botões faz com que eles pareçam inativos; (d) Falta legenda nos campos obrigatórios; (e) Falta legenda indicativa de quantos dígitos a senha deve conter; (f) O tamanho da fonte é pequeno; (g) Os asteriscos demonstram campo obrigatório. No entanto, podem ser confundidos como indicação de nota de rodapé, que traz outra informação, sobre navegador *popups*.

A quarta tarefa pretendeu verificar se os participantes eram capazes de identificar quais são as instituições parceiras. Para isso, lhes foi pedido que localizassem a Universidade Federal do Triângulo Mineiro / UFTM dentro das instituições parceiras da biblioteca. Da maneira descrita abaixo, a tarefa foi apresentada.

• Tarefa 4

Você se interessa por um tema que está sendo pesquisado na Universidade Federal do Triângulo Mineiro, UFTM.

Tarefa: Verificar se a Universidade Federal do Triângulo Mineiro / UFTM é uma das instituições parceiras da Biblioteca Digital Brasileira de Teses e Dissertações.

As dificuldades encontradas foram: (a) Leiaute da tela confuso; (b) Falta de ordenação dos resultados e ausência de opção para ordená-los; (c) Inexistência de caixa de busca.

Na quinta tarefa, tinha-se como objetivo que os participantes localizassem, dentro do *menu* “indicadores” (“resumo geral”), as instituições que registraram maior número de defesas. A tarefa foi apresentada da seguinte maneira:

• Tarefa 5

Você é um pesquisador e gostaria de saber quais as 3 instituições que tiveram maior número de defesas. **Tarefa:** Verificar quais são essas instituições.

Os principais problemas apontados pelos participantes foram: (a) Rotulação dos *links* pouco explicativos; (b) Os participantes se sentiram confusos com a grande quantidade de *links*; (c) A tela foi considerada pequena, não havendo opção para abri-la em tela cheia (não permite ver *fullscreen*).

Ao terminarem de executar as tarefas, foi pedido aos participantes que preenchessem um questionário de satisfação de uso da interface. O questionário com oito perguntas, divididas em três blocos que buscavam colher a opinião dos usuários sobre o leiaute do sistema, a terminologia e a facilidade de aprendizado do mesmo. Para cada um dos tópicos avaliados, foi solicitado ao participante que atribuísse uma nota em uma escala entre 1 (pior nota) e 9 (melhor nota) ou N/A (não aplicável).

A primeira parte do questionário, bloco 1 (questões 1 a 5), buscou avaliar o leiaute do sistema. Nesta oportunidade, os participantes avaliaram as (1) imagens dos comandos, se eram indutivas e visíveis; (2) quanto às fontes utilizadas, verificaram se apresentavam tipo e tamanho legíveis; (3) a respeito das cores, se eram adequadas; (4) sobre a quantidade de informações, avaliou-se a pertinência



ou a falta dela; e (5) acerca da organização das informações, se eram lógicas ou ilógicas. A nota média dada a esses quesitos pelos participantes variou entre 6 e 7. A segunda parte do questionário, bloco 2 (questão 6) pretendeu avaliar a terminologia. Sendo assim, nesse tópico, foi pedido aos usuários que avaliassem a consistência das siglas, códigos e abreviações. Já na terceira parte, último bloco (questões 7 e 8) analisou-se a facilidade de aprendizado do sistema. Nessa circunstância, os participantes avaliaram (7) a facilidade para começar a utilização do sistema, bem como (8) o tempo necessário para aprender a utilizá-lo.

Com relação à média geral da satisfação de uso, foi possível perceber que superou as notas esperadas pelos avaliadores, uma vez que, durante a execução das tarefas, os participantes demonstraram muitas dificuldades e sugeriram diversas mudanças. Apesar de as notas dadas pelos participantes no questionário serem relativamente altas, tendo média entre 6 e 7, alguns deles não conseguiram cumprir todas as tarefas previstas, o que torna esses resultados um pouco contraditórios. Somente na tarefa 3 todos conseguiram cumprir o pedido.

Na medida em que os usuários se deparavam com os problemas, eles já cogitavam possíveis soluções para que a interação com a interface se tornasse mais agradável, fato que tornou o teste bastante proveitoso. Dessa maneira, os participantes sugeriram que as seguintes mudanças fossem realizadas na interface: (a) Ao clicar em “início”, espera-se um retorno à página inicial, sem necessidade de limpar a procura; (b) O *menu* geral deve possuir uma arquitetura da informação que represente bem o conteúdo da biblioteca; (c) Manter a opção cadastro e *login* em todas as telas; (d) Renomear os *links* e deixá-los em destaque; (e) Alterar a aparência dos botões da tela de cadastro, para que não pareçam inativos; (f) Acrescentar uma caixa de busca na tela “instituições parceiras”; (g) Texto da sessão “indicadores” poderia estar dentro da sessão “BDTD”; (h) A sessão “Destques” deveria aparecer na página inicial.

O teste empírico contribuiu para que os participantes propiciou oportunidades para apresentarem suas dificuldades, ressaltarem os pontos que, a partir de suas análises, deveriam ser melhorados e sugerirem mudanças que, se implementadas, podem otimizar a usabilidade da interface, facilitando a interação.

7 CONSIDERAÇÕES FINAIS

A partir das análises realizadas, chega-se à constatação de que os casos de uso mais comuns e importantes em uma biblioteca digital são aqueles relacionados à interação entre os usuários e seu conteúdo, ou seja, o acesso aos documentos. Com a aplicação das avaliações empírica e analítica foi possível mapear os problemas de usabilidade da BDTD mantida pelo IBICT e, ainda, foram feitas várias propostas para solução dos problemas identificados.

Comparando os dois tipos de avaliação não se pode considerar que um método de avaliação se sobrepõe ao outro. De fato, nota-se que cada um deles se aplica melhor em diferentes contextos, podendo haver complementação entre eles.



Durante a avaliação heurística, os avaliadores receberam uma lista com algumas tarefas para cumprir. No entanto, era permitido que percorressem outras telas, identificando problemas em todo *site*. Na realidade, as tarefas serviram como norteador para a avaliação, mas os avaliadores não se limitaram a resolvê-las. Eles ultrapassaram essas tarefas com vistas à identificação de todos os problemas. O inverso disso ocorreu nos testes empíricos, quando os usuários focaram-se na resolução das tarefas propostas. Durante a execução das tarefas relativas a este teste, os participantes identificaram alguns problemas, sem observarem as outras telas. Nesse sentido, o resultado apresentado foi bastante específico para cada tarefa proposta. Diante disso, vale ressaltar a necessidade de maior cuidado no planejamento desse tipo de teste, pois, se as tarefas não refletirem as principais atividades a que o *site* se destina, a avaliação ficará comprometida e não abrangerá todas as funcionalidades que o sistema, potencialmente, pode oferecer aos seus usuários.

Como já mencionado, este trabalho se restringiu a avaliar o *site* da BDTD, mantida pelo IBICT. No entanto, outros fatores externos acabam se relacionando com a BDTD e podem causar frustração aos usuários, como, por exemplo, a inexistência de serviços anunciados no *site*² do IBICT, estante digital, serviço de alerta, gerenciador de repositórios e indicadores de produção e de pesquisa, como na BDTD. O anúncio dessas funcionalidades geram expectativas no usuário que essa biblioteca ainda não oferece ou que não estavam disponíveis durante o período em que foi avaliada.

Abstract: *The digital libraries aim to make available information to their users in a faster and easier ways. In these libraries, the interactions of users with the collections are carried out by means of their interfaces as a link between user and system. This study proposes a case study about the usability of the Brazilian Digital Library of Theses and Dissertations, which is maintained by the Brazilian Institute of Science and Technology (BDTD-IBICT) pointing out the contribution of usability principles to the improvement of interfaces of digital libraries; mapping the problems of usability of the interface chosen for this study; and proposing solutions for the problems that emerged. It is discussed how the assessments of usability can be applied in the evaluations of digital libraries in order to improve the human-computer interaction, enabling users a more efficient information retrieval. The interface was assessed by means of an analytical method (heuristic assessment) and an empirical method (test with users). In the heuristic assessment, the data were collected using three appropriate forms for each of the three stages of assessment. The use of these documents guided the evaluators, and avoided losses of information during the assessment. For the empirical assessment, we have used questionnaires and notes about the observations made during the assessment. The data collected in this qualitative research were assessed inductively by the evaluators. The theoretical and methodological basis of this research resorted to the literature about digital libraries, usability and interfaces. The results generated a list of problems related to the usability which interferes directly in the interaction of the user with the BDTD-IBICT.*

Keywords: Digital Library. Usability. Analytical evaluation of usability. Test with users.



REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro, 2002.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002. 24 p.

Brasil. Ministério da Saúde. Conselho Nacional de Saúde. Normas **de pesquisa envolvendo seres humanos**. Res. CNS 196/96. BioÈtica 1996; 4. Suppl:15-25. Disponível em: <<http://www.sbh.com.br/pdf/etica/PesqSeresHumanos.pdf>> Acesso em: 09 mar. 2010.

BLANDFORD, A. BUCHANAN, G. Usability of digital libraries: a source of creative tensions with technical developments. **IEEE-CS Technical Committee on Digital Libraries' on-line newsletter**. 2003. Disponível em: <<http://discovery.ucl.ac.uk/16648/1/16648.pdf>> Acesso em: 5 jul. 2011.

BOHMERWALD, Paula. Uma proposta metodológica para avaliação de bibliotecas digitais: usabilidade e comportamento de busca por informação na Biblioteca Digital da PUC/Minas. **Ciência da Informação**, Brasília, v. 34, n. 1, p. 95-103, 2005.

DIAS, Tania Mara. **Bibliotecas virtuais/digitais**: suas ferramentas e contribuições na educação superior – O caso da PUCPR. 2008. 59 f. Dissertação (Mestrado em Educação). Pontifícia Universidade Católica do Paraná, Paraná. 2008. Disponível em: <http://www.pergamum.pucpr.br/redepergamum/trabs/2008_02.pdf>. Acesso em: 03 fev. 2010.

DIEZ, Maria Luisa Alvite. Interfaces y funcionalidades de bibliotecas digitales. **Anales de documentacion**, n. 12, 2009. p. 7-23.

FERNANDES, Adrian. INSFRAN, Emilio. ABRAHÃO, Silvia. Usability evaluation methods for the web: A systematic mapping study. **Information and Software technology**. v. 53. p. 789-817. 2011. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0950584911000607>

FUHR, Norbert. et al. Evaluation of digital libraries. **Int J Digit Libr**. v.8, p. 21-38. 2007. Disponível em: <http://link.periodicos.capes.gov.br/ez27.periodicos.capes.gov.br/sfx?sid=metalib:METAPRESS_XML&id=doi:&genre=Journal%20Article&isbn=&issn=&date=2007&volume=8&issue=1&spage=&epage=&aulast=Fuhr&aufirst=Norbert&aunit=&title=International%20Journal%20on%20Digital%20Libraries&atitle=Evaluation%20of%20digital%20libraries&sici=&__service_type=&pid=%3Cmetalib_doc_number%3E062575079%3C/metalib_doc_number%3E%3Cmetalib_base_url%3Ehttp://svrperiodicos2.periodicos.capes.gov.br%3C/metalib_base>

ISO 9241-11 Part 11: **Guidance on usability**. In: Ergonomic requirements for office work with visual display terminals, 1998.

JENG, Judy. What is usability in the context of the digital library and how can it be measured? **Information technology and libraries**. Jun. 2005. Disponível em: <http://inventory.pbworks.com/f/Usability_DL.pdf> Acesso em: 5 jul. 2011.

KAFURE, Ivette. **Usabilidade da imagem na recuperação da informação no catálogo público de acesso em linha**. 2004a. 311 f. Tese (Doutorado em Ciência da Informação). Universidade de Brasília, Brasília. 2004a.



KAFURE, Ivette; CUNHA, Murilo Bastos. Usabilidade de ferramentas tecnológicas para acesso à informação. **Revista ABC**, Florianópolis, v.11, p. 273-282, 2006. Disponível em: <<http://revista.acbsc.org.br/index.php/racb/article/view/483/619>>. Acesso em: 01 fev. 2010.

MARCELINO, S. C. **Estudo de usuários e usabilidade de sites de bibliotecas especializadas: o caso da Biblioteca On-line do Instituto Nacional de Pesquisas Espaciais (INPE)**. 2008. 184 f. Dissertação (Mestrado em Semiótica, Tecnologias de Informação e Educação). Universidade Braz Cubas (UBC), Mogi das Cruzes. 2008.

NASCIMENTO, José Antônio Machado do. **Usabilidade no contexto de gestores, desenvolvedores e usuários do website da Biblioteca Central da Universidade de Brasília**. Brasília, 2006. 215 f. Dissertação (Mestrado em Ciência da Informação). Universidade de Brasília. Disponível em: <http://bdtd.bce.unb.br/tesedsimplicado/tde_busca/arquivo.php?codArquivo=276> Acesso em: 04 fev. 2010.

NASCIMENTO, José A. M. do; AMARAL, Sueli A. do. **Avaliação de usabilidade na internet**. Brasília: Thesaurus, 2010. 142p.

NIELSEN, Jakob. 2003. **Use it**. Disponível em <<http://www.useit.com/alertbox/20030825.html>>. Acesso em: 18 Jan. 2010.

NIELSEN, Jakob. **Heuristic Evaluation**, 2005. Disponível em: <<http://www.useit.com/papers/heuristic/>>. Acesso em: 18 Jan. 2010.

NIELSEN, Jakob. **Severity ratings for usability problems**. Useit.com, 2005. Disponível em: <<http://www.useit.com/papers/heuristic/severityrating.html>> Acesso em: 17 mar. 2010.

NIELSEN, Jakob; LORANGER, Hoa.; FURMANKIEWICZ, Edson. **Usabilidade na Web: projetando websites com qualidade**. Rio de Janeiro: Elsevier, Campus, 2007. 406 p.

NONATO, R. dos S.; Borges, G. S. B.; Maculan, B. C.; Lima, G. A. B. de O. Arquitetura da informação em bibliotecas digitais: uma abordagem da Ciência da Informação e da Biblioteconomia. **Informação & Informação**, v. 13, n. 2, p. 125-141, 2008. Acesso em: <http://dlib.libh.uoc.grwww.sun.com/offers/docs/digital_libraries.pdfwww.uel.br/revistas/uel/index.php/informacao/article/viewFile/1812/1682> Acesso em: 10 fev. 2008.

PADUA, Clarindo Isaías Pereira da Silva e. **Teste de usabilidade com usuários**. Material didático. Disponível em: <<http://homepages.dcc.ufmg.br/~clarindo/arquivos/disciplinas/eu/material/transparencias/topicos/12-2-teste-com-usuario.pdf>> Acesso em: 03 nov. 2010.

PADUA, Clarindo Isaías Pereira da Silva e. **Avaliação heurística**. Disciplina: Engenharia de usabilidade. Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. 2009. Disponível em: <<http://homepages.dcc.ufmg.br/~clarindo/arquivos/disciplinas/eu/material/>> Acesso em: 03 nov. 2010.

PRATES, Raquel Oliveira.; BARBOSA, S. D. J. Avaliação de interfaces de usuários: conceitos e métodos. In: Juan Manuel Adán Coello; Sandra C. P. Ferraz Fabbri. (Org.). **Jornada de Atualização em Informática (JAI), XXIII Congresso da Sociedade Brasileira de Computação**. Campinas: SBC, 2003, v. 2, p. 245-293. Disponível em: <http://www.inf.puc-rio.br/~inf1403/docs/JAI2003_PratesBarbosa_avaliacao.pdf> Acesso em: 6 out. 2010.



ARQUITETURA DA INFORMAÇÃO: REPRESENTAÇÃO DA INFORMAÇÃO DE PRONTUÁRIO ELETRÔNICO DO PACIENTE

*Virginia Bentes Pinto, Hamilton Rodrigues Tabosa,
Silvana Ap. B. Gregorio Vidotti*

Resumo: Com olhar na estética e na funcionalidade, o termo arquitetura aparece, ressignificado, por volta do final do século XX, cunhado por Wurman (1975), não mais para esse domínio de conhecimento, porém, na perspectiva de negar a entropia do mundo informacional, antes mesmo do surgimento da Web. A Arquitetura da Informação vem ao encontro das atividades tradicionais da pragmática biblioteconômica e, mais recentemente, da epistemologia da Ciência da Informação e da Ciência da Computação, refletindo não somente aspectos estéticos da categorização da informação no ciberespaço, mas também visando a estruturação e a representação da informação objetivando favorecer o acesso a essa informação com qualidade. Nessa perspectiva, este trabalho traz a seguinte questão-problema: que compreensão se deve ter da Arquitetura da Informação como ferramenta tecnológica para a representação da informação nos prontuários eletrônicos do paciente? Objetiva-se com essa pesquisa compreender a Arquitetura da Informação como ferramenta tecnológica para a representação informacional e na sua aplicabilidade ao ambiente do Serviço de Arquivos Médicos e Estatística (SAME) da Universidade Federal do Ceará visando favorecer uma melhor comunicação intra e entre a equipe de saúde, pacientes e pesquisadores. Trata-se de um estudo exploratório, apoiado na pesquisa bibliográfica, a fim de se compreender esse conceito e sua aplicabilidade. O lócus do estudo empírico é o SAME, tendo sido analisados os prontuários eletrônicos do Paciente (PEPs). Os resultados demonstram, inicialmente, que o conceito de Arquitetura da Informação precisa ser apropriado pela área. Esse conceito é entendido como uma tecnologia importante para a representação da informação dos PEPs, tanto no que diz respeito a sua estrutura física e lógica, como também para o sítio do SAME na *Web*, neste caso incluem-se os elementos de recuperação/busca, navegação, usabilidade e acessibilidade. No entanto, em se tratando da espacialidade, do armazenamento, do acesso e da preservação das informações do PEP, há que se levar em conta as questões relativas à legislação, ética e bioética.

Palavras-chave: Arquitetura da Informação. Prontuário Eletrônico do Paciente.



1 INTRODUÇÃO

“As estruturas de informação influenciam interações no mundo da mesma forma que as estruturas dos edifícios estimulam ou limitam as interações sociais”. (WURMAN, 1996).

Já não é mais novidade que o desenvolvimento científico e tecnológico atinge não apenas o ambiente industrial de transformação da matéria-prima em produtos maquinais de última geração, muito pelo contrário, afeta igualmente outras esferas, como é o caso das organizações que prestam serviços, a exemplo dos espaços ou unidades de documentação. Esse desenvolvimento tem proporcionado, sobremaneira, o aparecimento de tecnologias, ferramentas ou alternativas de produção, armazenamento, tratamento, organização e recuperação de informação. Entre essas tecnologias, destacam-se a Internet, que surgiu na década de 1960 e, mais recentemente, o sistema *Web*, em 1989, que viabiliza avanço cada vez mais acelerado da produção de conhecimentos tanto em suporte tradicional (analógico) como no meio digital. Consequentemente, percebe-se que nesse novo ambiente, a informação não para de crescer, independentemente de campos de saberes, viabilizando a integração de texto verbal e não verbal em único documento multimidiático. Em termos de informação, as possibilidades ofertadas por essas tecnologias são inimagináveis, indo desde a produção até o acesso em todos os territórios informacionais, mais especificamente no ciberespaço, que é um ambiente desterritorializado.

No contexto da saúde não poderia ser diferente, além da exponencialidade na produção de documentos tradicionais (livros e artigos científicos), o crescimento de outros tipos de documentos igualmente se verifica, destacando-se os prontuários do paciente. Na concepção de Bentes Pinto (2005), esses documentos registram todas as informações referentes ao estado de saúde da pessoa doente, “agregam em único documento midiático desde as narrativas da anamnese dos pacientes até as imagens capturadas pelos dispositivos de ponta, que fotografam nosso corpo fazendo-o transparente.” O Prontuário do Paciente é definido no Artigo 1º. da Resolução de nº 1.638/2002, do Conselho Federal de Medicina (CFM), como sendo um documento único constituído pelo

conjunto de informações, sinais e imagens registradas, geradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente e a assistência a ele prestada, de caráter legal, sigiloso e científico, utilizado para possibilitar a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo (BRASIL. CFM, 2002, p. 124).

O prontuário do paciente, antes em versão impressa, já se encontra em edição eletrônica e digital, sendo denominado de Prontuário Eletrônico do Paciente (PEP). Isso, certamente, provoca outro olhar por parte dos gestores dos Serviços de Arquivos Médicos e Estatísticas (SAME), local



onde são armazenados os prontuários analógicos e cujo acesso (a quem de direito) se efetiva ainda de forma manual, seguindo-se uma arquitetura informacional de organização física, tanto no que diz respeito ao *layout* das estantes ou gôndolas quanto à cronologia de disposição desses documentos, racionalidade organizacional que, certamente, dificulta o acesso a uma informação específica para resolver um problema pontual.

Com a chegada do PEP, as alternativas para viabilizar a organização e o acesso racional a esses documentos são inúmeras, sendo uma delas a aplicabilidade da Arquitetura da Informação, conceito que foi cunhado pelo arquiteto Richard Saul Wurman na década de 1970, e “[...] diz respeito à organização de padrões inerentes aos dados de modo a transformar o que é complexo em algo mais claro por meio da criação de estruturas ou mapas informacionais que viabilizem o alcance do conhecimento” (WURMAN, 1975, p. 25).

Embora entendamos que a compreensão da Arquitetura da Informação como uma tecnologia que pode ser aplicada ao contexto do prontuário do paciente, para a representação desses documentos, é preciso ter em conta que tal aplicabilidade demanda inúmeros desafios, pelas características *sui generis* desses documentos. Entretanto, tendo apoio sobre as experiências biblioteconômicas na representação de informação de coleções de livros e periódicos, em categorias de textos verbais e não verbais, podemos usufruir dessas experiências a fim de criar arquiteturas de informação para PEPs no SAME. Portanto, é nesse panorama que esta pesquisa se insere, tendo a seguinte questão-problema: Que compreensão se deve ter da Arquitetura da Informação como ferramenta tecnológica para a representação dos PEP?

A pesquisa aqui proposta tem como objetivos: compreender a metáfora Arquitetura da Informação como ferramenta tecnológica para a representação dos PEPs e sua aplicabilidade ao ambiente do SAME, visando favorecer o acesso à informação e a comunicação intra e entre a equipe de saúde, pacientes e pesquisadores.

Esclarecemos que a preocupação teórico-empírica desta pesquisa diz respeito à compreensão da arquitetura e sua aplicabilidade (ideias) aos espaços informacionais do prontuário do paciente e, por extensão, do SAME. Eis aí o porquê desta comunicação, que se constitui em uma pesquisa bibliográfica e vem ao encontro dos estudos que estamos realizando desde o ano de 2003 (PQ), 2005 (pós-doc), 2007 (universal), 2009 (PQ), 2010 (universal).

2 ARQUITETURA DA INFORMAÇÃO: UMA TECNOLOGIA DE REPRESENTAÇÃO, ROTULAGEM E ORGANIZAÇÃO DE INFORMAÇÃO

As tecnologias de produção, representação e organização do conhecimento estão presentes na vida do ser humano desde a Antiguidade Clássica, em que os desenhos eram registrados nas grutas ou murais seguindo uma ordem lógica de importância da mensagem que pudesse ser comunicada e entendida pelas gerações futuras. A invenção da escrita, do alfabeto e da imprensa aos poucos veio contribuir para o aumento e a reprodução acelerada do conhecimento e da informação. Esse



fato demandou olhares mais acurados para essas tecnologias, visando à organização informacional dos documentos produzidos, na perspectiva de favorecer o acesso à informação. Porém, nunca tinha sido visto um crescimento da produção de conhecimentos ter alcançado índices altíssimos e, conseqüentemente, as dificuldades de acesso à informação também foram se fazendo presentes. Tendo ciência dessa nova realidade, com as iniciativas de representação das coisas e objetos do mundo, novas propostas de organização da informação surgiram, destacando-se as classificações bibliográficas, com abordagens alternativas para categorizar informações no ambiente de bibliotecas, principalmente. Esse ambiente, mesmo que seja visto por alguns como sendo estático, movimentase pela etnografia dos profissionais da informação que, com suas representações reais e simbólicas, estabelecem estratégias que possam ir ao encontro das representações (igualmente simbólicas e reais) dos usuários, de modo que eles tenham possibilidades de acesso à informação. Realmente, esses profissionais, mesmo sem saber, já estavam trabalhando com Arquitetura da Informação, embora o termo ainda não existisse. Assim, efetivava-se Arquitetura da Informação, das representações mais gerais até as mais específicas, com as classes paternas apoiando sua prole de relação semântica de modo ontológico, como se faz na Arquitetura, onde os alicerces e as estruturas constituem os grandes pilares sobre os quais uma casa ou edifício se sustentam, para que paredes, janelas, teto, entre outros elementos, possam se apoiar em uma sintonia morfossintático-semântica, possibilitando uma estética harmoniosa e funcional. Além disso, também propunham uma Arquitetura da Informação espacial, por meio da distribuição dos mobiliários de modo que fosse possível aos usuários o deslocamento semântico nesses espaços.

Todo esse esforço foi muito bem visto durante séculos e ainda na contemporaneidade tem sua importância, contudo, devido ao aparecimento das Tecnologias Digitais de Informação e de Comunicação (TDIC), que contribuem para a produção de documentos e espaços multimidiáticos, novos olhares se voltam para a representação da informação, que possibilitem a neutralização do paradigma do excesso de informação *versus* acesso à informação, a fim de que os “andarilhos” desse novo território possam se encontrar. Nesse território, aquelas tradicionais tecnologias de tratamento e organização da informação, assim como de espacialidade do mobiliário, já não são mais tão eficazes. Precisam ser revistas e propostas novas alternativas ou ressignificações, tanto do ponto de vista estético, funcional como também de comunicação. É, pois, neste contexto que ressurge o conceito de arquiteto da informação (e por extensão, de Arquitetura da Informação), concebido por Richard Saul Wurman, como uma metáfora para compreender a organização física e espacial da informação, antes mesmo da popularização da *Internet* e do surgimento da *Web*. O conceito de arquiteto da informação que foi proposto por Wurman na década de 1960 aplica-se perfeitamente ao bibliotecário, como profissional preocupado com a coleta, tratamento, organização, disseminação e recuperação da informação. Acontece que o não conhecimento do fazer biblioteconômico impede que esse profissional seja percebido, nessa visão, como sendo um arquiteto da informação.



Não podemos deixar de perceber que, ao se falar da metáfora Arquitetura da Informação, há que se ter em mente a complexidade dos dois conceitos, arquitetura e informação, que demandam certos esclarecimentos. Etimologicamente, o termo “arquitetura” é oriundo de duas palavras gregas “arché”, que significa “primeiro” ou “principal”, e tékton, que possui o significado de “construção”. Informação, por sua vez, origina-se de três palavras (in-form-ação) “em cuja semântica estão embutidas a subjetividade, a forma e a dinamicidade” (informação verbal)¹. Assim, embora existam inúmeros conceitos para a palavra informação, ainda assim, nenhum vai se apresentar de forma linear e concreta. “Então, como fazer uma mistura com essas cinco palavras (arché + tékton + in + form + ação) de modo que resulte em um conceito para se pensar em organização da informação de maneira a apresentar não apenas uma estética” (informação verbal)¹ porém, conforme bem diz Wurman, (2005, p. 23) que traga “clareza ao que é complexo”, como por exemplo, os prontuários do paciente? “Quer dizer, qual é o sentido desse mixordismo (archétéktoninformação) como uma ferramenta que venha facilitar o processo de comunicação entre os sistemas informacionais e seus usuários, como no caso dos SAMEs e dos PEPs?” (informação verbal)¹

Richard Saul Wurman utiliza a metáfora Arquitetura da Informação para chamar atenção daqueles que trabalham a informação, tratando-a, organizando-a, disseminando-a, para torná-la acessível, inteligível e compreensível para o ser humano. Em sua proposta, o autor, mesmo sem intenção, ressignifica essa pragmática que a Biblioteconomia faz desde suas origens. No entanto, a popularização do conceito Arquitetura da Informação foi promovida pelos bibliotecários Louis Rosenfeld e Peter Morville (1998), como sendo a “arte e a ciência de organizar informações para auxiliar os indivíduos a satisfazerem as suas necessidades informacionais”. Por sua vez, Hagedorn (2000 apud SIQUEIRA, 2008, p. 31), diz que Arquitetura da Informação é “a arte e ciência da organização da informação para a satisfação de necessidades de informação, que envolve os processos de investigação, análise, desenho e implementação”. Já no entendimento de Macedo (2005, p. 132) a

‘Arquitetura da Informação’ é uma metodologia de ‘desenho’ que se aplica a qualquer ‘ambiente informacional’, sendo este compreendido como um espaço localizado em um ‘contexto’; constituído por ‘conteúdos’ em fluxo; que serve a uma comunidade de ‘usuários’. A finalidade da Arquitetura da Informação é, portanto, viabilizar o fluxo efetivo de informações por meio do desenho de ‘ambientes informacionais’.

A partir dessas reflexões, podemos perceber que os conceitos existentes sobre Arquitetura da Informação gravitam em torno de pelo menos quatro grandes eixos. O primeiro deles diz respeito ao estudo de *layouts* e interfaces, incluindo-se nesse primeiro eixo os estudos sobre usabilidade. O segundo eixo engloba os estudos que se relacionam com as questões de organização da informação e está muito ligado com o terceiro eixo, que abrange os conceitos que se relacionam com estudos sobre

¹ BENTES PINTO, V.; RAULINO NETO, F. H. R.; GIRÃO, I. P. T., MENESES, B. C. Discussão oriunda do Grupo de Estudos em Representação Cognitiva da Informação. Set. 2011. (<http://www.gpriufc.net>).



a representação da informação. Todos eles desembocam no quarto eixo, que é a comunicação, pois, de nada adianta se pensar em Arquitetura da Informação sem que se tenha a compreensão do processo de comunicação cuja estrela maior é o receptor ativo, portanto, somente com seu *feedback* é que a Arquitetura da Informação se efetivará, “de per si”.

Do mesmo modo que na organização da informação tradicional é necessário planejamento, não poderia ser diferente em relação à metáfora da Arquitetura da Informação, pois se trata de um esquema complexo que envolve questões como o que, para quem, para que, como, onde e quando organizar. Nas palavras de Rosenfeld e Morville (2006) ao se pensar Arquitetura da Informação, é preciso ter em mente três aspectos fundamentais: usuários, conteúdos e especificidades, conforme a Figura 1, tomada emprestada (literalmente) desses autores.

- a) Usuários: suas necessidades, aspectos culturais e comportamentais;
- b) Conteúdos: características do que será apresentado (objetivo, uso, volume, formato, estrutura, governança, dinamismo); e
- c) Especificidades do contexto de uso do sistema de informação: proposta de valor de *website*, cultura e política da organização, restrições tecnológicas, localização etc..

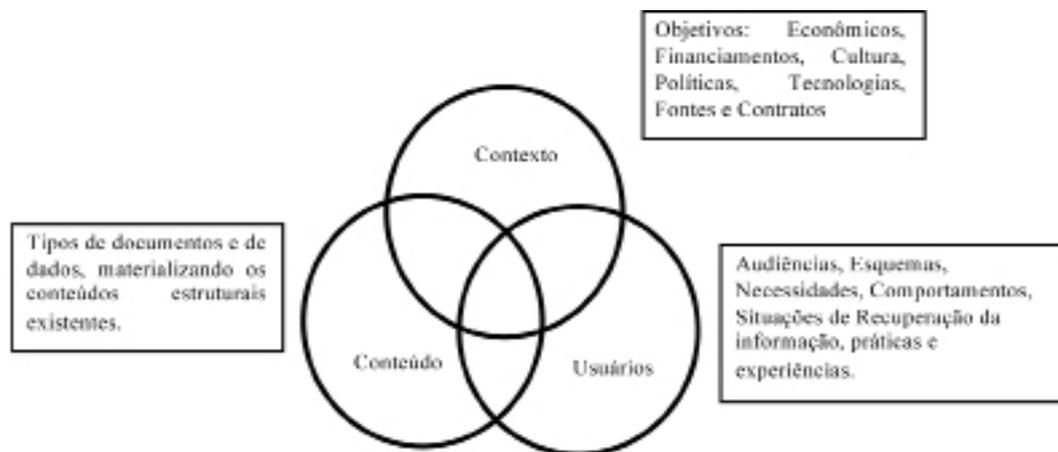


Figura 1- Os três círculos da Arquitetura da Informação
Fonte: MORVILLE, P; ROSENFELD, L. *Information Architecture for the World Wide Web*.
3. ed. Sebastopol: O'Reilly Media, 2006. p. 26

Acrescentamos a esses aspectos mais dois que consideramos de fundamental importância: o *feedback* e a avaliação, pois a Arquitetura da Informação, em nosso entendimento, configura-se como um processo de comunicação. Conforme Wurman (2005), o que faz a comunicação possível é a possibilidade de identificar, em seu receptor, aspectos não compreendidos por ele, verificar se o receptor tem algum interesse em compreendê-lo e descobrir a estrutura mais eficaz para transmitir a informação. Daí a necessidade de relacionar informação com conceitos que o usuário compreende, para que ele se sinta próximo do ambiente informacional, se veja nele e, conseqüentemente, o valorize. Para isso, o autor sugere as categorias que já há muito vêm sendo aplicadas pela Biblioteconomia, e



mais recentemente, pela Ciência da Informação, hierarquizadas em categorias, por ordem alfabética, lugar e tempo. Inclusive, o autor propõe a sigla LATCH: Local (mapas), Alfabeto, Tempo (períodos históricos e linhas de tempo), Categoria e Hierarquia. Corroborando esse autor, por outro lado, Evernden e Evernden (2003) argumentam que, no processo de criação de Arquitetura da Informação, é fundamental que sejam construídos mapas conceituais com ao menos oito facetas: categorias, entendimento, apresentação, evolução, conhecimento, responsabilidade, processo e metainformação, que são obtidas a partir do diagnóstico do uso de informação nas organizações.

Outro elemento da Arquitetura da Informação proposto por Morville e Rosenfeld (1998, p.105) é a ergonomia do sistema e apontam que as principais questões a serem consideradas em um projeto de Arquitetura da Informação são: nível de profundidade da pesquisa; tipo de informação a ser disponibilizada, tipo de informação a ser pesquisada e o tempo de resposta. Ainda segundo esses autores, há dois modos de representar a informação no território do ciberespaço: por grupos de palavras e por ícones. Além desses dois modos de representação da informação, podemos identificar outros, como por exemplo, o tempo ou a cronologia, proposição que não é diferente do que já vem sendo feito há séculos pelos bibliotecários. Esses autores também discutem os princípios de base da Arquitetura da Informação, quais sejam: sistemas de organização de conteúdos, sistemas de etiquetagem ou rotulagem (pontos de acesso), sistemas de navegação (estratégias de busca), sistemas de busca e redes semânticas ou terminológicas, por exemplo: tesauros, cabeçalhos de assuntos, vocabulário controlado, metadados.

- a) sistemas de organização de conteúdos: listam todos os conteúdos que farão parte do *layout*, de maneira a refletir o modo como eles possam ser agrupados, quer dizer: listar os conteúdos, categorizar as informações e estruturar os grupos de conteúdos;
- b) sistema de navegação: apresentam a espacialidade visual e interativa dos conteúdos, (elementos) de navegação (global, local, contextual, transversal), por exemplo, lugar das etiquetas, ícones a serem utilizados. Sua finalidade é ajudar o usuário a “navegar” nos conteúdos.
- c) sistema de busca: reflete o funcionamento das ferramentas de busca e recuperação da informação que precisa ser planejado, levando em consideração como os conteúdos foram representados, qual estratégia de busca o usuário deve propor, local do motor de busca, como mostrar os resultados, como organizá-los.
- d) Sistema de rotulagem/etiquetagem: enuncia o modo de representação dos conteúdos e as remissivas que são utilizadas na Arquitetura da Informação. Deve ser expresso em uma linguagem mais próxima da utilizada pelo usuário.

Esses princípios estabelecidos por Morville e Rosenfeld (2006) se configuram nos mesmos princípios dos sistemas de recuperação da informação, pois têm como finalidade básica a recuperação da informação com maior valor agregado.



A Arquitetura da Informação aplicada na criação de ambientes informacionais digitais da Web conduz melhor o usuário às informações desejadas e torna o acesso a elas mais eficaz e preciso, com o planejamento e a organização virtual da informação digital que facilita a construção e navegação neste sistema hipertextual.

De acordo com Camargo (2010, p. 52),

A Arquitetura da Informação é uma área do conhecimento que oferece uma base teórica para tratar aspectos informacionais, estruturais, navegacionais, funcionais e visuais de ambientes informacionais digitais, por meio de um conjunto de procedimentos metodológicos a fim de auxiliar no desenvolvimento e no aumento da usabilidade de tais ambientes e de seus conteúdos.

Ainda considerando os elementos da Arquitetura da Informação apresentados por Morville e Rosenfeld (2006) e os princípios de usabilidade e de acessibilidade, Vidotti, Cusin e Corradi (2008, p. 182) apresentam a seguinte definição:

A Arquitetura da Informação enfoca, organização de conteúdos informacionais e as formas de armazenamento e preservação (sistemas de organização), representação, descrição e classificação (sistema de rotulagem, metadados, tesouro, 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 hiperídia informacional digital.

A Arquitetura da Informação deve objetivar a criação de estruturas que priorizem a organização descritiva, temática, representacional, visual e navegacional de informações, em consonância com o conteúdo, o contexto e o usuário, adequando o dimensionamento e o direcionamento dos serviços e dos produtos informacionais aos usuários potenciais.

Qualquer processo deve possuir uma arquitetura eficiente, econômica, confiável, estável, cultural e esteticamente significativa e que se traduza em realização plena nos espaços da Web, tanto para produtores de conteúdos informacionais quanto para consumidores que usam a Web e deve considerar os aspectos aqui abordados, pois eles permitem melhores condições de usabilidade e de “encontrabilidade” de informação.

3 PRONTUÁRIO ELETRÔNICO DO PACIENTE

O prontuário do paciente, independentemente de ser analógico ou digital, é o documento de referência para o campo da saúde, cuja característica de leitura é hipertextual, mesmo que seu registro seja em formato impresso. Para Bentes Pinto (2005, p. 36), o prontuário do paciente é um documento que diz respeito à memória registrada da “história da pessoa doente, sendo, portanto, indispensável, para a comunicação intra e entre a equipe de saúde e o paciente, a continuidade, a segurança, a eficácia e a qualidade de seu tratamento, bem como da gestão das organizações hospitalares”. Com outras palavras, nesse documento são registradas “todas as informações relativas a uma pessoa doente



e as ações de cuidados e tratamentos a ela dispensados, a fim de que seja possível gerenciar o curso da patologia identificando os sintomas, causas e os remédios para solucioná-los”. São ricas fontes de informação e de comunicação, tanto para a equipe de saúde e os pacientes, como também para os estudos e pesquisas.

A gênese desses documentos remonta à Antiguidade Clássica, quando foram encontradas em cavernas anotações sobre as moléstias daquele período histórico. Conforme Calder ([19-], p. 50), localizaram-se, em vários sítios arqueológicos, crânios da Idade da Pedra que evidenciavam, por meio de orifícios, os tratamentos de moléstias como epilepsia e tumores. Outros registros pictográficos datados de 4.200 a.C. localizaram nas regiões da Suméria e Quixe, onde os textos referentes à Medicina vieram gravados nas “tabuinhas de *greda*”. Desse período identificaram-se, armazenados na Biblioteca de Nínive, 3000 ladrilhos de *greda*, sendo que, desse total, 800 continham informações referentes a enfermidades. São desse período também as chamadas listas de medicamentos, receitas e tratamento usados pelos sacerdotes-médicos da Babilônia: eis aí mais um embrião de prontuário do paciente. (CALDER, [19-], p. 57). Na Babilônia, as anotações informacionais sobre o domínio de conhecimentos da área de saúde estão no famoso Código de Hamurabi. Tais registros se referem, principalmente, aos custos terapêuticos e às penalidades (aos médicos não sacerdotes), desde que o enfermo não fosse escravo, naturalmente.

A redação desses documentos é feita de modo particular, pois seus “redatores” adotam expressões constituídas por nomes próprios, vocabulários do cotidiano dos pacientes, terminologia das diversas áreas de saúde, entre outras. Ainda concernente às características do prontuário, Friedman (2005) e Uzuner et al. (2008) ensinam que, na prática, o prontuário se caracterizava como uma coleção de documentos muito heterogêneos do ponto de vista do formato, do conteúdo e da semântica, conforme a seguir:

a) formato – texto semiestruturado ou não estruturado, como os atos de intervenção, consulta, anatomopatologia, imagens referentes a tomografia, ressonância magnética etc;

b) conteúdo – texto em linguagem natural que exprime fatos perceptíveis, decisões médicas de prevenção, diagnóstico e tratamento, que não seguem regularidades gramaticais ou lexicais e apresentam valores estruturais escondidos no texto (dados demográficos, medidas de análise como a taxa de colesterol, a taxa de glicose, a pressão arterial etc.); e

c) semântica – a inteligibilidade de um documento é dependente da história terapêutica do paciente contida em outros documentos do mesmo paciente.

Ante tal realidade, entende-se a organização de informações registradas em prontuário do paciente se faz necessária e urgente. Há de ser diferente, todavia, daquele efetuado em outras fontes, pois é peça *sui generis*, daí porque apenas tratar a informação nele contida não basta. É preciso, acima de tudo, que se tenha consciência de que esse fazer se constitui em um esquema de ações complexo visando à comunicação entre o usuário e o sistema de recuperação de informação. Sua efetivação somente acontecerá se a linguagem estiver em conformidade com a recepção dos participantes



desse processo, quer dizer, indexadores, equipe de saúde e paciente (receptor). Justamente por todas essas características é que o prontuário constitui documento complexo, e, naturalmente, pensar em Arquitetura da Informação como uma tecnologia para a organização das informações que ele porta, também perfaz um esquema intrincado, de modo que o acesso e a recuperação de informações neles contidos possam se processar com maior qualidade.

4 ESTUDO EMPÍRICO

Este estudo caracteriza-se como sendo de caráter exploratório, por dois aspectos, principalmente. Primeiramente porque a metáfora da Arquitetura da Informação ainda se constitui em um tema recente, inclusive no contexto da Ciência da Informação e, segundo, por não termos encontrado qualquer arquitetura referente ao prontuário eletrônico do paciente e ao ambiente do SAME. Por isso, esse assunto precisa ser mais estudado e discutido, para que os profissionais da área se apropriem desse conceito.

Em razão de estarmos estudando os prontuários do paciente e os SAMEs se enquanto fenômenos da realidade do mundo, optamos pela fenomenologia, como apoio para a compreensão desse fenômeno a fim de interpretá-los e nos apropriarmos dos conhecimentos registrados nesses documentos. Segundo Husserl (1980, p. 21), o método fenomenológico se baseia na percepção sensível daquilo que é dado à consciência, quer dizer, “o ato do conhecimento se fundamenta na percepção. [...] A palavra não está ligada a uma intuição isolada, mas pertence a uma multiplicidade infinita de intuições possíveis”. Ainda para esse filósofo

Toda a intuição em que se dá algo originariamente é um fundamento de direito do conhecimento; que tudo que nos brinda originariamente (por assim dizer, em sua realidade corpórea) na intuição, deve ser tomado simplesmente como se dá, mas também só dentro dos limites em que se dá. (HUSSERL, 1949 apud ZILLES, 1998, p. 127).

Assim, estudamos a estrutura física e lógica de 20 prontuários do paciente, bem como o espaço físico do SAME, a fim de podermos propor algumas ideias para a Arquitetura da Informação nesses espaços.

5 RESULTADOS E DISCUSSÕES

A partir do estudo dos prontuários, constatamos que, ao se pensar em Arquitetura da Informação para sítios relacionados a esses documentos e tomando por base o pensamento de Peter Morville e Louis Rosenfeld, temos que os usuários do PEP são todos aqueles que fazem parte da equipe de saúde, os pacientes, os pesquisadores e a equipe do SAME. Cada um com suas necessidades específicas de informação e com suas culturas e, portanto, com comportamentos diferentes. Porém, para que a concretização da Arquitetura da Informação seja possível, faz-



se necessário que inicialmente se faça um estudo de usuários de modo que venha refletir suas necessidades informacionais.

No que diz respeito aos conteúdos do sítio, percebemos que eles podem ser constituídos pelos componentes da estrutura física e lógica dos PEPs, sendo que a estrutura deve atentar para questões de ética e segurança. Já no que diz respeito a especificidades do contexto de uso do sistema de informação, é preciso levar em conta a cultura organizacional dos SAMES e suas políticas, além das restrições tecnológicas, como é o caso dos prontuários que foram digitalizados e que precisam ser atualizados, pois, a cada retorno do paciente, novas informações são agregadas ao seu prontuário e isso demanda tecnologias específicas e uma política direcionada para essas atualizações.

Em que concerne aos cinco princípios estabelecidos por Morville e Rosenfeld (2006), refletimos sobre o sistema de organização e fizemos um rol com os seguintes conteúdos: dados relativos ao paciente, anamnese, exames, dados relativos às contas médicas etc., de modo a categorizar as informações, estruturando-as por grupos de conteúdos. No que diz respeito ao sistema de navegação percebemos que, no caso dos PEPs, a navegação pode ser conforme as categorias estabelecidas anteriormente, porém, tendo-se o cuidado de seguir a legislação vigente sobre o acesso a esse documento.

O sistema de busca para os PEPs pode viabilizar certos tipos de informação e estabelecer regras de quem poderá ou não ter acesso e as estratégias de busca a serem adotadas pelos usuários, que dados poderão ser mostrados e como mostrá-los.

No que diz respeito ao sistema de rotulagem/etiquetagem, os conteúdos exibíveis precisam ser representados de modo claro e em uma linguagem mais próxima da utilizada pelo usuário. Também, deverão ser utilizados links de remissivas, por exemplo, nomes populares para nomes científicos das doenças e vice-versa.

Todos esses princípios são de fundamental importância a serem observados na Arquitetura da Informação de PEPs no contexto do SAME. No entanto, em se tratando da espacialidade e acesso às informações do PEP, há que se levar em conta todas as questões relativas à legislação, ética e bioética.

6 REFLEXÕES CONCLUSIVAS

A pesquisa, cujos resultados estão refletidos nesta comunicação, buscou compreender a metáfora Arquitetura da Informação como uma ferramenta tecnológica para a organização espacial dos PEPs e sua aplicabilidade ao ambiente do SAME, visando favorecer o acesso à informação e a comunicação intra e entre a equipe de saúde, pacientes e pesquisadores.

Por meio da revisão de literatura percebemos que, em realidade, a Arquitetura da Informação se constitui em um Sistema de Recuperação da Informação (SRI) no ambiente do ciberespaço e, portanto, visa oferecer alternativas que possam apresentar representações de informações de modo claro para os usuários. Logo, possui as unidades de entrada, processamento, saída e feedback.



Os estudos da literatura sobre o tema nos permitem inferir que, do mesmo modo como acontece na arquitetura espacial de prédios, onde forma, cores e textura buscam estabelecer harmonia e funcionalidade ao mundo e aos sujeitos do mundo, igualmente, a Arquitetura da Informação visa construir espaços informacionais semânticos visando não somente o compartilhamento de informação, mas também o acesso a esse bem que, como aqueles de primeira necessidade são fundamentais para as sociedades e para os sujeitos sociais. No caso dos PEPs, esse acesso é de importância ímpar.

Deduzimos ainda, que é possível se planejar Arquitetura da Informação no ambiente do SAME e no espaço informacional dos prontuários do paciente. Contudo, precisa ser apoiada em uma legislação específica com relação à integridade e ao sigilo das informações, previstos. Finalmente, a Arquitetura da Informação traz em sua linhagem a ideia de “cartografia” utilizada para a visualização de dados espaciais no território ciberespacial dos PEPs e do SAME.

INFORMATION ARCHITECTURE: REPRESENTATION OF INFORMATION ELECTRONIC PATIENT RECORD

Abstract: *The term architecture appears with new meaning, the aesthetic look and functionality, by the end of the twentieth century, coined by Wurman (1975), no further for that domain knowledge, however, the prospect of denying the entropy of the informational world, even before the emergence of the Web. The Information architecture meets the traditional activities of the pragmatic in the librarianship, more recently, the epistemology of information science and computer science, reflecting not only the aesthetic aspects of categorization of information in cyberspace, but also aiming at structuring and representation of information aimed at facilitating access to the data quality. From this perspective, this article aims to answer this question-problem: what must be understood if the information architecture as a technological tool for the representation of information in electronic patient records? The objective with this research to understand the information architecture as a tool for representing informational technology and its applicability to the environment of the Department of Archives and Medical Statistics (SAME), at Federal University of Ceará in order to foster better communication within and between the team health, patients and researchers. It is an exploratory study, supported in the literature, in order to understand this concept and its applicability. The locus of the empirical study is the SAME, with analyze of the electronic patient record (EPRs). The results show, at first, that the concept of information architecture needs to be appropriate for the area. This concept is understood as an important technology for the representation of information from EPRs, both as regards their physical and logical structure, but also for the SAME web site, in this case includes the elements of retrieval/search navigation, usability and accessibility. However, in the case of spatiality, storage, access and preservation of information of the EPR, it is necessary to consider the issues that involves a set of laws, ethics and bioethics.*

Keywords: Information Architecture. Electronic Patient Record.



REFERÊNCIAS

BENTES PINTO, V. Prontuário eletrônico do paciente: documento técnico de informação e comunicação do domínio da saúde. Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, n.21, 1º sem. 2006. Disponível em: <<http://redalyc.uaemex.mx/redalyc/pdf/147/14702104.pdf>> Acesso em; 10 ago. 2009

BENTES PINTO, V. Prontuário eletrônico do paciente: O documento de comunicação do domínio da saúde In: CONGRESSO LUSO-BRASILEIRO DE ESTUDOS JORNALÍSTICOS, 2, 2005, Porto-Portugal. Actas do II Congresso Luso-Brasileiro de Estudos Jornalísticos/IV Congresso Luso-Galego de Estudos Jornalísticos. Porto: , 2005.

BRASIL. CONSELHO FEDERAL DE MEDICINA. Resolução nº. 1.331 de 25 de Setembro de 1989. Diário Oficial [da] República Federativa do Brasil, Brasília, 25 set. 1989.

_____. Resolução nº. 1.638 de 10 de Julho de 2002. Diário Oficial [da] República Federativa do Brasil, Brasília, 10 jul. 2002. Seção 1, p. 124-5.

_____. Resolução nº 1.639 de 10 de Julho de 2002. Diário Oficial [da] República Federativa do Brasil, Brasília, 10 jul. 2002. Seção 1, p. 124-5. BRASIL. CONSELHO FEDERAL DE MEDICINA. Resolução nº. 1.331

_____. Resolução nº. 1.638 de 10 de Julho de 2002. Diário Oficial [da] República Federativa do Brasil, Brasília, 10 jul. 2002. Seção 1, p. 124-125.

CALDER, R. O Homem e a Medicina: mil anos de trevas. São Paulo: Hemus, [19-].

CAMARGO, L. S. A. Metodologia de desenvolvimento de ambientes científicos digitais baseada em princípios da Arquitetura da Informação. 2010. 322f. Tese (Doutorado em Ciência da Informação – Orientadora: Silvana A. B. G. Vidotti) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

EVERNDEN, R., EVERNDEN, E. Third-generation information architecture. Communications of the ACM, v. 46, no. 3, p. 95-98, mar. 2003.

FRIEDMAN, C. Semantic Text Parsing for Patient Records. New York: Springer, 2005.

HUSSERL, Edmund. Sexta investigação: elementos de uma elucidação fenomenológica do conhecimento. São Paulo: Abril Cultural, 1980, 184p.

MACEDO, F. L. O. Arquitetura da Informação: aspectos epistemológicos, científicos e práticos. 190 f. 2005. Dissertação (Mestrado)-Universidade de Brasília – Departamento de Ciências da Informação e Documentação, Brasília, 2005.

MORVILLE, P.; ROSENFELD, L. Information architecture for the world wide web. Sebastopol: O'Really, 1998.



_____. Information architecture for the world wide web. 3. ed. Sebastopol: O'Really, 2006.

OMS. Mieux communiquer pour mieux soigner. Disponível em: <<http://www.sante.gouv.fr/htm/actu/telesante/telesante1.htm>>. Acesso: 10 set. 2004.

SIQUEIRA, André Henrique de. A lógica e a linguagem como fundamentos da Arquitetura da Informação. 143. f. 2008. Dissertação (Mestrado)-Universidade de Brasília. Departamento de Ciências da Informação e Documentação, Brasília, 2008.

SABATTINI, R.M.E.. Preservando a confiabilidade médica na Internet. Revista Check-up. 2002. Disponível em: <http://www.nib.unicamp.br/papers/checkup-10.htm>. Acesso: 07 nov. 2004.

_____. Telemedicina: A Assistência à Distância. Revista Médico Repórter. Fev, 1999. Disponível em <http://www.nib.unicamp.br/papers/reporter-medico-03.htm>. Acesso: 20 dez. 2004.

SILVA, Patrícia Maria da. DIAS, Guilherme Ataíde. A Arquitetura da Informação centrada no usuário: estudo do website da biblioteca virtual em saúde (BVS). Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, n. 26, 2º sem. 2008.

UZUNER, O. et al. Identifying patient smoking status from medical discharge records, JAMIA, v. 15, n. 1, p. 14-24, jan. 2008.

VIDOTTI, S. A. B. G.; CUSIN, C. A.; CORRADI, J. A. M. Acessibilidade digital sob o prisma da Arquitetura da Informação. In: GUIMARÃES, J. A. C.; FUJITA, M. S. L. Ensino e pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar. São Paulo: Cultura Acadêmica, 2008.

WURMAN, R. S. Ansiedade da Informação. São Paulo: Cultura Editores Associados, 1991.

_____. Ansiedade da Informação 2. São Paulo: Cultura Editores Associados, 2005.

ZILLES, Urbano. Teoria do conhecimento. Porto Alegre: EDIPUCRS, 1998.



ANÁLISE DE RISCO: UM MÉTODO PARA A PRESERVAÇÃO DIGITAL

*Vildeane da Rocha Borba, Fanny do Couto Ribeiro, Marcos Galindo Lima,
Májury Karoline Fernandes de Oliveira Miranda*

Resumo: Através do universo de ativos digitais, o uso de novas tecnologias permitiu avanços consideráveis no tocante ao poder de armazenamento e transferência da informação, entretanto a garantia de acesso e interpretação dessas informações registradas em suportes digitais encontram-se ameaçadoramente desconhecida, o que torna a preservação digital um dos principais desafios do nosso tempo. Neste sentido, este artigo descreve o desenvolvimento de um método de análise de risco para a área da preservação digital e este estudo tomou como experimento a Biblioteca Digital de Teses e Dissertações da Universidade Federal de Pernambuco, com vistas a quantificar e qualificar riscos e ameaças concernentes à preservação da memória registradas em mídias digitais, assim como de monitorar o impacto e complexidade das ameaças relacionadas a salvaguarda do patrimônio digital. Para tanto, a investigação optou por uma pesquisa qualitativa e exploratória e por um referencial teórico de cunho bibliográfico (livros e periódicos em meio digital), que nos ajudou a descrever as práticas e as estratégias políticas de preservação digital e, também, ofereceu base teórico-metodológica necessária ao uso e aplicação de métodos de análise, que possibilitassem um constante controle de riscos. Os resultados apresentam que cerca de 75% representam riscos de maior probabilidade de ocorrência e que implicam em danos mais graves ao estoque dessas informações, sendo que cerca de 80% destes estão classificados como “catastrófico” e também “Muito freqüente”. Apoiados nesses dados foi possível comprovar a hipótese levantada nesse trabalho, de que não existe uma cultura de preservação digital para a produção científica no Brasil.

Palavras-chave: Preservação Digital. Análise de Risco. Memória. Tecnologias de Informação.

1 INTRODUÇÃO

Esta pesquisa¹ desdobra-se das atividades do projeto *Patrimônio Digital Ameaçado*² realizado no âmbito do grupo de pesquisa *Memória e Sociedade*, projeto que tem por objetivo investigar as práticas laborativas de preservação digital nos Repositórios Institucionais (RIs) desenvolvidos por instituições públicas de ensino e pesquisa de âmbito nacional. A pesquisa procurou refletir teoricamente sobre as noções de memória, experimentando métodos que tornem possível novas formas de administração de recursos e conteúdos em meio digital no campo da memória, com foco na preservação e no acesso.

1 Resultado parcial de dissertação de mestrado desenvolvida no PPGCI-UFPE.

2 Projeto de pesquisa ligado à temática da memória e preservação da informação digital. O projeto iniciou em 2008, sendo aprovado pela Câmara de Pesquisa da UFPE. Neste período produziu cinco dissertações e dois projetos de iniciação científica. Em 2010 foi aprovado pelo Edital MCT/CNPq 14/2010 Universal - CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO.



Ocorre que não obstante a preciosa contribuição da tecnologia da informação no campo do registro e da comunicação da informação científica, a perecibilidade e a rápida obsolescência desses registros e de sua tecnologia têm levado a perda permanente de considerável volume de documentos resultantes de pesquisa no Brasil. Esse fenômeno é o agente motivador da investigação que resultou no presente artigo.

A investigação optou pela pesquisa qualitativa e exploratória e por um referencial teórico que nos ajudou a descrever as características da população das instituições e serviços de informação no nível das práticas e das estratégias políticas de preservação digital. A motivação para pesquisar a temática surgiu da percepção dos desafios que envolvem a preservação do patrimônio informacional custodiados por instituições públicas de ensino e pesquisa de âmbito nacional, entre os quais se destacam a ausência de infra-estrutura adequada e pessoal especializado para a preservação de seus acervos, além da ausência de política dessas instituições.

No desenvolvimento da pesquisa, percebemos que a construção de um mapa-diagnóstico das práticas de preservação digital no ambiente dessas respectivas instituições e serviços de informação, pedia a aplicação de métodos que nos permitissem enxergar para além da mera identificação de estratégias de conservação. Era necessário abrir a ferida e expor o risco a que os estoques de informação científica estão expostos de modo que, diante das ameaças a consciência do perigo emergisse, gerando assim políticas, marcos legais e outros instrumentos de controle e preservação aplicados aos acervos digitais.

Neste sentido, este artigo descreve o desenvolvimento de um método de análise de risco para a área da preservação digital e este estudo tomou como experimento a Biblioteca Digital de Teses e Dissertações da Universidade Federal de Pernambuco, com vistas a quantificar e qualificar riscos e ameaças concernentes à preservação da memória registradas em mídias digitais, assim como de monitorar o impacto e complexidade das ameaças relacionadas a salvaguarda do patrimônio digital.

2 PRESERVAÇÃO DIGITAL

“[...] a memória não é sequer capaz de conservar todas nossas próprias idéias, e é muito importante que possamos conservar as de outros homens.”

(ORTEGA Y GASSET, 2006, p.16)

O olhar do filósofo espanhol Ortega y Gasset (2006), nos ajuda a compreender a necessidade humana de registrar suas ações cotidianas ao longo da história, como forma de se manter viva sua memória e transmiti-las as gerações futuras. Esta necessidade de registro e comunicação da memória de um povo, como legado às gerações ulteriores, tem provocado ao longo da história da humanidade a criação de inúmeros suportes, instrumentos e técnicas desenvolvidas pelo homem a fim de conservar a memória e resgatar o passado às futuras gerações.

A preocupação com a preservação da memória parece ter nascido com o desenvolvimento da



escrita e dos primeiros registros da produção do conhecimento. Esta prática permitiu às sociedades o acúmulo seguro de conhecimentos, bem como a sua transmissão as novas gerações.

Os meios de comunicação informacional foram emergindo ao longo do tempo, passando por vários suportes de informação como a pedra lascada, o pergaminho, voz humana, o papel, o computador e, a Internet. Esta materialização da informação em suportes é necessária à sua comunicação assíncrona em tempos e espaços diferentes. Para Kuramoto (2007, informação verbal)³, os meios de comunicação dominantes estão mudando da forma impressa para a forma digital, recorrendo a dispositivos/plataformas que constituem o(s) sistemas tecnológicos de informação (STI).

Atualmente, vive-se uma época marcada por grandes transformações na organização, tratamento e acesso à informação, onde os avanços das tecnologias digitais e o uso intensivo de tecnologias de informação surgem como instrumentos que além de propiciar um novo sistema de comunicação, possibilitam a criação de diversos suportes informacionais.

Esse momento caracteriza uma sociedade cada vez mais baseada na cooperação e convergência de tecnologias, onde a expansão das informações registradas em meio digital prossegue à grande velocidade. Portanto, falar dos novos meios de comunicação é falar das diferentes formas de comunicação tornadas possíveis pela convergência entre a informática e os novos registros do conhecimento, sendo a revolução digital a principal responsável pelas transformações ocorridas no modo como a sociedade registra e comunica sua memória.

Não obstante os avanços permitidos pelos avanços tecnológicos no tocante ao poder de armazenamento e transferência, a durabilidade das informações registradas em suportes digitais permanece ameaçadoramente desconhecida. Ou seja, o advento das novas tecnologias como suporte da informação passa por um processo constante de adaptação para atender as necessidades de armazenagem e “[...] apesar da sua reprodutibilidade tornar a informação digital teoricamente invulnerável aos estragos do tempo, os suportes no qual elas estão armazenadas estão longe de ser eternos.” (ROTHENBERG, 1999, p. 2, tradução nossa).

Devido às novas tecnologias, é possível afirmar que a informação digital não se encontra tão permanente quanto se credibilizava. Sendo assim, apesar dos diversos benefícios trazidos pelos novos meios de armazenamento, a preservação dos conteúdos em formatos digitais surge como um dos principais desafios do nosso tempo, visto que a fragilidade intrínseca nesses meios é a principal responsável por colocar em risco a sua longevidade. Neste mesmo sentido, Sayão (2010) reafirma

A preservação digital, enquanto um conjunto de atividades voltadas para garantir o acesso aos conteúdos digitais por longo prazo, é, ao mesmo tempo, um desafio técnico e organizacional que se desenrola permanentemente no tempo e no espaço [...]. As ameaças que cercam os objetos digitais são engendradas pela sua própria condição física, não fixada em suportes e fortemente dependente de contextos tecnológicos específicos e fugazes. (SAYÃO, 2010, p. 7)

³ Trata-se de informação fornecida por Hélio Kuramoto em conferência no XXII Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação – CBBDD, em Brasília, em julho de 2007.



É neste ambiente renovado pelas tecnologias, que a preservação digital surge como ferramenta capaz de assegurar que os meios digitais permaneçam acessíveis às gerações futuras, através de um conjunto de processos e atividades que garante armazenamento de longo prazo, sustentando acesso e interpretação da informação digital. E é com esse objetivo de garantir o pronto acesso aos recursos de informação, que a prática de preservar registros informacionais às novas gerações tem atraído cada vez mais esforços nos espaços destinados à memória, sendo considerada um dos grandes desafios do século XXI.

Nesse mesmo sentido, Ferreira (2006) define a preservação digital como

[...] a capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação. (FERREIRA, 2006, p.20)

Diante dos aspectos levantados, é possível entender a razão da crítica de Innarelli (2007, 2011), ao afirmar que a Sociedade da Informação nos traz uma facilidade imensa na geração de dados, informações e documentos, o mesmo acontece com a perda destas informações, pois a humanidade ainda não tem prática e nem experiência para a memória digital.

O pensamento de Innarelli é corroborado por novas reflexões sobre os aspectos sociais da preservação digital, assim como novos questionamentos quanto ao destino da humanidade em detrimento da perda de tudo o que foi produzido e colocado em bases digitais. A esse respeito, Rossi (2010) defende que a memória tem algo a ver não só com o passado, mas também com a identidade e, assim, com a própria persistência no futuro. Estas reflexões também são defendidas por Borba (2009, p.13), em resposta a esses questionamentos e as suas conseqüências em longo prazo para a história e a identidade de uma nação, quando a autora destaca que “a preservação digital é um ato de responsabilidade social com aqueles que ainda não nasceram, não reconhecem fronteiras de tempo e de espaço”.

A perda das informações registradas em mídias digitais permite o estabelecimento de uma lacuna histórica e cultural no desenvolvimento de uma sociedade, fato percebido por Rossi (2010, p.35) ao afirmar que “ressurgir de um passado que foi apagado é muito mais difícil que lembrar coisas esquecidas [...]”. Ainda nesse enfoque, diante da “inundação” das novas Tecnologias da Informação e Comunicação, Innarelli (2007, 2011) e Balbi (2010) nos alerta para o fato de que a história, a economia e cultura nacional de uma sociedade correm um sério risco de perder registros fundamentais para sua evolução e preservação.

Os documentos digitais, cada vez mais presentes no cotidiano das instituições e do próprio ser humano, são perdidos com a mesma facilidade que são gerados. Diante do exposto, é possível perceber que a tecnologia por si só não soluciona todos esses problemas, pelo contrário, na busca por preservar a informação em sua forma original surge paradoxalmente o desejo de acessá-la através de métodos cada vez mais novos e eficientes possíveis, ou seja, através da melhor tecnologia disponível. Então,



apesar do volume de informações registradas em mídias digitais tenha aumentado consideravelmente nas últimas décadas, essa situação, assinalada por Thomaz e Soares (2004), nos remete a problemática central da preservação digital, que é caracterizada pela instabilidade dos suportes responsáveis pelo transporte e armazenamento dessas informações.

Esse novo cenário traz consigo novos problemas para a gestão da documentação digital, os quais dependem diretamente da interferência humana e de políticas de preservação digital para serem solucionados. Segundo a ASSOCIATION FOR LIBRARY COLLECTIONS & TECHNICAL SERVICES (2007, tradução nossa), a preservação digital compreende políticas, estratégias e ações para garantir, a longo prazo, a reprodução precisa e assim o acesso a conteúdos “*nato-digitais*” ou aos convertidos para o formato digital, independentemente dos desafios de uma falha de mídia, corrupção de arquivos, e as mudanças tecnológicas.

Diante dessas abordagens, a construção de políticas e estratégias de preservação digital surge como dispendioso esforço por parte dos profissionais envolvidos com a gestão de documentos digitais nos diversos segmentos que compõem os serviços de informação, demandando desses gestores a capacidade de garantir sustentabilidade, legitimidade e fortalecimento de ações com base em uma proposta para a preservação do patrimônio digital de uma instituição.

Muitas estratégias e políticas de informação estão sendo concebidas para ajudar a gerir a preservação de materiais digitais, porém nenhuma delas é apropriada para todos os tipos de instituições ou dados. Sendo assim, sua escolha vai depender do dado, da natureza do material e quais aspectos devem ser mantidos, podendo variar de uma instituição para outra. Essas escolhas demandam por essência uma tomada de decisão racional de “o que preservar” e “como preservar”, atividades que demandam motivações intensas para a sua realização. Para tanto, essas decisões necessitam tomar por base algum tipo de previsão, visto que a incerteza constitui um fator complicador dos processos decisórios, assim como a ausência desse levantamento informacional como peça inicial dos processos de decisão.

Nos diversos setores do patrimônio informacional os fatores responsáveis pelo risco de segurança nas atividades de preservação digital ainda permanecem desconhecidos, o que resulta na inevitável sujeição de seus planejamentos e projetos a diversas incertezas, sejam elas relacionadas a aspectos do tipo: operacionais; geofísicos; políticos; entre outros.

Ao reconhecer a importante contribuição que os métodos estatísticos podem oferecer aos processos de preservação do patrimônio digital, é possível fazer uma avaliação do grau de incerteza existente nesses processos, e assim estabelecer uma estimativa do risco envolvido no cenário dessas instituições. Dessa, urge estimar com maior precisão as alterações sofridas por esses suportes digitais ao longo do tempo, o que a torna ações preventivas e curativas.

Apesar dos grandes avanços alcançados pelas pesquisas de preservação digital, observam-se, no entanto, que os mesmos são operados simploriamente por técnicas dirigidas à elaboração e aplicação eficiente de estratégias metodológicas, técnicas pouco preventivas e impossibilitadas de



diagnosticarem os riscos que envolvem os serviços de informação.

Assim, entende-se que trabalhar com a preservação digital é lidar com constante controle de riscos. O reconhecimento de que perdas são inevitáveis dá relevância aos modos de evitá-las. Essa idéia é corroborada por Galindo (em fase de elaboração)⁴, quando afirma que “um sistema bem protegido não é aquele melhor equipado – em todos os recursos – mas aquele que possui regras mais claras, eficientes e produtivas para proteção de seus ativos.” Dessa forma, quanto mais *padrões* e *boas práticas* forem debatidos e adotados no desenvolvimento e implementação de práticas de preservação digital por uma gestão de recursos digitais, maior probabilidade haverá da preservação e sustentabilidade do patrimônio digital.

3 GESTÃO E ANÁLISE DE RISCO

“Organizações de todos os tipos e tamanhos enfrentam influências e fatores internos e externos que tornam incerto se e quando elas atingirão seus objetivos. O efeito que essa incerteza tem sobre os objetivos da organização é chamado de ‘risco’”.

(ABNT ISO 31000:2009, 2009, p. 4)

Nesse cenário, a gestão e a análise de risco surgem como “atividades coordenadas para dirigir e controlar uma organização no que se refere ao risco” (ABNT ISO GUIA 73:2009 apud ABNT ISO 31000:2009, 2009, p.7) e que, ao se constituírem em um conjunto de técnicas e métodos, permitem detectar, analisar, quantificar, qualificar, além de justificar a prioridade e importância dos controles dos riscos e ameaças que comprometem a continuidade dos processos objetivados por uma organização.

Esse processo de gestão oferece subsídios para o controle dos riscos ao possibilitar o estabelecimento de prioridades e orientação nas tomadas de decisão de uma organização, baseando-se em estimativas científica que detectem a magnitude de impactos futuros sobre os seus objetivos.

Na tentativa de identificar de forma preventiva as situações que colocam a organização à exposição de diversos eventos ou perigos, a gestão e análise de riscos pode ser aplicada a toda uma organização, em suas várias áreas e níveis. Entretanto, esses tipos de análises têm significados e peças diferentes e são variadamente aplicadas em diferentes disciplinas como engenharia, estatística, economia, medicina e também no campo investigativo das ciencias sociais, o que torna sua aplicação, por vezes, confusa, visto que suas terminologias e noções são comumente transferidas de uma disciplina para outra sem as devidas modificações. (GANOULIS, 2009, tradução nossa).

Outro agravante no desenvolvimento e aplicação eficiente de técnicas de gestão e análise de risco é que atualmente diferentes sociedades tem desenvolvido suas próprias modalidades e percepções para interagir com a problemática das incertezas e assim gerenciar incidentes imprevistos no seu desenvolvimento e progresso.

Em ciência é comum o uso de termos que possuem significado diferente daqueles empregados no cotidiano. O risco pode ser considerado um exemplo desses termos, visto que seu conceito é, hoje

⁴ Artigo intitulado “O Dilema do Pharmakon”, de autoria de Marcos Galindo. No prelo.



em dia, amplamente utilizado em várias áreas do conhecimento, o que não nos permite uma única forma de conceituação e avaliação para ele. Portanto, considera-se neste estudo que, conceitualmente, o risco compreende a possibilidade de efeitos adversos, como perda e danos causados por exposição a um perigo. Definição que, segundo (GANOULIS, 2009, tradução nossa), traz consigo dois elementos essenciais para a descrição de risco: a gravidade do perigo e a suscetibilidade do sistema para sustentar o perigo, que são estudados aqui como a consequência de um cenário acidental e a sua frequência de ocorrência.

Partindo dos objetivos dessa pesquisa de quantificar a magnitude dos perigos que envolvem o patrimônio informacional registrado em mídias digitais, busca-se nesse estudo o aprofundamento do primeiro passo dessa gestão, que compreende a identificação de risco, de modo que seja possível identificar as diferentes situações em que as incertezas podem gerar risco de fracasso, além de possibilitar a geração de análises e quantificação desses riscos.

4 ANÁLISE PRELIMINAR DE RISCO E A PRESERVAÇÃO DIGITAL

O que se tem dito e escrito sobre preservação (e conservação, ou restauro...) enquadra-se naturalmente num paradigma tradicional e em crise, a que temos dado o nome de patrimonialista e custodial. É a prevalência do documento sobre a informação e do suporte sobre o conteúdo, encarando o objeto físico como um bem que é preciso proteger (conservar, restaurar) para que o seu valor patrimonial não se perca nem a possibilidade de manuseá-lo (sentido a sua textura, o seu cheiro...) se esgote. (PINTO, 2009, p.7)

As afirmações de Pinto (2009) remetem a necessidade de novas abordagens a ações interdisciplinares no tratamento dos desafios ocorridos com essa “nova era digital”, a qual, por viver novos problemas antes não sentidos, demanda respostas que já não são possíveis no quadro do paradigma tradicional, definido aqui como patrimonialista e custodial.

Diante desses aspectos, este estudo propõe uma nova perspectiva e um novo olhar para os métodos de preservação. Para isso, partindo do entendimento que lidar com a preservação digital é lidar com constante controle de riscos, busca-se o desenvolvimento de novos métodos de monitoramento. Esta ferramenta possibilita um constante controle de riscos, além de antecipar e diagnosticar os diversos perigos que incidem nos serviços de informação, como também mostra-se capaz de identificar os acidentes significativos ocorridos durante o processo de guarda e acesso da memória em meio digital.

Essa pesquisa propõe a utilização de análises preliminares de riscos com vistas a proporcionar uma melhor compreensão dos contextos perigosos que envolvem o patrimônio informacional registrado em mídias digitais. Para tanto, tomou-se como parâmetro de atuação o método proposto por Aguiar (s.d.), denominado Análise Preliminar de Perigos⁵ – APP. A escolha por esse tipo de análise é

5

Metodologia desenvolvida com vista a ser aplicada por pesquisas das Ciências Exatas.



justificada por se tratar de um método capaz de quantificar e qualificar riscos e ameaças concernentes à preservação da memória registradas em mídias digitais, assim como de monitorar o impacto e complexidade das ameaças concernentes a salvaguarda do patrimônio digital. Além disso, segundo Aguiar (s.d.), o reconhecimento antecipado dos riscos que envolvem os processos de preservação digital economiza tempo e recurso oriundos de modificações posteriores da instalação/sistema, assim como dos processos operacionais.

Em virtude do que foi mencionado, entende-se que a identificação dos potenciais perigos decorrentes do ambiente digital nos processos de guarda e preservação da memória tem por objetivo permitir, antecipadamente, a adoção de medidas preventivas a fim de eliminar as causas ou reduzir os impactos e conseqüências dos cenários de acidentes identificados. Assim, a utilização de métodos de análise preliminar de riscos tem por finalidade propor proteção e guarda ao patrimônio informacional gerenciado por sistemas de informação, na eventualidade de um possível acidente.

O método de análise proposta nessa pesquisa, ao tomar por base os estudos realizados por Aguiar (s.d.), compreende a execução das seguintes tarefas: Definição dos objetivos e do escopo da análise; Análise das instalações e local de acondicionamento oferecido por instituições e serviços ligados à guarda do patrimônio informacional; Coleta de informações sobre os perigos envolvidos no processo de preservação digital; Realização da análise propriamente dita (preenchimento da planilha); Elaboração das estatísticas dos perigos identificados por categorias de frequência e de severidade nos cenários de instituições que prestam serviços de informação; Análise dos resultados e preparação do relatório.

O escopo dessa análise abrange alguns dos principais eventos geradores de perigos encontrados no ambiente dessas instituições, englobando as falhas intrínsecas de sistemas, assim como eventuais erros operacionais (ações humanas). Além disso, abrange as análises dos eventos perigosos causados por agentes externos, tais como: inundações, tremores de terra, descarga elétrica.

Para esta pesquisa, as principais informações requeridas para a realização de uma análise preliminar dos perigos no ambiente dessas instituições e serviços de informação, quanto à guarda e preservação de sua memória institucional, são as seguintes:

Sobre as instalações: Analisar as condições de acondicionamento que essas instituições dispõem para salvaguarda destes materiais.

Sobre os processos: descrição dos processos envolvidos no gerenciamento dessas instituições; e principalmente os processos ligados a guarda e preservação do patrimônio informacional. Sobre os materiais e suportes: conhecer o tipo de material trabalhado, assim como sua durabilidade. Além das práticas de preservação desses materiais digitais custodiados por essas instituições. Sobre a equipe operacional: Verificar a existência de uma equipe específica responsável pelo gerenciamento dessas instituições, assim como, sua formação e capacitação.

O estabelecimento dessas informações possibilita a análise e descrição dos principais aspectos que impactam a segurança operacional da atividade de preservação digital nesses ambientes, assim como destaca o forte inter-relacionamento entre os fatores geradores de riscos (Ver Figura1)



FIGURA 1 - Fatores de relacionamento responsáveis pelo risco de segurança operacional nas atividades de preservação digital em instituições e serviços de informação.

FONTE: (RIBEIRO, 2011)

A realização da análise preliminar de risco para a preservação digital desse estudo é feita através do preenchimento de uma planilha constituída por 8 colunas, as quais são preenchidas conforme a descrição apresentada abaixo:

1ª coluna: Perigo - Esta coluna deve conter os perigos identificados para o sistema em estudo, ou seja, eventos que podem causar danos às instalações, materiais e ao patrimônio institucional. Portanto, os perigos referem-se a eventos tais como: terremotos e inundações; obsolescência tecnológica; falta de preparo dos profissionais; entre outros.

2ª coluna: Causa (s) - As causas de cada perigo devem ser discriminadas nesta coluna. Estas causas podem envolver falhas provocadas por agentes internos e externos.

3ª coluna: Efeito (s) - O resultado de uma ou mais causas é definido como efeito. Os possíveis efeitos danosos de cada perigo identificado devem ser listados nesta coluna. Os principais efeitos dos acidentes envolvendo a guarda da memória registrada em mídias digitais incluem, entre outros: Ilegibilidade da informação e Danificação das mídias.

4ª coluna: Severidade - No âmbito desta análise, um cenário de acidente é definido como o conjunto formado pelo perigo identificado, suas causas e cada um dos seus efeitos. Para fins dessa pesquisa, os cenários de acidentes encontrados no ambiente das instituições e serviços de informação devem ser classificados em quatro categorias de Severidade, as quais fornecem uma indicação qualitativa do grau de severidade das conseqüências desses acidentes. As categorias de Severidade utilizadas no presente trabalho estão reproduzidas no quadro abaixo:

QUADRO 1 - Categorias de Severidade para Análises de Risco

CATEGORIAS DE SEVERIDADE	
CATEGORIA	DESCRIÇÃO
I – Baixa	Sem danos, ou a falha não irá resultar em um dano mensurável nas instalações e processos de um repositório institucional. Dessa forma, não contribui para o aumento do risco na preservação da sua memória científica.
II – Moderada	A falha irá resultar em danos controláveis no ambiente do RI, ou seja, que não irá comprometê-lo seriamente.
III - Crítica	A falha causará danos substanciais ao processo de gestão do RI, assim como em seus equipamentos, provocando lesões e resultando em risco inaceitável. Esta categoria exige ações corretivas imediatas para evitar seu desdobramento em situações mais desastrosas.
IV - Catastrófica	A falha irá produzir severa degradação ao RI, possibilitando também danos irreparáveis aos seus equipamentos. Esses impactos significativos irão resultar na perda parcial e até mesmo total da memória dessas instituições, tendo o tempo de recuperação elevado. Exige pronta resposta à emergência para a minimização dos impactos.

FONTE (adaptada): (AGUIAR, [s.d])



5ª coluna: Frequência – O método de análise de risco usada no decorrer deste estudo tem seus cenários de acidentes classificados em categorias de frequências, as quais possibilitam uma indicação qualitativa da frequência esperada de ocorrência em cada um dos cenários identificados.

DESCRIÇÃO DE FREQUÊNCIAS	
CATEGORIA	DESCRIÇÃO
A - Remota	Conceitualmente possível, porém não é esperado acontecer durante o ciclo de vida útil do RI quanto ao processo de preservação da sua memória científica.
B - Ocasional	Pouco provável de ocorrer, a ocorrência do cenário depende de uma única falha interno-externa.
C - Provável	Esperado acontecer pelo menos uma vez a ocorrência do cenário no ciclo de vida do RI.
D - Frequente	Esperado acontecer algumas vezes a ocorrência do cenário durante o ciclo de vida do RI.
E - Muito Frequente	Esperado acontecer várias vezes durante o ciclo de vida útil do RI quanto ao processo de preservação da sua memória científica.

QUADRO 2 - Categorias de Frequência para Análises de Risco
FONTE (adaptada): (AGUIAR, [s.d])

6ª coluna: Valor do Risco - Na busca por caracterizar os fatores mais importantes que devem ser identificados e controlados para reduzir os níveis desses riscos ao patrimônio informacional registrado em mídias digitais, combina-se as categorias de frequências com as de severidade. Portanto, entende-se que para avaliar um risco é necessário estimar a probabilidade de que o evento venha a ocorrer e a extensão dos danos que o mesmo pode causar.

A composição desses dois fatores irá definir o nível de risco que envolve o ambiente das instituições e serviços de informação, o qual pode ser classificado em desprezível, menor, moderado, sério e crítico. Diante disso, conclui-se que riscos de maior probabilidade de ocorrência e que impliquem em danos mais graves devem ser, obviamente, confrontados em primeiro lugar, em qualquer plano de controle de riscos.



QUADRO 3 - Matriz Referencial dos Riscos

		SEVERIDADE			
		I - Baixa	II - Moderada	III - Crítica	IV - Catastrófica
FREQUÊNCIA	Remota(A)	1	1	1	2
	Ocasional(B)	1	1	2	3
	Provável(C)	1	2	3	4
	Frequente(D)	2	3	4	5
	Muito Frequente(E)	3	4	5	5

VALOR DO RISCO
1 - Desprezível
2 - Baixo
3 - Moderado
4 - Sério
5 - Crítico

FONTE (adaptada): (AGUIAR, [s.d])

7ª coluna: Medidas Preventivas/ Mitigadoras - Esta coluna contém as medidas mitigadoras recomendadas, propostas que podem ser utilizadas para evitar / minimizar o evento indesejável e suas conseqüências. Esta coluna deve ser preenchida com as medidas preventivas/corretivas ou quaisquer observações adequadas para a redução dos riscos.

ANÁLISE PRELIMINAR DE RISCOS							
INSTITUIÇÃO:							
GESTOR RESPONSÁVEL:				DATA:			
OBSERVAÇÕES:							
PERIGO	CAUSA	EFEITO	FREQUÊNCIA	SEVERIDADE	VALOR DE RISCO	MEDIDAS PREVENTIVAS	HIPÓTESES

8ª coluna: Número da Hipótese - Esta coluna é preenchida com o número da hipótese correspondente

QUADRO 4- Planilha de Análise de Risco utilizada na pesquisa

CATEGORIA DE SEVERIDADE	CATEGORIA DE FREQUÊNCIA	VALOR DO RISCO
I - Baixa	A - Remota	1 - Desprezível
II - Moderada	B - Ocasional	2 - Baixo
III - Crítica	C - Provável	3 - Moderado
IV - Catastrófica	D - Frequente	4 - Sério
	E - Muito Frequente	5 - Crítico

FONTE (adaptada): (AGUIAR, [s.d])



5 PLANEJAMENTO DE CENÁRIOS DE RISCOS

O planejamento e gestão de risco com base em cenários de acidentes permitem que organizações vislumbrem e preparem-se para as diversas situações e incertezas futuras, facilitando o processo decisório. Nesse sentido, segundo Schoemaker (1995) o planejamento de cenários de riscos compreende um método disciplinado para imaginar futuros possíveis, aplicável a uma grande variedade de questões, principalmente em situações em que essas explanações de possíveis futuros sejam necessárias nos processos de tomada de decisão. (SCHOEMAKER,1995)

Segundo SCHWARTZ (2000), o desenvolvimento dessa técnica de planejar cenários ocorreu após a Segunda Guerra Mundial, através de um planejamento militar da Força Aérea Norte-americana na tentativa de imaginar o que seus inimigos pretendiam fazer para assim definir e preparar estratégias alternativas e eficientes. A partir disso, entende-se por planejamento de cenário a técnica que nos permite traçar e ampliar o leque de possibilidades futuras e fatos incertos envolvidos na gestão de em uma determinada organização ou sistema.

Para este estudo, os cenários são desenvolvidos com vistas a estimar responsabilidade e empenho nos profissionais envolvidos com a gestão de documentos digitais ao longo prazo, objetivando antecipar cuidados com a preservação e com os custos despendidos durante o processo de guarda e acesso de uma memória digital. Além disso, o planejamento de cenários também possibilita um controle regulamentar das consequências de mudanças e obsolescência de suportes digitais, assim como propõe ajudar esses profissionais a visualizar as mudanças futuras nos serviços de informação.

Apesar da técnica de planejar acidentes permitir a captura de uma riqueza e variedade de possibilidades necessárias à tomada de decisão racional, ao mesmo tempo, o planejamento desses cenários organiza esses acidentes em narrativas que são mais fáceis de entender, simplificando a avalanche de dados em um número limitado de estados possíveis. Para tanto, esse estudo propõe o levantamento de alguns dos possíveis acidentes ocorridos durante o processo de guarda e acesso de uma memória digital, os quais possibilitam o “despertar” para a importância de se aumentar e melhorar as estratégias de preservação dessa memória, visto que a memória digital apresenta desafios diversos, que vão além do seu suporte.



CENÁRIO DE ACIDENTES			
ACIDENTES	PERIGOS	CAUSAS	EFEITOS
ACIDENTE Nº1	Obsolescência de hardware e software	Devido ao mercado altamente competitivo da tecnologia da informação	Levando a ciclos de renovação de tecnologia a cada 3 a 5 anos (HEDSTRON, 1998)
ACIDENTE Nº2	A falta de especialização e capacidade de domínios técnicos daqueles que lidam com nossa herança digital	Devido à falta de investimento e recursos oferecidos pelas suas instituições	Resultando na insegurança desses profissionais em trabalhar com novas tecnologias, originando sentimentos como o medo
ACIDENTE Nº3	A falta de administração dos riscos que envolvem a preservação digital da produção intelectual de uma instituição a longo prazo	Devido à falta de cooperação com outras iniciativas de preservação digital, assim como a falta de elaboração de manuais que ofereça orientações gerais quanto ao tratamento de objetos digitais e o gerenciamento dos riscos envolvidos na sua preservação	Resultando na vulnerabilidade dos seus processos de preservação digital e elaboração de projetos redundantes, além da minimização de esforços de preservação.
ACIDENTE Nº4	A ausência de clareza no papel de cada indivíduo envolvido no processo de gestão de instituições e serviços de informação	A falta de equipe específica para desempenhar as diferentes atividades realizadas nessas instituições e serviços de informação	Resultando na incapacidade de identificar e apontar responsabilidades, entre elas a de manter o acesso a longo prazo das informações contidas nessas instituições
ACIDENTE Nº5	Degradação de mídias digitais	Devido ao mau acondicionamento dos materiais digitais nas instalações dessas instituições	Resultando na legibilidade da informação registradas nessas mídias
ACIDENTE Nº6	Instabilidade de suporte por longo prazo	Devido à falta de uso preferencial de padrões e formatos de arquivos de dados abertos, com amplo acesso e assistência técnica (THOMAZ E SOARES, 2005)	Resultando na complexidade no momento de selecionar e aplicar estratégias de preservação, minimizando sua efetividade
ACIDENTE Nº7	Condições inadequadas do ambiente no qual estão inseridos e do manuseio dos mesmos.	Devido às condições ambientais, econômicas e políticas no território latino-americano, que se apresentem de formas mais adversas a prática de preservação digital. (GALINDO, 2005)	Levando ao desgaste e até mesmo a perda de leitura das informações registradas em objetos digitais. Sendo considerado um dos principais causadores da perda da informação contidas em hardwares e softwares.
ACIDENTE Nº8	A falta de clareza na definição de quais elementos dos objetos digitais e quais informações serão efetivamente preservadas	Devido à ausência de planejamento do que será selecionado para ser preservado, ou seja, de uma criteriosa política de seleção	Prejudicando a consolidação de uma coleção digital e a compreensão da mesma, por parte dos usuários.
ACIDENTE Nº9	A falta de autenticidade do objeto digital preservado	Devido à falta de identificação da origem e do histórico do objeto digital durante processo de preservação	Levando a dificuldade de confirmar a integridade e a preservação do objeto na sua forma original
ACIDENTE Nº10	O dilema da escolha das estratégias e ações de preservação	Devido à ausência de uma solução prática aplicável universalmente ao problema da obsolescência tecnológica e da degradação dos objetos digitais	Resultando na adoção de estratégias pouco significativas, a qual compreende uma decisão não ponderada e sem base nos vários fatores que envolvem a preservação do objeto digital
ACIDENTE Nº11	A escolha não ponderada de estratégias para a preservação dos materiais digitais	Devido à falta de conhecimento técnico daqueles que lidam com nossa herança digital	Levando a maximização de esforços desnecessários à preservação de objetos digitais
ACIDENTE Nº12	A falta de métodos organizados para a realização das atividades de preservação	Devido à falta de políticas institucionais voltadas à guarda e preservação de objetos digitais	Resultando na ausência de organização e clareza dos objetivos, diretrizes, práticas e intenções organizacionais que servem para fortalecer as decisões locais, ou seja, o caminho para alcançar o consenso corporativo

QUADRO 5 - Cenário de acidentes para a preservação digital
 FONTE: (RIBEIRO, 2011)



6 A APLICAÇÃO DA ANÁLISE DE RISCO: BDTD/UFPE

QUADRO 6 - Aplicação da Análise de Risco na BDTD/UFPE

ANÁLISE PRELIMINAR DE RISCOS							
INSTITUIÇÃO: BDTD-UFPE							
GESTOR RESPONSÁVEL: MARIA DALVA PEREIRA				DATA: 15/07/2011			
OBSERVAÇÕES:							
PERIGO	CAUSA	EFEITO	FREQUÊNCIA	SEVERIDADE	VALOR DE RISCO	MEDIDAS PREVENTIVAS	HIPÓTESES
1 - Obsolescência de hardware e software	Obsolescência tecnológica	Vulnerabilidade da gestão de documentos.	Muito frequente	Catastrófica	Crítico	O estabelecimento de políticas e estratégias de preservação digital.	A qualidade e integridade da informação digital são extremamente vulneráveis a obsolescência tecnológica.
2 - Ausência de recursos financeiros destinados a PD	Falta de investimento	Incapacidade de guarda e acesso as informações.	Muito frequente	Catastrófica	Crítico	O apoio institucional às práticas de preservação digital.	Existe uma carência de recursos necessários às atividades de PD.
3 - A falta de capacitação de domínios técnicos	Falta de apoio institucional	Profissionais e serviços inseguros.	Muito frequente	Catastrófica	Crítico	O apoio institucional aos seus servidores.	Ausência de apoio e políticas de capacitação.
4 - Indefinição de função e responsabilidades profissionais	A falta de equipe específica	Incapacidade de identificar e apontar responsabilidades	Remota	Baixa	Desprezível	Estabelecer uma equipe específica, com atividades e funções específicas	Faz-se necessário uma equipe específica na gestão de documentos digitais.
5 - Degradação de mídias digitais	Mau acondicionamento	Ilegibilidade da informação	Muito frequente	Catastrófica	Crítico	Preparar ambientes específicos para a guarda de documentos digitais.	O mau acondicionamento compreende um dos principais desafios da PD
6 - Instabilidade de suporte por longo prazo	Falta de uso de padrões e formatos de arquivos de dados abertos.	Complexidade na seleção e aplicação de estratégias de PD	Remota	Baixa	Desprezível	Estabelecer politicamente o uso de padrões de arquivos abertos	O uso desses padrões garante a efetividade da gestão e preservação da informação
7 - Condições inadequadas do ambiente e do manuseio de documentos digitais	Condições ambientais, econômicas e políticas	Desgaste e perda de leitura das informações registradas em objetos digitais.	Muito frequente	Catastrófica	Crítico	Garantir condições e recursos à prática de PD.	Os suportes tecnológicos possuem tempo de vida útil instável, quando expostos sob condições ambientais inadequadas.
8 - Ausência de políticas de seleção para PD	A ausência de planejamento	Prejudicando a consolidação e desenvolvimento de uma coleção digital	Muito frequente	Catastrófica	Crítico	Estabelecer políticas de seleção para os objetos a serem preservados	Políticas de seleção são tarefas fundamentais na gestão dos documentos de uma organização.
9 - A falta de autenticidade do objeto digital preservado	Falta de identificação histórica do objeto digital	Dificuldade de confirmar a integridade e a preservação do objeto na sua forma original	Remota	Baixa	Desprezível	Desenvolver identificação para os objetos preservados	A identificação histórica do objeto digital garante integridade e autenticidade ao documento preservado
10 - O dilema da escolha das estratégias de PD	Ausência de solução prática aplicável universalmente ao problema de PD	Adoção de estratégias pouco significativas	Muito frequente	Catastrófica	Crítico	Estabelecer estudos prévios a escolha e aplicação dessas estratégias	O conhecimento garante a escolha eficiente de estratégias de PD
11 - A escolha não ponderada de estratégias de PD	Falta de conhecimento técnico	Maximização de esforços	Muito frequente	Catastrófica	Crítico	Promover capacitação profissional àqueles que gerenciam documentos digitais	A falta de capacitação profissional compreende um dos principais desafios da PD
12 - A falta de métodos organizados para a PD	Falta de políticas institucionais voltadas à PD.	Ausência de organização dos objetivos organizacionais	Muito frequente	Catastrófica	Crítico	Desenvolver políticas de preservação	As políticas de PD não estão frequentemente presentes tanto no setor privado quanto no setor público.
13 - A falta de administração dos riscos que envolvem a PD	Falta de elaboração de manuais e de cooperação com outras iniciativas de PD	Vulnerabilidade e redundância de projetos e esforços para a PD	Muito frequente	Catastrófica	Crítico	Construir parcerias e maximizar os esforços de preservação através de processos cooperativos de PD	Participar de projetos cooperativos de PD proporciona diversos benefícios para a gestão de documentos

FONTE: (RIBEIRO, 2011)



Em 2007 foi implementada a Biblioteca Digital de Teses e Dissertações da Universidade Federal de Pernambuco (BDTD/UFPE), integrando-o a Biblioteca Digital de Teses e Dissertações coordenada, em âmbito nacional, pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), do Ministério da Ciência e Tecnologia-MCT.

Na busca por caracterizar o nível de risco que envolve o ambiente da BDTD/UFPE, com vistas a propor proteção e guarda ao patrimônio científico, com esta análise foi possível expor o risco a que esse estoque de informação está exposto.

Através desse experimento observamos que a biblioteca apresenta, em totalidade, todos os possíveis cenários de acidentes de riscos levantados na presente pesquisa. Entre os riscos analisados, cerca de 75% representam riscos de maior probabilidade de ocorrência e que implicam em danos mais graves ao estoque dessas informações, sendo que cerca de 80% destes estão classificados como “catastrófico” (Ver Quadro 1) e também “Muito freqüente” (Ver Quadro 2).

Diante das principais informações levantadas para a realização da análise preliminar dos perigos no ambiente das instituições e serviços de informação (Ver Figura 1), aqui representados pela BDTD/UFPE, de início, foi possível identificarmos que as condições de acondicionamento que essas instituições dispõem para salvaguarda de seus materiais digitais constituem um dos principais desafios à preservação de seu patrimônio digital (perigos 5, 7), caracterizando-se como risco “crítico” e de maior probabilidade de ocorrência.

No que diz respeito aos processos envolvidos no gerenciamento dessa instituição, quanto à guarda e preservação de seu patrimônio informacional, com esta forma de análise foi possível alcançar resultados claros de que os perigos (2, 3), que correspondem respectivamente à ausência de apoio e investimento institucional às iniciativas de PD à memória em meio digital, ainda são identificados como práticas não contempladas pelo repositório da IES analisada, apesar da confiança que a sociedade acadêmica tem depositado neste tipo de repositório digital, reconhecido e qualificado pelos pares.

Ainda na análise dos processos de gerenciamento deste repositório, a análise de risco permitiu constatar que a falta de políticas institucionais voltadas à PD (perigos 8 e 12) caracteriza-se como risco que pode acontecer várias vezes durante o ciclo de vida útil do repositório, além de configurar-se como responsável por impactos significativos que irão resultar na perda parcial e até mesmo total da memória dessa instituição. Esta situação é abordada por Ribeiro (2006) quando afirma que a questão da preservação das memórias em instituições de pesquisa ainda é incipiente devido à falta de vontade política.

Em pesquisa recente elaborada por Pinto (2009), a autora investiga, sistematiza e divulga os diversos projetos relacionados com a preservação digital no âmbito internacional. A partir daí, a autora apresenta resultados inequívocos e que sustentam as idéias de cooperação e complemento das pesquisas desenvolvidas em prol da preservação em longo tempo. Entretanto, com a análise desenvolvida no ambiente da BDTD/UFPE, identificamos que a ausência desse repositório na



construção e participação em projetos cooperativos de PD (perigo 13) configura-se com um risco “crítico”, tendo como conseqüências a maximização, vulnerabilidade e redundância de projetos e esforços para a PD.

A análise desenvolvida na presente pesquisa também possibilitou um diagnóstico da equipe operacional desse repositório, onde identificamos que, apesar da existência de uma equipe específica responsável pelo gerenciamento dessas instituições (perigo 4), a questão da formação e capacitação desses profissionais ainda é um dos principais desafios para a gestão de seus acervos digitais (perigos 3, 11), sendo caracterizados como riscos “críticos”. Estes resultados nos levam a entender outros aspectos que impactam a segurança operacional da atividade de preservação digital, que são aqueles que envolvem os materiais e suportes digitais custodiados por esse repositório. O forte inter-relacionamento entre esses fatores geradores de riscos nos leva a entender a característica “crítica” dos perigos (1, 5, 10, 11).

Esses resultados nos apresentam a magnitude dos perigos que envolvem o patrimônio científico registrado em mídias digitais, assim como nos permite inferir que, apesar do valor cultural, científico e econômico dos conteúdos científicos, a sociedade acadêmica encontra-se insuficientemente capaz de *preservar* e *salvaguardar* esse conhecimento para a posteridade, o que comprova a ausência de uma cultura de preservação digital para a produção científica no Brasil. Diante dessas ameaças, a consciência do perigo se faz necessária, gerando assim políticas, estratégias e outros instrumentos aplicados a preservação de acervos digitais, como medidas preventivas capazes de reduzir dos impactos e conseqüências dos cenários de acidentes identificados.

7 CONSIDERAÇÕES FINAIS

Apesar dos grandes avanços alcançados pelos estudos realizados em pesquisas de Preservação Digital, grande parte dos projetos que envolvem a gestão do patrimônio informacional ainda permanece nas mãos do destino. Desde o início dessa “nova era digital” que diversos projetos tem resultado em verdadeiros desastres devido a eventos sinistros que poderiam ser evitados caso houvesse elaboração e aplicação eficiente de estratégias metodológicas de preservação. Esse fato se deve a motivos históricos e culturais, que levam a grande maioria dos profissionais a ignorarem o uso de métodos de monitoramento de risco, entre esses motivos está a discriminação da própria instituição aos profissionais que ousam apontar as incertezas e que envolvem o alcance eficiente dos objetivos de uma instituição.

Ao entendermos que lidar com a preservação digital é lidar com constante controle de riscos, a necessidade de aplicação de métodos de análise e gestão de riscos no domínio da CI surgem como instrumental eficiente na aferição de risco, na administração do problema do risco (mitigação de fatores causais) e no desenvolvimento de mecanismos de observação permanentes que se antecipam protegendo o sistema interno de impactos contra eventos sinistros da perda de registros do patrimônio digital.



Além disso, ao discutirmos os conceitos e aplicações de novos métodos de monitoramento de risco no âmbito da CI, foi possível concluir que muitos profissionais ainda não intuíram as enormes implicações que as mudanças e incertezas futuras podem provocar no processo de gestão de suas informações, o que demonstra a falta de conhecimento conceitual das técnicas metodológicas que permitem a análise de incertezas e riscos que envolvem o patrimônio informacional custodiados por essas instituições e serviços de informação.

Espera-se que a abordagem feita nesta pesquisa sirva como ponto de partida para futuras reflexões e para o “despertar” na importância de se aumentar e melhorar as práticas e aplicações de métodos de monitoramento de risco no campo da Ciência da Informação.

Abstract: *Across the universe of digital assets, the use of new technologies enabled significant advances for the power storage and transfer of information, however the guarantee of accessibility and interpretation of information recorded in digital media are ominously unknown, which makes the preservation Digital one of the main challenges of our time. Therefore, this article describes the development of a method of risk analysis for the area of digital preservation and this study focused experience the Digital Library of Theses and Dissertations at the Federal University of Pernambuco, in order to quantify and qualify the risks and threats concerning the preservation of memory recorded in digital media, as well as to monitor the impact and complexity of threats related to safeguard digital heritage. To this end, opted for a research and exploratory qualitative research and a theoretical nature of literature (books and journals in digital form), which helped to describe the practices and political strategies for digital preservation and also offered toric base and methodological required the use and application of analytical methods, which would allow a constant control of risks. The results show that about 75% represent risks most likely to occur and involve more serious damage to the stock of such information, and about 80% of these are classified as “catastrophic” and also “very frequent”. Supported in these data was possible to prove the hypothesis in this work, that there is a culture of digital preservation for the scientific production in Brazil.*

Keywords: Digital preservation. Risk Analysis. Memory. Information Technology.

REFERÊNCIAS

AGUIAR, L. A. de. **Metodologias de análise e riscos APP e Hazop**. Rio de Janeiro, RJ: [s.n., s.d.]. Disponível em: <http://professor.ucg.br/SiteDocente/admin/arquivosUpload/13179/material/APP_e_HAZOP.pdf>. Acesso em: 20 set. 2010.

ASSOCIATION FOR LIBRARY COLLECTIONS & TECHNICAL SERVICES. Definitions of Digital Preservation. In.: ALA ANNUAL CONFERENCE, 2007, Washington, D.C. [**Anais eletrônicos...**]. Washington: American Library Association's, 2007. Disponível em: <<http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/defdigpres0408.pdf>>. Acesso em: 04. mar. 2011.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ISO 31000:2009**: Risk management -- Principles and guidelines. New Zealand: Standards, 2009.



BALBI, G. **Doing Media History in 2050**. Lugano, Suíça: [s.n.: 2010]. Disponível em: <<http://web.mit.edu/comm-forum/mit6/papers/Balbi.pdf>>. Acesso em: 20 mar. 2011.

BORBA, V. da R. **Modelo orientador para construção de estratégias de preservação digital: estudo de caso do Banco de Teses e Dissertações da UFPE**. 2009. 134 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal da Paraíba, 2009.

FERREIRA, M. **Introdução à preservação digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006.

GALINDO, M. Tragédia da Memória. **Massangana**, Recife, v. 2, n. 1, p. 29-31, 2005

GANOULIS, J. **Risk Analysis of Water Pollution**. 2. ed. [s.n.]: [WILEY-VCH Verlag GmbH & Co. KGaA], 2009. Disponível em: <<http://pt.scribd.com/doc/55784093/Risk-Analysis-of-Water-Pollution-Wiley>>. Acesso em: 10 jul. 2011

HEDSTROM, M. Digital preservation; a time bomb for digital libraries. **Computer and the Humanities**, v.31, n.3, 1998, p.189-202. Disponível em: <http://deepblue.lib.umich.edu/bitstream/2027.42/42573/1/10579_2004_Article_153071.pdf>. Acesso em: 03 nov. 2010.

INNARELLI, H. C. Preservação digital e seus dez mandamentos. In: SANTOS, V. B.; INNARELLI, H. C.; SOUSA, R. T. B. (Orgs.). **Arquivística: temas contemporâneos: classificação, preservação digital, gestão do conhecimento**. Distrito Federal: SENAC, 2007.

INNARELLI, H. C. Preservação digital: a influência da gestão dos documentos digitais na preservação da informação e da cultura. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 8, n. 2, p. 72-87, jan./jun. 2011. Disponível em: <http://www.sbu.unicamp.br/seer/ojs/index.php/sbu_rci/article/view/487>. Acesso em: 01 jun. 2011.

ORTEGA Y GASSET, J. **Missão do Bibliotecário**. Tradução e posfácio de Antônio Agenor Briquet de Lemos. Brasília, DF: Briquet de Lemos, 2006.

PINTO, M. M. **Preservmap: um roteiro da preservação na era digital**. Porto: Edições Afrontamento, 2009.

RIBEIRO, F. do C. **Análise de Risco: uma metodologia a serviço da preservação digital**. 2011. Dissertação. (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2011. No prelo.

ROSSI, P. **O passado, a memória, o esquecimento: seis ensaios da história das ideias**. São Paulo: Editora Unesp, 2010.

ROTHENBERG, J. **Ensuring the Longevity of Digital Information**, 1999. Disponível em: <<http://www.clir.org/pubs/archives/ensuring.pdf>>. Acesso em 13 set. 2010.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. Enc. Bibli: R. Eletr. Bibliotecon. **Ciência da Informação**, Florianópolis, v. 15, n. 30, p. 1-31, 2010. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/12528/14379>>. Acesso em: 03 jan. 2011.



SCHOEMAKER, P. J. H. Scenario Planning: A Tool for Strategic Thinking. **Sloan Management Review**, v. 36, n. 2, p. 25-40, 1995. Disponível em: < <http://www.favaneves.org/arquivos/scenarioplanning.pdf>>. Acesso em: 10 jul. 2011

SCHWARTZ, P. **A arte da Visão de Longo Prazo**: planejando o futuro em um mundo de incertezas. São Paulo. Editora Best Seller, 2000.

THOMAZ, K. P.; SOARES, A. J. A preservação digital e o modelo de referência Open archival Information System (OAIS). **DataGramZero**: Revista de Ciência da Informação, v. 5, n. 1, fev. 2004. Disponível em: <http://www.dgz.org.br/fev04/Art_01.htm>. Acesso em: 13 set. 2010.



SERVIÇOS DE RECOMENDAÇÃO EM BIBLIOTECAS DIGITAIS CIENTÍFICAS

Flavia Maria Bastos, Silvana Aparecida Borsetti Gregorio Vidotti

Resumo: O presente estudo apresenta como tema o serviço de recomendação no contexto das bibliotecas digitais com uso dos recursos da *Web 2.0*, em destaque, os serviços de recomendação aplicados ao uso das informações científicas produzidas em ambientes informacionais digitais, como o das bibliotecas digitais científicas. Assim, pretende-se ressaltar com base na literatura como a força e a colaboração da comunidade acadêmica podem gerar recomendações através do uso e da produção de conteúdo social. Para o desenvolvimento desse estudo observou-se o funcionamento de recursos da tecnologia colaborativa (recursos da *Web 2.0*) utilizados em ambientes informacionais digitais que se estendem ao conceito de Biblioteca 2.0. Assim, com base na literatura, foi retratada a importância dos serviços de recomendação no contexto das bibliotecas digitais científicas. Optou-se pela pesquisa qualitativa com observação direta não participante para realizar a identificação e descrição do funcionamento das técnicas amplamente difundidas no contexto político, econômico, e social, como *sites* de busca, de *download*, de compra, de entretenimento, de relacionamento (comunidades sociais), aplicadas ao ambiente das bibliotecas universitárias digitais. A inserção dessas tecnologias apresenta-se como inovação, mas sua natureza está vinculada à tradição e à missão das bibliotecas. No entanto, com a explosão informacional, os usuários muitas vezes apresentam dificuldades para realizarem suas escolhas e descobertas em função da quantidade de recursos e fontes informacionais disponibilizadas no ambiente das bibliotecas digitais científicas, por exemplo. Os recursos da *Web 2.0* estão propiciando um ambiente informacional estruturado de forma mais flexível, atrativa e dinâmica, onde o usuário torna-se o ator principal na construção do seu ambiente, possibilitado pelos recursos de customização e personalização. Além disso, permite a descoberta do conteúdo informacional científico e a geração de informações pelos próprios usuários, que seguem retroalimentadas ao sistema novamente, pois o serviço de recomendação acompanha as relações humanas, o que enriquece a estrutura de informação existente nesses ambientes.

Palavras-chave: Bibliotecas Digitais. *Web 2.0*. Tecnologia Colaborativa. Serviço de Recomendação.

1 INTRODUÇÃO

O presente estudo apresenta como tema o serviço de recomendação em bibliotecas digitais no contexto dos recursos da *Web 2.0* aplicados ao uso das informações científicas produzidas em ambientes informacionais digitais, como o das bibliotecas digitais científicas.

Não obstante as bibliotecas digitais utilizarem recursos da *Web 2.0*, ainda faltam práticas colaborativas efetivas que absorvam ou reconheçam a importância da participação e contribuição do usuário nesses ambientes digitais a ponto de interferir tanto no uso realizado por outros usuários como



na retroalimentação ou reestruturação do próprio sistema. Desta forma, com as constantes mudanças ocorridas em função das inovações tecnológicas que influenciam profundamente o ciclo da informação na sociedade, observamos e aplicamos à biblioteca algumas dessas inovações tecnológicas, no intuito de intensificar sua função enquanto socializadora do conhecimento através da armazenagem, da organização e do acesso à informação. No entanto, mesmo com a evolução dos recursos utilizados pela biblioteca tradicional e sua respectiva evolução para o ambiente de biblioteca digital, notamos a falta de uma melhor integração e apropriação entre os recursos da tecnologia colaborativa e os serviços oferecidos pelas bibliotecas digitais.

Assim, questionamos como é possível promover uma efetiva interação entre bibliotecas digitais e usuários, bem como a integração dos conteúdos sociais produzidos por eles aos ambientes e serviços dessas bibliotecas.

A biblioteca tem por característica ser um ambiente informacional controlado, e por este motivo apresenta dificuldades para entrada e validação de informações externas, dificultando assim sua efetiva interação e integração dos conteúdos socialmente produzidos pelos usuários.

Desta forma, nos propomos a apresentar reflexões sobre o uso dos serviços de recomendação em bibliotecas digitais, bem como, uma análise teórica da incorporação do conteúdo social produzido pelo usuário no ambiente das bibliotecas digitais. Com objetivo de tornar a biblioteca digital um ambiente colaborativo a partir da interação e participação do usuário, bem como, um ambiente de compartilhamento de informações.

Com as influências proporcionadas pelas tecnologias de informação e comunicação sobre o ambiente das bibliotecas, essas passam contar com recursos tecnológicos para lidar com a explosão informacional que atinge os usuários, desta forma, a partir dos serviços de recomendação temos a possibilidade de obter um relato objetivo do uso e da expressão de preferências da comunidade, pois a força destes serviços esta justamente na comunidade contribuindo e participando a ponto de formarem uma enorme massa de dados com condições de representá-los diante de suas relações sociais.

Para o desenvolvimento dessa pesquisa optou-se pela pesquisa qualitativa com observação direta não participante para realizar a identificação e descrição do funcionamento de recursos da *Web 2.0* aplicados em ambientes digitais informacionais, como *The University of Iowa libraries* (2011), o repositório digital *Merlot* (2011), a comunidade virtual *LibraryThing* (2011), e o site comercial da *Amazon* (2011). Além disso, foi realizada revisão de literatura sobre serviços de recomendação de modo a serem agregados ao ambiente das bibliotecas digitais científicas.

Assim, o presente estudo aborda o uso dos recursos da *Web 2.0* no ambiente de bibliotecas digitais científicas, ancorado no contexto da produção científica, salientando como a interação e participação, a partir das recomendações, podem tornar o ambiente das bibliotecas algo complexo



baseado na multiplicidade e diversidade das fontes de informação armazenadas, tais como livros, documentos eletrônicos, jornais eletrônicos, e perfis de usuários, entre outros.

Em vista disso, são apresentadas algumas reflexões sobre o uso e aplicação dos serviços de recomendação ao ambiente de bibliotecas digitais científicas, voltados para fins de colaboração entre os usuários da biblioteca, especialmente pesquisadores.

2 TECNOLOGIAS COLABORATIVAS EM AMBIENTES INFORMACIONAIS : O Caminho para as Bibliotecas 2.0

A evolução das bibliotecas tem-se destacado ao longo destes anos quanto ao desenvolvimento e uso de novas tecnologias potencializadas por meio dos novos recursos de acesso e formatos de intercâmbio.

Diante das constantes mudanças ocorridas em função das inovações tecnológicas que influenciam profundamente o ciclo da informação na sociedade, observamos e aplicamos à biblioteca algumas dessas inovações tecnológicas no intuito de intensificar sua função enquanto socializadora do conhecimento através da armazenagem, da organização e do acesso à informação. No entanto, mesmo com a evolução dos recursos utilizados pela biblioteca tradicional e sua respectiva evolução para o ambiente de biblioteca digital, notamos a falta de uma melhor integração e apropriação entre os recursos da tecnologia colaborativa e os serviços oferecidos pelas bibliotecas digitais.

Hoje, é expressiva a presença de aspectos sociais no ambiente da biblioteca digital e durante alguns eventos, como o “*Workshop sobre os Aspectos Sociais das Bibliotecas Digitais*”, (BORGMAN, 1996), foram discutidas questões que têm as bibliotecas digitais como ambientes construídos, coletados e organizados, por (e para) uma comunidade de usuários, e suas capacidades funcionais sustentando as necessidades de informação e usos dessa comunidade.

Ainda, de acordo com as reflexões desse workshop, temos no ambiente das bibliotecas digitais indivíduos e grupos interagindo uns com os outros, usando sistemas e fontes de dados, informação e conhecimento. Neste sentido, representam extensão, aumento e integração de uma variedade de instituições de informação, assim como locais físicos, onde recursos são selecionados, coletados, organizados, preservados, e acessados em função de uma comunidade de usuários.

A evolução das bibliotecas tem merecido grande destaque ao longo destes anos em relação ao desenvolvimento e ao uso de tecnologias, que se intensificam por meio dos novos recursos de acesso e formatos de intercâmbio. Desde 1440, antes do surgimento da imprensa, as bibliotecas contavam com acervos formados por diversos tipos de materiais, como o tablete de argila, o papiro e o pergaminho, mas foi a partir da introdução da prensa móvel de Gutenberg que o papel passou a ser o suporte de registro principal da informação. O conceito tradicional das bibliotecas, como organização que seleciona, coleta, organiza, conserva, preserva e fornece acesso à informação para uma determinada comunidade de usuários, tem prevalecido ao longo dos anos.



O avanço proporcionado pelas tecnologias informacionais e comunicacionais (TICs) influencia a prática dos profissionais da informação no direcionamento e execução de suas atividades, visto que sem a devida reflexão e consciência do uso dos recursos tecnológicos, muitas vezes os profissionais acabam adotando a tecnologia e esquecendo os subsídios teóricos que fundamentam suas produções por meio dos serviços, ferramentas e técnicas existentes há anos. Por consequência, fica a impressão de que serviços tradicionais sumiram ou ficaram obsoletos, quando na realidade ocorreu o contrário, pois apesar da inserção dessas tecnologias afigurar-se como inovação, sabemos que sua natureza está vinculada à tradição e a missão das bibliotecas, como no caso do serviço de Disseminação Seletiva da Informação (DSI) e os recursos utilizados com a tecnologia RSS aplicada ao ambiente de bibliotecas.

Grande parte das bibliotecas tradicionais é dependente do fator humano e, com a atual explosão informacional, os recursos tecnológicos são ferramentas eficientes no processamento de um grande volume de dados, permitindo um melhor aproveitamento dos recursos humanos, de forma mais interativa e participativa nos ambientes informacionais digitais.

Temos assim a possibilidade de se romper com o modelo tradicional de biblioteca, onde somente se consulta e se retira material, para transformar as bibliotecas digitais em ambientes fomentados pelas comunidades de usuários e bibliotecários e para criar uma rede informacional a partir dos relacionamentos desses atores. Em outras palavras, pretende-se socializar a gestão dos ambientes informacionais nos quais as bibliotecas digitais estão inseridas.

Atualmente, há redes digitais favoráveis ao movimento dos softwares livres, à formação das iniciativas *wiki* e trabalhos em *creative commons* que permitem ações colaborativas em ambientes informacionais digitais. Nesse contexto, os ambientes informacionais digitais tornam-se ambientes colaborativos por estarem imersos na criação de múltiplos subsídios para a disseminação da informação. Isso possibilita a configuração de uma inteligência coletiva e a inclusão dos indivíduos na sociedade da informação, a partir do uso das tecnologias de informação e da comunicação, como meios para a construção do conhecimento.

A convergência de diferentes mídias, a utilização da *web* como plataforma, o surgimento de tecnologias que estimulam o compartilhamento de informações e as práticas colaborativas, são algumas das características de um novo sistema, chamado de *Web 2.0*.

O conceito de *Web 2.0* foi criado em 2004 por Tim O'Reilly, em um momento de crise das empresas ligadas à internet, onde alguns *sites* se diferenciavam na web por se pautarem nos pilares da *web* social, quais sejam:

- A *web* como plataforma;
- Maior interação entre a *web* e a arquitetura da participação.



Sendo assim, essas tecnologias colaborativas ocorrem na *web*; são serviços, não produtos; não são limitados a um único produto de software ou a uma única máquina; estão abertos e compartilhados; em grupo, usuários interagem socialmente e tornam-se parte integrante de suas organizações.

No contexto dos ambientes informacionais de bibliotecas, temos a atitude da *Web 2.0* e as tecnologias aplicadas às coleções e aos serviços bibliotecários, formando assim, a Biblioteca 2.0, que envolve:

- Atitudes: melhora contínua, confiança radical e aproveitamento da inteligência coletiva;
- Ferramentas: *blogs*, *wikis*, *RSS*, *software social*, *mashups*;
- Conteúdo social: comentários, classificação, etiquetas e outras contribuições.

No contexto da Biblioteca 2.0, a maioria dos pesquisadores concordaria com o fato de que muito do que foi aprovado pelas bibliotecas na primeira revolução da *web* é estático. Por exemplo, catálogos on-line de acesso público (OPAC) exigem que os usuários busquem a informação (Casey; Savastinuk, 2006) - embora muitos estejam começando a incorporar técnicas da *Web 2.0* relativas à pesquisa de dados, eles não respondem com recomendações, tal como ocorre no *site* da *Amazon*, por exemplo, que se apresenta com maior dinamismo. Do mesmo modo, a primeira geração de biblioteca on-line foi elaborada por meio de textos tutoriais estáticos e que não respondiam às necessidades dos usuários, nem permitiam que interagissem uns com os outros. Porém, com o uso de bancos de dados mais sofisticados, as bibliotecas têm começado a evoluir para uma estrutura mais interativa, com meios de comunicação sociais e ricos em tutoriais, programação e animação.

Em oposição ao modelo tradicional, onde o usuário é quem tinha que se aproximar da biblioteca, a Biblioteca 2.0 apresenta-se como um ambiente voltado não só às necessidades desse usuário em consumir, mas também em produzir e trocar informações de forma colaborativa, reestruturando o sistema através do *feedback* estabelecido com a biblioteca.

Desse modo, fica clara a transformação da relação entre a informação e o conhecimento no ambiente das bibliotecas digitais, em oposição aos ambientes tradicionais, que só permitiam ao usuário uma única forma de se relacionar com o conteúdo armazenado fisicamente. Hoje, não só o usuário participa como também adiciona conteúdos, que ao serem compartilhados em outros recursos colaborativos, passam a receber comentários dos membros vinculados à sua comunidade, agregando valor à informação compartilhada, além de criar vários caminhos para a localização do próprio recurso.

Para o desenvolvimento deste estudo foram escolhidas fontes de informação baseadas na literatura utilizada para fundamentação teórica de minha pesquisa de doutorado em andamento, onde se observou o funcionamento de recursos da tecnologia colaborativa utilizados no ambiente informacional digital da *The University of Iowa libraries* (2011), com o repositório digital *MERLOT* (2011), a comunidade virtual *LibraryThing* (2011), e o *site* da *Amazon* (2011). Com base nos recursos identificados e coletados, foi verificado como são aplicados os recursos da tecnologia colaborativa, conhecida como



recursos da *Web 2.0*. A inserção dessas tecnologias apresenta-se como inovação, mas sua natureza está vinculada à tradição e à missão das bibliotecas. O ambiente científico informacional observado, em função das tecnologias da *Web 2.0*, estão estruturados de uma forma mais flexível, atrativa e dinâmica, no qual o usuário torna-se o ator principal na construção desse ambiente, possibilitado pelos recursos de customização e personalização.

Assim, com base nesses recursos exemplificamos o trabalho da *The University of Iowa libraries* (<http://www.lib.uiowa.edu/>):

- A Biblioteca da Universidade de *Iowa* propicia aos seus usuários a disseminação do conteúdo científico por meio da utilização de recursos de *RSS*;
- Possibilita aos usuários enviarem, durante suas buscas, a informação recuperada via SMS (mensagens de texto);
- Disponibiliza espaço para comunicação via *blog*;
- A biblioteca também oferece e orienta seus serviços em um *chat on-line*, juntamente com outros serviços que podem ser selecionados no *Bookmark Share* que possibilita uma maior interação entre o usuário e o conteúdo armazenado;
- As informações, selecionadas pelo usuário na biblioteca, também podem ser visualizadas por suas *tags* personalizadas.

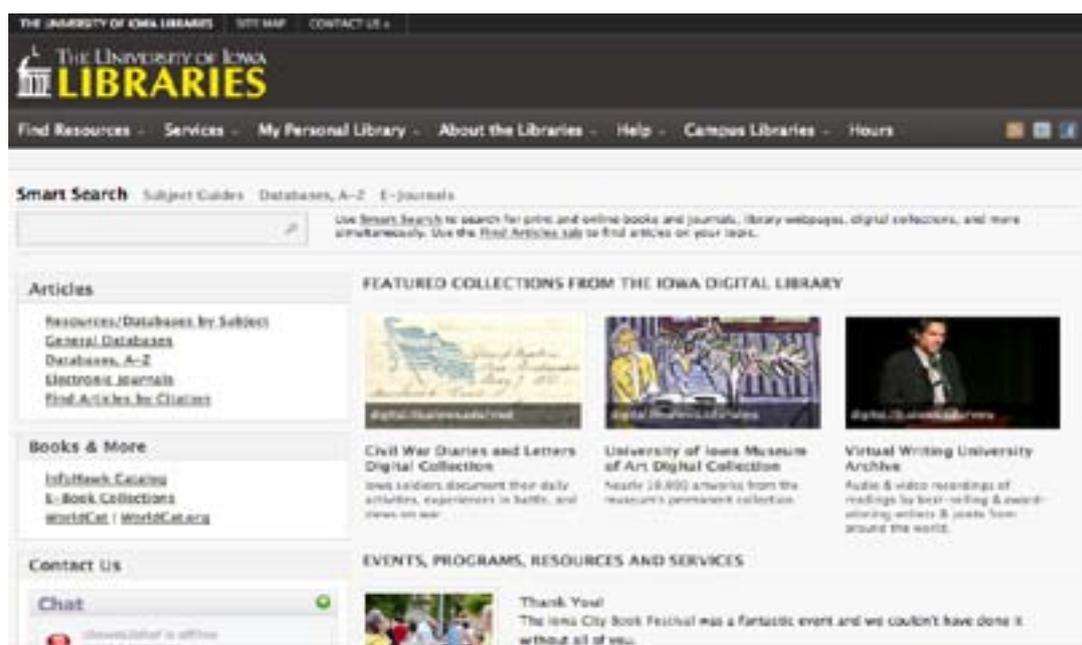


Figura 1: Site *The University of Iowa libraries*
Fonte: <http://www.lib.uiowa.edu/>



MERLOT – Multimedia Educational Resource for Learning and Online Teaching (<http://www.merlot.org/>). Esse repositório trabalha com a filosofia do autoarquivamento, permitindo uma interação entre os pesquisadores e o ambiente de depósito; os usuários do repositório podem postar comentários (*Peer Review*); apresenta links associados, além de informações referentes ao nível de contribuição de cada pesquisador, bem como os prêmios conquistados; o repositório também possibilita o envio de informações para outra pessoa.



Figura 2: *MERLOT – Multimedia Educational Resource for Learning and Online Teaching*

Fonte: <http://www.merlot.org/>

O *LibraryThing* – Catalogue seus livros *online* (<http://www.librarything.com/>) é uma comunidade virtual que reúne o maior clube literário formado por pessoas que gostam de livros e possibilita aos usuários:

- Catalogar livros (a partir dos dados da Biblioteca do Congresso dos EUA e da *Amazon API*);
- Relacionar livros, revistas, usuários, catálogos;
- Oferece serviços de recomendação;
- Atribuir *tagging*, bem como uso tradicional de cabeçalhos de assunto;
- Trabalhar com comunidade.



Figura 3: *LibraryThing* | *Catalogue seus livros online*

Fonte: <http://www.librarything.com/>

Amazon.com – O site da *Amazon* (<http://www.amazon.com/>) aproveita cada vez mais o conteúdo gerado pelo usuário no momento da aquisição de um produto.

Os produtos comprados na *Amazon* fornecem informações sobre os interesses dos clientes cadastrados, pois a empresa desenvolveu bases de dados com informações anônimas sobre os usuários para entender sua forma de interação com o site. Por exemplo, a partir da utilização do histórico de compras de um usuário é possível compará-lo com as compras realizadas por outros usuários com interesses semelhantes, possibilitando assim, fazer recomendações personalizadas (por exemplo, “os clientes que compraram este item também compraram ...”).



Figura 4: *Site da Amazon.com*

Fonte: <http://www.amazon.com/>



Os ambientes informacionais analisados intensificam o uso das ferramentas colaborativas por meio de atividades sociais eletrônicas que possibilitam a participação, bem como a adição de conteúdos e valores informacionais construídos pelos próprios usuários, que a partir desse processo estabelecem as regras a serem utilizadas pelas comunidades de prática. Estas, então, são incorporadas pelas bibliotecas, repositórios digitais, comunidades virtuais e/ou sites de comerciais, em um determinado nível de representação, descrição, recuperação e disseminação do conteúdo informacional.

A *Web 2.0* nas bibliotecas pode ser uma ferramenta que possibilite a gênese de uma base de conhecimento a partir da inteligência coletiva, como também uma ferramenta interativa e facilitadora para a gestão e descoberta do conhecimento.

Hoje, não só o usuário participa como também adiciona conteúdos que, ao serem compartilhados em outros recursos colaborativos, passam a receber comentários dos membros vinculados à sua comunidade, agregando valor à informação compartilhada, além de criarem vários caminhos para a localização do próprio recurso. Assim, temos uma nova relação e interação na qual aumenta o nível sofisticado de atitudes sociais eletrônicas.

3 SERVIÇO DE RECOMENDAÇÃO EM AMBIENTES INFORMACIONAIS DIGITAIS

Em função do aumento do fluxo das informações, desde o início da década de 1990, os sistemas de recomendação evoluíram. Esses sistemas ajudam usuários a identificar um subconjunto de itens em uma grande quantidade de informação. Contrários, portanto, aos sistemas tradicionais de busca, que trabalham com palavras-chave comuns.

Os sistemas de recomendação operam na tentativa de encontrar itens que combinem com as necessidades do usuário e, para Serrano-Guerrero (2011), eles podem ser considerados como um sistema que provê serviço de informação personalizada em diferentes modos, reduzindo a sobrecarga de informação, considerando as preferências do usuário.

O autor ainda pontua a necessidade de se estudar quatro dimensões que caracterizam os sistemas de recomendação. Elas podem variar de acordo com o contexto em que o sistema se encontra inserido, sendo elas: modelada e projetada; alvo (para individual, grupo ou tópico); construído e mantido (*on-line* verso *off-line*).

Os serviços de recomendações estão amplamente difundidos em nosso contexto social, tais como: *sites* de busca, de *download*, de compra, de entretenimento, de relacionamento (comunidades sociais), entre outros. Por meio desses serviços, acompanhamos o fundamento existente na relação social entre seres humanos, reforçando a capacidade e a eficácia deste processo de indicação.

As relações sociais presentes nos recursos tecnológicos atuais investem na circulação das informações e num tempo fluente em que as coisas não ocorrem ao mesmo tempo. Assim, despertamos para a autonomia, a liberdade e um maior domínio informacional sobre o mundo.



Os serviços de recomendação permitem uma retroalimentação do sistema, tanto com informações internas como externas. Esse processo é alimentado pelo conteúdo organizado e processado pela biblioteca e também pela entrada do conteúdo social produzido pelos usuários a partir da atribuição de comentários ou avaliações, recomendações e *tags*.

Assim, as pesquisas geradas no contexto da *Web 2.0* utilizam redes sociais e são elaboradas em espaços de trabalhos compartilhados ou abrangem a mineração de dados sociais e processos de inteligência coletiva para melhorar o processo de pesquisa. Dessa forma, sistemas de recomendações intensificam a busca por informações com base nas opiniões ou atividades de outras pessoas.

Uma biblioteca digital científica, além de guardar informações com o histórico das transações de um usuário, armazena o próprio perfil e a identidade desse usuário, bem como as informações agregadas como as avaliações, comentários, recomendações e *tags*, que podem enriquecer a estrutura informacional existente no ambiente dessas bibliotecas.

Com os recursos tecnológicos disponíveis, a elaboração do perfil de usuário tornou-se mais enriquecedora, pois a tecnologia nos permite rastrear ações a partir do uso que se faz, bem assim registrar de forma identificada todas as suas ações, ou seja, as bibliotecas digitais não só gravam dados sobre o uso de seus usuários mas também os reutilizam, não importando o volume de dados a ser processado.

De acordo com Foulonneau e Riley (2008), os sistemas de recomendação e revisões são considerados uma categoria dentro dos sistemas de *Web 2.0* e *mashups* baseados na participação dos usuários, podendo ser utilizados nas aplicações de bibliotecas digitais.

Segundo as autoras, esses sistemas podem trabalhar com dois métodos: no primeiro, o sistema solicita um comentário textual e/ou numérico classificado pelo usuário; no segundo método, são utilizados padrões de uso para inferir que o usuário recomenda um produto.

Rastrear as informações dentro do ambiente acadêmico também favorece a segurança e a confiança nas informações repassadas, também a credibilidade de um recomendador torna-se bastante reforçada, se forem informações relativas à experiência pessoal com o material utilizado.

No ambiente acadêmico científico, os pesquisadores acadêmicos fazem uso substancial de recomendações referendadas por seus pares. Assim, nos estudos relacionados com a comunicação científica, é ressaltada a participação desses pesquisadores em conferências da área onde ocorrem reuniões informais que remetem a recomendações importantes.

Novos alunos, ainda não inseridos nesse contexto de relações estabelecidas por grandes pesquisadores, podem se beneficiar de um sistema de recomendação no ambiente da biblioteca digital ao utilizarem o sistema.

Ressaltamos, ainda, a importância do esclarecimento para o usuário da maneira como sua participação e colaboração são incorporadas ao sistema, para que tenha consciência dos benefícios da sua interação e do seu relacionamento com esse sistema, que gera recomendações a partir de seu próprio uso.

No contexto acadêmico, as recomendações trazem uma noção implícita do uso, pois para atribuir



uma crítica a um determinado livro, tem-se como base a leitura da obra pelo revisor. Assim, dados sobre o uso podem ser incorporados a uma recomendação com mais credibilidade e isso fortalece a ênfase aplicada aos artigos de revistas.

Para Grosser (1991), as recomendações incorporam a experiência e o julgamento de uma comunidade e são os meios pelos quais os membros podem se beneficiar das experiências anteriores de outros usuários.

A relação desses serviços pode ser identificada na similaridade entre itens e usuários, na influência das escolhas por meio de sugestões, na utilização de técnicas inteligentes e no aumento do consumo de produtos e serviços.

Por estarem inseridas nesse contexto, as bibliotecas estão agregando esses novos recursos a fim de proporcionar melhorias nos serviços prestados a seus usuários, que passam a colaborar com o próprio sistema.

Além disso, com a explosão informacional, as escolhas dos usuários, muitas vezes, tornam-se difíceis em função da quantidade de recursos e fontes informacionais que são disponibilizadas nas bibliotecas universitárias, por exemplo.

Os autores Resnick e Varian (1997) apresentam as principais técnicas de recomendação, sendo identificadas como as filtragens colaborativas (*collaborative filtering*), também conhecidas como *social filtering*, baseadas no compartilhamento de experiências entre indivíduos que apresentam interesses semelhantes; a filtragem baseada em conteúdo (*content-based filtering*), onde predomina a recomendação de objetos, baseando-se na correlação entre o conteúdo de itens e preferências dos usuários em relação a esses itens; e a técnica baseada nos algoritmos de filtragem, baseada em conteúdo, que trabalha com uma combinação das técnicas citadas acima, resgatando e reduzindo as vantagens e desvantagens de cada técnica.

Os sistemas de recomendação estão relacionados na identificação de similaridade entre itens e usuários, na influência das escolhas por meio de sugestões, na utilização de técnicas inteligentes e no aumento do consumo de produtos e serviços.

Cazzella et al. (2010), também reconhece o crescente número de técnicas com intuito de identificar padrões de comportamento (consumo, pesquisa e outros), bem como a utilização destes padrões na personalização do relacionamento com os usuários.

Camargo e Vidotti (2008) também identificam as técnicas de recomendações dentro dos processos específicos de personalização e customização através do item – Recomendar informações aos usuários – sendo relacionados à arquitetura da informação para repositórios científicos digitais.

A característica comum de todas essas abordagens é que a atividade de um usuário pode ter influência sobre os futuros usuários, reforçando, assim, a rede de relacionamento estabelecida entre os pesquisadores e os sistemas de recomendação.



4 CONCLUSÕES

Essas novas ferramentas de *software*, aliadas ao crescimento e à expansão das comunicações, oferecem uma oportunidade única para que especialistas em Ciência da Informação construam, juntamente com os usuários, serviços capazes de atender as necessidades de informação e também preservar e manter atualizada a memória da humanidade.

A utilização dessas tecnologias e aplicações *Web 2.0*, constitui uma mudança significativa e essencial na história das bibliotecas. A biblioteca tornar-se-á mais interativa e totalmente acessível.

Acreditamos que a melhor concepção da Biblioteca 2.0, neste momento, seria uma rede social cuja interface fosse construída pelo usuário, relacionada com a funcionalidade e demonstrada por sua popularidade, de característica multifacetada, social, flexível, dinâmica, rápida, simples e pronta para uso.

A *Web 2.0* nas bibliotecas mostra-se uma ferramenta essencial na concepção de uma base de conhecimento a partir da inteligência coletiva, como também uma ferramenta interativa e facilitadora na descoberta e na gestão do conhecimento.

Esses sistemas passam a ser mais uma possibilidade no auxílio à seleção de informação, por reunirem tanto o histórico de uso do sistema, como o histórico de usuários parecidos, que juntos, representam os pontos fortes existentes no conteúdo informacional disponibilizado à comunidade de usuários.

Destacamos também os serviços de recomendação que, dependendo da técnica aplicada em ambientes informacionais, podem ser considerados um relato objetivo do uso e expressão da preferência de sua comunidade, que contribuem e participam a ponto de formarem uma enorme massa de dados com condições de representá-la diante de suas relações sociais.

A adoção de serviços de recomendação em bibliotecas digitais possibilita maior integração entre os membros de uma comunidade acadêmica, estabelecendo vínculos com o sistema a partir da efetiva colaboração dos usuários. Essa interação, baseada na produção de conteúdo social (comentários, avaliações e atribuição de *tags*, etc.), é importante, pois regenera o próprio sistema.

RECOMMENDATION SERVICES IN SCIENTIFIC DIGITAL LIBRARIES

Abstract: *This study presents the theme of the service recommendation in the context of digital library resources using Web 2.0 services featured recommendation applied to the use of scientific information produced in digital information environments, such as scientific digital libraries. Thus, it is intended to highlight based on the literature as the strength of the academic community and collaboration can generate recommendations through the use and production of social content. To develop this study, we observed the operation of collaborative technology resources (Web Resource 2.0) used in digital information environments that extend the concept of Library 2.0, so based on the literature was portrayed the importance of services in the recommendation scientific context of digital libraries, we opted for qualitative research with direct observation of non-participant to perform the identification and description of the functioning of the techniques in widespread political, economic, and social, as search engines, download, purchase, entertainment, relationships (social communities) applied to the digital*



environment university libraries. The integration of these technologies presents itself as an innovation, but its nature is bound to tradition and mission of libraries. However, with the information explosion users often find it difficult to carry out their choices and discoveries in the quantity of informational resources and sources available in the environment of scientific digital libraries, for example. The features of Web 2.0 are providing an information environment more flexibly structured, attractive and dynamic, where the user becomes the main actor in the construction of their environment, enabled by customization and personalization features. It also allows the discovery of information content and the generation of scientific information by the users who follow the system fed back again because the recommendation service accompanies human relationships, which enriches the information structure exists in these environments.

Keywords: Digital Library. Web 2.0. Collaborative Technology. Recommendation Service

REFERÊNCIAS

Amazon.com: Online Shopping for Eletronics, Apparel, COmputers, Books, DVD & more, 2011, disponível em: <<http://www.amazon.com/>>. Acesso em 22 de julho de 2011.

BORGMAN, C. L. Social aspects of digital libraries. IN: UCLA/NFS Workshop on Social Aspects of Digital Libraries, 1996, Los Angeles. **Final Report**. Los Angeles, ULCA/NFS, 1996. Disponível em <http://www-lis.gseis.ucla.edu/DL/>. Acessado em: 22 de julho de 2011

CAMARGO, L.; VIDOTTI, S. Arquitetura da informação para biblioteca digital personalizável. **R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, Santa Catarina, 2008.

CAZELLA, S. C., NUNES, M. A. S. N. & REATEGUI, E. **A ciência do palpite:** estado da arte em sistemas de recomendação, PucRIO, 2010

CASEY, M.; SAVASTINUK, L.. Library 2.0, 2006, **Library Journal**. Disponível em: <<http://www.libraryjournal.com/article/CA6365200.html>> Acesso em 10 de março de 2011.

FOULONNEAU, Muriel; RILEY, Jenn. The future of metadata. In: _____. **Metadata for digital resourcers:** implementation, systems design and interoperability. Oxford: Chandos publishing, 2008. cap.12 p.187-197

GROSSER, K. Human networks in organizational information processing, **Annual Review of Information Science and Technology**, 26, 349-402, 1991.

LibraryThing.com, 2011. Disponível em:<<http://librarything.com>> . Acesso em 22 de julho de 2011

MERLOT: Multimedia Educational Resource for Learning and Online Teaching, 2011. Disponível em: <<http://www.merlot.org/>>. Acesso em: 22 de julho de 2011.

PEDRONETTE, D; TORRES, R. da S. Uma Plataforma de Serviços de Recomendação para Bibliotecas Digitais. In: Simpósio Brasileiro de Banco de Dados (SBBDD), 23, 2008, Campinas, São Paulo. **Proceedings**. [S.l.: s.n.], 2008.

REATEGUI, E. B.; CAZELLA, S. C. Sistemas de Recomendação. In ENIA 2005: Encontro Nacional de Inteligência Artificial. **Anais**, 2005 p. 306–349.



RESNICK, P.; VARIAN, H. R. Recommender Systems. **Communications of the ACM**. v.40, n.3, 1997. Disponível em: <http://delivery.acm.org/10.1145/250000/245121/p56-resnick.pdf?ip=200.145.3.29&CFID=33532844&CFTOKEN=40486464&__acm__=1312053348_e4d9c8d975d4cdc71c9c8a83dccb7f5e> Acesso em 22 de julho de 2011.

SERRANO-GUERRERO, J. et. al. A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0. **Information Sciences**. v.181, p. 1503-1516, 2011

The University of Iowa libraries, 2011, disponível em: <<http://www.lib.uiowa.edu/>>. Acesso em 15 de março de 2011.



CLIO-I : PRIMANDO PELA USABILIDADE E ACESSIBILIDADE EM UM SISTEMA PARA GERENCIAMENTO E INTEROPERABILIDADE DE REPOSITÓRIOS DIGITAIS

*Marcos Cardoso Jr., Sandra de Albuquerque Siebra, Marcos Galindo, Vildeane Borba,
Majóry Miranda*

Resumo: O interesse na criação de Repositórios Digitais cresceu significativamente a partir do surgimento e disseminação da Web, que trouxe consigo a necessidade de ferramentas que facilitassem a publicação, gerenciamento e recuperação da informação digital. Neste contexto, este artigo apresenta o Clio-i, um sistema para gerenciamento e interoperabilidade de Repositórios Digitais que possui um visualizador de documentos que prima por características de usabilidade e acessibilidade. Também, são apresentados estudos de caso de uso do sistema mostrando alguns contextos onde ele pôde ser aplicado com sucesso. Adicionalmente, é apresentado o resultado de uma pesquisa de satisfação com usuários do sistema. O resultado da pesquisa mostra que mesmo pequenas modificações já podem implicar em benefícios diretos para o usuário, tendo como consequência um maior interesse pelo uso do sistema.

Palavras-chave: Repositórios Digitais. Interoperabilidade. Usabilidade.

1 INTRODUÇÃO

De acordo com o glossário do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) / Ministério da Ciência e Tecnologia, repositórios digitais são sistemas de informação que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidades científicas. Incentivam e gerenciam a publicação pelo pesquisador (auto-arquivamento), utilizam tecnologia aberta e podem ser acessados por diversos provedores de serviços nacionais e internacionais. Em geral, repositórios digitais são desenvolvidos com a intenção de atender necessidades urgentes de uma comunidade específica, e assim, poucos desses serviços são planejados como sistemas que possam ser reusados em diferentes contextos (CARDOSO JUNIOR, 2007). Assim, torna-se importante a criação de sistemas de propósito geral que diminuam os custos de instalação e manutenção, tenham uma boa usabilidade, primem pela acessibilidade às informações e facilitem a implantação e gerenciamento de repositórios digitais.

Neste contexto, este artigo objetiva apresentar o Clio-i, um software livre para gerenciamento e acesso a Repositórios Digitais que oferece ao usuário uma ferramenta de indexação e busca automática por documentos e um módulo de visualização que permite operações sobre os documentos eletrônicos para melhorar sua visualização. Essas funcionalidades aumentam significativamente a usabilidade dos



documentos digitais. Adicionalmente, o Clio-i proporciona interoperabilidade com outros repositórios digitais que façam uso do protocolo OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), que é um padrão internacional para interoperabilidade de repositórios digitais.

O Clio-i começou a ser desenvolvido em 2005, no Líber (Laboratório de Tecnologia do Conhecimento da Universidade Federal de Pernambuco)¹, com apoio do Ministério de Educação (MEC) e convênio firmado entre o Centro de Documentação e de Estudos da História Brasileira (CEHIBRA) da Fundação Joaquim Nabuco (FUNDAJ)² e a UFPE. Dentro do contexto do projeto houve uma dissertação de mestrado defendida (CARDOSO JUNIOR, 2007) e em 2006 foi lançada a primeira versão do protótipo funcional que foi aprimorada e atualizada em 2007, com a inclusão dos protocolos de interoperabilidade. Desde então, não havia sido feita mais nenhuma modificação no sistema e o projeto ficou suspenso até junho de 2011, quando começou a ser retomado por uma nova equipe do Liber em parceria com o Centro de Informática da UFPE (Cin/UFPE)³. Entre os aprimoramentos a serem feitos no sistema estão a implementação das regras de acessibilidade sugeridas pela W3C (<http://www.w3.org/standards/webdesign/accessibility>), o aprimoramento da usabilidade de acordo com a ISO 9241 (ISO 9241, 1998 apud SANTOS, 2008), a taxonomia de Ferreira e Leite (2003) e as recomendações de Presmann sobre requisitos não-funcionais de usabilidade da interface de um sistema (PRESSMAN, 1992 apud FERREIRA; LEITE, 2003).

Atualmente, o Clio-i é utilizado em algumas instituições como, por exemplo, a FUNDAJ (Fundação Joaquim Nabuco) e a instituição Save The Children (<http://www.savethechildren.se/>) da Suécia e em diversos projetos como, por exemplo, o acervo sobre o Museu Murilo Lagreca (<http://www.liber.ufpe.br/lagreca>), o acervo do memorial popular Liedo Maranhão (<http://www.liber.ufpe.br/liedo>) e o acervo de obras holandesas sobre o Brasil (<http://liber.ufpe.br/visaoholandesa>).

O restante deste artigo está dividido como segue. Na seção 2 é dada a fundamentação teórica necessária para a completa compreensão do artigo. Na seção 3 são descritos: a arquitetura escolhida para desenvolvimento do Clio-i e o diferencial do Clio-i com relação a outras ferramentas e sistemas similares para a criação de repositórios digitais. Na seção 4 é apresentado o protótipo desenvolvido para avaliar a arquitetura proposta. Na seção 5 são apresentados dois estudos de caso de uso do Clio-i em diferentes contextos de utilização. Na seção 6 é descrita a pesquisa de satisfação realizada com usuários do Clio-i. Por fim, na seção 7 são apresentadas as conclusões deste artigo e são apontados melhorias e ajustes que estão programados para serem realizados no sistema, de forma a aprimorar sua usabilidade, enquadrá-lo dentro das regras de acessibilidade da W3C e torná-lo mais robusto.

1 O Liber (www.liber.ufpe.br) é um espaço de pesquisa ligado ao Departamento Ciência da Informação da UFPE, criado em 1996, como grupo de pesquisa e formalizado em 2006 como Laboratório de Tecnologia do Conhecimento. O Liber dedica-se a investigações no campo do gerenciamento eletrônico do conhecimento. Para este fim vem desenvolvendo, em ambiente controlado, repositórios e ferramentas que permitem a disponibilização, gerenciamento e pesquisa de conteúdos em formato digital.

2 URL: www.fundaj.gov.br/

3 URL: <http://www.cin.ufpe.br>



2 REPOSITÓRIOS DIGITAIS

Segundo Viana, Márdero Arellano e Shintaku (2005), um repositório digital é uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado. Essa estratégia foi possibilitada pela queda nos preços no armazenamento, pelo uso de padrões como o protocolo de coleta de metadados da Iniciativa dos Arquivos Abertos (OAI-PMH)⁴. (SOMPEL et al, 2004) e pelos avanços no desenvolvimento dos padrões de metadados que dão suporte ao modelo de comunicação dos arquivos abertos. Viana, Márdero Arellano e Shintaku (2005) e Souza, Cruz e Braga (2008) evidenciam ainda que os repositórios digitais incentivam a publicação na Rede gerenciada pelo pesquisador (auto-arquivamento); utilizam tecnologias abertas (open source); seu conteúdo está em regime de acesso aberto; há garantia de preservação digital do conteúdo a longo prazo (memória da produção científica); têm sistema de gestão integrado com outros serviços; promovem interoperabilidade com sistemas; usam padrões universais, como o protocolo OAI e as informações ficam disponíveis para serem acessadas, permanentemente, por diversos provedores de serviços a nível nacional e internacional.

Segundo o site OpenDOAR (Directory of Open Access Repositories)⁵ nos últimos anos, tem aumentado o número de repositórios de acesso aberto, principalmente, dentro da comunidade universitária. É de responsabilidade e interesse das universidades que seus membros disponibilizem os seus trabalhos nesses repositórios institucionais. Esse interesse é também partilhado pela comunidade científica, pois ao disponibilizarem os seus trabalhos, estão contribuindo para que a sua instituição possa ganhar prestígio e credibilidade no meio da comunidade científica, além de contribuir para o progresso científico. (CARVALHO et al, 2006). Outros tipos de repositórios, além dos científicos/institucionais, também são fontes importantes de informação, porém, com finalidades e público alvo diferentes. Na verdade, os repositórios digitais podem ser implantados e operacionalizados em organizações do setor público ou privado, podendo ser acadêmicos, administrativos, técnicos ou mistos, mantendo a mesma confiabilidade de informação. (SHINTAKU; MEIRELLES, 2010)

2.1 Sistemas para criação de repositórios digitais

Existem diversos sistemas que aplicam-se à ampla maioria dos contextos institucionais, permitindo uma variedade de funcionalidades e recursos para gestão de documentos digitais. Entre eles, podemos citar: EPrints (<http://software.eprints.org/>); Fedora (<http://www.fedora.info>); Archimède (<http://archimede.bibl.ulaval.ca/>); MyCoRe (<http://www.mycore.de>); Digital Commons (<http://digitalcommons.bepress.com/>); Greenstone Digital Library Software (<http://www.greenstone.org/>); dLibra (<http://dlibra.psnc.pl/>) e DSpace (<http://www.DSpace.org>).

Segundo o OpenDOAR, o sistema mais utilizado, atualmente, é o DSpace (Ver Figura 1). O DSpace pertence à categoria dos Repositórios Digitais (RI) e é o resultado de um esforço conjunto do Massachusetts Institute of Technology (MIT) e da Hewlett-Packard (HP) e apoiado no Brasil pelo IBICT,

4 Maiores informações em: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

5 URL: <http://www.opendoar.org>



desenvolvido com a finalidade de coletar, preservar e disseminar documentos digitais produzidos por docentes e pesquisadores do MIT. Posteriormente, adaptado por um conjunto de oito universidades americanas, o DSpace é hoje disponibilizado sob a forma de um produto open source, que pode ser livremente adaptado e expandido funcionalmente, nos termos da BSD (Berkeley Standard Distribution License open source license). Utiliza metodologia de arquivos abertos (Open Access Initiative) e padrões internacionais, como o Dublin Core (CARVALHO et al, 2006). Por ter como unidade os objetos digitais, o DSpace permite vários tipos de objetos e tipos de arquivos, tais como: texto, imagem, áudio e vídeo para os tipos de objetos digitais e todos os tipos mais comuns de arquivos como: PDF, DOC, TXT (para texto), JPG, GIF (para imagens), etc.

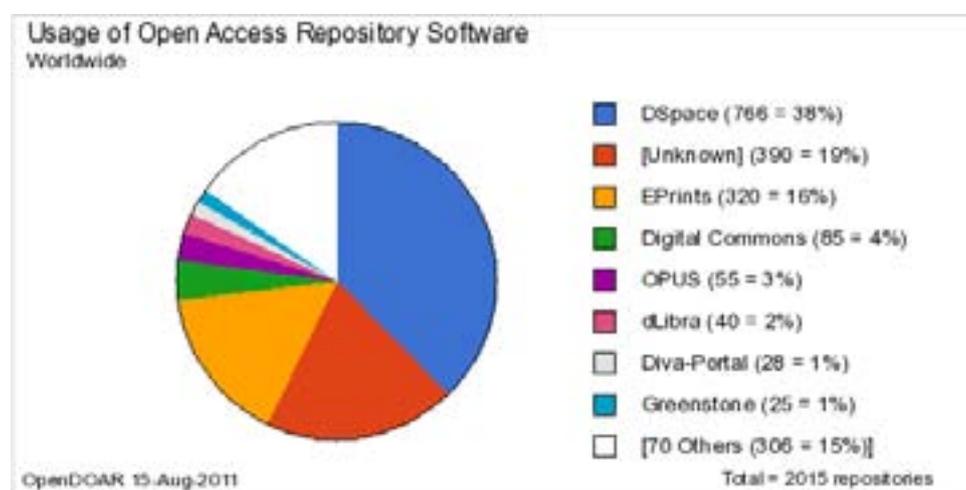


Figura 1: Estatísticas de uso dos softwares para criação de repositórios digitais
Fonte: OpenDoar - 15/08/2011

O DSpace é o software adotado como referência no Brasil pelo IBICT e, a partir dessa adoção, várias iniciativas brasileiras passaram a utilizar o software DSpace como ferramenta de implementação de seus repositórios institucionais, tais como: Biblioteca Digital do Senado (<http://www2.senado.gov.br/bdsf/>); Repositório Institucional da Universidade de Brasília (<http://repositorio.bce.unb.br/>); Repositório Digital da UFRGS (<http://www.lume.ufrgs.br/>); Infoteca-e da EMBRAPA (<http://www.infoteca.cnptia.embrapa.br/>); entre outras.

Alguns problemas que puderam ser identificados no DSpace a partir da vivência dos autores com relação ao DSpace são:

- O DSpace foi desenvolvido com base na comunicação científica, seu projeto foi embasado na disseminação de literatura científica em formato, principalmente, de artigos que foram publicados anteriormente em periódicos. Por essa razão, sua organização está intimamente ligada a sua origem acadêmica. Assim, alguns conceitos orientadores dos repositórios remetem à comunicação científica. Logo, não há ferramentas diferenciadas para visualização e manipulação de, por exemplo, imagens digitalizadas (o que poderia ser útil, por exemplo, no contexto do acervo de museus e instituições).
- Como um software altamente configurável, o DSpace possui várias opções que lhe permitem ajustar-



se às diversas necessidades de uma instituição. Esse benefício é também uma dificuldade. Visto que a instalação e configuração do ambiente não é trivial e exige conhecimento técnico da equipe envolvida.

- As tarefas dos administradores requerem conhecimentos específicos do DSpace que exigem leitura de manuais ou treinamento apropriado.
- Não houve preocupação com acessibilidade e usabilidade na criação do DSpace
- É um software com muitas opções e muito amplo para atender a construção de qualquer tipo de repositório institucional. Por isso mesmo, se torna um pouco complexo para ser utilizado por iniciativas menores, como o acervo digital de museus, bibliotecas e instituições ou empresas públicas ou privadas.

3 O PROJETO DO SISTEMA CLIO-I

O Clio-i - sistema para gerenciamento e interoperabilidade entre Repositórios Digitais, oferece ao usuário uma ferramenta de busca automática por documentos, e um módulo de visualização que permite operações sobre os documentos eletrônicos para melhorar sua usabilidade. Outra característica importante do Clio-i é a adoção do protocolo OAI-PMH para prover a Interoperabilidade com outros repositórios. O uso do protocolo OAI-PMH implica que qualquer repositório que atenda as exigências do protocolo pode ter seus metadados coletados pelo Clio-i. Além disso, os metadados de um repositório local gerenciado pelo Clio-i poderão ser exportados para Provedores de Serviços OAI. A principal idéia por trás da criação do CLIO-I é que ele fosse um sistema fácil de instalar, configurar e manter, que não exigisse muita preparação da equipe envolvida na sua utilização e manutenção.

Na Figura 2, é possível visualizar o fluxo de informações do Clio-i. O sistema se comunica com uma base de dados, que armazena documentos digitais e seus metadados. Os metadados dos documentos podem ser inseridos localmente por um administrador ou podem ser coletados de outros repositórios (através de requisições a Provedores de Dados OAI externos). A definição dos repositórios acessados pelo Clio-i é realizada pelo Administrador do sistema. Os metadados e documentos inseridos do banco de dados são apresentados aos usuários através de serviços do Clio-i (e.g., módulo de Recuperação de Informação). Por fim, o sistema também pode disponibilizar os metadados depositados na sua base, para coleta por parte de Provedores de Serviço OAI.

3.1 Arquitetura Geral

A arquitetura geral do sistema Clio-i conta com quatro módulos principais (Ver Figura 3) que serão descritos a seguir.

Recuperação de Informação - O módulo de Recuperação de Informação tem a finalidade de responder as consultas dos usuários por documentos referenciados na base de dados do Clio-i. Os registros são retornados em uma ordem de relevância, considerando tanto consultas aos metadados



que descrevem os documentos, como ao conteúdo dos documentos (no caso de documentos de texto).

Visualização de Documentos - Após o usuário realizar a pesquisa no Clio-i, os registros recuperados são listados com os metadados de cada documento. A partir daí, o usuário pode visualizar um documento específico de sua escolha, através do Visualizador de Documentos. O Visualizador de Documentos do Clio-i trabalha com arquivos do tipo áudio, imagem, vídeo e texto, e fornece ao usuário mecanismos para manipular o documento de texto e imagem de maneira rápida e simples.

Clio-i Data Provider - Um Provedor de Dados OAI é um sistema que exporta os metadados da sua base seguindo o protocolo OAI-PMH. Assim, esse módulo tem como função básica exportar os metadados da base de dados do Clio-i de acordo com o protocolo OAI-PMH. Os metadados poderão, assim, ser coletados por Provedor de Serviços OAI externos.

Clio-i Service Provider - Um Provedor de Serviços OAI realiza a coleta de diferentes repositórios que estejam de acordo com o protocolo padrão OAI-PMH (versão 2.0). Assim a função principal desse módulo é coletar metadados disponibilizados por outros Provedores de Dados OAI externos e armazená-los na base de dados do Clio-i.

4 O PROTÓTIPO IMPLEMENTADO

Esta seção descreve alguns detalhes de implementação do protótipo atual do Clio-i. Este protótipo foi utilizado nos estudos de caso apresentados na seção 4. O protótipo foi implementado em PHP (GILMORE, 2009), uma linguagem adequada para criação de páginas dinâmicas na Web. O sistema Clio-i pode ser utilizado facilmente em computadores que possuam servidores Web (e.g., Apache Server 2.0). Por se tratar de um sistema WEB, houve a preocupação de o sistema funcionar em diversos browsers, como o Internet Explorer e o Mozilla FireFox.

4.1 Base de Dados

O banco de dados escolhido para o armazenamento dos metadados foi o MySQL(<http://www.mysql.com>), um dos bancos de dados open-source mais populares do mercado. Esse software fornece recursos simplificados e apropriados para as suas aplicações, tendo um custo extremamente reduzido. Desta forma, destacamos as seguintes características para a escolha do MySQL: confiabilidade e bom desempenho, facilidade de utilização e distribuição, recursos e suporte para as mais diversas plataformas (WELLING; THOMSON 2004).



O sistema Clio-i usa como base de seu conjunto de metadados os campos descritores definidos no padrão Dublin Core (DC) (ROSETTO; NOGUEIRA, 2002) (ver quadro 1). O DC é um padrão internacional criado com o objetivo de fornecer um conjunto simples e intuitivo de metadados para a descrição de documentos digitais em contextos de um determinado domínio. Adicionalmente, o Dublin-Core é o padrão de metadados sugerido pelo protocolo OAI-PMH para permitir a unificação de metadados de diferentes repositórios. Além dos metadados do DC, cada documento é associado ainda a informações específicas dos objetos digitais como resolução, duração (no caso de áudio e vídeo) e seqüência de páginas (para texto e imagem).

Elemento Dublin Core	Descrição
<i>Title</i>	Título do documento
<i>Creator</i>	Autor do documento
<i>Description</i>	Resumo do documento
<i>Type</i>	A natureza ou gênero do conteúdo. Descreve categorias gerais de documentos.(e.g., iconografia, cordéis, etc.).
<i>Coverage</i>	Onde o documento foi produzido / impresso
<i>Publisher</i>	A entidade responsável por tornar o documento disponível na sua forma atual.
<i>Date</i>	A data da produção / impressão do documento
<i>Subject</i>	As palavras-chaves que identificam o assunto documento
<i>Contributor</i>	Entidade responsável por contribuições ao conteúdo do recurso (e.g., tradutor, revisor, ilustrador, etc.).
<i>Source</i>	Documento do qual o presente documento foi derivado.
<i>Language</i>	Principal idioma do conteúdo do documento.
<i>Format</i>	Manifestação física ou digital do documento (e.g., jpeg, tiff, mp3).
<i>Rights</i>	Copyright, direitos do autor, propriedade intelectual.
<i>Relation</i>	Referência para um documento relacionado.
<i>Identifier</i>	Uma referência não ambígua que identifica o documento (e.g., uma URL).

4.2 Módulo de Recuperação de Informação

Nesse módulo, foi utilizado um recurso do banco de dados MySQL, chamado MySQL Full-Text Search (versão 3.23.23). Esse recurso permite o casamento de palavras-chave de consulta e

expressões booleanas com os campos do banco de dados. Os resultados da busca são ordenados segundo um critério de relevância. A Figura 4 apresenta um exemplo do uso das funções MATCH() e AGAINST() em uma cláusula WHERE de uma consulta SQL. Nesse exemplo, os campos de metadados pesquisados (i.e. título e resumo) são definidos na função MATCH() e as palavras-chave (i.e. Brasil) na função AGAINST().

Para cada resultado da consulta, a função MATCH() retorna um valor de relevância. Essa relevância é medida através da frequência com que os termos de consulta são encontrados nos campos do documento, ponderados pela frequência com que os termos estão presentes na base de documentos (BAEZA-YATES; RIBEIRO-NETO, 1999).

```
SELECT * FROM documentos WHERE MATCH (título,resumo) AGAINST ('Brasil')
```

id	título	resumo
5	O Brasil descoberto	...no brasil, o pau-brasil foi...
10	Tempos de crise	O Brasil de 1964...
7	A fonte de riqueza	...em meados de 80, o Brasil...

Figura 4: Uma busca utilizando o MySQL Full-Text Search

Fonte: Os autores, 2011

O MySQL Full-Text Search permite ainda que se trabalhe com buscas full-text booleanas, utilizando o modificador IN BOOLEAN MODE. Na Figura 5, temos um exemplo de consulta que recupera todos os documentos que contenham a palavra Brasil, mas que não possuam o termo ditadura.

4.3 Visualizador de Documentos

O visualizador de documentos foi implementado utilizando DHTML (Dynamic HTML). O DHTML é um método interativo de visualização de informação. A tecnologia foi empregada tanto nas operações sobre os documentos, quanto na construção das janelas móveis resultando num módulo satisfatório em termos de usabilidade e velocidade.

Na Figura 6, é apresentada a interface do módulo de Visualização de Documentos usada no Acervo Digital da Fundação Joaquim Nabuco. Os documentos são apresentados na tela central. À esquerda da tela, uma barra de ferramentas oferece opções para o usuário realizar operações sobre a imagem do documento sendo visualizado (e.g., zoom, giro vertical e horizontal, entre outros.). Em muitos casos, essas operações são essenciais para possibilitar uma boa leitura dos documentos apresentados (e.g., imagens de documentos manuscritos ou mal conservados), contribuindo assim para aumentar a usabilidade do repositório digital.

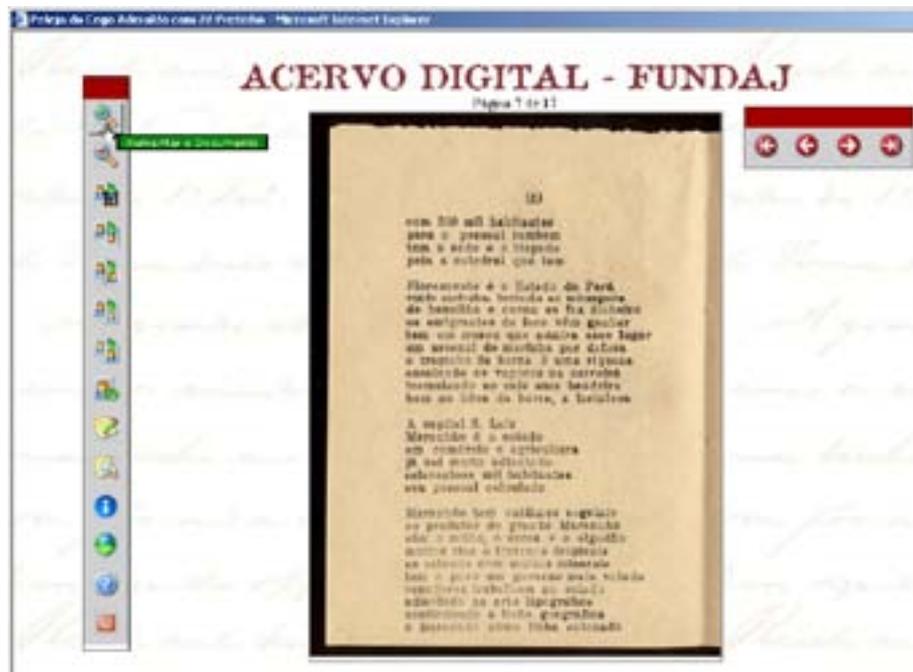


Figura 6: Apresentação de documento no *Visualizador*
Fonte: Os autores, 2011

Os efeitos visuais disponíveis podem ser aplicados sobre documentos do tipo texto ou imagem, e incluem as seguintes opções: (1) operações de Zoom, para aumentar ou diminuir progressivamente o tamanho da imagem do documento; (2) negativar, para uma melhor visualização, em alguns casos é necessário negativar o documento; (3) girar vertical ou horizontalmente, para facilitar a leitura, por exemplo, de documentos que foram digitalizados em posição invertida; (4) tornar mais ou menos opaco, permite clarear ou escurecer progressivamente o documento; (5) restaurar, após aplicar diversas operações, o usuário tem a opção de restaurar a configuração original do documento. A barra de ferramentas apresenta ainda outras opções aplicáveis para qualquer tipo de documento. Essas opções dizem respeito às possibilidades de interação do usuário com o sistema, e incluem inserir notas sobre o documento, ler as notas inseridas, fazer download em formato PDF (no caso de textos e imagens) e um menu de ajuda. Este módulo oferece ainda uma barra de navegação (canto superior direito da figura 6), utilizada para “folhear” os documentos de texto ou imagem que contenham mais de uma página.

4.4 Clio-i Data Provider

Para este módulo, um conjunto de funções foi implementado e as principais estão representadas na Figura 7, que apresenta um fluxograma do funcionamento deste módulo cujas principais funções são descritas a seguir

Parser. Conjunto de funções com a finalidade de validar as requisições em HTTP vindas de um Provedor de Serviços OAI.

Gerador XML. Caso a requisição HTTP tenha sido aprovada pelo parser, a função respostaXML retorna uma string no formato XML com todas as informações requisitadas. Como a quantidade de registros de metadados exportados pode ser grande, o módulo fornece a opção de retornar apenas 100 registros de cada vez na sua exportação. Ou seja, se a resposta da solicitação possuir mais do que esse limite estabelecido, o componente Controle de Fluxo do Clío-i Data Provider é acionado.

Controle de Fluxo. Como dito anteriormente, caso a requisição validada possua mais de 100 registros em sua resposta, esse componente é ativado através da função respostaResumption. Essa função retorna a mesma resposta em XML que é gerada pelo componente anterior. A diferença é que a função do Controle de Fluxo recebe um parâmetro a mais, que é justamente o resumptioToken, responsável por realizar a paginação dos resultados em XML das requisições.

Gerador de erros. Este componente é chamado caso o parser retorne falso para a validação da requisição HTTP. Através da função selecionaErro, é verificado que erro foi cometido e sua resposta em formato XML é apresentada.

Clío-i Service Provider. A implementação do Clío-i Service Provider foi feita com base nas funções apresentadas no fluxograma da figura 8 e detalhadas a seguir.

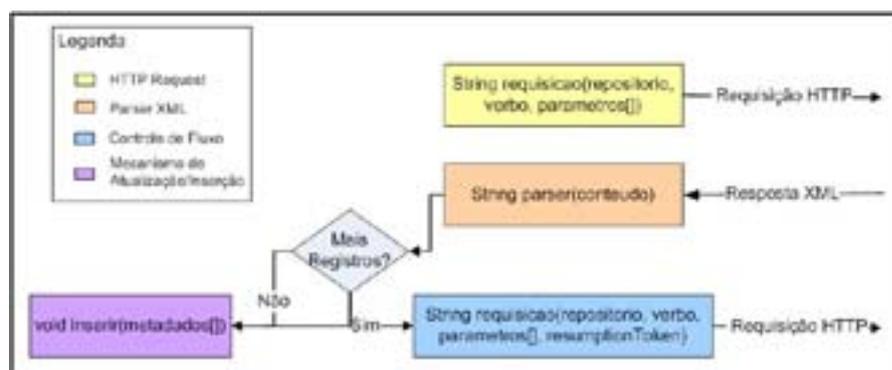


Figura 8: Principais funções e componentes do Clío-i Service Provider
Fonte: Os autores, 2011

HTTP Request. Passo inicial para a coleta dos dados, este componente realiza a requisição para um Provedor de Dados previamente cadastrado no sistema. O HTTP Request possui a função requisição, onde passamos como parâmetro o repositório escolhido, o verbo e seus possíveis argumentos. Dessa maneira, a função seleciona a URL Básica para coleta do repositório e monta a requisição com base nos parâmetros passados pelo administrador do sistema.



Parser XML. Este componente valida o conteúdo XML enviado por um Provedor de Dados, e faz uma chamada ao componente Mecanismos de Inserção/Atualização para inserir os dados no banco MySQL. O Parser XML ainda verifica se há mais requisições a serem feitas (i.e. se o Provedor de Dados envia metadados usando Controle de Fluxo). Caso ainda haja porções de registros a serem coletadas do repositório selecionado, o componente Controle de Fluxo é chamado.

Controle do Fluxo. Após a confirmação que ainda existem registros a serem coletados, a função requisição é acionada, essa do componente Controle de Fluxo. Essa função é similar a encontrada no HTTP Request, tendo como diferença o parâmetro adicional `resumptionToken` que controla o fluxo de registros coletados.

Mecanismos de Atualização/Inserção. Recebe do parser XML os valores necessários para o componente realizar a atualização da base de dados do Clio-i Service Provider. Através da função `inserir`, os metadados coletados são inseridos ou atualizados, caso o registro coletado já tenha sido armazenado na base alguma outra vez. Ressaltamos que essa atualização é necessária para que a base do Clio-i Service Provider esteja de acordo com a base do Provedor de Dados em questão.

5 ESTUDOS DE CASO

Versões anteriores do Clio-i foram aplicadas em diferentes estudos de caso de Repositórios Digitais (CARDOSO JÚNIOR, 2007). A diferença do estado atual do Clio-i para suas versões anteriores é o uso do protocolo OAI-PMH para interoperabilidade de repositórios. Desta forma, nessa seção apresentamos dois os estudos de caso em que o Clio-i foi usado com as funcionalidades dos módulos Clio-i Data Provider e Clio-i Service Provider.

5.1 Estudo de Caso 1: Integrador de Repositórios Científicos

O primeiro estudo de caso corresponde ao Integrador de Repositórios Científicos (ver Figura 9). O objetivo principal desse estudo de caso foi testar o módulo de coleta de dados (Clio-i Service Provider). Para isto, foram coletados metadados sobre artigos científicos oriundos de diversos provedores de dados, formando uma base de metadados centralizada.

Para tanto, foi feita inicialmente uma consulta à página dos Provedores de Dados oficiais da OAI, a fim de selecionarmos os repositórios científicos que seriam coletados pelo Clio-i Service Provider. Ao todo foram selecionados 19 Provedores de Dados OAI de diferentes nacionalidades. A lista dos 19 repositórios pode ser vista em (CARDOSO JÚNIOR, 2007)

Após essa seleção manual de repositórios científicos, deu-se continuidade à construção da



base de metadados. Cada um dos 19 repositórios selecionados foi cadastrado no sistema, e em seguida foi realizada a coleta de todos os metadados disponíveis nos repositórios. Após a coleta de todos os metadados, um total de 138.510 registros de artigos científicos foi reunido na base de dados do Clio-i. Essa base pode ser consultada pelos usuários através do módulo de Recuperação de Informação na URL: <http://www.liber.ufpe.br/clioi>.



Figura 9: Clio-i Integrador de Repositórios Científicos
Fonte: Os autores, 2011

Destacamos que o Provedor de Serviços desse estudo de caso foi registrado oficialmente no Open Archives Initiative, que conta atualmente com apenas 26 Provedores de Serviços cadastrados⁶. Vale salientar que o Clio-i foi o primeiro da América Latina a integrar a lista de sistemas que seguem o protocolo OAI-PMH.

5.2 Estudo de Caso 2: Integrador de Repositórios Multimídia

O segundo estudo de caso avaliou mais especificamente o módulo Clio-i Data Provider. Para isto, o Clio-i Data Provider foi utilizado para expor os metadados de três Bibliotecas Digitais que usam o Clio-i. Foram elas:

⁶ Open Archives Initiative Service Providers - <http://www.openarchives.org/service/listproviders.html>



Acervo Digital FUNDAJ: O acervo digital da Fundação Joaquim Nabuco utiliza os serviços de Recuperação da Informação e Visualização de Documentos do Clio-i desde o início do ano de 2006. Adicionamos ao sistema o módulo Clio-i Data Provider para expor os seus metadados e os seus recursos.

Holandeses na Bahia: Base composta essencialmente por documentos do tipo texto sobre a conquista holandesa de Salvador no Brasil em 1624, e a reconquista da cidade por uma armada luso-espanhola em 1625.

Escrito nas Estrelas: Este repositório agrega vídeos e áudio de uma exposição sobre o Brasil na época do início do desenvolvimento dos transportes aéreos, e dos correios no início do século XX. Para o presente trabalho, foram selecionados alguns desses vídeos, utilizados em um Clio-i Data Provider.

Os metadados dessas três bases foram integrados por um Clio-i Service Provider com o nome de Integrador de Repositórios Multimídia. Dessas três bases, o Acervo Digital da FUNDAJ foi o primeiro a ser registrado, oficialmente, como Provedor de Dados no site oficial do Open Archives Initiative⁷, que conta com mais de 500 repositórios registrados. Os outros dois repositórios foram registrados posteriormente.

6 ESTUDO DE SATISFAÇÃO DO USUÁRIO

Em qualquer sistema que vá ser utilizado por diferentes perfis de usuário e que seja importante o acesso a informações através dele, torna-se ainda mais relevante aliar recursos de acessibilidade e usabilidade à interface, na medida em que essa ação pode aproximar mais os usuários potenciais e reais para uso efetivo do mesmo. Isso porque quando estão realizando pesquisas, os usuários buscam respostas concretas e esperam interagir com o sistema de maneira rápida, simples e efetiva, sem necessitar de treinamentos, consultas à ajuda ou da leitura de manuais. Assim, foi feito um estudo inicial, simplificado, da usabilidade do sistema CLIO-I e da satisfação dos usuários com o mesmo.

O ajuste da usabilidade e acessibilidade permite o uso eficiente, produtivo, satisfatório, útil e seguro para o usuário do computador e destinatário da informação do conhecimento disponibilizado na rede. Ou seja, as condições de usabilidade e de acessibilidade podem assegurar que a informação disponibilizada na internet possa ser compreendida e utilizada por todos, independente de limitações (IGNÁCIO; CARVALHO, 2008).

O teste da interface foi realizado com 20 estudantes de graduação de uma Instituição de Ensino Superior (IES) que foram convidados a participar do teste (nenhum deles tinha tido

⁷ Open Archives Initiative Service Providers - <http://www.openarchives.org/Register/BrowseSites>



contato anterior algum com o sistema) e com 10 usuários dos acervos do Clio-i, atualmente, em funcionamento. Inicialmente, foi solicitado que cada usuário realizasse quatro atividades no Clio-i: 1) buscar documentos que tivesse a palavra Brasil; 2) acessar uma base de dados fora do sistema clio-i (disponibilizada através do protocolo OAI-PMH); 3) Manipular algum documento do tipo imagem com a ferramenta de Visualização e 4) fazer o download de algum documento recuperado.

O teste foi realizado individualmente por cada usuário e foi acompanhado por um observador que se valeu do método “pensando em voz alta”. Tal método consiste em uma observação direta de cada usuário utilizando o site. De acordo com Nielsen “ouvir os ‘pensamentos’ de cada usuário permite entender as razões de suas ações e essas informações são muito importantes”. (NIELSEN, 1993). Ao final do teste, foi feita uma entrevista com os usuários para levantar o grau de satisfação desses usuários. Foi pedido que eles pontuassem com notas de 0 a 10: 1) a qualidade dos resultados da busca usando o Clio-i; 2) a facilidade de baixar um arquivo pesquisado; 3) a forma de apresentação e manipulação das informações recuperadas; 4) a facilidade de utilização do sistema como um todo.

Todos os usuários envolvidos no teste conseguiram executar as ações solicitadas e comentaram que acharam a interface simples. Considerando a média aritmética das quatro questões pontuadas ao final do uso do sistema, o Clio-i ficou com média 8,69, que foi considerada muito boa, especialmente porque o sistema não sofre atualizações desde 2008 e possui diversas melhorias a serem feitas. Durante o teste os usuários comentaram que acharam muito diferente e interessante o módulo de visualização de documentos e as possibilidades que ele oferece. Alguns usuários durante a entrevista solicitaram novas funcionalidades e deram sugestões para o sistema que serão levadas em conta no novo ciclo de implementação que está apenas começando. O foco na simplicidade, usabilidade e acessibilidade, sem deixar de lado as funcionalidades básicas necessárias a uma ferramenta para criação de repositórios digitais são o grande diferencial do CLIO-I que pretendem ser mantidas na próxima versão a ser desenvolvida. Isso pode torná-lo interessante para ser utilizado por pessoas experientes e sem experiência na criação e utilização de repositórios digitais, especialmente no contexto de pequenas e médias iniciativas.

7 CONCLUSÕES E TRABALHOS FUTUROS

O Clio-i é uma ferramenta adequada, especialmente, para pequenas e médias iniciativas de repositórios digitais e para gerenciamento de acervos de museus. Isso devido a sua simplicidade de instalação e utilização para criação do repositório, a disponibilização de busca simples e avançada, a apresentação diferenciada da informação, ao seu módulo visualizador de documentos com opções de manipulação do mesmo e por seu desenvolvimento ter primado pela usabilidade. Adicionalmente, devido ao Clio-i ser um sistema implementado usando protocolos de interoperabilidade (OAI-PMH), é possível disponibilizar também, através dele, o acervo de outros repositórios que façam uso do protocolo OAI-PMH, o que amplia ainda mais as possibilidades de recuperação da informação.



O projeto CLIO-I está sendo retomado pela equipe do laboratório Líber em parceria como Centro de Informática da UFPE, a fim de aprimorar a ferramenta e torná-la uma opção simples e fácil de usar para pequenas e médias iniciativas que precisem montar o seu repositório digital. A idéia central será primar pela usabilidade, facilidade de uso e acessibilidade da ferramenta, tornando-a simples de usar inclusive por pessoas que não sejam da área de Ciência da Informação ou que não tenham grande conhecimento da tecnologia da informação e comunicação. Entre os trabalhos futuros a serem realizados no contexto da retomada do projeto CLIO-I estão:

- Possibilitar a leitura de metadados de imagens no formato JPEG. Isso se torna relevante a medida que várias imagens utilizadas hoje são originadas em câmeras ou filmadoras digitais que já imprimem na imagem os dados sobre o equipamento utilizado para obtenção da fotografia.
- Possibilitar a leitura de metadados nos formatos EXIF e IPTC
- Aprimorar o mecanismo de busca integrando conceitos de Web Social, Web Semântica e Findability. Neste contexto, instrumentos como ontologias ou taxonomias, possibilidade de criação de redes sociais dentro do repositório e variações das estratégias de pesquisa e uso da informação podem ser bastante úteis.
- Possibilitar a geração de tipos variados de relatórios gerenciais (ex: relatório de acesso e visualização de documentos) e infográficos (ex: pesquisas realizadas por perfil de usuário)
- Implementar a possibilidade de criptografar determinados arquivos que se classifiquem como confidenciais ou de acesso restrito, como medida de segurança da informação.
- Implementar as sugestões de acessibilidade da W3C.

Abstract: *The interest in creating digital repositories has grown significantly from the emergence and spread of the Web, which brought with it the need for tools to facilitate the publication, management and retrieval of digital information. In this context, this paper presents the Clio-i, a system for management and interoperability of digital repositories that has a document viewer that excels in usability and accessibility features. It also presents case studies of use of the system showing some contexts where it could be successfully applied. We also present the results of a survey of satisfaction with system users. The research result shows that even small changes may already result in direct benefits to the user, resulting in an increased interest in using the system.*

Keywords: Digital Repositories. Interoperability. Usability.



REFERÊNCIAS

BAEZA-YATES, R. A.; RIBEIRO NETO, B. **Modern Information Retrieval**. Massachusetts: Addison-Wesley, 2010.

CARDOSO JUNIOR, Marcos José de Menezes. **Clio-i: Interoperabilidade entre repositórios digitais utilizando o protocolo OAI-PMH**. 2007. 137f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-graduação em Ciência da Computação, Centro de Informática, Universidade Federal de Pernambuco, Recife, 2007. Disponível em: < <http://www.cin.ufpe.br/~rbcp/dissertacoes/dissertacaoCARDOSO.pdf>>. Acesso em: 15 ago. 2010.

CARVALHO, Maria Carmen Romcy de et al. O Repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília. In: CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA, 1. 2006. **Anais...** Disponível em: < http://DSpace.ibict.br/dmdocuments/Maria_Carmen_Romcy_de-Carvalho.pdf>. Acesso em: 12 jun. 2011.

FERREIRA, Simone B. L.; LEITE, Julio C. S. do P. Avaliação da usabilidade em sistemas de informação: o caso do sistema submarino. **Revista de Administração Contemporânea – RAC**, São Paulo, v.7, n.2, p.115-137, abr./jun. 2003. Disponível em: <http://www.anpad.org.br/rac/vol_07/dwn/rac-v7-n2-sbf.pdf>. Acesso em: 10 jun. 2011.

GILMORE, W. J. **Dominando Php e Mysql: do Iniciante ao Profissional**. Rio de Janeiro: Alta Books, 2009.

IGNÁCIO, Edilson A.; CARVALHO, José O. F. de. Avaliação da acessibilidade de sites oficiais de pesquisa no Brasil por pessoas com deficiência. **Encontros Bibli**, Florianópolis, v. 13, n.26, 2008. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/7193/6641>>. Acesso em: 10 jun. 2011.

NIELSEN, J. **Usability Engineering**. New York, NY: Academic Press, 1993.

OPEN DOAR. Disponível em: < <http://www.opendoar.org/>>. Acesso em: 10 jul. 2011.

ROSETTO, M.; NOGUEIRA, A. Aplicação de elementos metadados Dublin Core para descrição de dados bibliográficos on-line da biblioteca digital de teses da USP. In: SIMPÓSIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 12. São Paulo. **Anais...** 2002. Disponível em: < <http://www.sibi.ufjf.br/snbu/snbu2002/oralpdf/82.a.pdf>>. Acesso em: 10 Mar. 2010.

SANTOS, Rodrigo C. dos. Revisão das métricas para avaliação de usabilidade de sistemas. In: GLOBAL BUSINESS AND TECHNOLOGY ASSOCIATION CONFERENCE, 2008, Madrid. **Anais...** Espanha: GBTA, 2008. Com o título original de: Systems Usability Evaluation Metrics Review. Disponível em: <http://www.marcelomoraes.com.br/conteudo/marcelo/metricas_usabilidade.pdf>. Acesso em: 10 jun. 2011.

SHINTAKU, Milton; MEIRELLES, Rodrigo. **Manual do DSpace** : administração de repositórios. Salvador : EDUFBA, 2010.



SOMPEL, H. V. de; NELSON, M.L.; LAGOZE, C.; WARNER, S. Resource Harvesting within the OAI-PMH Framework. **D-Lib Magazine**, v. 10, n. 12, dez. 2004. Disponível em: <<http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html>>. Acesso em: 04 jan. 2010.

SOUSA, M. C. P.; CRUZ, M. A. L.; BRAGA, M. F. A. Acesso livre e repositório institucional: uma ferramenta indispensável nas Instituições de Ensino Superior. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 15., 2008, São Paulo. **Anais**. Disponível em: <<http://www.repositorio.ufma.br:8080/jspui/bitstream/1/256/1/artigo%20R.I.pdf>>.

Acesso em: 20 jan. 2009.

VIANA, C. L. M.; MÁRDERO ARELLANO, M. A.; SHINTAKU, M. **Repositórios institucionais em ciência e tecnologia**: uma experiência de customização do DSpace. 2005. Disponível em: <<http://eprints.radis.org/archive/00005563>>. Acesso em: 30 maio 2008.

WELLING, Luke; THOMSON, Laura. **Tutorial MYSQL**: Uma Introdução Objetiva aos Fundamentos do Banco de Dados MYSQL. Rio de Janeiro: Ciência Moderna, 2004.

W3C. **Recomendações de Acessibilidade para Conteúdo Web (WCAG) 2.0**. 2008. Tradução de Everaldo Bechara. Disponível em: <<http://www.ilearn.com.br/TR/WCAG20/>>. Acesso em: 10 jun. 2011.



O SISTEMISMO DE BUNGE: FUNDAMENTOS, ABORDAGEM METODOLÓGICA E APLICAÇÃO A SISTEMAS DE INFORMAÇÃO

Vinicius Medina Kern

Resumo: Sistemas de informação muito complexos são entendidos muitas vezes de forma fragmentada, parcial, até por seus protagonistas. Este artigo apresenta o sistemismo de Mario Bunge, uma abordagem sistêmica, como fundamento teórico-metodológico para a concepção e estudo de sistemas de informação. O artigo contrasta a visão de mundo individualista, característica da ciência analítica, com o holismo característico das abordagens “sistêmicas” anteriores a Bunge e introduz o sistemismo bungeano como abordagem que nega e, ao mesmo tempo, conjuga as visões individualista e holista. São apresentados os postulados do sistemismo, seu modelo de sistema e as regras metodológicas gerais, um conjunto de etapas metodológicas bastante amplas para a abordagem de sistemas – em especial, sistemas de informação. Uma aplicação dessa abordagem à Plataforma Lattes é oferecida como ilustração. A abordagem descrita e ilustrada pretende servir como roteiro metodológico aos interessados em empreendimentos de pesquisa em sistemas de informação com visão sistêmica. Adicionalmente, o artigo delinea alguns problemas de investigação sobre os mecanismos do crescimento e manutenção da Plataforma Lattes.

Palavras-chave: Sistemismo. Sistemas de informação. Plataforma Lattes. Metodologia da pesquisa.

1 INTRODUÇÃO

O estudo de sistemas complexos, tais como unidades e sistemas de informação, requer abordagens amplas, sistêmicas, mas a efetividade das teorias de sistemas mais conhecidas é duvidosa. Como compreender grandes sistemas de informação?

Este artigo apresenta o sistemismo de Bunge (1997, 2000, 2003, 2004, 2006) como fundamento ontológico-metodológico para a construção e estudo de sistemas de informação. O sistemismo é uma ontologia (no sentido filosófico do termo) elaborada nos últimos 30 anos, aproximadamente, com desenvolvimento metodológico mais recente, a partir da publicação de um conjunto de regras metodológicas amplas (BUNGE, 1997).

O artigo discute as falhas das teorias de sistemas, que lhes conferiram um substancial descrédito, em especial sua característica **holística** e não sistêmica. O sistemismo de Bunge é apresentado como abordagem que nega e conjuga, ao mesmo tempo, o individualismo característico da ciência analítica com o holismo característico das teorias de sistemas.

São apresentados (i) os postulados, (ii) o modelo de sistema do sistemismo, bem como (iii) um roteiro metodológico para abordar o estudo de sistemas de informação. Nenhum texto em português aborda o assunto e, mesmo em inglês, nenhuma publicação reúne os 3 elementos do sistemismo apresentados neste artigo. Uma aplicação à Plataforma Lattes ilustra a abordagem e apresenta a formulação de problemas de pesquisa nesse contexto metodológico.



2 TEORIAS DE SISTEMAS: USO NA ÁREA, DESCAMINHOS, PERSPECTIVA

Uma teoria de sistemas deve permitir o estudo e a descrição de qualquer sistema ou coleção estruturada de objetos que em algum aspecto se comporta como um todo. As teorias de sistemas, no entanto, são muitas – segundo Bunge (2004, p. 191), “há quase tantas teorias de sistemas quantos são os teóricos”.

Há enganos e disparidade. Bertalanffy (1968, p. 18), autor da primeira teoria de sistemas amplamente disseminada, critica a ciência analítica e aponta suas limitações de uma forma que sugere seu abandono completo e a substituição por estudos que tentem descrever apenas o todo, sem analisar as partes e suas relações.

Skyttner (1996, p. 30), por outro lado, afirma que é paradoxal e “impossível tornar-se consciente de um sistema como um todo sem analisar suas partes” (1996, p. 42). Essa afirmação é invertida na máxima holista de que não se pode conhecer o todo a partir das partes, sintetizada por Keller e Golley (2000, p. 171) como “o conhecimento das partes não é necessário nem suficiente para o conhecimento do todo”.

2.1 Teorias de sistemas na ciência da informação

As teorias de sistemas são ocasionalmente citadas na literatura nacional da ciência da informação, mas sua aplicação é bastante limitada, da mesma forma que nas demais áreas do conhecimento. A teoria geral dos sistemas de Bertalanffy (1968) é implicada no artigo sobre a teoria da biblioteconomia e da ciência da informação de Shera (1973), ao mencionar um afastamento da ciência da informação – das humanidades em direção às ciências –, embora o artigo não nomeie Bertalanffy ou qualquer teórico de sistemas.

Souza (2007) alude à teoria de sistemas sem apontar autor. Também Assis (1981) usa a teoria geral dos sistemas para discutir a biblioteca universitária como sistema. Marcial et al. (2007) mencionam a teoria de sistemas de Luhmann (1995), da mesma forma que Lima, Carvalho e Lima (2011).

Borges (2000) cita Bertalanffy e Churchman (1972), entre outros, para fundamentar o enfoque sistêmico que adota. Sayão (2001), ao revisar modelos teóricos, menciona a idéia de universo como “grande conjunto interconectado” (p. 84) de Bertalanffy, que também é citado por Robredo (2003) em seu estudo da epistemologia da área.

Araújo (2009) apresenta a teoria sistêmica como uma das seis correntes teóricas da ciência da informação, citando Bertalanffy. Por outro lado, vale assinalar que não constam teorias de sistemas no estudo de Freire (2008) sobre a epistemologia da ciência da informação.

Passaram-se seis décadas e meia desde o surgimento da teoria geral dos sistemas¹, consolidada em livro há quase meio século (BERTALANFFY, 1968), mas seu impacto parece estar ainda por vir. Esta apreciação de Shera, de há 38 anos, parece atualíssima:

¹ A concepção de teoria geral dos sistemas apareceu originalmente no artigo “Zu einer allgemeinen Systemlehre”, na revista *Blätter für deutsche Philosophie*, 3/4, 1945.



O valor da teoria geral dos sistemas para os bibliotecários está por ser testado, mas certamente ela promete uma revolução profunda na ciência e em outras áreas do pensamento e, agora, aparenta poder dar aos bibliotecários insights e compreensão que estiveram faltando por muito tempo (SHERA, 1973, p. 94).

Não é mera coincidência que Shera se refira aos potenciais benefícios da teoria geral de sistemas “desde um quadro de referência holístico” (p. 94). De fato, a confusão entre holismo e visão sistêmica é antiga, causada pelos próprios autores “sistêmicos”, como se discute a seguir.

2.2 Os “sistêmicos” não são sistêmicos, mas holistas

As teorias de sistemas servem para fazer a crítica da ciência analítica – fundamentalmente, apontar a incapacidade de lidar com problemas complexos, isto é, aqueles que resistem à estratégia da decomposição em problemas cada vez menores. Como fundamento para pesquisa, as teorias de sistemas dispõem de conceitos importantes, tais como as idéias de hierarquia (níveis), modelos e estados de sistemas, bem como de métodos, tais como as descrições de sistemas por meio de sistemas de equações diferenciais. Porém, essas ferramentas intelectuais não levaram as teorias de sistemas a uma adoção ampla.

As falhas das teorias de sistemas levaram Bunge (2003, p. 113) a afirmar que “quando cientistas sociais rigorosos contemporâneos ouvem a palavra ‘sistema’, ficam propensos a sacar suas armas intelectuais”. Abordagens sistêmicas tendem a ser vistas, nesse sentido, como “alternativas”, com a conotação de fracas, pouco rigorosas.

Bunge (2003) qualifica teorias de sistemas muito citadas, como as de Bertalanffy (1978) e de Laszlo (1972), como holistas. O holismo aborda sistemas como todos (objetos complexos), “mas se recusa tanto a analisá-los quanto a explicar a emergência e a análise das totalidades em termos de seus componentes e das interações entre eles” (BUNGE, 2003, p. 38). Como abordagem filosófica ampla, contrapõe-se ao individualismo, característico das ciências analíticas, que “enfoca a composição de sistemas e se recusa a admitir quaisquer entidades supra-individuais ou suas propriedades” (p. 38) – ou seja, reconhece apenas indivíduos (átomos compostos por partículas, na física atômica, moléculas compostas por elementos, na química, organismos compostos por micro- e mesossistemas físicos e químicos na biologia etc.)

2.3 Nova perspectiva na lacuna holismo-individualismo

Bunge (2000, 2003) propõe conjugar as virtudes do holismo (reconhecer que existem supra-entidades ou todos) e do individualismo (a tese de que, em última análise, só existem particulares) no **sistemismo**. Individualismo e holismo são vistos como os “principais rivais” (BUNGE, 2003, p. 81) do sistemismo e como abordagens opostas que o sistemismo nega e, ao mesmo tempo, conjuga. O Quadro 1 apresenta individualismo, holismo e sistemismo como categorias ontológicas que têm contrapartes



epistemológicas e metodológicas, bem como associa um slogan e cita alguns pensadores e correntes filosóficas afiliados.

Ontologia	Epistemologia	Metodologia	Slogan	Autores e idéias (exemplos)
Individualismo	Racionalismo ou empirismo	Análise: microrredução	<i>Divide et impera.</i>	Cartesianismo; Adam Smith – o interesse individual é que move a economia e a sociedade; Karl Popper – o conhecimento humano resulta de conhecimentos individuais.
Holismo	Intuicionismo	Síntese: macrorredução	<i>Conjuga et impera.</i>	Parmênides (“tudo é um”); Bertalanffy e Laszlo, autores de teorias de sistemas; Luhmann – eliminação do indivíduo e reificação das instituições; gestaltismo.
Sistemismo	Realismo científico	Análise e síntese	<i>Divide et conjuga.</i>	Newton, a quem couber reconhecer que o sistema solar é mais do que um agregado de corpos celestes; Tocqueville, analista histórico com visão sistêmica.

Quadro 1 – As 3 ontologias, suas epistemologias e metodologias.

Fonte: Adaptação do autor a partir de Bunge (2003, p. 126).

O sistemismo “não é uma teoria para substituir outras teorias, mas uma estratégia para desenhar projetos de pesquisa cujo intuito é descobrir algumas características de sistemas de um tipo particular” (BUNGE, 2004, p. 191). A próxima seção apresenta os fundamentos do sistemismo para embasar a abordagem, desenvolvida na seção 4, à construção de visões sistêmicas, em especial no âmbito da ciência da informação.

3 ELEMENTOS DO SISTEMISMO

“Problemas vêm em pacotes ou sistemas”, segundo Bunge (2003, p. 92), inclusive os problemas práticos. O autor exemplifica o argumento com o problema da drogadição: não se resolve punindo os traficantes, muito menos os viciados. É preciso atacar também as “raízes econômicas e culturais do abuso de substâncias, como a pobreza, o mercado competitivo de drogas, a anomia e a ignorância”.

A abordagem “uma coisa de cada vez” não funciona para problemas complexos. O sistemismo se propõe como visão abrangente para abordar esses problemas complexos:

A visão de que o conceito de sistema é central para a ciência e a tecnologia social e, com efeito, para todas as ciências e tecnologias, pode ser chamada **sistemismo**. O sistemismo tem dois componentes: ontológico e metodológico. O sistemismo ontológico afirma que o mundo é um sistema antes que uma coleção de indivíduos ou um bloco sólido. Seu parceiro metodológico é a visão de que os todos (wholes) são melhor entendidos por meio da análise (top-down) e indivíduos por meio da síntese (bottom-up) (BUNGE, 2006, p. 12-13, grifo do autor).

Os postulados do sistemismo, elemento ontológico-metodológico, são apresentados a seguir. Depois, o modelo (ontológico) de sistema de Bunge é brevemente conceitualizado e ilustrado. As regras gerais para a pesquisa sistemista, elemento metodológico, complementam esta descrição do sistemismo.



3.1 Postulados do sistemismo

O sistemismo fundamenta-se nestes postulados (BUNGE, 2000, p. 149, 2003, p. 114):

- Toda coisa, seja concreta ou abstrata, é um sistema ou um componente ou potencial componente de sistema.
- Sistemas têm características sistêmicas (emergentes) que seus componentes não têm.
- Todos os problemas deveriam ser abordados de forma sistêmica em vez de em forma fragmentada.
- Todas as idéias deveriam ser unidas em sistemas (preferencialmente teorias); e
- O teste de qualquer coisa, seja idéia, método ou artefato, supõe a validade de outros itens que são tomados como pontos de referência (*benchmark*) provisoriamente.

Com esses postulados, Bunge declara a ubiqüidade dos sistemas, a crença de que não há nada permanentemente isolado ou errante, e aconselha a adoção de uma visão de mundo sistêmica. Para realizar uma visão sistêmica de um sistema específico, é possível construir um modelo de sistema, como se discute a seguir.

3.2 O modelo de sistema CESM

Qualquer sistema concreto pode ser descrito, segundo Bunge (2003, p. 34-38), por meio do modelo CESM, segundo sua **composição** (*Composition*) – coleção de partes ou elementos componentes, **ambiente** (*Environment*) – coleção de itens que não fazem parte do sistema mas atuam ou sofrem ação por algum componente, **estrutura** (*Structure*) – coleção de ligações entre componentes e entre esses e itens do ambiente e **mecanismo** (*Mechanism*) – coleção de processos que geram a novidade qualitativa (BUNGE, 1997), i. e., que promovem ou obstruem as transformações, causando a emergência ou o desmantelamento do sistema ou de alguma de suas propriedades. O Quadro 2 mostra exemplos sucintos de sistemas descritos segundo o modelo CESM. Os sistemas podem ser naturais, sociais, técnicos ou mistos.

Sistema	C Composição	E Ambiente	S Estrutura	M Mecanismo
Organismo vivo	Micro- e mesossistemas físicos e químicos (esp. água, carboidratos, lipídios, proteínas e ácidos nucleicos).	Meio rico em nutrientes e fluxos de energia, com variáveis (pressão, temperatura etc.) restritas a pequenos intervalos.	Ligações diretas e indiretas, físicas ou químicas, covalentes e não-covalentes que mantêm unidos os componentes, mais as ligações químicas, físicas e biológicas com o ambiente.	Processos que mantêm o organismo vivo, e.g., síntese de algumas moléculas, captura e armazenagem de energia, transporte, rearranjo, montagem e desmantelamento de componentes que acompanham o metabolismo.



Empresa	Os funcionários (<i>personnel</i>) e a gerência.	Mercado e governo.	As relações de trabalho entre membros da empresa e entre esses e o ambiente.	As atividades que resultam em produtos da empresa.
Comunidade linguística	As pessoas que falam a mesma língua.	A cultura na qual a linguagem é usada.	A coleção de relações de comunicação linguística.	A produção, transmissão e recepção de símbolos.
Sistema solar	Sol, planetas e asteróides.	Demais corpos celestes.	Forças gravitacionais.	Translação dos componentes em órbitas que permitem certa continuidade do sistema (sem dispersão ou colapso) pela inércia.

Quadro 2 – Exemplos de modelos CSM de sistemas existentes.

Fonte: Composição do autor a partir de Bunge (2003).

Em termos simbólicos, um modelo μ de um sistema σ pode ser escrito segundo a quádrupla $\mu(\sigma) = \langle C(\sigma), E(\sigma), S(\sigma), M(\sigma) \rangle$. Qualquer objeto de estudo, desde que concreto, pode ser modelado assim, dado que tudo no universo é, foi ou será um sistema ou componente de um sistema. Não há elementos permanentemente isolados ou perdidos.

Todos os sistemas, exceto o universo, têm um ambiente. As ligações (estrutura) são fundamentais para a auto-organização e sua força tem usualmente uma relação inversa com o tamanho do sistema (quanto maior o sistema, mais fracas as ligações). As ligações entre componentes formam a endoestrutura e as ligações entre componentes e itens do ambiente formam a exoestrutura, com dois subconjuntos peculiares: entradas (*inputs*) e saídas (*outputs*).

Composição, ambiente e estrutura são também características de sistemas abstratos, porém apenas os concretos têm mecanismo, pois “a mutabilidade (ou a energia) é a propriedade definidora” das coisas concretas, sejam físicas, químicas, vivas, sociais ou técnicas (BUNGE, 2004, p. 191-192), mas não é propriedade das idéias (abstrações).

Sistemas altamente complexos, tais como células vivas e escolas, têm muitos mecanismos concorrentes (BUNGE, 2004). Em uma escola, por exemplo, o controle disciplinar e a difusão do conhecimento são mecanismos que podem, ocasionalmente, estar em conflito.

Ao contrário da composição, ambiente e estrutura, o mecanismo é usualmente invisível e precisa ser conjecturado. O mecanismo é o *modus operandi* do sistema (BUNGE, 2003, p. 20). O Quadro 3 apresenta alguns exemplos de sistemas e seus mecanismos.

Sistema	Função específica	Mecanismo(s)
Rio	Drenagem	Fluxo de água
Organismo	Manutenção	Metabolismo
Fábrica	Produção de mercadorias	Trabalho, gestão



Corte de justiça	Busca da justiça	Litígio
Escola	Aprendizagem	Ensino, estudo, discussão
Laboratório científico	Avanço do conhecimento	Pesquisa
Comunidade científica	Controle de qualidade	Revisão por pares (<i>peer review</i>)

Quadro 3 – Alguns sistemas, suas funções específicas e mecanismos.

Fonte: Bunge (2003, p. 30).

Embora simples, a aplicação generalizada do modelo CESM é inviável devido à impossibilidade de se descrever plenamente os detalhes desses 4 aspectos de um sistema concreto. Na prática, os modelos são construídos em certo nível _a da composição, nível _b do ambiente, nível _c da estrutura e nível _d do mecanismo, formando o modelo CESM reduzido, ou $_{abcd}(\sigma) = \langle C_a(\sigma), E_b(\sigma), S_c(\sigma), M_d(\sigma) \rangle$.

São exemplos de delineamento no nível da composição: átomos em uma molécula, células em um tecido vivo ou organismo, indivíduos em uma sociedade. Um exemplo de delineamento no nível da estrutura é: ao conceber os componentes de um sistema social como pessoas, faz sentido limitar a endoestrutura a relações interpessoais, desprezando relações gravitacionais e eletromagnéticas entre as pessoas. Pode-se construir diferentes modelos de uma sociedade variando os níveis _{abcd} da análise, resultando modelos de subsistemas (e. g., família, empresa, escola, clube social, partido político, comunidade religiosa). Procedimento similar pode ser adotado, postula Bunge (2003, p. 37), em todos os campos do conhecimento.

3.3 Regras gerais da pesquisa sistemista

As seguintes “regras metodológicas gerais” para a abordagem sistêmica de problemas (BUNGE, 1997, p. 458) complementam os postulados e o modelo CESM apresentados acima:

1. Colocar todo fato social em seu contexto mais amplo (ou sistema).
2. Dividir cada sistema em sua composição, ambiente e estrutura.
3. Distinguir os vários níveis de sistema e exibir suas relações.
4. Procurar os mecanismos que mantêm um sistema funcionando ou levam à sua decadência ou crescimento.
5. Ter razoável certeza de que o mecanismo proposto é compatível com as leis e as normas relevantes e conhecidas e, se possível, verificar a hipótese ou teoria mecanística manipulando experimentalmente as variáveis referidas.
6. Mantidas as demais condições, preferir hipóteses, teorias e explicações mecanísticas (dinâmicas) às fenomenológicas (cinemáticas) e, em seu turno, preferir essas descrições cinemáticas aos modelos de equilíbrio e às descrições de dados.



7. Em caso do mau funcionamento do sistema, examinar todas as quatro fontes possíveis – composição, ambiente, estrutura e mecanismo – e tentar reparar o sistema alterando alguma ou todas as fontes.

Essas “regras metodológicas gerais” foram enunciadas no contexto das Ciências Sociais, a partir do postulado de que tudo o que essas ciências estudam são sistemas sociais (BUNGE, 1997). Bunge (2000, p. 157) propõe que o sistemismo, mais provavelmente do que seus rivais individualismo e holismo, pode “inspirar e defender políticas que combinem competição e colaboração e que promovam o bem-estar e a liberdade individuais ao mesmo tempo que fortalecem ou reformam as instituições necessárias”.

A regra 1 trata de identificar ou reconhecer o sistema em foco. Fatos sociais são modos de pensar, sentir e agir que se impõem sobre indivíduos a partir de regras, princípios e práticas socialmente construídos (SCOTT; MARSHALL, 2009). Uma discussão detalhada sobre fato social, como definido por Émile Durkheim, está fora do escopo deste artigo, mas “colocar todo fato social em seu contexto mais amplo” significa não isolar o fato social – por exemplo, um valor, crença ou norma social – do sistema social em que pode ser estudado.

A regra 2 trata da descrição dos aspectos usualmente observáveis (mas não necessariamente palpáveis) de um sistema. A regra 3 expressa a identificação de macro- e microssistemas relevantes do sistema em foco e suas relações. A motivação para essa atividade é a construção de conjeturas de mecanismos ou explicações mecanísticas que envolvem fenômenos causais entre diferentes níveis de sistemas.

A regra 4 trata de buscar (ou, mais apropriadamente, conjeturar) os mecanismos que criam, sustentam, fazem crescer, reformam ou destroem um sistema ou alguma de suas propriedades. A regra 5 recomenda testar, verificar as conjeturas.

A regra 6 é uma recomendação de “melhores práticas”, buscando promover formas mais profundas de explicação. Finalmente, a regra 7 é uma sugestão de intervenção para promover efeitos sistêmicos desejados sem incorrer nos riscos das intervenções simplistas que, em sistemas complexos, frequentemente têm resultados contraproducentes.

Dado que tudo o que existe é sistema ou componente de sistema (primeiro postulado do sistemismo), esse roteiro metodológico fundamental para abordar problemas em sistemas sociais pode ser adaptado para investigar sistemas de informação. Toma-se a expressão “sistemas de informação” num sentido amplo e não apenas em seu aspecto técnico, como é frequente. São sistemas de informação as unidades de informação e outros grandes ecossistemas informacionais. A próxima seção ilustra as possibilidades de pesquisa sistemista nesse sentido.



4 PERSPECTIVAS DE PESQUISA SISTEMISTA EM SISTEMAS DE INFORMAÇÃO

Esta seção discute brevemente a existência de visões parciais ou setoriais (mormente as tecnocêntricas) de fenômenos informacionais muito complexos, exemplifica a abordagem sistemista proposta na seção anterior com uma aplicação à Plataforma Lattes e cogita a aplicação da abordagem sistemista de duas maneiras: na descrição e na investigação de mecanismos de sistemas de informação.

4.1 Nem todos vêm os mesmos sistemas

Fuchs (2005) observa que a internet é definida por agências como a Internet Engineering Taskforce e o Federal Networking Council norte-americano como um sistema global puramente tecnológico, com processos mecânicos (“mecânicos”, maquinais – e não mecânicos, termo sistemista relativo à explicação da emergência de sistemas ou suas propriedades). Fuchs advoga a concepção da internet como um sistema auto-organizado.

O Conselho Nacional de Pesquisa e Desenvolvimento Tecnológico (CNPq) mantém, pelo menos desde 2007 (na percepção deste autor), a seguinte definição da Plataforma Lattes²: “[...] é a base de dados de currículos, instituições e grupos de pesquisa das áreas de Ciência e Tecnologia”. O entendimento já foi diferente. Segundo uma declaração preexistente nas páginas da agência na web, a Plataforma Lattes representaria

a integração de seus sistemas de informação gerenciais, instrumento fundamental não somente para as atividades de financiamento, mas também no tratamento e na difusão de informações necessárias para a formulação e a gestão de políticas de ciência e tecnologia (SABBATINI, 2001).

Percebe-se a diferença entre uma definição sistêmica, que dá conta de “colocar todo fato social em seu contexto mais amplo” (BUNGE, 1997, p. 458) e outra filosoficamente individualista, que “enfoca a composição de sistemas e se recusa a admitir quaisquer entidades supra-individuais ou suas propriedades” (BUNGE, 2003, p. 38). São paradigmas diferentes, concepções ontologicamente díspares. É difícil o diálogo entre quem esposa uma e os praticantes da outra.

Uma visão de mundo individualista aceita as seguintes assertivas corretas, porém parciais, limitadas:

- Uma biblioteca coleciona conteúdos e os cede em consulta e/ou empréstimo.
- Um arquivo é um órgão que acumula e preserva documentos.
- Um repositório institucional é um depósito de conteúdos de uma instituição.



- O controle bibliográfico universal (CBU) é uma coleção de instrumentos que inclui a identificação de conteúdos monográficos e em série, o depósito legal e a bibliografia nacional.
- O YouTube é uma videoteca de filmes curtos.
- Governo eletrônico é a oferta de serviços governamentais via digital.

A visão sistemista, no entanto, não aceita essas assertivas senão como definições parciais, enviesadas, que podem valer em uma situação ou contexto muito limitado, mas nunca como afirmações gerais. Uma definição adequada de biblioteca, arquivo, repositório institucional ou qualquer sistema de informação deve levar em conta seu papel num contexto mais amplo, incluindo o sistema social a que serve e/ou em que se insere.

O sistemismo bungeano, em geral, e o modelo CESM, em particular, pareceram ao autor deste artigo um referencial teórico promissor para explicar sistemas de informação muito complexos. A próxima seção prepara uma explicação sistemista da Plataforma Lattes mostrando evidências da visão sistêmica na sua concepção.

4.2 Plataforma Lattes: Visão sistêmica na concepção

O lançamento do Sistema de Currículo Lattes em agosto de 1999³ é um marco inicial da Plataforma Lattes. O sistema de currículo e mais de uma centena de outros artefatos foram desenvolvidos pelo extinto Grupo Stela da Universidade Federal de Santa Catarina, sob contrato com o CNPq. Num dos primeiros relatos publicados sobre a Plataforma Lattes, Pacheco e Kern (2003) evidenciaram a visão sistêmica que a dirigiu. A Figura 1 dá indícios dessa visão sistêmica.



Figura 1 – Ligações que envolvem *stakeholders* (partes interessadas) da Plataforma Lattes.

Fonte: Pacheco e Kern (2003).

A visão sistêmica na concepção explica algumas decisões de projeto. O currículo vitae (CV) é do indivíduo, o que implica que coautores registrarão itens da produção intelectual em duplicata. Isso cria redundância, violando uma regra clássica da teoria de bancos de dados (CODD, 1970), mas uma opção de projeto pela normalização (i. e., eliminação da redundância) implicaria inviabilizar o fluxo de informação, não ter currículos na base curricular. O fluxo de informação curricular é tão viável que a Plataforma tem 12 anos de crescimento exponencial, como mostra a Figura 2 – outro indício da visão sistêmica na concepção.

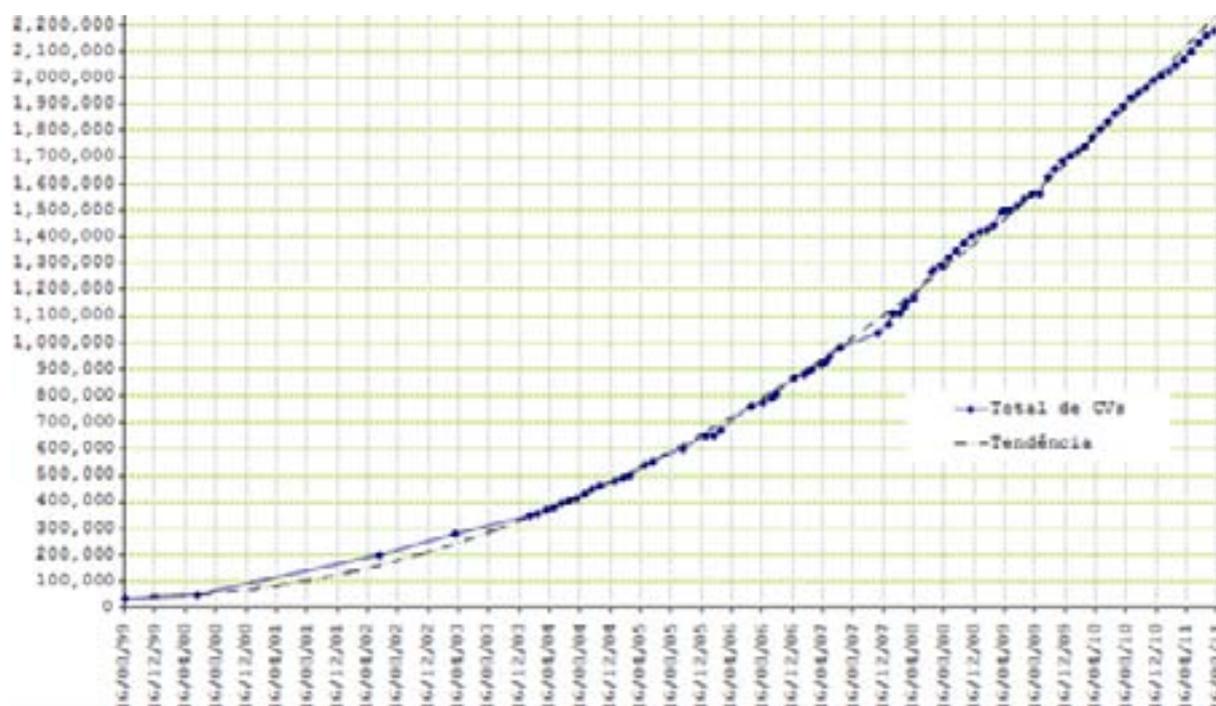


Figura 2 – Crescimento quadrático do número de currículos na Plataforma Lattes.

Fonte: Composição do autor a partir de dados da Plataforma Lattes, até 2004, e do Portal Inovação, desde 2005.

As palavras-chave registradas nos itens da produção curricular também são individuais. Coautores podem ter palavras-chave distintas, para evidenciar sua contribuição individual. Isso violaria um princípio da recuperação da informação bibliográfica, se o CV tivesse esse objetivo. Palavras-chaves registradas com esse fim são as associadas aos itens bibliográficos nas bases de publicações. No CV, servem como um rastro da competência individual, insumo para a localização de especialistas e o fomento à cooperação tecnológica, processos importantes no sistema nacional de inovação ou SNI (OECD, 1997), um complexo de agentes e fatores que levam ao desenvolvimento econômico de uma nação.

A visão sistêmica na concepção evidenciada no relato citado (PACHECO; KERN, 2003) tem formulação informal, não baseada em teoria de sistemas. Kern (2010) apresenta uma descrição sistemista da Plataforma Lattes. A próxima seção delinea a Plataforma segundo essa perspectiva sistemista.



4.3 Plataforma Lattes: Descrição e agenda de pesquisa sistemista

Bunge (1997, p. 457) pleiteia para o sistemismo um “poder heurístico e sistematizador”, uma utilidade prática que tem sido reconhecida por este autor e colaboradores na concepção de sistemas de larga escala em alto nível de abstração, bem como em seu estudo. Esta seção objetiva ilustrar esse poder heurístico por meio de uma descrição da Plataforma Lattes.

Um contexto mais amplo da Plataforma Lattes é o sistema nacional de inovação. Dessa forma, cumprindo o primeiro dos 7 itens do roteiro metodológico sistemista apresentado na seção 3.3, pode-se compreender o papel assumido pelo CNPq – o de uma agência inserida no SNI, patrocinador da Plataforma, com atendimento privilegiado, mas não o “dono do sistema” (PACHECO; KERN, 2003) a ponto de tomar decisões sem considerar os demais interessados.

O item 2 da metodologia geral sistemista recomenda delinear a composição, o ambiente e a estrutura do sistema. Kern (2010) descreve essas 3 dimensões a partir da concepção sociotecnológica de Fuchs (2005), o que implica reconhecer que há agentes tecnológicos que atuam em **colaboração dinâmica** com as pessoas, fazendo análises, sínteses e estabelecendo vínculos – por exemplo, os indexadores, os construtores de *résumés*, os aplicativos gráficos de redes sociais e os criadores de links entre um CV e outras bases (Scielo, por exemplo).

Ainda que seja usual e correto referir-se a “ambiente de software” (i.e., virtualização de um espaço), a tecnologia, no caso da Plataforma Lattes, não é apenas ambiente. Alguns agentes tecnológicos são componentes, **fazem** o sistema. A composição inclui, além desses agentes especializados, as pessoas detentoras de currículo (pesquisadores, docentes, estudantes, técnicos, especialistas, profissionais et al.) e os servidores do staff do CNPq, bem como o artefato essencial do sistema – o currículo Lattes, componente que liga o subsistema social ao subsistema técnico. Sem o artefato essencial como componente, perdura a cisão sociotécnica, a lacuna entre os vieses técnico e social que a abordagem sociotécnica não conseguiu vencer (SAWYER; CROWSTON, 2004).

O ambiente da Plataforma Lattes é constituído por atores e fatores, elementos que não compõem o sistema, mas o influenciam ou recebem sua ação. A identificação de itens do ambiente pode basear-se na investigação das possíveis origens dos fatos sociais propostas por Bunge (2003): ambiental, biopsicológica, econômica, política e cultural.

Incluem-se no ambiente: a Web como sistema de conteúdos interconectados (mais que como tecnologia), a infraestrutura de tecnologias de informação e comunicação (TIC) que permite o fluxo via Web, as fontes de informação que trocam dados e links com a Plataforma (e. g., Scielo, ISI, INPI), o público consultante (jornalistas, empresários et al., inclusive os próprios acadêmicos no papel de consultantes), as organizações de ciência, tecnologia e inovação (ICT –em especial, as universidades) por meio de seu pessoal técnico e administrativo, as agências de fomento à pesquisa e outras agências estatais, bem como a cultura na qual a Plataforma Lattes está inserida (como em qualquer sistema social, a cultura é fator ambiental essencial).



A estrutura da Plataforma Lattes é feita de ligações entre componentes (endoestrutura) e entre esses e itens do ambiente (exoestrutura). Kern (2010) apresenta um extenso mapa de ligações que evidencia o papel estruturante do CV como componente do sistema sociotecnológico – é o componente mais conectado, tanto com outros componentes quanto com itens do ambiente. O Quadro 4 apresenta uma síntese das ligações mais relevantes no sistema, bem como da composição e do ambiente – resultado da aplicação da regra metodológica 2.

C Composição	E Ambiente	S Estrutura
<ul style="list-style-type: none">• indivíduos detentores de currículo (pesquisadores, docentes, estudantes, técnicos, especialistas, profissionais et al.),• servidores do staff do CNPq que administram a Plataforma e suas bases de dados,• agentes tecnológicos (de análise, síntese ou estabelecimento de vínculos),• currículos Lattes (artefato essencial do sistema).	<ul style="list-style-type: none">• a Web,• a infraestrutura TIC,• as fontes de informação relacionadas,• o público consultante (jornalistas, empresários et al., inclusive acadêmicos),• as ICT (em especial, universidades),• as agências de fomento à pesquisa e outras agências estatais,• a cultura na qual a Plataforma Lattes está inserida.	<p>Ligações de</p> <ul style="list-style-type: none">• criação/atualização de CV,• propriedade do CV (pessoa-CV),• acesso e visualização,• extração de CVs em XML (pelas organizações conveniadas com CNPq),• extração, transformação e carga de arquivos de CV (por agentes),• vinculação de CVs com fontes de informação e outros CVs (por agentes),• etc. (uma longa lista).

Quadro 4 – Composição, ambiente e estrutura da Plataforma Lattes.

Fonte: O autor.

A regra metodológica 3 recomenda apontar outros níveis de sistema e suas relações com o sistema em foco. A Plataforma Lattes está inserida no SNI, sistema de nível superior, para o qual representa uma fonte de informação crítica e de alto valor agregado – afinal, permite identificar a competência acadêmica nacional de forma simples e rápida. O SNI, por sua vez, é a fonte do financiamento de todo o setor acadêmico, incluindo o CNPq, patrocinador da Plataforma.

No nível micro, o indivíduo é o principal elemento. As relações mais importantes são a entrada e a saída de dados entre a Plataforma Lattes e os indivíduos detentores de currículos ou consultantes. As ICT, elemento do ambiente da Plataforma, representam um microssistema importante para a Plataforma Lattes na medida em que empregam grande parte dos detentores de CVs e exercem pressão sobre o CNPq – a ponto de se criar uma política de convênios para cessão gratuita dos dados de currículos, da Plataforma Lattes para as organizações afiliadoras.

Os macro- e microssistemas identificados com a aplicação da regra 3 permitem articular hipóteses mecanísticas, tentativas de explicação do funcionamento da Plataforma Lattes como sistema. Isso pode ser feito por meio de diagramas de Boudon-Coleman (BUNGE, 1997, 2000, 2003, 2004, 2006). Essa conjectura, correspondente à regra metodológica 4 da abordagem sistêmica de Bunge, trata de buscar a explicação de fenômenos no nível do sistema de interesse recorrendo a cadeias causais que



envolvem mais de um nível de sistema. As figuras 3 e 4 ilustram hipóteses mecanísticas formuladas por Kern (2010).

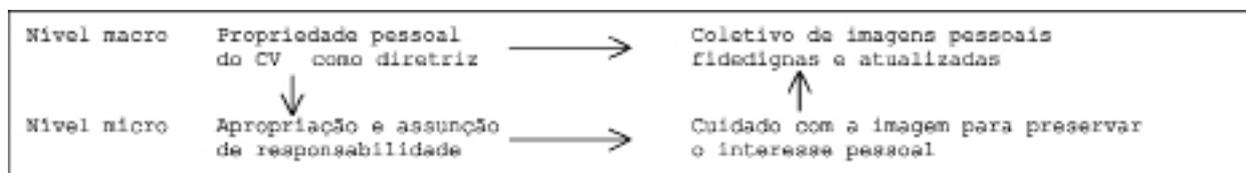


Figura 3 – Mecanismo “Apropriação individual do currículo vitae”.

Fonte: Kern (2010).

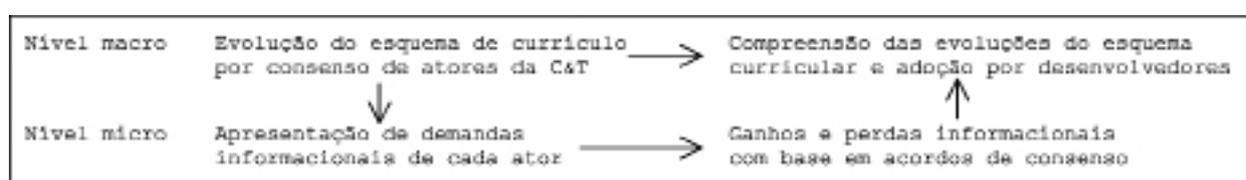


Figura 4 – Mecanismo “Evolução do esquema curricular por consenso da comunidade”.

Fonte: Kern (2010).

A apropriação individual leva à assunção individual de responsabilidade sobre a informação curricular. A evolução curricular por consenso permite a integração de informações de diversos sistemas técnicos, conforme Pacheco e Kern (2001), que estão além dos limites da Plataforma Lattes. Além desses, outros mecanismos propostos por Kern (2010) são a inversão da demanda informacional, a sinergia com outras iniciativas de Estado e o depósito curricular mandatório em processos de fomento. A inversão da demanda – do CNPq sobre os pesquisadores para as ICT sobre o CNPq – resultou na assinatura de acordos de extração de dados curriculares do CNPq para as ICT e a pressão dessas sobre seus afiliados pelo preenchimento do currículo, em vez de depender da iniciativa desses afiliados.

A sinergia com outras iniciativas inclui acordos com o Portal Inovação, o Portal Sinaes e outros, com o uso dos CVs como fonte de informação para essas iniciativas, o que criou motivos para a entrada de novos currículos no sistema. O depósito obrigatório do currículo é a exigência de possuir currículo para participar de processos de seleção de propostas de pesquisa.

A formulação de hipóteses mecanísticas é ponto de partida para o desenho de experimentos que verifiquem essas hipóteses – regra metodológica 5 da abordagem sistêmica. As regras 6 e 7 representam estágios avançados do processo de pesquisa, nos quais se consolidam conclusões sobre as hipóteses mecanísticas e se promovem ou se propõem intervenções no sistema.

4.4 Estudos de descrição e de investigação

A descrição da Plataforma Lattes, contemplada nas etapas 1 a 3 do roteiro metodológico sistemista, requer técnicas de pesquisa descritiva, envolvendo coleta e sistematização de informações, classificação e síntese. Quaisquer outras descrições sistemistas de sistemas de informação requerem essa mesma modalidade de pesquisa descritiva para a qual parecem especialmente talhados os pesquisadores da área da informação. Assim como a ontologia precede a epistemologia, a investigação aprofundada



do funcionamento de um sistema de informação muito complexo requer que, antes, as características essenciais (composição, ambiente e estrutura) do sistema sejam elucidadas.

Por outro lado, a investigação dos mecanismos de funcionamento de grandes sistemas de informação, como exemplificada na seção anterior, parece especialmente apropriada para abordagem pela ciência da informação, já que outras áreas do conhecimento que tratam de sistemas de informação não se dedicam a explicar o sistema. A ciência da computação, por exemplo, busca ver a dimensão social do sistema de informação como uma caixa-preta da qual o profissional de computação busca extrair “requisitos” que são entrada para construir o sistema técnico. Equaciona, assim, “sistema de informação” a “sistema técnico”, “software”.

5 CONSIDERAÇÕES FINAIS

Este artigo pretende permitir um vislumbre do “poder heurístico e sistematizador” (BUNGE, 1997, p. 457) do sistemismo, desfrutado por este autor e colaboradores na concepção e estudo de sistemas de alta complexidade em alto nível de abstração – como se relata, por exemplo, em Kern (2010). O artigo pretende ser, também, um guia introdutório aos interessados nos fundamentos e na metodologia da investigação sistemista, em especial na ciência da informação.

Bunge não é um dos filósofos mais conhecidos da área. O autor do *Treatise on Basic Philosophy* (8 volumes, 1974-89) não é citado no recém-lançado livro da ASIS&T (DAVIS; SHAW, 2011) *Introduction to Information Science and Technology*, que junta aos nomes de Aristóteles, Bacon, Boole, Descartes, Diderot, Kant, Leibnitz, Mill, Pascal e Peirce citações aos filósofos ainda vivos (como Bunge) Capurro, Floridi e Habermas, bem como aos falecidos em décadas recentes Foucault, Kuhn e Popper.

Na literatura nacional, Bunge também não é citado no amplo panorama filosófico descrito por Robredo (2007). Uma pesquisa na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI, <http://www.brapci.ufpr.br/>) por “Bunge” revela apenas 2 ocorrências. A primeira é a apropriação de Bazi e Silveira (2007) da epistemologia bungeana, numa época da produção do filósofo (1980) anterior à consolidação de seu modelo (BUNGE, 2003) e metodologia geral (BUNGE, 1997) sistêmicas. A segunda publicação é um trabalho orientado por este autor (MORETTO; GALDO, KERN, 2010). Este artigo busca contribuir para tornar mais acessível e conhecida a filosofia bungeana.

As principais contribuições do artigo são: apresentar uma abordagem metodológica ao estudo de sistemas de informação baseada no sistemismo de Bunge, reunindo num único documento 3 elementos importantes do sistemismo (postulados, modelo, regras metodológicas), e oferecer o primeiro texto em português sobre essa abordagem.

Abstract: *Large, complex information systems are often seen in a fragmented, partial way, even by its protagonists. This paper presents the systemism of Mario Bunge, a systemic approach, as a theoretical and methodological framework for the design and study of information systems. The article contrasts the individualist worldview, characteristic of*



analytical science, with the holism that underlies the “systemic” approaches prior to Bunge’s. Systemism is introduced as an approach that denies and, at the same time, conjugates individualism and holism. The postulates of systemism are presented, together with its system model and general methodological rules for approaching the design and study of systems – especially information systems. An application of this approach to Plataforma Lattes is offered as illustration. The approach is aimed as a methodological guide for those interested in information systems research with a systemic view. Additionally, the paper outlines some research problems about the mechanisms of growth and maintenance of Plataforma Lattes.

Keywords: Systemism. Information systems. Plataforma Lattes. Research methodology.

REFERÊNCIAS

ARAÚJO, Carlos Alberto Ávila. Correntes teóricas da ciência da informação. **Ciência da Informação**, v. 38, n. 3, p. 192-204, dez. 2009.

ASSIS, Edvaldo de. A biblioteca universitária e as teorias dos sistemas. **Revista Brasileira de Biblioteconomia e Documentação**, v. 14, n. 3/4 p. 174-178, jul./dez. 1981.

BAZI, R. E. R.; SILVEIRA, M. A. A. Constituição e institucionalização da ciência: apontamentos para uma discussão. **Transinformação**, v. 19, n. 2, p. 129-137, maio/ago. 2007.

BERTALANFFY, Ludwig von. **General system theory: Foundations, development, applications**. Revised edition. New York: George Braziller, 1968. 295 p.

BORGES, Maria Alice Guimarães. A compreensão da sociedade da informação. **Ciência da Informação**, v. 29, n. 3, p.25-32, dez. 2000.

BUNGE, Mario. A systemic perspective on crime. In: WIKSTRÖM, Per-Olof H.; SAMPSON, Robert J. **The explanation of crime: context, mechanisms, and development**. Cambridge: Cambridge University Press, 2006, p. 8-30.

BUNGE, Mario. **Emergence and convergence: Qualitative novelty and the unity of knowledge**. University of Toronto Press, 2003. 330 p.

BUNGE, Mario. How does it work? The search for explanatory mechanisms. **Philosophy of the Social Sciences**, v. 34, n. 2, p. 182-210, 2004.

BUNGE, Mario. Mechanism and explanation. **Philosophy of the Social Sciences**, v. 27, n. 4, p. 410-465, 1997.

BUNGE, Mario. Systemism: the alternative to individualism and holism. **Journal of Socio-Economics**, v. 29, n. 2, p. 147-157, 2000.

CHURCHMAN, C. West. **Introdução à teoria dos sistemas**. Rio de Janeiro: Vozes, 1972. 309 p.

CODD, Edgar F. A relational model of data for large shared data banks. **Communications of the ACM**, v. 13, n. 6, p. 377-387, 1970.

DAVIS, Charles H.; SHAW, Debora (Orgs.). **Introduction to information science and technology**. ASIST monograph series. Medford-NJ: Information Today, 2011. 272 p.

FREIRE, I. M. Um olhar sobre a produção científica brasileira na temática epistemologia da ciência da informação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 1, n. 1, p. 1-31, 2008.

FUCHS, Christian. The internet as a self-organizing socio-technological system. **Cybernetics and Human Knowing**, v. 12, n. 3, p. 57-81, 2005.



ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO Brasília, Distrito Federal 23 a 26 de outubro de 2011

- KELLER, David R.; GOLLEY, Frank B. **The philosophy of ecology**: from science to synthesis. University of Georgia Press, 2000. 366 p.
- KERN, V. M. Plataformas e-gov como sistemas sociotecnológicos. In: ROVER, A. J.; GALINDO, F. (Orgs.). **O governo eletrônico e suas múltiplas facetas**. LEFIS (Legal Framework for the Information Society) Series, v. 10. Zaragoza/Espanha: Prensas Universitarias de Zaragoza, 2010, p. 39-67.
- LASZLO, Ervin. **Introduction to systems philosophy**. New York: Gordon and Breach, 1972.
- LIMA, C. R. M.; CARVALHO, L. S.; LIMA, J. R. T. Notas para uma administração discursiva das organizações. **DataGramZero**, v. 11, n. 6, p. 1-14, dez. 2010.
- LUHMANN, Niklas. **Social systems**. Palo Alto-CA: Stanford University Press, 1995. 627 p.
- MARCIAL, E. C.; RAMOS, H. S. C.; SHINTAKU, M.; RODRIGUES, R. C.; VASCONCELOS, W. Epistemologia da Ciência da Informação: a presença do paradigma social de Capurro na literatura. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007. **Anais...** Salvador: ANCIB, 2007.
- MORETTO, L. A. M.; GALDO, A. M. R.; KERN, V. M. Uma análise sistêmica sociotecnológica da engenharia de requisitos. **Encontros Bibli**, v. 15, n. esp. 2, p. 26-40, 2010.
- OECD (Organisation for Economic Co-operation and Development). **National innovation systems**. Paris: OECD, 1997.
- PACHECO, Roberto C. S.; KERN, Vinícius M. Arquitetura conceitual e resultados da integração de sistemas de informação e gestão da ciência e tecnologia. **DataGramZero**, v. 4, n. 2, 2003.
- PACHECO, Roberto C. S.; KERN, Vinícius Medina. Uma ontologia comum para a integração de bases de informações e conhecimento sobre ciência e tecnologia. **Ciência da Informação**, v. 30, n. 3, p. 56-63, 2001.
- ROBREDO, Jaime. Epistemologia da Ciência da Informação revisitada. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003. **Anais...** Belo Horizonte: ANCIB, 2003.
- ROBREDO, Jaime. Filosofia da ciência da informação ou Ciência da informação e filosofia? In: TOUTAIN, Lídia Maria Batista Brandão (Org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007. p. 36-73.
- SABBATINI, M. Lattes, o cómo gestionar la ciencia brasileña en la red. **e-ciencia** (Portal de Ciencia y Tecnología), 18 jun. 2001.
- SAWYER, Steve; CROWSTON, Kevin. Information systems in organizations and society: Speculating on the next 25 years of research. In: **Information systems research**. Boston: Springer, 2004. p. 35-52.
- SCOTT, J.; MARSHALL, G. **Oxford dictionary of sociology**. 3rd ed. revised. Oxford Paperback Reference series. Oxford University Press, 2009. 736 p.
- SHERA, Jesse H. Toward a theory of Librarianship and information science. **Ciência da Informação**, v. 2, n. 2, p. 87-97, 1973.
- SKYTTNER, Lars. **General systems theory**: An introduction. McMillan Press, 1996. 290 p.
- SOUZA, Maria da Paixão Neres de. Abordagem inter e transdisciplinar em ciência da informação. In: TOUTAIN, Lídia Maria Batista Brandão (Org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007. p. 75-90.



O IMPACTO DO *DESIGN* EMOCIONAL NA RECUPERAÇÃO DA INFORMAÇÃO NO PORTAL DO INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT

Luciana dos Santos Nahuz

Resumo: O trabalho procurou identificar se a *interface* do portal do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) desperta sentimentos agradáveis ou não e como isso afeta na recuperação da informação. A pesquisa buscou analisar a *interface* com base na literatura existente sobre *design* emocional, satisfação do usuário, e apresentou sugestões que podem ser utilizadas para melhor interação emocional e agradabilidade do portal com seus usuários. Os dados foram coletados por meio de questionários aplicados na *interface* principal do Ibict.

Palavras-chave: Palavra *Design* Emocional. *Interfaces* digitais. Recuperação da Informação. Usabilidade.

1 INTRODUÇÃO

As pesquisas sobre o *design* emocional utilizadas em *interfaces online* de informação têm interferência direta no usuário, despertando sentimentos de felicidade, tristeza, dúvida, tranquilidade, irritação, medo, paixão e mistério. A interação emocional que uma *interface* pode causar, no primeiro momento, em um usuário torna-se decisiva no que se refere ao uso divertido e prazeroso do serviço e/ou produto oferecido por esta *interface* (portal). O não despertar de emoções agradáveis pode comprometer a interação do usuário com produtos e/ou serviços, prejudicando assim o uso e a recuperação da informação.

A agradabilidade é uma dimensão que considera o nível emocional na interação das pessoas com a *interface*. Para Memória (2005), existem questões que vão além de uma boa usabilidade, pois, apesar de toda a metodologia a ser seguida, torna-se necessário pensar em questões emocionais.

2 OBJETIVO GERAL

Esta pesquisa tem por objetivo geral analisar o impacto do *design* emocional na interação do usuário com a interface do portal do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).



3 OBJETIVOS ESPECÍFICOS

- Identificar a emoção que a *interface* do portal do Ibiict desperta em seus usuários;
- Avaliar a usabilidade da *interface* principal do Portal;
- Diagnosticar o nível de agradabilidade dos usuários no uso da *interface* do portal.

4 JUSTIFICATIVA

Estudos recentes comprovam que o *Design Emocional* da Informação tem papel relevante no momento em que o usuário realiza suas pesquisas. Informações com imagens, sons, cores e vídeos estimulam os usuários a voltarem sempre em determinada *interface* motivando a pesquisa. Nos portais que o recurso da recuperação da informação se dá mais por textos, com cores sombrias, neutras sem vida, acaba gerando o afastamento e, conseqüentemente, o não uso daquela *interface*.

O Portal do Ibiict tem recebido inúmeras reclamações em relação a sua *interface*. Os usuários que costumam acessá-la apontam que, geralmente, não conseguem localizar nem recuperar a informação de que necessitam. Comentam do tamanho das letras, da falta de atualização das informações, do leiaute, das cores e *links* que lá estão.

Esta pesquisa justifica-se pela necessidade de verificar a veracidade de tais reclamações, buscando por meio de questionários identificar o impacto do *design* emocional na recuperação das informações no portal e oferecer sugestões de melhorias relacionadas à agradabilidade com base nas respostas dos questionários aplicados. Enfim, esta pesquisa vem contribuir na interação do usuário com a *interface* do Ibiict, levando em consideração as solicitações do usuário.

5 REVISÃO DE LITERATURA

Antes de tratar sobre o *design* emocional, faz-se necessário primeiramente abordar o *design* de interação. Este, por sua vez, significa uma ação multidisciplinar que visa identificar as necessidades do usuário e depois projetar *interfaces* usáveis e agradáveis ao usuário.

Para que isso aconteça, devem-se considerar as potencialidades das pessoas e pensar no que pode ser feito para proporcionar experiências de qualidade aos usuários ouvindo o que eles querem. E após o conhecimento das características do público ao qual se destinam as informações, deve-se trazer isso para o *design*, lembrando-se de utilizar técnicas baseadas nestes usuários durante o desenvolvimento do *design*.

Assim como o de *design* de interação, o conceito de usabilidade tem também relevância e é de suma importância no momento de avaliar o uso de *interface*. Para Pezzini (2009), usabilidade “é a medida na qual um produto pode ser usado por usuários específicos para alcançar objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso”.



A usabilidade de uma *interface* compõe-se de aspectos **objetivos** – a produtividade na interação – e **subjetivos** – o prazer do usuário em sua experiência com a *interface* – direcionados à maneira como o usuário pensa e se comporta. Para isso a *interface*, além de eficaz e eficiente, deve ser segura no uso, de boa utilidade, fácil de aprender e de lembrar como se usa.

Para Andrade (2009), os testes que avaliam a usabilidade de *interfaces* devem verificar o grau com que esse usuário consegue realizar uma tarefa. Para isso, são medidos itens de funcionalidade correta, eficiência de uso, facilidade de aprendizagem, facilidade de lembrar, tolerância a erro e satisfação subjetiva.

Para realização dos testes de usabilidade mais precisos, utilizam-se métodos de inspeção como: Avaliação Heurística, Revisão de *Guidelines*, Percurso Pluralístico, Inspeções de Consistência, Inspeções de *Standards* e Percurso Cognitivo. Vale lembrar que as 7 Regras Básicas de Usabilidade na *Web* são: clareza na arquitetura da informação; facilidade de navegação; simplicidade; a relevância do conteúdo; manter a consistência e o padrão; tempo suportável e foco nos usuários.

Andrade (2011) indica que os testes podem ser feitos com até 5 usuários, pois após aplicar o teste para o primeiro usuário e captar as avaliações, os conhecimentos sobre os problemas de usabilidade da aplicação, atingem cerca de um terço do total. Explica que outros problemas surgirão em menor número, já que o segundo usuário observa itens distintos na *interface*, mas repetindo alguns já encontrados pelo primeiro; o mesmo ocorre com o terceiro usuário, relatando muitas ações já observadas com o primeiro e segundo, o que diminui as informações de dados novos. Com a aplicação dos testes, o gráfico de aprendizado de novos problemas de *interface* reduz cada vez mais.

Para Barros (2011), “a Agradabilidade é a dimensão que considera o nível emocional na interação das pessoas com a interface”. O prazer é provocado por um conjunto de experiências emocionais positivas, fisiológica, social, psicológica e ideológica. Para abordar aspectos subjetivos, podemos utilizar problemas de mensuração de emoções, características não verbalizáveis, mesclada com processamentos cognitivos. E, para chegarmos ao resultado, foram utilizadas as técnicas de investigação combinadas, como: grupos de foco, entrevistas, questionários e observação visual, tátil e de uso.

Diante deste contexto, observa-se que o usuário consciente quer mais, quer *interfaces* que sejam satisfatórias, agradáveis, divertidas, interessantes, úteis, motivadoras, esteticamente apreciáveis, criativas, compensadoras e emocionalmente adequadas. Quando se trata da emoção provocada por estímulos exteriores e se atinge aspectos fisiológicos, comportamentais e cognitivos, chega-se ao *design* emocional – um projeto de sistemas que proporciona personalização, estética, interação fluída e prazerosa, conforto psicológico, lembrança de memória afetiva, significado social simbólico e compartilhamento de valores morais.

Existem dois tipos de emoções: as primárias – medo, raiva, alegria etc. – e as secundárias que são culturais – a culpa, solidão, saudade etc. Conforme Normam (2008), tem-se também três níveis



de processamento destas emoções formuladas pela mente humana, as quais revelam e possibilitam a interpretação das experiências do usuário:

- **Visceral** – É a relação do impacto emocional imediato. No *design* estão ligados ao primeiro impacto da aparência de um produto no cérebro humano.
- **Comportamental** – É o uso, o desempenho, a *performance*, o contato físico e o prazer no uso do produto. O que importa é a função, a usabilidade e sensação física.
- **Reflexivo** – É a mensagem, a cultura. São lembranças, a auto-imagem e as mensagens que um produto envia às outras pessoas.

Para esta pesquisa, faz-se necessário também esclarecer o conceito de recuperação da informação, já que consta como uma das reclamações mais frequentes dos usuários que acessam ao *site* do Ibcit. Entende-se por recuperação da informação, segundo Silva (2009), o processo de procurar conteúdos que estão armazenados em um espaço com o propósito de acessá-los e, assim, atender às necessidades de informação dos usuários.

Quando se pensa no portal de Informação em Ciência e Tecnologia, imagina-se que esse *site* utiliza o que há de mais avançado no que se refere à informação digital; entretanto, este fato não procede. A recuperação da informação na página do portal do Ibcit precisa urgentemente ser modificada. Já foram apresentadas algumas propostas de reestruturação da *interface*, mas as mesmas ainda não foram implantadas. Para Furquim (2004), permanece a dúvida: *Sobre o que os usuários da web pensam?*

Os usuários da *web* detestam quando não conseguem encontrar no *sítio* uma informação que eles sabem estar lá.

Um fator determinante na interação com usuário e a interface é a estética. Entende-se assim que os usuários são motivados pela estética, podem usar um *sítio* somente por ser esteticamente agradável. Os usuários visitam e retornam a um *sítio* que julgam ser útil e, por fim, oferece a possibilidade de obter informações customizadas e personalizadas.

6 METODOLOGIA

Para realização desta investigação, utilizou-se de pesquisa bibliográfica e de campo, para analisar os dados e propor uma nova *interface*, conforme as informações coletadas dos usuários analisados em interação com o Portal do Ibcit.

Para tanto, percorreu-se os seguintes passos:

- Realizou-se uma breve introdução sobre o *Design Emocional da Informação*,
- Avaliou-se o impacto que o *design* emocional tem na recuperação das informações no portal e a necessidade da interação do usuário com o portal de forma positiva e agradável, diagnosticando estes aspectos por meio da aplicação de questionários.
- Utilizou-se uma pesquisa qualitativa que resultou na análise dos dados coletados.



7 ANÁLISE DOS DADOS

A avaliação da *interface* principal do portal do IbiCT foi realizada de acordo com a interação individual do usuário com o portal e aplicação de entrevista com cada um. Os sujeitos da pesquisa que contribuíram para a concretização da pesquisa caracterizam-se por cinco alunos da disciplina *Seminários em Organização da Informação: Design Emocional da Informação* do Curso de Mestrado em Ciência da Informação da UnB como usuários do portal IbiCT.

Os resultados dos questionários revelam-se nos dados encontrados e analisados a seguir:

- 1. Quanto o estado afetivo (agradabilidade)** – 40% dos entrevistados ficam tranquilos, 40% ficam duvidosos e 20% ficam tristes.
- 2. Quanto ao sentimento ao navegar no sítio (agradabilidade)** – 40% se sentem irritados, 20% hipnotizados, 20% tristes e 20% não responderam.
- 3. Quanto à navegabilidade** – 80% acham regular e 20% insuficiente.
- 4. Quanto à arquitetura do conteúdo** – 80% acham regular e 20% insuficiente.
- 5. Quanto à usabilidade:**
 - 20% acham insuficiente, porque não encontram o que estavam procurando e pensaram estar o sítio desatualizado;
 - 60% acham regular, porque precisam aprender a mexer no sítio, ficaram meio confusos quanto à disponibilização das informações. Os menus não geram expectativa nem se caracterizam como atrativos aos usuários questionados, o que gera uma baixa interatividade entre os sujeitos da pesquisa e os usuários. Relataram também que, ao entrar em um *link*, perderam a navegação para outras áreas do sítio, obrigando-os a retornar a página anterior ou inicial; e
 - 20% acham boa.
- 6. Quanto à funcionalidade dos serviços:** 40% acham regular, 40% insuficiente e 20% boa.
- 7. Quanto às cores (agradabilidade):** 40% acham sombrias, 20% desagradáveis, 20% tristes e 20% acham agradáveis.
- 8. Quanto a mudanças no design da interface:**
 - 80% responderam que gostam quando a interface muda, pois as mudanças de visual são interessantes, representa inovação, mudança, atualização, novidade, dando a impressão de que o portal tem manutenção; caracteriza também atendimento à qualidade evolutiva na usabilidade do sítio; e
 - somente 20% responderam que não gostam quando a interface muda, porque produz uma sensação de desconhecimento, de “estar perdido”.
- 9. Quanto ao aspecto agradável, se o portal motiva o usuário a acessá-lo:** 60% responderam que sim e 40% não.
- 10. Quanto ao design da interface, se é um fator que faz com que o usuário se torne assíduo:** 100% responderam que sim.



A análise do impacto do *design* emocional da *interface* principal do portal do IbiCT nos usuários foi realizada com a participação de cinco usuários, questionados e em interatividade individual. Em relação à busca de informação no Portal, 40% dos usuários consideram o Portal agradável, com bom conteúdo e fácil de usar. Mas a emoção que a *interface* provoca gera tristeza, dúvida e, até mesmo, sensação de tranquilidade, ou seja, nada motivador, prazeroso ou que desperte o estímulo de continuar navegando no portal.

É relevante também registrar que os participantes que responderam ao questionário conhecem as técnicas que devem ser utilizadas, para que uma *interface* seja agradável e assim proporcionar uma prazerosa e motivadora recuperação da informação.

Em relação às perguntas abertas do questionário, os registros estão concentrados na melhoria do layout e da usabilidade do Portal, tomando como possíveis respostas as opções de regular a insuficiente.

8 CONCLUSÃO

Considera-se que os objetivos propostos foram alcançados, visto que ao final foi diagnosticado o impacto do *design* emocional na recuperação da informação do portal do IbiCT. A pesquisa possibilitou a identificação de alguns pontos a serem considerados para o aprimoramento da *interface* do portal quanto ao layout, usabilidade e a agradabilidade.

Constatou-se, conforme as respostas da maioria dos usuários, que as emoções geradas ao acessar a página principal do portal não são agradáveis. Identificou-se que a emoção dos usuários diante da interface provocada foi desde o estado de irritação, tranquilidade à dúvida.

Entende-se, assim, que o Instituto analisado deva implantar o aprimoramento da apresentação visual da interface principal do portal, visando a interação motivadora e agradável com o usuário, a manutenção do conteúdo do portal, para que o mesmo seja atualizado e o desenvolvimento de um *designer* da *interface* possibilite a busca no portal de forma intuitiva e fidelizada.

Espera-se que esta pesquisa seja útil para a melhoria da *interface*. Torna-se urgente o aprimoramento da página principal do portal do IbiCT, para que haja eficiência, eficácia e, principalmente, a efetividade de um portal de Informação em Ciência e Tecnologia.

REFERÊNCIAS

ANDRADE, Milton. **Testes de usabilidade, como efetuar**. Disponível em: <http://miltonandrade.com/novo/?category_name=usabilidade>. Acesso em: 6 jun. 2011.

BARROS, Tiago. **Agradabilidade**. Disponível em: <https://public.me.com/tiagobarros>. Acesso em: 04 jun. 2011. (PPT)

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro, 2002.



ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002. 24 p.

DAMASCENO, A. **Webdesign**: teoria e prática. Florianópolis, SC: Visual Books, 2003.

FURQUIM, Tatiana Almeida. Fatores motivadores de uso de *sítio web*: um estudo de caso. **Ci. Inf.**, Brasília, v. 33, n. 1, p. 48-54, jan./abril 2004.

GABRIELI, Leandro V.; Cortimiglia, Marcelo N.; RIBEIRO, José Luís D. Modelagem e avaliação de um sistema modular para gerenciamento de informação na Web. **Ci. Inf.**, Brasília, v. 36, n. 1, p. 35-53, jan./abr. 2007.

HACK, Josias Ricardo; SANTOS, Juliano Alves dos. Influência do design emocional na interação homem/computador. **Liinc em Revista**, v.6, n.2, setembro, 2010, Rio de Janeiro, p. 411- 421.

KRUG, S. **Não me faça pensar**: uma abordagem do bom senso à navegabilidade da Web. São Paulo: Alta Books, 2000.

MEMÓRIA, Felipe F. P. **Design para Internet**: projetando a experiência perfeita. Rio de Janeiro: Elsevier, 2005.

_____. Usabilidade de Interfaces e Arquitetura da Informação: Navegação Estrutural. Disponível em: <http://www.fmemoria.com.br/artigos/nav_estr.pdf>. Acesso em: 23 maio 2011.

NORMAN, D. A. **Design Emocional**: por que adoramos (ou detestamos) os objetos do dia a dia. Rio de Janeiro: Rocco, 2008.

PEZZINI, Marina. **Usabilidade**. Disponível em: <http://www.slideshare.net/marinapzn/ergo-3-aula-2-3321054>>. Acesso em: 19 maio 2011.

SILVA, Fábio Marques Brito da. **O impacto do design emocional na recuperação da informação no catálogo público de acesso em linha**. 2009. 61 f. Monografia (Bacharelado em Biblioteconomia) - Universidade de Brasília, Brasília, 2009.

THOMPSON, J. B. **A mídia e a modernidade**: uma teoria social da mídia. 10. ed. Petrópolis, RJ: Vozes, 2008.

TORRES, Elisabeth Fátima; MAZZONI, Alberto Angel. Conteúdos digitais multimídia: o foco na usabilidade e acessibilidade. **Ci. Inf.**, Brasília, v. 33, n. 2, p. 152-160, maio/ago. 2004

USABILIDADE. Disponível em: <<http://www.avellareduarte.com.br/projeto/interface/interface5/interface5.htm>>. Acesso em: 20 maio 2011.

VYGOTSKY, L.S. **Pensamento e linguagem**. São Paulo: Martins Fontes, 1993.

ANEXO

Questionário aplicado

O questionário foi desenvolvido para atender os objetivos da presente pesquisa sobre o tema: Avaliação do impacto do design emocional da informação na interface do portal do IbiCT.



1 Quando você está diante da interface do portal seu estado afetivo fica?



feliz triste duvidoso tranquilo irritado hipnotizado apaixonado

Fale um pouco sobre o motivo que o faz se sentir dessa forma?

2 Como você se sente ao navegar na interface?



Hipnotizado Irritado Apaixonado Dúvida Triste

3 Como você avalia a navegabilidade no portal?

O sítio deve permitir que diferentes recursos facilitem o acesso do conteúdo sem desconsiderar que as tecnologias utilizadas não podem estar incompatíveis com as mais comuns encontradas na rede.



insuficiente regular boa muito boa excelente

4 Qual sua opinião sobre a arquitetura do conteúdo?

Diz respeito à forma como o conteúdo do sítio é disposto, tornando sua navegação facilitada no que tange à utilização de recursos cognitivos como a informação é redigida ou a utilização de ilustrações.



insuficiente regular boa muito boa excelente

5 Avalie a usabilidade do portal.

Capacidade que apresenta um sistema interativo de ser operado, de maneira eficaz, e eficiente, em um determinado contexto de operação, para realização das tarefas de seus usuários. standards ISO 9241-11.



insuficiente



regular



boa



muito boa



excelente

Por quê?

6 Como você avalia a funcionalidade dos serviços *online* disponibilizados?



insuficiente



regular



boa



muito boa



excelente

7 Como avalia o uso das cores utilizadas no sítio?

agradáveis sombrias tristes vivas alegres desagradáveis

8 Você recebe bem uma mudança de design de interface de um sítio ao qual está acostumado a acessar?

Sim Não.

Por quê?

9 O aspecto emocional é um dos fatores que o faz acessar o portal?

Sim Não

10. Você considera o design da interface de um portal algo importante para que se torne usuário assíduo do mesmo?

Sim Não



REPOSITÓRIO ELETRÔNICO INSTITUCIONAL: PRESERVANDO O CONHECIMENTO ATRAVÉS DE TECNOLOGIAS DA INTELIGÊNCIA

Rosilene Fernandes Machado, Guilherme Ataíde Dias

Resumo: Esta pesquisa aborda a criação de um mecanismo que possibilite o acesso online, aberto e irrestrito aos documentos acadêmicos produzidos no Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba (CCEN/UFPB) à comunidade acadêmica. As Tecnologias de Informação e Comunicação (TIC) proporcionam o acesso e a disseminação da informação, e trazê-las para o meio acadêmico é colocar o indivíduo diante do mundo. Esta é uma pesquisa aplicada com o propósito de gerenciar a produção de Trabalhos de Conclusão de Curso (TCC) - CCEN/UFPB. O objetivo desta pesquisa é o de implantar uma política de preservação dos trabalhos monográficos do CCEN, através do uso de um Repositório Eletrônico Institucional (REI), com os seguintes objetivos específicos: divulgar o projeto na comunidade científica; prover o banco de dados com os TCC; divulgar a produção dos TCC do CCEN/UFPB. Metodologicamente, a pesquisa é de natureza descritiva, de abordagem quantitativa e qualitativa, e descreve as ações que as bibliotecas que fazem parte do Sistema de Bibliotecas da UFPB estão implementando para a implantação dos seus repositórios. Esta pesquisa está sendo conduzida nas dez bibliotecas setoriais que estão no Campus I da UFPB, além das bibliotecas setoriais situadas nas cidades de Areia (*Campus II*), Bananeiras (*Campus III*), Rio Tinto e Mamanguape (*Campus IV*), e seu universo será composto pelos bibliotecários responsáveis pela implantação do repositório.

Palavras-chave: Repositório Institucional. DSpace. Gestão da informação.

1 INTRODUÇÃO

As universidades e os institutos de pesquisa desempenham um papel sobremaneira importante na sociedade, porquanto são responsáveis pela geração e pela produção de saberes, pela formação de competências, pelo compartilhamento de conhecimentos e por atender às necessidades educacionais, culturais, tecnológicas e científicas da sociedade. Nesse contexto, a biblioteca universitária é a responsável pelo patrimônio informacional da universidade e tem como função educativa orientar os usuários na utilização da informação.

Com o advento do Programa de Apoio ao Plano de Reestruturação e Expansão das Universidades Federais (REUNI), nas Instituições Federais de Ensino Superior (IFES) brasileiras, especificamente na Universidade Federal da Paraíba (UFPB), houve um acréscimo do número de vagas dos cursos de graduação. Esse fato implicará o aumento significativo da quantidade de TCC, produção acadêmica dos cursos de graduação. Portanto, é fundamental a criação de um serviço voltado para a organização,



a disseminação e o acesso dos TCC para a comunidade acadêmica.

Atualmente, esses trabalhos são entregues às Coordenações de Cursos, que os encaminham às Bibliotecas Setoriais, em formato impresso. Assim, a tendência desse processo é de que haja uma proporção maior do acervo em relação ao espaço físico. Isso, provavelmente, incidirá em dificuldades ao acesso e à recuperação da informação.

Diante dessas questões, indagamos: **como gerenciar o processo de produção, armazenamento e disseminação de TCC do Centro de Ciências Exatas e da Natureza da UFPB?** Nessa perspectiva, é necessário implantar um mecanismo por meio do qual o acesso online a essas publicações seja favorável a um maior número de usuários. Dessa forma, alunos, professores, servidores e pesquisadores serão beneficiados com o uso das TIC, o que garantirá o acesso à informação de modo eficaz. Com isso, na Biblioteca Setorial do CCEN/UFPB, o referido projeto foi implantado como projeto-piloto.

A escolha deste trabalho se justifica pelo fato de a pesquisadora trabalhar como bibliotecária na biblioteca do CCEN/UFPB, razão por que pode acompanhar de perto a evolução do projeto, identificar possíveis problemas, criar mecanismos para resolvê-los e adotar políticas de compartilhamento com as demais bibliotecas setoriais que fazem parte do Sistema de Bibliotecas da UFPB (SISTEMOTECA) para implantar os seus repositórios.

2 O MOVIMENTO DO ACESSO LIVRE À INFORMAÇÃO

No final dos anos 90, verificou-se um aumento da produção de periódicos científicos, sendo o custo um dos fatores que dificultava o acesso a estes documentos. Conforme Kuramoto (2006), essa problemática iniciou um processo de movimentos sobre o acesso à informação científica e técnica, através de manifestos como o de Bethesda, o de Budapeste, o de Berlim e o do Brasil. Surgiram também outras manifestações de segmentos governamentais e internacionais, com o intuito de diminuir as barreiras para o compartilhamento da informação.

Em seu manifesto sobre o livre acesso à literatura científica e aos documentos de pesquisa, a *International Federation of Library Association and Institutions (IFLA)* declara:

O livre acesso à Internet, oferecido pelas bibliotecas e serviços de informação, contribui para que as comunidades e os indivíduos atinjam a liberdade, a prosperidade e o desenvolvimento. As barreiras para a circulação da informação devem ser removidas, especialmente aquelas que favorecem a desigualdade, a pobreza e o desespero (IFLA, 2002 apud MELO, 2009).

Entende-se que o acesso à informação proporciona o aumento da produtividade, levando o país ao desenvolvimento social e econômico. De acordo com Rodrigues (2004), o acesso livre, de maneira geral, contribui para que o usuário possa acessar amplamente a Internet e disponibiliza a literatura de caráter acadêmico ou científico a qualquer pessoa que queira ler, descarregar, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral dos documentos.

Esses novos serviços têm sido estimulados pelos órgãos de fomento e por universidades para manter repositórios institucionais de acesso livre e levar os profissionais da informação a reverem



iniciativas de reforma do sistema de comunicação da ciência. Assim, maximiza a visibilidade, consente e incentiva o acesso à produção científica das universidades, através do estabelecimento de políticas institucionais que empreendam esforços cooperativos e ações facilitadoras para a organização e estruturação dos recursos de informação em arquivos abertos.

3 REPOSITÓRIOS ELETRÔNICOS INSTITUCIONAIS: software DSpace

As TIC estão presentes no nosso cotidiano e criam instrumentos indispensáveis à comunicação humana, propiciando meios de acesso e disseminação da informação técnica e científica, através de arquivos abertos com grande efetividade para os serviços cooperativos, o intercâmbio de informações e o ordenamento delas.

Esses novos serviços estimulam as instituições a manterem repositórios institucionais de acesso livre e conduzem os profissionais da informação a reverem iniciativas de reforma do sistema de comunicação da ciência. Os repositórios eletrônicos são sistemas de informação que têm como objetivo gerir e disseminar os trabalhos gerados na própria instituição. A aplicação desses serviços facilita o fluxo da informação e contribui para o processo de construção do conhecimento para um indivíduo que dela necessite, ampliando seu processo cognitivo.

Das possibilidades tecnológicas existentes para a criação de repositórios institucionais, optou-se por software que não estivessem vinculados a licenças proprietárias. Após uma análise inicial de viabilidade, foram pré-selecionados os seguintes produtos: DSpace (<http://www.dspace.org/>), Fedora (<http://fedora-commons.org/>) e Eprints (<http://www.eprints.org/>). A escolha recaiu sobre o DSpace, em virtude de já ser utilizado por outras instituições federais de ensino superior e por sua reconhecida estabilidade e performance.

O sistema DSpace foi produzido para possibilitar a formação de repositórios digitais com funções de captura, distribuição e preservação da criação intelectual, a fim de que outras instituições possam adotá-lo em forma consorciada federada. Esse sistema já foi criado com características adaptáveis para várias instituições. Ele torna possível o gerenciamento da produção científica em vários suportes digitais, oferece mais visibilidade e garante a sua acessibilidade de maneira duradoura.

O repositório digital DSpace foi idealizado pelo Instituto Tecnológico de Massachussets (MIT) e pela Hewlett-Packard (HP). Trata-se de um software livre, por meio do qual as organizações podem gerenciar e conservar tudo o que é perceptível digitalmente. Além disso, auxilia na recuperação da informação, de maneira simples, e constrói uma grande rede de serviços de informação (SHINTAKU; MEIRELLES, 2010).

4 POLÍTICA DE IMPLANTAÇÃO

A política de informação começa com o diagnóstico do acervo de monografias existentes na biblioteca. Com isso, definem-se as técnicas e os processos a serem seguidos para o seu desenvolvimento, o que implica em um conjunto de ações que iniciam com uma análise situacional,



estrutural e organizacional que fazem parte do Sistema de Bibliotecas.

A implantação do projeto-piloto está sendo desenvolvida na Biblioteca Setorial do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba, o qual abrange os cursos de Física, Química, Matemática, Estatística, Biologia, Engenharia da Computação, Ciência da Computação, Geociência e Sistemática e Ecologia.

As diretrizes para a política de implantação do Repositório começaram em setembro de 2010, com a instalação do DSpace no servidor do Núcleo de Tecnologia da Informação (NTI) da UFPB. As fases relacionadas ao processo de construção do Repositório Institucional foram:

- a) Análise dos software disponíveis para a construção de repositórios institucionais;
- b) Escolha do software DSpace com versão do ambiente operacional Java 1.6.0_16, um banco de dados PostgreSQL e sistema operacional Linux;
- c) Avaliação da interface Manakin, criada pela Biblioteca da Universidade do Texas;
- d) Estudo dos atributos utilizados para a descrição da informação (Dublin Core);
- e) Análise das ferramentas de software livre associadas ao projeto;
- f) Avaliação da estrutura do sistema;
- g) Estruturação das comunidades:
 1. Configuração da Comunidade: CCEN (Centro de Ciências Exatas e da Natureza);
 2. Subcomunidade: Trabalhos de Conclusão de Curso – TCC;
 3. Coleções: Cursos de Física, Química, Matemática, Estatística, Biologia, Engenharia da Computação, Ciência da Computação, Geociência e Sistemática e Ecologia;
- h) Implantação de repositório-piloto;
- i) Avaliação do repositório-piloto;
- j) Disponibilização definitiva do repositório.

As condições favoráveis à implantação do projeto são a disponibilidade de equipamentos solicitados e necessários para a implementação do projeto e a aceitação da comunidade acadêmica que se beneficiará com o projeto, além de parcerias com professores e funcionários de outros centros da instituição.

Entre as possíveis barreiras para o desenvolvimento dos repositórios nas bibliotecas que fazem parte do Sistema de Bibliotecas da UFPB, estão a falta de recursos humanos e de um analista de sistemas voltado especificamente para atender às necessidades do sistema de bibliotecas.

O referido projeto será desenvolvido em parceria com as Coordenações de Graduação do Centro de Ciências Exatas e da Natureza da UFPB, e sua expansão contará com a parceria das bibliotecas setoriais da UFPB que fazem parte do Sistema de Bibliotecas da UFPB, visando aos seguintes benefícios: melhor visibilidade da produção científica; facilidade para gerir a produção científica; redução de custos de gestão da informação científica e disponibilidade do acesso permanente por meio de qualquer dispositivo *Web*.



Para isso, tem como propósito: Implantar e alimentar o Repositório da Biblioteca do CCEN com os TCC produzidos pelos alunos de graduação do Centro de Ciências Exatas e da Natureza entregue em DVD, como projeto-piloto; conscientizar os bibliotecários que fazem parte do Sistema de Bibliotecas da UFPB da importância de haver integração entre os componentes do SISTEMOTECA, com a finalidade de implantar uma nova cultura de disseminação da produção científica dos graduandos e incentivar a cooperação entre os dirigentes de cada centro para a implementação dos repositórios em seus respectivos centros.

5 PROCEDIMENTOS METODOLÓGICOS

Os procedimentos metodológicos envolvem um conjunto de abordagens técnicas e processos a serem seguidos na realização da pesquisa. A classificação da pesquisa será realizada a partir de uma análise descritiva sobre as ações que estão sendo implementadas para a implantação dos repositórios das bibliotecas que fazem parte do Sistema de Bibliotecas da UFPB.

A pesquisa será desenvolvida nas dez bibliotecas setoriais, que estão nos *Campi I* (João Pessoa), *II* (Areia), *III* (Bananeiras) e *IV* (Rio Tinto e Mamanguape), e os bibliotecários serão os responsáveis pela implantação dos repositórios nas bibliotecas setoriais de cada Centro e/ou Campus analisado.

Os instrumentos de pesquisa a serem utilizados serão questionários e entrevistas semiestruturadas, tendo em vista que se trata de uma pesquisa de natureza quantitativa, porquanto identifica o quantitativo de TCC defendidos em comparação com os disponibilizados no Repositório, com o intuito de identificar o nível de aceitação da comunidade discente; é, também, uma pesquisa qualitativa, em que serão identificados o processo de desenvolvimento de implantação dos Repositórios dos Centros e Campus envolvidos e a aceitação e as necessidades dos bibliotecários no desenvolvimento dos seus respectivos Repositórios.

A aplicação dos questionários será via e-mail para todas as bibliotecas envolvidas. Quanto às entrevistas, no primeiro momento, serão aplicadas em todas as bibliotecas setoriais do *Campus I*; em seguida, serão agendadas as entrevistas no Campus de Rio Tinto e no de Mamanguape; depois, no Campus de Areia e no de Bananeiras, por estarem situados em cidades mais distantes de onde se desenvolve o projeto-piloto.

Depois que o projeto for implantado efetivamente, serão desenvolvidas métricas para avaliar o acesso aos documentos pela comunidade acadêmica. Essa ação será implementada através do módulo gerador de estatísticas disponível no próprio DSpace.

6 CONSIDERAÇÕES FINAIS

Este projeto de pesquisa encontra-se em processo de mobilização, sensibilização e incentivo na comunidade envolvida. O modelo proposto está sendo socializado, a fim de minimizar as dificuldades no processo de disponibilização dos TCC.

O que se deseja na implantação desse repositório, como projeto de pesquisa, é contribuir para



a criação e a regulamentação de uma política de implantação do REI, que colabore para que todas as bibliotecas setoriais da UFPB se engajem para coletar, gerir, divulgar e preservar a produção dos TCC da instituição de forma ampla.

Como resultado parcial da pesquisa, mencionamos uma ação aplicada, com a implantação de um piloto do repositório, cujos atores envolvidos em seu processo definitivo de criação estão tendo suas impressões iniciais colhidas e avaliadas para que sejam feitos os devidos ajustes na configuração final da iniciativa.

REFERÊNCIAS

BARRETO, A.R. Implantação de gestão do conhecimento a partir dos repositórios de conhecimento explícito. Disponível em: <<http://badinfo.apbad.pt/congresso8/comm3.pdf>>. Acesso em: 08. set. 2010.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. In: **Ciência da Informação**. Brasília, v. 35, n. 2, p. 91-102. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/831/678>> Acesso em: 04. jun. 2011.

LAKATOS, E.M. Metodologia científica. 4.ed. São Paulo: Atlas, 2006.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto**. Brasília: IBICT, 2009.

MARCONI, M. de A.; LAKATOS, E.M. **Técnicas de pesquisa**. 2. ed. São Paulo: Atlas, 1990.

MARQUES, A. M. N.; MAIO, S. R. S. Repositórios Institucionais. Disponível em: <<http://repositoriosdigitais.web.simplesnet.pt/PDF'S/Artigo%20%20Repositorios%20Institucionais.pdf>> Acesso em: 25. set. 2010.

MELO, L. B.; SAMPAIO, M. I. C.; PIRES, C. **A questão do acesso aberto em Portugal e no Brasil**. Disponível em <www.sbu.unicamp.br/snbu2008/anais/site/pdfs/2796.pdf>

RODRIGUES, E. Acesso livre ao conhecimento: a mudança do sistema de comunicação da ciência e os profissionais de informação. In: **Cadernos BAD**. ISSN 0007- 9421 :1 (2004), 24-35. Disponível em: <<http://hdl.handle.net/1822/670http://repositorium.sdum.uminho.pt/bitstream/1822/670/1/Cadernos%20BAD%202004.pdf>>. Acesso em: 11. set. 2010.

_____. Repositório: criação e desenvolvimento do Repositório Institucional da Universidade do Minho. Disponível em: <<http://badinfo.apbad.pt/congresso8/com14.pdf>>. Acesso em: 06. set. 2010.

ROSA, F.; MEIRELLES, R. F.; PALÁCIOS, M. Repositório Institucional da Universidade Federal da Bahia: implantação e acompanhamento. **Inform. & Soc.:** Est, João Pessoa, v.21, n.1, p. 129-141, jan./abr. 2011. Disponível em: <http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/4058/5603> Acesso em: 02. jul. 2011.

ROZADOS, H. B. F. Uso de indicadores na gestão de recursos de informação. **Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas, v.3, n.1, p.60-76. jul/dez. 2005.

RODRIGUES, E. Acesso livre ao conhecimento: a mudança do sistema de comunicação da ciência e os profissionais de informação. In: **Cadernos BAD**. ISSN 0007- 9421 :1 (2004), 24-35. Disponível em: <http://hdl.handle.net/1822/670>. Acessado em 22 jun. 2010.



METADADOS PARA A REPRESENTAÇÃO DAS IMAGENS DIGITAIS

Ana Carolina Simionato

Resumo: As redes se configuram como um grande centro de circulação de informações de rápido acesso, este universo digital é constituído por várias informações, entre elas as imagens digitais, que são capturadas e divulgadas por diversos tipos dispositivos diariamente em todo o mundo. Assim, destaca-se a importância do Tratamento Descritivo da Informação para a imagem digital na construção de formas para representação e apresentação, a fim de torná-las disponíveis, acessíveis e recuperáveis. Neste sentido, apresenta-se como problema de pesquisa: a unicidade exigida para a localização e recuperação destes recursos informacionais em ambientes digitais, deve ser por elementos que descrevem sua forma, de maneira a tornar possível a identificação de um esquema para construção de formas de representação da imagem de modo a torná-la única, o que exige um tratamento para além de conteúdo. Portanto, o objetivo desta pesquisa constitui-se em determinar quais os atributos necessários para a representação da imagem digital, tendo como base os instrumentos de descrição pertencentes ao Tratamento Descritivo da Informação (TDI) em sua padronização. Assim, esta pesquisa pontuará a importância dos metadados na descrição das propriedades da imagem digital e os elementos responsáveis para a recuperação da informação. A metodologia constitui-se em uma análise exploratória e descritiva da literatura disponível sobre o tema proposto. O que permitirá a construção de um conhecimento teórico sobre os instrumentos que orientam a construção de representações na área de Ciência da Informação.

Palavras-chave: Informação e Tecnologia; Catalogação de imagens digitais; Imagens digitais; Metadados para imagens digitais; Padrões de metadados.

1 Introdução

As redes se configuram como um grande centro de circulação de informações de rápido acesso e o crescimento exponencial de recursos é uma das características desses ambientes. Transformar o volume de dados disponíveis em informações significativas para os usuários requer um processo de persistência de dados, que gere o instanciamento e a recuperação de informações cada vez mais eficiente. Neste sentido, com crescente fluxo das informações imagéticas pensa-se em considerar as



formas de tratamento para a utilização e reutilização dos recursos e informações disponíveis.

Nesse sentido, ressalta-se na descrição das imagens digitais¹ o catalogador precisa se preocupar com a maneira pela qual o conteúdo é visualizado, analisado, interpretado e apresentado. O catalogador tem como principal tarefa: analisar sistematicamente e descrever os elementos de uma imagem para a sua recuperação e utilização, tornando possível ao usuário ter acesso à informação que procura. O processo de construção de formas de representação permite que as informações esquematizadas e estruturadas ofereçam a identificação, a localização e a recuperação automática das informações, e tais representações devem propiciar, por meio dos padrões de descrição, a persistência de dados, o instanciamento de recursos e a interoperabilidade entre sistemas de informação (SANTOS, 2008).

Assim, o propósito desta pesquisa é a análise e a verificação das contribuições do Tratamento Descritivo da Informação (TDI)² para a construção de formas de representação para a informação imagética, no intuito de construir uma proposta de metadados padronizados para a representação da imagem digital.

2 A relação da Ciência da informação com a recuperação de imagens digitais

A área de Ciência da Informação se depara com um ponto crítico a ser solucionado, que se constitui na essência e no ponto de origem dessa ciência: a necessidade de tornar acessível e disponível o crescente volume de informações produzidas buscando soluções estratégicas e metodológicas, para os problemas relacionados à recuperação de informações e de recursos informacionais. O que exige que os aspectos tecnológicos envolvidos no processo de representação da informação para a otimização de sua utilização. Um diálogo interdisciplinar se faz necessário na proposta de criar novas formas de tratamento dos recursos informacionais no contexto web em uma proposta de uso conjugado de uso tecnologias disponíveis e os métodos de representação da informação e sua aplicação em recursos imagéticos digitais.

2.1 Metadados

O termo metadados adquire uma maior amplitude semântica para denominar as novas formas de representação e descrição dos recursos informacionais, como exemplo os dados EXIF

1 Uma unidade bibliográfica no ambiente digital não é diferente da unidade bibliográfica no contexto analógico, porém apresenta-se com propriedades físicas distintas e exige uma infraestrutura de descrição que permita a identificação, escolha, localização e recuperação do recurso a partir da produção de registros bibliográficos com um nível profundo de granularidade em uma estrutura de representação pautada na utilização de normas específicas que permitam a extração e/ou a geração automática de metadados (JONSSON, 2002).

2 O universo da pesquisa se constitui dos instrumentos de representação presentes no TDI, como o AACR2 (Código de Catalogação Anglo-Americano), ISBD (Descrição Bibliográfica Internacional Normalizada), *Grafic Materials*, RDA (Descrição de Recursos e Acesso), *Schema della struttura dei dati* (Shelda F), MARC21 (Catalogação Legível por Computador) e *Dublin Core*.



(informações técnicas de captura da imagem gravadas junto ao arquivo da imagem na forma de metadados etiquetados). Assim, por meio da necessidade de denominar o tratamento da informação em meio eletrônico; a palavra metadados passa a ser utilizada com intensidade e em diversas áreas com o mesmo objetivo: dar tratamento às informações tendo em vista o gerenciamento informacional. Segundo a definição, os metadados são:

atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação (ALVES, 2010, p.47).

Neste sentido, a função dos metadados é descrever recursos informacionais, apresentando uma divisão em cinco níveis: *administrativos* (metadados usados no gerenciamento e administração dos recursos informacionais), *descritivos* (metadados usados para descrever e identificar informações sobre recursos), *conservação* (metadados relacionados à conservação de recursos de informação), *técnicos* (metadados relacionados ao funcionamento do sistema e comportamento dos metadados) e *uso* (metadados relacionados com o nível e tipo de uso dos recursos informacionais). Já os padrões de metadados estabelecem regras para a definição dos atributos, para a) obter coerência interna entre os elementos por meio de semântica e sintaxe; b) promover necessária facilidade para esses recursos serem recuperados pelos usuários; c) permitir a interoperabilidade dos recursos de informação (ALVES, 2010, p.47).

Conforme aponta Zeng e Qin (2008) existe uma intrínseca ligação no estabelecimento de metadados e formatos de metadados. O estabelecimento de um formato de metadados exige a construção padronizada de metadados (atributos que representarão uma entidade), assim como os metadados precisam estar em uma estrutura de descrição padronizada (esquema de metadados) que irá compor o formato para serem úteis e válidos em um sistema de informação.

De modo geral, pode-se dizer que o objetivo e a função dos metadados são os mesmos da catalogação: representar as características e o conteúdo de um recurso informacional de forma padronizada, facilitando a identificação, a busca, a localização e recuperação desses recursos (HSIEH-YEE, 2006). A diferença encontra-se na forma de representação do recurso informacional determinado pelo ambiente tecnológico.

Desta forma, destaca-se a importância do Tratamento Descritivo da Informação para a imagem digital, com o intuito de construir formas mais eficientes para representação e apresentação é o modo de torná-las disponíveis, acessíveis e recuperáveis. Esforço especialmente necessário para a padronização das orientações e para a descrição do material imagético digital, assim explicitando as diferenças do recurso imagético analógico.



3 Representação e recuperação das imagens digitais

Do mesmo modo, a estrutura de representação de recursos informacionais encontra-se na camada intangível dos sistemas de informação e ao usuário o acesso se dá na camada aparente do sistema, na camada de apresentação. É nesta camada que se estabelecem as solicitações e o uso dos marcadores de memória, são apresentados os resultados de busca em diferentes formas de apresentação para a solução de problemas específicos.

A exigência para este problema são sistemas gerenciadores de formas de representação, que sejam capazes de reconhecer atributos e relacionamentos caracterizados pelos metadados e pelas ontologias para localizar as informações solicitadas disponíveis em ambientes informacionais distribuídos, como também gerar apresentações adequadas. Nos últimos anos, vê-se que no sentido das divergências, a concepção e a implementação de sistemas de metadados são uma das questões-chave no campo de pesquisa da representação da imagem digital.

Na concepção pelos metadados de imagens, a questão mais importante é a definição e representação dos metadados de imagens, e como consequência a extração de imagens dinamicamente de acordo com impressão do usuário e o conteúdo da imagem (HAYAMA; KIYOKI; KITAGAWA, 1994). Se os profissionais da informação produzem e padronizam metadados desde as primeiras tentativas de organização da informação, recentemente, os profissionais de outras áreas têm criado novos métodos de tratamento da informação, sem conhecer o que já foi feito.

Simone Garcia (1999) propõe um desenvolvimento de uma modelagem conceitual para a descrição e recuperação de imagens estáticas digitalizadas (tipo fotografia, pintura ou gravura qualquer), baseadas em metadados para a descrição de características técnicas e semânticas de imagens, segundo abordagens estudadas na Ciência da Computação e na Ciência da Informação. Este esquema permite explicitar um rico conjunto de elementos que pode ser utilizado em arquiteturas de metadados existentes, tais como *Warwick*, *RDF* e *MCF*, visto que as arquiteturas não se preocupam em especificar elementos descritores e sim descrever o ambiente de forma a prover a interoperabilidade entre os padrões de metadados, porém a catalogação em si é pouco apontada pela pesquisadora. Portanto, ainda há o que investigar sobre o que já temos realizado nesta área como estudar os instrumentos necessários para a construção de representações informacionais das imagens digitais no âmbito do tratamento descritivo da informação.

4 Considerações

Esta pesquisa está em desenvolvimento, porém apresentam-se algumas considerações: a representação das informações imagéticas digitais deve estar de acordo com um padrão de representação do seu domínio (campo de estudo). E que a descrição deve ser construída a partir de uma observação sistemática qualitativa dos materiais, por meio dos parâmetros de descrição imagética e das traduções



intersemióticas.

E como o Tratamento Descritivo da Informação desenvolvido no domínio bibliográfico o que favorecerá e agilizará o processo de recuperação de informações nos ambientes informacionais digitais, e ainda auxiliará sistemas digitais de busca, ou mesmo, sites de armazenamento de imagens digitais já disponíveis, como: *Flickr*, *Picasa* e o Banco de Imagens da *Library of Congress*.

REFERÊNCIAS

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 134f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

GARCIA, S. S. **Metadados para documentação e recuperação de imagens**. 1999. 152f. Dissertação (Mestrado em Sistemas e Computação) – Instituto Militar de Engenharia, Rio de Janeiro, 1999.

HAYAMA, T.; KIYOKI, Y.; KITAGAWA, T. A metadatabase system for semantic image search by a mathematical model of meaning. In: **ACM SIGMOD Record**, v.23, n.4, p. 34 -41, 1994.

HSIEH-YEE, I. **Organizing audiovisual and electronic resources for access: a cataloging guide**. Washington: Library of Congress, 2006. 375p.

JONSSON, G. **The basis for a record in major cataloguing codes and the relation to FRBR**. In: 68th IFLA Council and General Conference. 2002. Disponível em: <http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED472894&ERICExtSearch_SearchType_0=no&accno=ED472894> . Acesso em: 08/07/2010.

SANTOS, P. L. V. A. C. Redes informacionais como ambiente colaborativo e de empoderamento: a catalogação em foco. In: FUJITA, M. S. L.; GUIMARÃES, J. A. C. (Org.) **Ensino e Pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar**. Marília: FUNDEPE, 2008.

ZENG, M. L.; QIN, J. **Metadata**. New York: Neal-Schuman Publishers, 2008.



USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO MULTIMODAL

Emilio Evaristo de Sousa, Cláudio Gottschalg Duque

Resumo: Recuperar informação disponibilizada em vídeos longos pode ser uma experiência demorada. A recuperação da informação em vídeos apresenta alguns problemas, pois é realizada, principalmente, através das descrições textuais feitas pelos usuários. O objetivo desta pesquisa é promover, com maior índice de precisão (Cleverdon, 1962), a recuperação da informação disponibilizada em vídeos por meio da indexação multimodal e uso de ontologia. Foi desenvolvida uma ontologia de domínio, associada com uma indexação multimodal para avaliação da precisão em um sistema de recuperação da informação audiovisual. Observou-se nos experimentos um aumento significativo da precisão nas configurações que utilizaram a ontologia para recuperar informação audiovisual. A utilização de uma ontologia de domínio da Doutrina Espírita alinhada a indexação multimodal, em um sistema de recuperação de trechos de vídeos de palestras sobre esse tópico, permitiu melhorar a eficiência do mesmo em termos de precisão.

Palavras-chave: Recuperação da Informação Audiovisual. Indexação Multimodal. Ontologia.

1 INTRODUÇÃO

A *World Wide Web* já se tornou uma grande biblioteca virtual, onde a informação sobre qualquer assunto está disponível a qualquer hora e em qualquer lugar, com ou sem custo, e em diferentes formatos como textos, áudios, vídeos, etc.

No atual estágio da internet, a quantidade de informação e conteúdo disponibilizado por meio de vídeos vem crescendo a cada dia. Segundo dados da *ComScore*¹, somente em setembro de 2010, foram realizadas 10,6 bilhões de buscas no *Google*. O *site* do *Yahoo* chegou a receber 2,7 bilhões de buscas no mesmo período.

No âmbito dos conteúdos multimídias, o esforço volta-se para conseguir a recuperação da informação com eficiência através desses motores de busca e outras ferramentas automáticas. Repositórios de vídeos e de imagens estão entre os sites mais acessados da Internet. O *YouTube*, que oferece busca e compartilhamento de vídeos, já é o terceiro site mais acessado do mundo, segundo *Alexa, The Web Information Company*².

¹ Pesquisa mensal realizada nos Estados Unidos em setembro de 2010.

² Os dados da *Alexa* são atualizados diariamente. Acesso em 23/05/2011.



Diante de tal realidade, encontrar a informação precisa no *YouTube* é uma tarefa desafiante devido ao tamanho do repositório e à sua grande base de dados textuais (Sureka et al, 2010).

Assim, Brascher (2002) afirma que:

A necessidade de recuperação de informações armazenadas em grandes repositórios de informação disponíveis na Internet e de responder com maior precisão às buscas realizadas diretamente pelos usuários finais, têm levado a um esforço no sentido de adicionar informação semântica às páginas Web. (BRASCHER, 2002,p.1).

Os sistemas de recuperação da informação atuais dependem do uso de termos de indexação na busca para obter os documentos relevantes. Sendo as coleções de documentos indexadas com base no conteúdo textual, os documentos não textuais (imagens, vídeos) dependem de uma indexação manual (associação de termos de indexação) após a interpretação de seu conteúdo.

Um método para melhorar a recuperação seria indexar o conteúdo visual e o áudio, e usar formas mais elaboradas de extração do significado. O sistema de recuperação deve ser capaz de processar documentos e criar descrições que relacionem os objetos, usando palavras, ontologias subjacentes e a análise audiovisual.

Avançar para além da recuperação de texto requer a análise de segmentos de áudio e vídeo. De que forma a utilização de descritores audiovisuais, conjuntamente com ontologias de domínio, poderá auxiliar na recuperação da informação disponibilizada em vídeos?

Essa preocupação pode-se apresentar nos termos seguintes: como armazenar e manipular (processar, gerir) o conhecimento nos sistemas de informação, de maneira formalizada de tal modo feita que possa ser usada por mecanismos para realizar uma tarefa determinada?. (ROBREDO, 2005).

O objetivo dessa pesquisa é promover, com maior índice de precisão, a recuperação de uma informação específica existente em um repositório de vídeos por meio da indexação multimodal e uso de ontologia.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os conceitos teóricos sobre indexação audiovisual, multimodalidade e ontologia, estabelecendo uma base para a verificação proposta neste trabalho.

2.1 Indexação da Informação Audiovisual

Com o excesso de informação já existente, somada às milhares que são produzidas diariamente no contexto da web, enfrentamos questões sérias relativas a organização e recuperação do conteúdo, e a indexação é um fator determinante para que um usuário que busca por informação recupere documentos que satisfaçam suas necessidades.



A informação audiovisual, de acordo com Barreto (2007), apresenta grandes desafios relativos ao armazenamento, indexação, formulação de consultas e recuperação de conteúdo semântico. Para o autor, um caminho seria a construção de sistemas híbridos, com indexação automática e análise de conteúdo supervisionada, pois existem sérias limitações ao uso de indexadores manuais, uma vez que requerem anotação individual, dificultando seu uso em grandes repositórios de informação, e que sofrem influência tanto do domínio de aplicação quanto do conhecimento da pessoa que realiza a tarefa.

Segundo Lancaster (2004),

A capacidade de armazenar, em formato digital, em bases de dados, qualquer tipo de imagem, e especialmente de poder acessar milhões delas na Rede, causou impressionante ressurgimento do interesse por imagens em geral, e, em particular, por modos de indexá-las (LANCASTER, 2004, p.214).

Para Barreto (2007, p.20) a recuperação eficaz do conteúdo visual e sonoro só é possível com uma indexação significativa e discriminante, e que deve estar relacionada com intenções e procedimentos do usuário quando faz a consulta no ambiente real.

Tal afirmação indica que o usuário é fator determinante para a indexação, e que esta deverá ser baseada em suas necessidades informacionais. A precisão da busca no processo de recuperação da informação vai demonstrar a utilidade do sistema, utilidade esta determinada pelo usuário, como percebemos nas palavras de Araújo Júnior (2005),

O papel da precisão em um processo de busca e recuperação da informação é dar a noção exata se o que está sendo recuperado na base de dados é útil ao usuário. Desta conclusão uma série de decisões poderão mudar os rumos de qualidade da resposta que se obtém nos sistemas de recuperação da informação (ARAÚJO JR, 2005).

2.2 Indexação Multimodal

Analisando a partir da perspectiva geral do sistema, a multimodalidade é a capacidade do sistema para se comunicar com o usuário ao longo de diferentes tipos de canais de comunicação e extrair e transmitir um significado automaticamente (O'Halloran, 2004). Outra perspectiva, é a visão que a multimodalidade dos meios de comunicação de vídeo é a capacidade de um autor para expressar uma ideia pré-semântica, através da combinação de um *layout* com um conteúdo específico, usando pelo menos dois canais de informação, onde os canais podem ser visuais, auditivos ou textuais (Snoek e Worrying, 2005).

Como o sistema está utilizando várias técnicas de indexação de vídeos, a qualidade dessas técnicas irá certamente afetar a qualidade das informações de saída. A aplicação de técnicas multimodais e processos de avaliação podem ajudar a reduzir este problema.



2.3 Ontologia

Ontologia é uma descrição formal e explícita de uma conceitualização compartilhada (Gruber, 1996). Já outros autores acreditam que ontologia é uma teoria lógica que fornece um relato explícito e parcial de uma conceitualização (Guarino e Giaretta, 1995). Essa forma de organizar o conhecimento visa descrever estruturas conceituais de domínios específicos, e tornar viável o uso do vocabulário compartilhado de uma maneira coerente e consistente.

As ontologias são estruturadas de tal maneira que permitem um considerável ganho de qualidade quando empregadas num sistema de classificação. Elas oferecem maiores possibilidades estruturais (classes; instâncias; parte-todo; pai-filho; etc.) das que são oferecidas por outros sistemas, como por exemplo, thesauri. A ideia é que o índice, criado a partir de estruturas conceituais geradas por meio do resultado de extensa análise de linguagem natural, apresente um melhor desempenho para as respostas às consultas de usuários. (DUQUE, 2006).

Uma das características principais das ontologias é que elas permitem compartilhar informações de um domínio, podendo ser utilizada por diversas aplicações. Um requisito para compartilhamento é a padronização da linguagem de representação. Nesse sentido, existem diversas linguagens que surgiram para a representação de ontologias que podem ser utilizadas (Corcho e Gómez-Pérez, 2000). A Web Ontology Language, OWL (McGuinness e Harmelen, 2003), é uma recomendação apresentada pela W3C como uma linguagem de ontologias para a web. OWL aumenta a expressividade da linguagem pois permite identificar restrições de propriedade, de equivalência, e quantificadores.

Objetivando aumentar a precisão no processo de busca e recuperação da informação contida em vídeos, utilizando uma ontologia OWL e a indexação multimodal, propomos uma arquitetura inovadora, apresentada na seção a seguir.

3 ARQUITETURA

O protótipo construído combina muitos tópicos discutidos na investigação de recuperação da informação audiovisual, como ontologia, a multimodalidade, a indexação de imagens e indexação de áudio. O protótipo foi baseado em uma solução aberta desenvolvida na linguagem *java* para integrar, posicionar, arquivar, e distribuir gravações/apresentações audiovisuais que utilizam *slides* como apoio. O processo do uso de ontologia para recuperação de informação audiovisual por meio de análise multimodal incluiu os seguintes passos:

1. Segmentação do vídeo e indexação da imagem: Um dos módulos do protótipo criado possui a característica de segmentar os vídeos a cada novo slide identificado. A partir da segmentação do vídeo e identificação dos *slides* foi possível utilizar o Reconhecimento Óptico de Caracteres (OCR) para extrair o conteúdo imagético. O protótipo utiliza a tecnologia de OCR alinhada



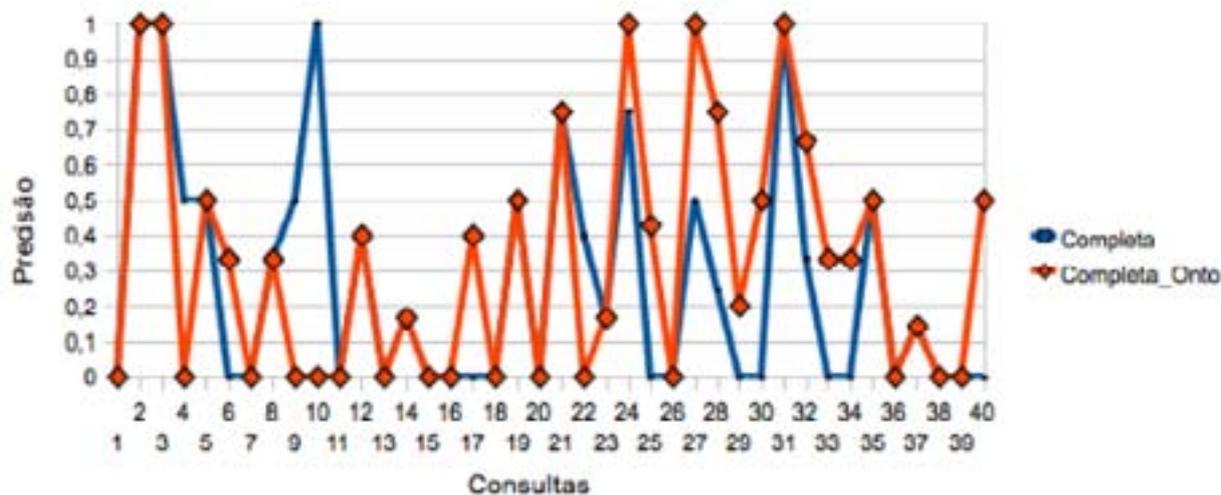
- a um dicionário pré-definido do Português Brasileiro para aumentar a taxa de acerto no reconhecimento das palavras. Assim, além de segmentar em *clips* de vídeo, todas as palavras constantes nos *slides* dos palestrantes nos vídeos foram indexadas de forma sincronizada.
2. Indexação do áudio: O áudio foi outro canal de comunicação indexado nesta pesquisa, e apesar de existirem algumas propostas, modelos, projetos e até soluções abertas para indexação automática do áudio, utilizamos a indexação manual. Esta escolha deve-se pela necessidade de adaptação computacional com o protótipo, o que demanda maior complexidade computacional. Assim, fizemos a indexação manual e adicionamos as palavras indexadas ao protótipo. O protótipo contém as palavras resultantes da indexação multimodal. É importante salientar que as palavras indexadas foram relacionadas em sincronia com a ordem na qual aparecem nos *slides* ou foram pronunciadas nos vídeos.
 3. Utilizando a ontologia: Recentemente, ontologias têm sido utilizadas na recuperação da informação para melhorar a revocação e a precisão. Sua principal utilização está relacionada com a expansão da consulta, que consiste em incrementar os termos relacionados pela ontologia com os termos da consulta. Utilizamos o Método 101 (Noy; McGuinness, 2001) no desenvolvimento da ontologia deste trabalho. A ontologia aborda conceitos relativos à Doutrina Espírita, com foco nos conceitos utilizados e difundidos atualmente. O propósito da ontologia é realizar o mapeamento dos conceitos da Doutrina Espírita para integração com o protótipo, que foi implementado para auxiliar na recuperação da informação. Deste modo, a ontologia procurou definir, de forma não ambígua, todos os conceitos e relações deste domínio. A ontologia desenvolvida foi usada com o intuito de expandir a rede semântica e diminuir ambiguidades.

4 RESULTADOS PRELIMINARES

Realizou-se um experimento objetivando a verificação de coerências na utilização de Multimodalidade e Ontologia em um Sistema de Recuperação de Informação Audiovisual. Nesse experimento foram utilizadas 6 horas de vídeos produzidos em língua portuguesa, e parte da coleção da televisão do Conselho Espírita Internacional. Para a realização do experimento, utilizou-se duas configurações distintas: uma somente com a indexação multimodal (Completa) e a outra com uso da ontologia associada à indexação multimodal (Completa_Onto).

Foram realizadas consultas à coleção, utilizando palavras-chave desejadas pelo usuário, e avaliação dos cinco primeiros documentos retornados pela busca das duas configurações, de acordo com a relevância do documento para a consulta. Para viabilizar esta comparação, utilizamos como métrica avaliativa o índice de precisão, uma medida objetiva e consagrada desde a sua proposta por Cleverdon (1962).

Gráfico 1 – Resultado de Precisão do experimento comparando as configurações Completa e Completa_Onto



Fonte: Dados da Pesquisa

No gráfico 1 observam-se os resultados de Precisão obtidos tanto para a configuração Completa quanto para a configuração Completa_Onto. Os resultados de Precisão apresentados pela Completa_Onto foram em média 18,51 % (dezoito, cinquenta e um por cento) superiores aos resultados apresentados pela configuração Completa nas consultas realizadas no experimento. Existem 4 consultas em que a configuração Completa, sem o uso da ontologia, obteve maior precisão que a configuração Completa_Onto. Nesse caso específico, o uso da expansão semântica apresentou documentos irrelevantes devido ao sentido da consulta estar diferente do determinado nas relações da ontologia.

5 CONCLUSÕES

A utilização de uma ontologia de domínio da Doutrina Espírita alinhada à indexação multimodal, em um sistema de recuperação de trechos de vídeos de palestras sobre esse tópico, permite melhorar a eficiência do mesmo em termos de índice de precisão.

Este trabalho consistiu na elaboração de um sistema de recuperação de informação (RI) que utiliza ontologias e indexação multimodal. O uso de ontologias para recuperação da informação disponibilizada em vídeos mostrou-se promissor. Alguns dos termos utilizados na versão da ontologia descrita são dependentes do contexto no qual será feito uso das informações, e isto deve ser melhor avaliado em futuras versões da ontologia.

O tamanho da coleção é um fator importante no desenvolvimento e, principalmente, na avaliação de sistemas de RI. A variedade de consultas que podem ser feitas é muito grande, e o sistema termina, em muitos casos, trazendo respostas não muito relevantes, pois não conta com respostas adequadas para a consulta do usuário. O conjunto de vídeos utilizado no teste ainda é pequeno. Isso faz com que um único documento irrelevante recuperado ou, um único documento relevante não recuperado, afete consideravelmente a precisão do sistema.



Abstract: *Retrieving information available on long videos can be a time consuming experience. Information retrieval on video presents some problems because users mainly do it through textual descriptions. The objective of this research is to promote, with more precision (Cleverdon,1962), information retrieval available on video through the use of multimodal indexing and use of ontology. It was developed a domain ontology, associated with an multimodal index for evaluation of accuracy in a retrieval system of audiovisual information. It was observed in the experiments a significant increase in accuracy in settings that used the ontology to retrieve audiovisual information. The use of a domain ontology of the Spiritist Doctrine aligned multimodal indexing in a retrieval system for sections of videos of lectures on this topic, allowed for increased efficiency in terms of the same precision.*

Keywords: Audiovisual Information Retrieval. Multimodal Indexing. Ontology.

REFERÊNCIAS

ARAÚJO JÚNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação.** Tese (Doutorado em Ciência da Informação) – Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2005.

BARRETO, Juliano Serra. Desafios e avanços na recuperação automática da informação audiovisual. **Ciência da Informação**, Brasília, v.36, n. 3, p.17-28, set./dez, 2007.

BRASCHER, M. A ambigüidade na recuperação da informação. **DataGramaZero**, Rio de Janeiro, v.3, n.1, fev. 2002.

CLEVERDON, C. W. **Report on testing and analysis of investigation into comparative efficiency of indexing systems.** Cranfield: Aslib, 1962.

CORCHO, O; GOMEZ-PEREZ, A. A roadmap to ontology specification languages. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT METHODS, MODELS AND TOOLS, 12., 2000, France. **Proceedings...** France, 2000. p.80-97.

DUQUE, C. G. **Uma abordagem ontológica para a indexação de documentos eletrônicos.** Brasília, 2006. Disponível em: <<http://www.asocarchi.cl/DOCS/53.PDF>>. Acesso em 07 abr. 2009.

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

McGUINNESS, D. L.; HARMELEN, F. V. **OWL web ontology language overview: W3C Candidate Recommendation.** Disponível em : <<http://www.w3.org/TR/2003/CR-owl-features-20030818/>>. Acesso em 05 fev. 2011.

NOY, F. N.; GUINNESS, D. L. **Ontology development 101: a guide to create your first ontology.** Disponível em: <<http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.doc>>. Acesso em 03 abr. 2011.



O'HALLORAN, Kay L. **Multimodal discourse analysis: systemic functional perspectives**. London: Continuum, 2004.

ROBREDO, J. **Documentação de hoje e de amanhã**. 4. ed. Brasília: Edição de autor, 2005.

SNOEK, C.G.M.; WORRING, M. Multimodal video indexing: a review of the state-of-the-art. **Multimedia tools and applications**, [S.I], v. 25, n. 1, p. 5-35, 2005.

SUREKA, A. Mining youtube to discover extremist videos, users and hidden communities. **AIRS**, v. 6458, p. 13-24, 2010.



WEB SEMÂNTICA E SUAS CONTRIBUIÇÕES PARA A CATALOGAÇÃO:

UM ESTUDO SOBRE ARQUITETURAS DE METADADOS E FRBR

Renata Eleuterio da Silva, Plácida Leopoldina Ventura Amorim da Costa Santos

Resumo: Os avanços tecnológicos provenientes da evolução da Internet proporcionaram um crescimento exorbitante das informações contidas na Web, dificultando o acesso e recuperação destas. Muitas se localizam em uma camada profunda da Web, na qual mecanismos de busca Web não conseguem chegar. É o caso das informações contidas nos bancos de dados e sistemas de bibliotecas, cuja recuperação de registros e documentos digitais somente podem ser encontrados através de suas próprias interfaces de busca. A proposta de uma Web Semântica, idealizada por Tim Berners-Lee no ano de 2001, surgiu como uma alternativa ao caos da Internet, ao passo que possibilitaria a interpretação das informações por máquinas, permitindo, assim, maior qualidade nas buscas e resultados mais relevantes. Para tanto, desenvolveram-se diversas tecnologias e mecanismos inteligentes, estruturados em metadados, padrões, linguagens, ontologias e agentes de busca. Objetivava-se identificar as contribuições que a Web Semântica pode oferecer à área de Ciência da Informação e, sobretudo à modelagem e arquitetura de metadados em catálogos online de bibliotecas, baseando-se nos conceitos contidos nos FRBR, um modelo conceitual desenvolvido para a representação do universo bibliográfico cuja estrutura se baseia na modelagem entidade-relacionamento. A proposta central será estudar diversas arquiteturas de metadados semântica, além de analisar o modelo FRBR e o desenvolvimento de uma ontologia baseada em seu vocabulário e conceitos.

Palavras-chave: Informação e Tecnologia. Catalogação. Web Semântica. Arquiteturas de metadados. FRBR.

1 INTRODUÇÃO

A constante evolução das Tecnologias de Informação e Comunicação (TIC) se destaca como fator fundamental para o desenvolvimento das pesquisas na área da Ciência da Informação, levando-se em conta o fato de que tais tecnologias avançam consideravelmente em curtos períodos de tempo. Há, então, a necessidade de pesquisar os impactos de tais tecnologias, principalmente, no que se refere ao uso destas na prática profissional, momento no qual o estudo da teoria e a investigação se configuram como etapas essenciais.

O aumento nos fluxos informacionais, gerados pela evolução da internet, torna fundamental o desenvolvimento e a otimização de mecanismos de busca e recuperação nesses ambientes informacionais.



Desde a criação dos primeiros códigos de catalogação, especialistas e pesquisadores buscam sua melhor utilização no Tratamento Descritivo da Informação (TDI), levando em consideração a eficácia da recuperação informacional nos ambientes informacionais. Em 1998, um grupo de estudos da IFLA (*International Federation of Librarian Association and Institutions*) publicou um relatório de pesquisa desenvolvida entre os anos de 1992 e 1995, que ficou mundialmente conhecido no campo da catalogação como os FRBR, sigla para *Functional Requirements for Bibliographic Records*, em português, Requisitos Funcionais para Registros Bibliográficos. (IFLA STUDY GROUP..., 2009)

No contexto da internet, a recuperação da informação se configura como um dos maiores desafios aos profissionais e pesquisadores da área, já que há um aumento exponencial de informações que precisam ser descritas e localizadas por mecanismos de busca. A Web Semântica, idealizada por Tim Berners-Lee e o W3C (*World Wide Web Consortium*) no início deste século, é uma extensão da Web atual, que busca integrar significado aos conteúdos Web. (BERNERS-LEE; HENDLER; LASSILA, 2001)

Com base nos FRBR, nas propostas que a Web Semântica tem com relação à recuperação da informação e ao fato das informações contidas em catálogos online de bibliotecas não serem cobertas por mecanismos de busca da Web, se apresenta como o problema desta pesquisa: os conceitos e arquiteturas de metadados propostos para a Web Semântica podem contribuir positivamente para a construção de catálogos online tendo por base ontologias pautadas no modelo conceitual FRBR?

Propõe-se, então, o desenvolvimento da pesquisa através de estudos e análises que busquem identificar como as tecnologias utilizadas para a Web Semântica podem contribuir com a Biblioteconomia e a Ciência da Informação, no que se refere à modelagem de OPACs, à luz dos conceitos pontuados nos FRBR, seu vocabulário e demais esforços relacionados ao modelo.

1.1 Objetivo geral

Esta pesquisa tem como objetivo geral identificar as contribuições que os conceitos, tecnologias, arquiteturas de metadados utilizados pela Web Semântica podem oferecer à área da Ciência da Informação, mais especificamente ao desenvolvimento, modelagem e arquitetura de metadados em catálogos online, tendo por base os conceitos definidos nos *Functional Requirements for Bibliographic Records* (FRBR), bem como estudos que possibilitam sua utilização como recurso ontológico. Visa, sobretudo, estudar meios de tornar as informações contidas em bases de dados e catálogos de bibliotecas (registros catalográficos e recursos informacionais digitais – *born-digital documents* – localizados na “Web profunda”) passíveis de recuperação na “Web superficial”.

1.2 Objetivos específicos

- Identificar os principais aspectos da Web Semântica, destacando suas características, padrões de



metadados, linguagens de ontologias e demais conceitos que a fundamentam;

- Analisar e identificar as arquiteturas de metadados semânticas mais indicadas a serem atribuídas na modelagem de OPACs, a fim de possibilitar a interoperabilidade das informações e recursos contidos no catálogo para a sua recuperação na Web;
- Analisar, por fim, a possibilidade de utilizar os FRBR como modelo ontológico, tendo por base as semelhanças do modelo conceitual com a estrutura de ontologias, focando no estudo do documento FRBR/CIDOC (*International Committee on Documentation*).

2 FUNDAMENTAÇÃO TEÓRICA

A enorme quantidade de informações disponível na Web faz com que hoje em dia exista uma ampla mobilização de recursos humanos e financeiros envolvidos em desenvolver maneiras de otimizar a recuperação de todas as informações dispostas nesta grande teia. No entanto, a maior parte das informações presentes na Web se encontram em camadas escondidas, nas quais os robôs de busca (Web buscadores, como Google e Yahoo) não podem ou não conseguem alcançar. Este não é um fato negativo, já que muitas das bases de dados contêm informações sigilosas que não devem ser acessadas sem restrições, como bases de empresas privadas, bancos ou órgãos governamentais, por exemplo. O foco desta pesquisa, porém, está nas informações contidas em bases de dados de bibliotecas, cujos conteúdos se encontram nas camadas da chamada “Web profunda”, as quais podem ser disseminadas e, sobretudo, encontradas. (BERGMAN, 2001 apud MARCONDES, 2006)

Assim como se deu o desenvolvimento e crescimento da Web, nas bibliotecas a quantidade de informações também cresceu substancialmente, porém num maior período de tempo, levando em conta o fato de que as bibliotecas existem desde os períodos mais remotos. O surgimento dos microcomputadores e da Internet no final dos anos 80 do século passado teve grande destaque neste crescimento, pois possibilitou organizar de maneira muito mais ágil os registros catalográficos antes elaborados manualmente e também automatizou os demais serviços da biblioteca. Com a evolução do paradigma orientado ao bibliotecário para um orientado ao usuário (LE COADIC, 2004), novas visões foram atribuídas ao processo de catalogação, bem como ao profissional catalogador.

Sob esse novo paradigma foram estipulados os FRBR: um documento elaborado por especialistas da IFLA cuja finalidade era definir diretrizes para a modelagem conceitual de bases de dados de catálogos de bibliotecas, pautadas nas tarefas do usuário definidas nos Princípios de Paris (1961), garantindo os requisitos fundamentais para a descrição de recursos informacionais presentes em bibliotecas e outras unidades de informação. Como modelo conceitual do tipo entidade-relacionamento, os FRBR são compostos por entidades, atributos e relacionamentos. (TILLET, 2004)

A proposta de atribuir significado às páginas Web para que sejam interpretadas por máquinas surgiu no ano de 2001, a partir da publicação de um artigo na revista americana *Scientific American*,



cujos títulos são: “Web Semântica: um novo formato de conteúdo para a Web com significado para computadores vai iniciar uma revolução de novas possibilidades”. Tal artigo foi publicado por Tim Berners-Lee (diretor do W3C e pesquisador do Instituto de Tecnologias de Massachusetts - MIT), James Hendler (professor da Universidade de Maryland) e Ora Lassila (pesquisador e membro do W3C), e é até hoje um texto referência sobre Web Semântica, pois nele os autores definem seus principais conceitos, estrutura e ilustram as situações que esse novo momento da Web pode proporcionar. (BERNERS-LEE; HENDLER; LASSILA, 2001)

A Web Semântica, segundo Alves (2005), tem como proposta estruturar e atribuir semântica aos dados representados com a finalidade de diminuir problemas de recuperação da informação na Web, apresentando, para tal, uma estrutura que permitirá a compreensão do conteúdo dos recursos informacionais, através da valorização semântica destes recursos e de agentes inteligentes capazes de processar informações e trocá-las com outros programas.

Os recursos propostos para a Web Semântica e os esforços intelectuais proporcionados pela Ciência da Informação por meio dos FRBR agindo conjuntamente poderão exercer transformações significativas no que tange à representação e recuperação de recursos na Web, possibilitando, assim, a visibilidade dos recursos localizados nos sistemas e bases de dados de bibliotecas. A atribuição de uma ontologia FRBR à arquitetura de metadados adequada, utilizando princípios da Web Semântica de Tim Berners-Lee, poderá proporcionar benefícios às unidades informacionais, aos usuários das informações e, sobretudo, às informações, que deixarão de estar, como afirma Grandmann (2005), “enterradas” na Web.

5 METODOLOGIA

A pesquisa se caracterizará por seu caráter teórico-exploratório, realizando-se por meio de análises descritivas sobre os assuntos que irá abordar, de modo a construir o conhecimento teórico sobre esses temas. Para tanto, a pesquisa será feita em livros e periódicos da área de Ciência da Informação e Tecnologias, anais de congressos, teses, dissertações, monografias, revistas eletrônicas, bases de dados textuais e referenciais, documentos eletrônicos disponíveis na Internet, bibliografias, catálogos, dentre outras fontes informacionais confiáveis que se fizerem necessárias ao desenvolvimento da etapa de busca.

DISCUSSÕES

Para que a Web Semântica seja possível é necessário que diversas ferramentas tecnológicas trabalhem de forma integrada em sua estrutura de implementação. Tais ferramentas podem ser resumidas em: metadados; linguagens de marcação; arquitetura de metadados; ontologias; e agentes inteligentes. (JORENTE; SANTOS; VIDOTTI, 2009)



Uma dos elementos essenciais para que a interoperabilidade semântica ocorra entre sistemas de informação é a elaboração de ontologias. Breitman (2005, p. 100) afirma que

[...] ontologias têm muito em comum com outras modelagens que utilizamos em nossa prática de software. Exemplos são modelos de Análise Estruturada, modelos de Entidade Relacionamento, de Análise Essencial e Orientação a Objeto. A construção desses modelos, bem como a de ontologias, envolve processos de descoberta, modelagem, validação e verificação da informação.

O W3C pontua que as ontologias devem prover descrições para os conceitos de “classes” (ou “coisas”), relacionamentos entre essas classes e propriedades (atributos). (BREITMAN, 2005, p. 31) Com base nas semelhanças existentes entre ontologias e modelagens de dados, como o aspecto conceitual que possuem e a utilização de relacionamentos, evidenciam-se os esforços da criação de uma ontologia baseada no modelo conceitual FRBR.

Fusco (2010) afirma que, dentre as muitas possibilidades de usos do modelo em questão, destaca-se a iniciativa dos estudos sobre FRBR_{oo}, realizado pelo CIDOC (*International Committee on Documentation*), a partir do documento já elaborado CIDOC CRM (*Conceptual Reference Model*), em conjunto com a IFLA, formando o *International Working Group on FRBR/CIDOC CRM Harmonisation*, responsável pela elaboração de uma ontologia formal dos FRBR para uso na alternativa proposta da modelagem orientada a objetos.

CONCLUSÃO

A internet se apresenta, no contexto atual, como o ambiente informacional que oferece o maior número de recursos informacionais e é, sem dúvidas, o que necessita de maiores cuidados com o tratamento descritivo de suas informações. Inseridos nesse ambiente estão os catálogos online de bibliotecas. A proposta de levantar e analisar as arquiteturas de metadados utilizadas pela Web semântica é uma forma de buscar informações sobre como modelar um catálogo online de biblioteca que possa interoperar suas informações com motores de busca, possibilitando tornar suas informações visíveis na Web.

Levando-se em conta que esta pesquisa está em sua fase inicial, espera-se que o estudo das arquiteturas de metadados identificados, juntamente à análise da ontologia FRBR proposta pela IFLA e pelo CIDOC, possam contribuir significativamente com os estudos da área, assim como para a implementação de catálogos online semânticos e mais úteis aos usuários que dele necessitam.



REFERÊNCIAS

ALVES, R. C. V. *Web Semântica: uma análise focada no uso de metadados*. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.

BERNERS-LEE, T.; HENDER, J.; LASSILA, O. *The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*. 2001. Disponível em: <<http://www.scientificamerican.com/article.cfm?id=the-semantic-Web>>. Acesso em: 14 set. 2011.

BREITMAN, K. *Web Semântica: a internet do futuro*. Rio de Janeiro: LTC, 2005. 190 p.
FUSCO, E. *Modelos conceituais de dados como parte do processo da catalogação: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais*. 2010. 249 f. Tese (Doutorado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2010.

GRANDMANN, S. rdfs:frbr – Towards an implementation model for library catalogs using semantic Web technology. In: LE BOEUF, P. *Functional Requirements for Bibliographic Records (FRBR): hype ou cure-all?* New York: The Haworth Information Press, 2005, p. 63-75.

IFLA STUDY GROUP on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records: final report*. 1998. As amended and corrected through February 2009. Disponível em: <http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf>. Acesso em: 14 set. 2011.

JORENTE, M. J. V.; SANTOS, P. L. V. A. C.; VIDOTTI, S. A. B. G. Quando as Webs se encontram: social e semântica – promessa de uma visão realizada? *Informação & informação*, Londrina, v.14, n. esp., p. 1-24, 2009.

LE COADIC, Y. F. *A Ciência da informação*. Tradução de Maria Yêda F. S de Figueiras Gomes. 2 ed. Brasília, DF: Briquet de Lemos/Livros, 2004, 124p.

MARCONDES, C. H. Metadados: descrição e recuperação de informações na Web. In: _____. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. 2. ed. Salvador: EDUFBA; Brasília: IBICT, 2006. p. 95-111.

TILLET, B. B. *What is FRBR?: A Conceptual Model for the Bibliographic Universe*. Library of Congress, Cataloging Distribution Service, 2004. Disponível em: <www.loc.gov/cds/downloads/FRBR.PDF>. Acesso em: 14 set. 2011.



MINERAÇÃO DE DADOS PARA APOIO À TOMADA DE DECISÃO DE DOCENTES NA AVALIAÇÃO DE ALUNOS DE GRADUAÇÃO A DISTÂNCIA

José Leonardo Oliveira Lima

Resumo: A presente comunicação é uma síntese do projeto de Doutorado do curso de Ciência da Informação e relato da fase inicial em que se encontra a pesquisa. Objetiva propor um processo de mineração de dados e desenvolver um módulo (protótipo) de apoio à tomada de decisão para avaliação, a ser integrado a um Ambiente Virtual de Aprendizagem (AVA), satisfazendo as necessidades de informação dos docentes na avaliação de alunos em cursos de graduação a distância. Os procedimentos metodológicos envolvem pesquisa descritiva, metodológica e aplicada com pesquisa bibliográfica e de campo. A pesquisa se encontra na fase inicial de revisão bibliográfica, construção do referencial teórico e comunicação com os pares e especialistas nos assuntos objetos de estudo, mas já com elementos relevantes sobre a temática.

Palavras-chave: Necessidades e uso da informação, mineração de dados, avaliação, educação a distância.

1 INTRODUÇÃO

A humanidade tem vivenciado um período histórico de mudanças e incertezas. Com o volume de informações e recursos disponíveis, o ser humano e as organizações apresentam sinais de dificuldade, confusão, frustração e impotência quanto ao uso do volume de dados e recursos tecnológico disponíveis, evidenciando a distância ainda existente entre a disponibilização de novas tecnologias ou recursos e a capacidade para o uso que possibilitem o acesso e a recuperação da informação com eficiência e eficácia.

A universidade, como o lócus do desenvolvimento e disseminação do conhecimento científico, é um espaço privilegiado para problematizar e também absorve os benefícios e impactos das referidas mudanças.

No Brasil, na última década, ocorreu um expressivo crescimento da educação superior na modalidade a distância, mesmo com um atraso de mais de 30 anos em relação a outros países (LITTO; FORMIGA, 2009; SANCHEZ, 2008). Na Educação a Distância (EaD) “em linha” - que utiliza os recursos e possibilidades das Tecnologias da Informação e Comunicação (TICs) - os Ambientes Virtuais de Aprendizagem (AVA) favorecem a interação entre alunos e professores e possibilitam a colaboração e discussão de idéias, disponibilização de tarefas, atividades e materiais de estudo (textos, fotos, vídeos, animações etc), avaliações, construção colaborativa por meio de fóruns, áudio e videoconferência etc (MOORE; KEARSLEY, 2007).



As TICs podem ser elementos potencializadores da atividade docente e importante apoio ao processo ensino-aprendizagem, mas também oferecem diversos desafios. Por exemplo: como avaliar efetivamente a participação do aluno, a originalidade dos materiais que o aluno posta, a qualidade e efetividade da sua participação em um AVA?

Do mesmo modo que existe a Internet para favorecer o acesso à informação, também existem os mecanismos usados pelos alunos para cometer pequenos delitos, como o “Control C + Control V”, os sites de disponibilização de trabalhos prontos etc. Esses são apenas alguns exemplos que permitem ilustrar a problemática que envolve a atividade docente de avaliação¹ dos alunos que utilizam os AVAs para participação, interação e resolução de atividades nos cursos a distância.

Muitos dos AVAs apresentam relatórios com informações sobre cada aluno, acesso aos recursos do sistema, *logs*² de uso do ambiente etc, porém essas são muito detalhadas, técnicas e confusas. Faltam informações sintéticas e significativas (inclusive com gráficos e elementos visuais) que facilitem ao professor detectar não conformidades e poder avaliar, de forma mais justa e consistente, a participação e o desenvolvimento do conhecimento do aluno em um curso por intermédio de um AVA.

Face ao exposto, temos como problema de pesquisa: quais são as necessidades de informação dos docentes dos cursos de graduação a distância no processo de avaliação dos alunos em um Ambiente Virtual de Aprendizagem (AVA) e como a mineração de dados (*data mining*) pode ser integrada para apoiar à tomada de decisão nesse processo de avaliação?

O objetivo geral do presente projeto é propor um processo de mineração de dados e desenvolver um módulo (protótipo) de apoio à tomada de decisão para avaliação, a ser integrado a um AVA, satisfazendo as necessidades de informação dos docentes na avaliação de alunos em cursos de graduação a distância.

2 REFERENCIAL TEÓRICO

Capurro e Hjørland (2007) colocam que a “sociedade da informação” passou a ser assim denominada em razão do surgimento da Tecnologia da Informação (TI) e dos seus impactos globais. Referidos autores também apresentam a dificuldade de se manter o termo “informação” formalmente definido, devido a sua utilização em diversos outros campos do conhecimento.

Segundo Capurro e Hjørland (*op. cit.*, 154-155): “quando usamos o termo informação em Ciência da Informação (CI), deveríamos ter em mente que informação é o que é informativo para uma determinada pessoa”. Para o presente projeto, definimos o termo “informação” considerando essa perspectiva.

1 Avaliação é um processo amplo e complexo que lida com diversos aspectos do processo ensino-aprendizagem. Para fins do presente projeto, consideramos os elementos disponíveis no AVA (dados, logs, postagens, participação etc.) que possam ser minerados, sintetizados e integrados e que sirvam de apoio ao processo de avaliação e tomada de decisão pelo professor para compor uma nota ou menção de avaliação, que poderá ser formativa ou somativa.

2 Logs – arquivo de registro automático de todas as operações realizadas por um usuário em um sistema ou ambiente.



A ênfase interdisciplinar é um aspecto que apresenta a importância da colaboração entre as diversas áreas de conhecimento e é preponderante na CI. Para Le Coadic (2004, p. 20), “a ciência da informação é uma das novas interdisciplinas [que colaboram entre si] a psicologia, linguística, sociologia, informática, matemática, lógica, estatística, eletrônica, economia, direito, filosofia, política e telecomunicações”. Saracevic (1995, p. 2) apresenta a abordagem complementar entre CI e as diversas ciências, dentre elas a Ciência da Computação (CC): “Ciência da Computação aborda os algoritmos relacionados à informação, enquanto Ciência da Informação abrange a natureza da informação e sua utilização pelos seres humanos. Os dois problemas não estão em competição, eles são complementares”.

Na CI, temos os estudos que envolvem a organização da informação e a complexidade dos mecanismos de armazenamento, disponibilização e recuperação da informação, em que precisa ser observado, desde os aspectos objetivos da informação-como-coisa (que a Tecnologia da Informação dá uma maior ênfase), passando pelos aspectos da informação-como-processo e informação-como-conhecimento, conforme as nomenclaturas propostas por Buckland (1991).

Capurro e Hjørland (2007, p. 193) fazem uma distinção importante entre informação como um objeto ou coisa (por exemplo: número de bits, dados armazenado em arquivos, banco de dados, dentre outros) e informação como conceito subjetivo, dependendo da interpretação de um sujeito cognitivo.

Na CI, desde meados da década de 80, os estudos de usuários enfatizam essa mudança de enfoque, saindo do “foco no sistema” para o “foco no uso”. Dervin e Nilan (1986) levantaram as seguintes questões: sistemas de informação poderiam servir melhor os usuários – aumentando sua utilidade (do sistema) para os clientes e ser mais responsável para com eles; para servir melhor a clientela, suas necessidades e uso de informação devem ser o objetivo central do sistema; um melhor atendimento à clientela pode exigir a implementação de um sistema com orientações remodeladas; os sistemas de informação não se utilizam da tecnologia para atender melhor à clientela.

Courtright (2006) coloca que a Tecnologia da Informação (TI) está envolvida em práticas de informação do dia-a-dia, mediando e modelando não só o tipo, o volume e a apresentação da informação disponível, como também as expectativas dos atores em relação ao tipo de informação que eles podem, devem e irão procurar.

O processo de mineração de dados (*data mining*), dentre outras possibilidades da CC, pode contribuir para oferecer informações significativas para o usuário. Fayyad et. al. (1996, p. 4) definem mineração de dados como “o processo de identificação de padrões válidos, novos, potencialmente úteis e compreensíveis, embutidos nos dados”. Consiste em um conjunto de ferramentas e técnicas que, por intermédio do uso de algoritmos, são capazes de explorar uma grande base de dados, auxiliando o homem a extrair as informações necessárias à obtenção do conhecimento (GOLDSCHMIDT; PASSOS, 2005).

A proposição de estudos de mineração de dados em um AVA envolve compreender a complexidade e a amplitude dos aspectos objetivos existentes nos dados armazenados no sistema, a



correlação entre eles para produzir novas informações que atendam as expectativas e necessidades do usuário de forma a propiciar o conhecimento (elemento subjetivo que ocorre no sujeito cognoscente).

Um Sistema de Apoio à Decisão (SAD) é definido por Stair (2005, p. 22) como um “conjunto organizado de pessoas, procedimentos, software, base de dados e dispositivos usados para dar apoio à tomada de decisões em problemas específicos”. Os elementos essenciais de um SAD incluem um conjunto de modelos usados para apoiar um tomador de decisão ou usuário (base de modelos), um conjunto de fatos e informações para apoiar à tomada de decisão (base de dados) e sistemas e procedimentos (interface com o usuário) que auxiliam os tomadores de decisão e outros usuários a interagir com o SAD.

Partindo das necessidades de informação dos docentes, das possibilidades da Mineração de Dados e da ênfase dos SADs é que propomos realizar uma síntese que denominamos “Sistema de Apoio à Decisão nos Processos de Avaliação”, a ser incorporado a um AVA, cujo processo e produto são os objetos do presente estudo.

3 METODOLOGIA

Utilizamos como referencial a taxionomia proposta por Vergara (2009), que apresenta o tipo da pesquisa quanto aos fins e quanto aos meio. Quanto aos fins a pesquisa é descritiva, metodológica e aplicada. Descritiva por que contempla a interpretação e descrição das percepções dos atores a luz do problema de pesquisa. Metodológica, pois envolve a proposição e desenvolvimento de um processo de mineração de dados em um AVA. Aplicada já que, ao final, será desenvolvido um módulo (protótipo de SAD) para ser integrado ao AVA. Quanto aos meios, a pesquisa é bibliográfica e de campo. Bibliográfica, para a construção do referencial teórico. De campo, pois envolve pesquisas com diversos atores do cursos de graduação a distância que utilizam o AVA.

Os sujeitos da pesquisa (atores) são os especialistas em avaliação, professores (ou tutores), coordenadores de curso e alunos de cursos de graduação a distância. Com os especialistas, professores, tutores e coordenadores de curso, levantaremos os aspectos referentes ao processo de avaliação de alunos e as necessidades de informação para a tomada de decisão sobre o processo de avaliação, bem como o comportamento no AVA. Com os alunos serão observados os mecanismos e o comportamento na utilização do AVA para levantamento dos costumes, boas e más práticas no uso do AVA, processo de interação etc.

A amostra tem caráter não probabilístico, tendo como critério de escolha a facilidade de acesso aos partícipes. São 3 cursos de graduação a distância previstos e, em cada um, contaremos com a participação de coordenadores, professores e/ou tutores (conforme as peculiaridades de cada projeto pedagógico e de cada curso) e alunos. O AVA escolhido para o presente estudo foi o Moodle pois, além de ser um software livre e de código aberto - podendo ser adaptado e personalizado conforme as necessidades de cada instituição, tem sido adotado por diversas Instituições de Educação Superior brasileiras.



Os instrumentos utilizados para a coleta dos dados são entrevistas (abertas e por pauta, dependendo da etapa e do ator), questionários (com questões objetivas e algumas abertas) e observação. Também estão sendo avaliados, para utilização, sistemas de software e ferramentas de mineração de dados (como por exemplo o WEKA), banco de dados MYSQL, linguagem de programação PHP (usada no AVA Moodle), geradores de gráficos e programa estatístico Sphinx Lexica.

Os procedimentos para a coleta de dados são os específicos para cada instrumento. O tratamento de dados está sendo feito de forma quantitativa e qualitativa. Na abordagem qualitativa, estamos utilizando os processos de análise e síntese tanto no estudo bibliográfico, em andamento, como também serão utilizados na análise dos discursos colhidos nas entrevistas e questionários com os diversos atores. Na abordagem quantitativa (ainda não iniciada), faremos o uso das ferramentas e técnicas estatísticas (frequência média, moda, desvio padrão), principalmente na redução dos questionários, bem como nas análises dos *logs* e dados referentes ao comportamento de uso de professores e alunos no AVA, e na montagem do processo de mineração de dados e módulo (protótipo) de suporte à decisão sobre avaliação.

4 CONSIDERAÇÕES FINAIS

Mesmo no primeiro semestre de execução, já percebemos, pelo levantamento bibliográfico realizado e estudos da fundamentação teórica em CI, a relevância do presente estudo para além do tema delimitado, podendo ser mais um elemento contributivo no processo de desenvolvimento do conhecimento tanto nos aspectos teóricos e interdisciplinares da CI, em conjunto com a CC e Ciência da Educação, bem como nos aspectos práticos de pesquisa aplicada, contribuindo para a superação do paradoxo que foi apresentado no relatório da Conferência Mundial sobre Educação Superior realizado pela UNESCO há mais 12 anos em Paris, mas ainda existente:

A comunidade científica é a origem de grande parte das soluções técnicas e metodológicas necessárias para o desenvolvimento das Novas tecnologias da Informação e comunicação - NTIC. Contudo, os setores da educação são os que menos aproveitam das contribuições das novas tecnologias. (CONFÉRENCE MONDIALE SUR L'ENSEIGNEMENT SUPÉRIEUR, 1998, p. 7, tradução nossa).

Abstract: This paper is a summary of a PhD project in Information Science and a reporting of the initial researches. It contemplates the use of data mining in Virtual Learning Environment (VLE) to achieve information needs of teachers when evaluating students of undergraduate distance courses in the VLE. It proposes to develop a process and a module (prototype software built into the VLE) to perform data mining on databases in order to compose a Decision Support System (DSS) covering the relevant aspects of evaluation in a distance learning course by teachers. The methodological procedures involves descriptive, methodological and applied research, literature review and field procedures. The research is in process of literature review and theoretical construction, communication with peers and experts, but already with relevant datas.

Keywords: Information needs and use, data mining, evaluation, distance learning.



REFERÊNCIAS

BUCKLAND, M. K. Information as thing. **Journal of the American Society of Information Science**, v. 42, n.5, p. 351-360, 1991.

CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em ciência da informação**, v. 12, n. 1, p. 148-207, 2007.

CONFÉRENCE MONDIALE SUR L'ENSEIGNEMENT SUPÉRIEUR. Debats thématiques: Du traditionnel au virtuel - les nouvelles technologies de l'information. In: **L'enseignement supérieur au XXIe siècle – vision e actions**. v. 4. Paris, França: UNESCO, 1998.

COURTRIGHT, Christina. The challenge of context in information behavior research. **Annual Review of Information Science and Technology (ARIST)**, v. 41, 2006.

DERVIN, Brenda; NILAN, Michael. Information needs and uses. **Annual Review of Information Science and Technology (ARIST)**, v. 21, p. 3-33, 1986.

FAYYAD, Usama M. **Information visualization in data mining and knowledge discovery**. 1996. 346 p. Disponível em:

<<http://books.google.com.br/books?id=rYFvnyPRwkgC&printsec=frontcover&dq=fayyad&ei=joDsSZrRMY-gkQTI4Y2aAQ&client=firefox-a#PPA3,M1>>. Acesso em 14/04/2010.

GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel. **Data mining: um guia prático**. Rio de Janeiro: Elsevier, 2005.

LE COADIC, Y. F. **A ciência da informação**. 2. ed. Brasília: Briquet de Lemos, 2004.

LITTO, F. M.; FORMIGA, M. M. M. (orgs.). **Educação a distância: o estado da arte**. São Paulo: Pearson do Brasil, 2009.

MOORE, M; KEARSLEY, G. **Educação a distância: uma visão integrada**. São Paulo: Thomson, 2007. p. 77-106.

SANCHEZ, F. (coord.). **Anuário brasileiro estatístico de educação aberta e a distância**. 4. ed. São Paulo: Instituto Monitor, 2008.

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da informação**, v. 24, n. 1, 1995.

STAIR, Ralph M. **Princípios de sistemas de informação**. 6. ed. São Paulo: Thomson, 2005.

VERGARA, Sylvia Constant. **Projetos e relatórios de pesquisa em administração**. 11. ed. São Paulo: Atlas, 2009.