

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PUC-SP

RENATE LANDSHOFF

Findability:

elementos essenciais para as formas de encontro da informação em Bibliotecas Digitais

MESTRADO EM TECNOLOGIAS DA INTELIGÊNCIA E DESIGN DIGITAL

São Paulo

2011

RENATE LANDSHOFF

Findability:

elementos essenciais para as formas de encontro da informação em Bibliotecas Digitais

MESTRADO EM TECNOLOGIAS DA INTELIGÊNCIA E DESIGN DIGITAL

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de Mestre em Tecnologias da Inteligência e Design Digital sob a orientação do Prof. Doutor Alexandre Campos Silva.

São Paulo

2011

ERRATA

LANDSHOFF, Renate. **Findability**: elementos essenciais para as formas de encontro da informação em Bibliotecas Digitais. 2011. 127 f. Dissertação (Mestrado em Tecnologias da Inteligência e Design Digital)–Pontifícia Universidade Católica de São Paulo, São Paulo, 2011.

Folha	Linha	Acrescentar
12	3	<p>Este estudo explora os fatores críticos no tratamento e recuperação da informação que dificultam os usuários de Bibliotecas Digitais na Web de encontrarem a informação desejada e tem por objetivo identificar elementos essenciais em projetos de construção de Bibliotecas Digitais do ponto de vista da entrada de dados e da interface de busca para promover a encontrabilidade da informação.</p> <p>A interface de busca é o principal instrumento de diálogo entre o gestor de uma biblioteca digital e seus usuários na Web. Existe um esforço enorme no tratamento da informação, ou seja, na entrada de dados e representação descritiva de seus recursos informacionais, mas são poucas as bibliotecas digitais que desempenham esforço semelhante no desenvolvimento de interfaces inteligentes, capazes de interagir de forma intuitiva com seus usuários virtuais.</p>
128	1	<p>GLOSSÁRIO</p> <p>CIBERESPAÇO: Termo foi criado em 1984 pelo escritor norte-americano Wiliam Gibson no seu livro de ficção científica Neuromancer, é o espaço não físico, constituído pelas redes digitais.</p> <p>ENIAC: Eletronic Numerical Integrator and Calculator. (BREVE, 2011).</p> <p>FEED: É um formato de dados usado em formas de comunicação com conteúdo atualizado frequentemente, como sites (sítios) de notícias ou blogs. (WIKIPÉDIA, 2011b).</p> <p>HTML: Abreviatura de HyperText Markup Language – linguagem composta por um conjunto de comandos de formatação e utilizada na criação de documentos.</p> <p>HTTP: Tradução de Hypertext Transfer Protocol, faz a comunicação entre o cliente e o servidor por meio de mensagem.</p> <p>LÓGICA BOOLEANA: A lógica booleana se refere ao sistema de lógica matemática denominada Álgebra booleana, cujo nome teve origem no matemático inglês George Boole. É utilizada para criar regras lógicas ou expressões. (INTRODUÇÃO, 2011).</p>

	<p>PLUG-IN: É um programa de computador usado para adicionar funções a outros programas maiores, provendo alguma funcionalidade especial ou muito específica. Geralmente pequeno e leve, é usado somente sob demanda. (WIKIPÉDIA, 2011a).</p> <p>PORTLETS: É um componente visual independente que pode ser utilizado para disponibilizar informações dentro de uma página.</p> <p>PROSUMIDOR: Termo que resulta da junção das palavras, produtor e consumidor, que salienta o crescente papel dos consumidores no processo produtivo. (CARDOSO, 2009). Web Semântica.</p> <p>XANADU: Theodor Holm Nelson, idealizador do hipertexto, propôs a criação de uma biblioteca eletrônica (World Publishing Repository _ Project Xanadu), onde seriam armazenados os textos integrais de documentos. Xanadu seria um "lugar mágico da memória literária onde nada seria esquecido". (BORGES, 2003).</p> <p>WWW: Sigla de World Wide Web, a Web é um Sistema hipertextual que funciona sobre a Internet.</p>
--	--

LANDSHOFF, Renate.

Findability: elementos essenciais para as formas de encontro da informação em Bibliotecas Digitais / Renate Landshoff, – 2011. 127 f.

Orientador: Prof. Doutor Alexandre Campos Silva
Dissertação (Mestrado em Tecnologias da Inteligência e Design Digital) – Pontifícia Universidade Católica de São Paulo, 2011.

1. Biblioteca digital. 2. Interface de busca. 3. Findability. 4. Metadados. 5. Vocabulário controlado. 6. Busca. 7. Tratamento da informação. I. Título.

Banca Examinadora

Dedico este estudo a todos que foram vitais para a conquista deste objetivo, e envolveram-se de forma decisiva para que eu conseguisse a determinação necessária para alcançá-lo.

Ao meu Orientador, **Alexandre Campos Silva**, pelo envolvimento, estímulo, palavras certas nas horas certas, objetividade e sensibilidade com que assumiu e conduziu esse processo em um momento tão delicado.

Ao querido **Sérgio Storch**, que foi, é e sempre será uma pessoa muito especial para mim. Sua paixão pelo conhecimento e seu entusiasmo pelo meu crescimento, instigaram-me a perseguir esse desafio e essa conquista.

À querida **Suely de Brito Clemente Soares**, minha irmã de alma e minha sócia, pelo apoio, carinho e compreensão incondicionais, vitais para que eu mantivesse o foco em momentos tão críticos.

À **Valéria Valls**, Coordenadora do curso de Biblioteconomia e Ciência da Informação da FESPSP, pelas oportunidades proporcionadas para que eu pudesse encontrar o equilíbrio necessário para manter-me determinada.

À minha mãe de alma, **Irene Sad**, pelo amor, acolhimento e sabedoria. Suas palavras e conselhos tornaram-se responsáveis por este feito.

À dedicada e competente **Edna Conti**, responsável pela secretaria do TIDD/PUC-SP, o meu profundo reconhecimento por sua importância e relevância em fazer com que tudo acontecesse no seu tempo. Suas valiosas e detalhadas informações sempre nortearam o meu sentido de urgência.

Aos meus queridos filhos, **Daniel e Débora**, que são fontes de eterna inspiração e energia e à minha família.

Com todo o meu amor e gratidão,

Renate Landshoff

RESUMO

Este estudo tem por objetivo recomendar elementos essenciais que impactam as formas de encontro da informação em ambientes de bibliotecas digitais, incluindo as funcionalidades encontradas na Web 2.0. A fundamentação teórica está baseada no conceito de *Findability*, cunhado por Morville, na contribuição expressiva de Manovich, no estudo de Santaella sobre os perfis cognitivos de leitores e internautas e na visão de Weinberg e Cunha sobre a missão de bibliotecas digitais. O desenvolvimento deste trabalho está baseado no método de estudos de casos qualitativos, ressaltando aspectos positivos e negativos de interfaces de busca a partir de elementos considerados relevantes para projetos de bibliotecas digitais e identificados com a contribuição de conceitos trazidos da área de Biblioteconomia e Ciência da Informação. O estudo destaca a contribuição das bibliotecas digitais para a pesquisa na *Web*, sem no entanto deixar de apontar para as suas fragilidades e necessidade de evolução. O tratamento da informação no momento da entrada de dados, o desenvolvimento de interfaces de busca mais inteligentes, os diversos perfis cognitivos de usuários, a participação de equipes multidisciplinares e o monitoramento constante do ambiente digital devem ser elementos considerados e perseguidos ao longo da existência de um projeto deste perfil. O estudo destaca por fim, as mudanças que devem ocorrer com a busca a partir da Web Semântica e o desenvolvimento da inteligência artificial, impactando na forma como estamos acostumados a buscar a informação, aumentando cada vez mais a sua “encontrabilidade”.

Palavras-chave: Biblioteca digital. Interface de busca. Findability. Metadados. Vocabulário controlado. Busca. Tratamento da informação.

ABSTRACT

This study aims to recommend key elements which impact the ways to find information in digital library environments, including features found in Web 2.0. The theory is based on the concept of Findability, coined by Morville, on the significant contribution of Manovich, on Santaella's study about cognitive profiles of readers and Internet users, and on Weinberg and Cunha's view on the mission of digital libraries. The development of this work is based on the method of qualitative case studies, highlighting positive and negative aspects of search interfaces from elements considered relevant to digital library projects and identified with the contribution of concepts brought from Library and Information Science. The study emphasizes the contribution of digital libraries for Web search, without however fail to point to their weaknesses and need for evolution. The processing of information at the time of data entry, the development of more intelligent search interfaces, the different cognitive profiles of users, the participation of multidisciplinary teams and constant monitoring of the digital environment elements must be considered and pursued over the existence of a design of this profile. The study points out finally the changes that must occur in searching from the Semantic Web and the development of artificial intelligence, impacting the way we are used to seek information, continually increasing its "findability."

Keywords: Digital library. Search interface. Findability. Metadata. Controlled vocabulary. Search. Information processing.

LISTA DE ILUSTRAÇÕES

Figura 1 – Busca avançada no Google	18
Figura 2 – Resultado de pesquisa no Google	19
Figura 3 – Pesquisa no Portal Senado Federal - Prodasen	20
Figura 4 – Resultado de busca no Portal Senado Federal - Prodasen	21
Figura 5 – Biblioteca Digital Unicamp, resultado de busca	24
Figura 6 – Interface de busca da BDTD	28
Figura 7 – Resultado de busca da BDTD	29
Figura 8 – Busca avançada no Google Acadêmico.....	30
Figura 9 – Resultado de busca no Google Acadêmico	30
Figura 10 – Busca na BDTD do IBICT	33
Figura 11 – Resultado na BDTD do IBICT	33
Figura 12 – Busca na BDTD pela instituição de ensino	34
Figura 13 – Resultado de busca na BDTD pelo nome completo da instituição de ensino.....	34
Figura 14 – Mapa do site da Comissão de Valores Mobiliários.....	37
Figura 15 – Apresentação do Tesouro da BVS	39
Figura 16 – Delicious.....	41
Figura 17 – Tela principal de busca da Brasileira USP	45
Figura 18 – Possibilidade de ordenação de resultados na SciELO.....	50
Figura 19 – A máquina Memex, de Vannevar Bush	51
Figura 20 – Página principal do Google Books	54
Figura 21 – Modelo tradicional de atendimento ao usuário.....	56
Figura 22 – Modelo de atendimento na <i>web</i>	56
Figura 23 – Figura representando o conceito de Biblioteca 2.0	57
Figura 24 – Google Books.....	59

Figura 25 – Página principal do blog da FFLCH.....	61
Figura 26 – Tela principal da Scielo	75
Figura 27 – Resultado de busca da SciELO pelo índice “Assunto” na letra M	76
Figura 28 – Pesquisa na SciELO de artigo no índice “Autor”	77
Figura 29 – Exemplo na SciELO de redundância no índice autor	79
Figura 30 – Busca avançada na SciELO a partir de formulário livre	80
Figura 31 – Orientação na SciELO para formulação de busca	81
Figura 32 – Resultado na SciELO de pesquisa com possibilidade de refinamento da busca.....	82
Figura 33 – Formatos de citação disponíveis.....	83
Figura 34 – Acesso às redes sociais a partir de login no ambiente da SciELO	84
Figura 35 – Página principal da BDTD Sapiientia – PUC-SP.....	85
Figura 36 – Resultado de busca da BDTD Sapiientia	86
Figura 37 – Página Secundária da BDTD Sapiientia	87
Figura 38 – Busca simples da BDTD Sapiientia.....	88
Figura 39 – Página de busca por Programa de Ensino da BDTD Sapiientia	89
Figura 40 – Página de resultado de pesquisa da BDTD Sapiientia.....	90
Figura 41 – Página de resultado da pesquisa da BDTD Sapiientia.....	91
Figura 42 – Resultado na SciELO de pesquisa por índice de autor	94
Figura 43 – Resultado na SciELO de pesquisa por índice de autor	94
Figura 44 – Refinamento da busca na SciELO	95
Figura 45 – Página principal do site	98
Figura 46 – Página principal: Anne Arbor Digital Library.....	98
Figura 47 – Taxonomia do site Flora Brasiliensis.....	100
Figura 48 – Interface de busca da Biblioteca Digital da Unicamp	101
Figura 49 – Visão de conjunto e busca por navegação.....	102
Figura 50 – Portal SEBRAE	105

Figura 51 – Busca avançada da SciELO.....	106
Figura 52 – Refinamento de busca SciELO	106
Figura 53 – Página principal da biblioteca.....	107
Figura 54 – Chat.....	108
Figura 55 – Mecanismo de ajuda detalhado.....	109
Figura 56 – Ambiente inicial da interface de busca do Google	110
Figura 57 – Resultado de busca no Google pelo Termo: CMS	112
Figura 58 – Ordenação de resultado de busca da Biblioteca da UFSC	113

SUMÁRIO

1 INTRODUÇÃO	12
2 FINDABILITY EM BIBLIOTECAS DIGITAIS	15
2.1 A diferença entre buscar e encontrar informação	17
2.2 A importância dos metadados no tratamento e recuperação da informação.....	22
<i>2.2.1 FRBR - Requisitos Funcionais para os Registros Bibliográficos.....</i>	<i>25</i>
<i>2.2.2 Dublin Core Metadata Element Set (DCMES)</i>	<i>26</i>
<i>2.2.3 MTD-Br</i>	<i>27</i>
2.3 Vocabulários controlados	31
<i>2.3.1 Taxonomia</i>	<i>36</i>
<i>2.3.2 Tesouro.....</i>	<i>38</i>
<i>2.3.3 Folksonomia</i>	<i>39</i>
2.4 A interface de busca como fator crítico para o encontro da informação .	42
2.5 A representação dos resultados de busca	46
3 BIBLIOTECAS DIGITAIS.....	51
3.1 A evolução das bibliotecas digitais para a Web 2.0.....	55
3.2 As bibliotecas nas redes sociais	60
4 OS PERFIS COGNITIVOS DO USUÁRIO NA WEB	62
4.1 A Linguagem hipermidiática	63
4.2 Os usuários no papel de internautas	67
4.3 Os usuários de bibliotecas digitais e a mudança de paradigma	71
5 ESTUDOS DE CASOS DE INTERFACES DE BUSCA EM BIBLIOTECAS DIGITAIS.....	73
5.1 SciELO - Scientific Electronic Library Online - www.scielo.br.....	74
<i>5.1.1 A interface de busca e seus elementos</i>	<i>75</i>

5.1.2 Considerações sobre a Biblioteca on-line da SciELO.....	84
5.2 Sapientia – Biblioteca Digital de Teses e Dissertações – PUC-SP	85
5.2.1 A interface de busca e seus elementos	85
5.2.2 Considerações sobre a BDTD Sapientia da PUC-SP.....	91
5.3 Comentário sobre os estudos de casos	92
6 ELEMENTOS ESSENCIAIS PARA O ENCONTRO DA INFORMAÇÃO EM BIBLIOTECAS DIGITAIS: RECOMENDAÇÕES.....	93
6.1 Metadados	93
6.2 Comunicação.....	96
6.3 Visão de conjunto	97
6.4 Busca simples	101
6.5 Busca avançada	103
6.6 Mecanismos de ajuda	107
6.7 Representação de resultados de busca.....	111
7 CONSIDERAÇÕES FINAIS	115
REFERÊNCIAS.....	118
REFERÊNCIAS CONSULTADAS	126

1 INTRODUÇÃO

“A informação será usada na proporção direta do quanto ela for fácil de ser encontrada.” (MORVILLE, 2005, tradução nossa).

A questão central desse estudo é um tema que me instiga desde o início de minha carreira, em 1983. Na época, ainda estávamos distantes da Internet, mas a dificuldade em encontrar uma informação já fazia parte de meu dia a dia.

Mesmo me apropriando de códigos e regras universalmente adotadas, a dificuldade persistia; hoje, em tempos de Web 2.0, as bibliotecas digitais buscam sanar essa dificuldade criando interfaces de busca aparentemente suficientes para que o usuário encontre a informação.

Como bibliotecária, vivenciei diversos paradigmas rompendo-se no universo das bibliotecas até a chegada das bibliotecas digitais e da Web 2.0.

A partir do início da automação das bibliotecas no Brasil, no começo da década de 80, as bibliotecas começaram a conviver com o desafio de transportar seus catálogos para ambientes digitais.

No início foram disponibilizados apenas os catálogos, permitindo o acesso a partir de redes de computadores, fisicamente conectados.

Com a chegada da Internet e em seguida da *Web*, os conteúdos disponíveis começaram a ser ampliados e “linkados”, potencializando o acesso à informação. Assim foram surgindo as bibliotecas digitais, que buscam se reinventar a cada tecnologia que surge.

Embora o acesso à informação tenha explodido com o surgimento da *Web*, não é de hoje que acompanhamos esse fenômeno. A quantidade de conhecimento acumulado vem aumentando exponencialmente há mais de meio século. O uso científico da Internet já é bastante difundido. Um exemplo citado em Renato Sabbatini (1997), aponta para a existência de apenas 10 revistas científicas de circulação mundial no início do século XX, saltando para 100 mil revistas em 1990 com uma projeção de mais de 200 mil para o início do século XXI.

O ato de encontrar uma informação na *Web* torna-se um desafio que aumenta a cada dia. Estima-se que estejamos em torno de mais de 1 bilhão de páginas em HTML.

A partir da Web 2.0, as bibliotecas começam a sentir mais ainda a pressão para mudanças, pois as funcionalidades e vantagens da colaboração, percebidas nas redes sociais, começam a ser possíveis também em soluções para bibliotecas digitais. Não há mais como manter-se hermética, com vocabulário próprio controlado e sem uma participação efetiva e colaborativa do usuário, agora invisível.

Com o surgimento do Google, as bibliotecas digitais passaram a ser questionadas, cobradas e até abandonadas. Será mais fácil encontrar informação no Google? Será que as interfaces de busca de bibliotecas digitais conseguirão alcançar e manter-se no estado da arte? O que está faltando para que as bibliotecas consigam atender as necessidades de seus usuários?

A partir desses e outros questionamentos ao longo do texto, o trabalho foca nos elementos essenciais necessários para as formas de encontrar a informação em bibliotecas digitais. O que pode ser feito para transformar a experiência de buscar em encontrar de fato a informação desejada?

Através de estudo de caso qualitativo de interfaces de busca em bibliotecas digitais, é possível analisar a intenção e o impacto das interfaces de busca na encontrabilidade da informação, conceito cunhado por Peter Morville (2005), denominado *findability*. Esse conceito apresentado no capítulo 2 busca elencar todos os fatores envolvidos no tratamento e recuperação da informação em uma biblioteca digital. Aspectos relacionados à entrada de dados e à interface de busca, somados, representam o conceito de *findability*.

Como o estudo está focado em ambientes de bibliotecas digitais, o capítulo 3 é dedicado à evolução das bibliotecas e suas diversas reinvenções desde o seu surgimento. Os vários paradigmas que foram rompendo-se, os usuários que foram migrando para a *Web* e as soluções interativas e colaborativas que surgiram com o Google e as redes sociais, foram provocando mudanças nas bibliotecas digitais. Não basta apenas estar na *Web*. Há que ser colaborativa, ter uma interface de busca intuitiva e ser eficaz em seu objetivo. Os usuários tornaram-se mais exigentes e seus perfis cognitivos estão tornando-se cada vez mais complexos, conforme apontado no capítulo 4.

Observa-se, assim, que as bibliotecas digitais estão ficando distantes de soluções inteligentes. Então, para recomendar elementos essenciais que possam contribuir com as formas de encontrar a informação e tornar suas interfaces de busca mais eficazes, foi desenvolvido no capítulo 5 um estudo de caso exploratório

com 2 bibliotecas digitais; a partir desse estudo, chegou-se à recomendação final de aspectos a serem considerados em projetos de bibliotecas digitais na *Web*, com o foco de encontrar a informação e, assim, alcançar o seu objetivo. Segundo Peter Morville (2005, tradução nossa), precisamos desenvolver um estilo para buscar informação (*wayfinding*), que envolve: “saber onde se está, saber o próprio destino, buscar o melhor caminho para chegar ao este destino, ser capaz de reconhecer o destino e ser capaz de retornar ao ponto de partida”. E a responsabilidade é de quem projeta as soluções, de quem quer ser encontrado.

O estudo termina supondo que a tecnologia facilitará muito esse processo, com a chegada da Web Semântica e com o desenvolvimento da inteligência artificial, mas a entrada de dados e a representação descritiva de um recurso informacional sempre serão feitas pelo seres humanos, e estes precisam ter em mente a complexidade de sua responsabilidade.

2 FINDABILITY EM BIBLIOTECAS DIGITAIS

O termo *findability* foi cunhado por Peter Morville, autor do livro *Ambient Findability* e está relacionado com a capacidade de encontrar informação em ambientes digitais.

Não há uma tradução oficial para o termo; portanto, neste estudo, será utilizada a expressão “encontrabilidade”, “formas de encontrar a informação” e “encontro da informação” para traduzir *findability*.

O estudo analisa a dificuldade na encontrabilidade da informação em bibliotecas digitais a partir de suas interfaces de busca e traz uma contribuição importante do ponto de vista da Ciência da Informação para valorizar alguns elementos essenciais ao seu desenvolvimento.

Com a explosão da informação a partir do período pós-guerra, nasce oficialmente, em 1962, em uma reunião do Georgia Institute of Technology, a Ciência da Informação, definida como

[...] a ciência que investiga as propriedades e o comportamento da informação, as forças que governam o fluxo da informação e os meios de processamento da informação para a acessibilidade e usabilidade ótimas (SHERA, 1977 apud BRAGA, 1995).

Segundo Saracevic (1996) e Oliveira (2005), sua gênese dá-se no cenário de revolução científica e técnica, que ocorre logo após a Segunda Guerra Mundial, tendo sofrido inúmeras influências da Documentação e da Recuperação da Informação.

Pode-se dizer que a Ciência da Informação cresceu e ganhou forma como área de estudos teóricos e aplicados, pela necessidade humana de lidar e superar problemas informacionais – principalmente pela adoção de métodos e tecnologias – motivados principalmente pelo aumento crescente da quantidade de informação na sociedade, seja na área científica ou fora dela (SARACEVIC, 1996).

Sendo assim, rompe-se o paradigma da organização da informação, e inicia-se uma grande preocupação em localizar este universo informacional.

O nascimento da Ciência da Informação coincide com o surgimento dos computadores, que foram sendo rapidamente substituídos por novas gerações de máquinas, cada vez mais potentes.

Com o surgimento da automação de bibliotecas na década de 80, surge um movimento de migração de acervos físicos para acervos digitais, rompendo-se um paradigma em relação ao usuário, que modifica seus hábitos, começando a frequentar cada vez menos o ambiente físico da biblioteca, favorecido pelos catálogos *on-line* e pelo surgimento das bibliotecas digitais com o advento da Internet e da *Web*. Assim, o usuário foi aos poucos tornando-se invisível aos olhos do profissional bibliotecário, com hábitos e necessidades desconhecidos, mas com carência cada vez maior de orientação por falta de competências informacionais no meio digital.

Pierre Lévy ressalta que,

Assim que penetramos no universo da *Web*, descobrimos que ele constitui não apenas um imenso “território” em expansão acelerada, mas que também oferece inúmeros “mapas”, filtros, seleções para ajudar o navegante a orientar-se. (LÉVY, 2003, p. 85).

Os elementos essenciais para a formulação da busca, disponíveis em um ambiente digital a partir de sua interface, passam a ser um fator crítico de sucesso na recuperação da informação, e, além da interface de busca, os metadados e os vocabulários controlados ganham importância nessa nova configuração, pois o profissional bibliotecário já não está disponível fisicamente para orientá-lo.

No contexto do ciberespaço¹, onde um turbilhão de dados disputa entre si a chance de ser encontrado, as bibliotecas digitais são uma espécie de oásis no deserto, pois procuram reunir um conjunto de informações sobre um determinado tema com uma interface de busca mais personalizada, facilitando, assim, teoricamente, o encontro da informação a partir de instrumentos e códigos padronizados universalmente.

¹ Termo foi criado em 1984 pelo escritor norte-americano William Gibson no seu livro de ficção científica *Neuromancer*, é o espaço não físico, constituído pelas redes digitais.

Como ir da teoria à prática, propiciando um resultado de busca adequado às necessidades do usuário, é o que pretendo abordar neste estudo, identificando os elementos essenciais à realização desse objetivo.

2.1 A diferença entre buscar e encontrar informação

Encontrar a informação é uma experiência muito diferente de buscar informação. Segundo Peter Morville (2005), a experiência de encontrar nos transforma. E de fato, quando encontramos uma informação útil e a internalizamos, passamos a modificar nosso conhecimento e, portanto, nos transformamos. Passamos a agir com alguns elementos a mais de informação, e isso nos faz diferentes.

O ato de buscar não presume o encontro da informação. A quantidade de informação resultante de uma busca, muitas vezes, leva o usuário a desistir e a se confundir diante de tantos resultados. Já o ato de encontrar é um sentimento perceptível, pois gera ganho de tempo, de produtividade, além de provocar um sentimento de satisfação e reconhecimento.

Morville (2005) diz que o conceito de “findability” está intrinsecamente ligado à forma como buscamos informação. O encontro da informação é o que realmente importa quando fazemos uma busca na *Web*. Buscar e confiar são passos anteriores a encontrar. A busca é o meio e não o fim. Só é possível encontrar, se confiamos na informação. O fim é encontrar alguém ou algo que esteja em qualquer lugar a qualquer momento. Muitas vezes, achamos que encontramos, mas, se não confiamos na fonte de informação, podemos estar cometendo um equívoco e essa sensação é perceptível a partir de um sentimento de insegurança em relação à fonte de informação consultada. Em um ambiente digital, a diferença entre a busca e o encontro da informação está nos elementos que compõem a sua estrutura: seus metadados, sua linguagem, sua interface de busca e a tecnologia que disponibiliza essa busca. Quando estamos diante de uma interface de busca, tentando traduzir nossa busca utilizando os parâmetros oferecidos pela interface. Na prática estamos buscando informação, pesquisando, procurando interpretar as orientações de um sistema para inserir os termos adequados, delimitando melhor o campo de busca,

escolhendo o formato do documento; enfim, estamos nos adaptando a um ambiente digital para nos fazer entender e, de fato, encontrar o que buscamos. Portanto, a busca ocorre antes do encontro da informação. As duas experiências são distintas. Buscar não significa encontrar, buscar significa tentar encontrar, e isso é o que mais ocorre, de fato, em ambientes de busca na *Web*. Os usuários, em geral, acumulam muito mais experiências em buscar que, realmente, em encontrar.

Para exemplificar a diferença entre buscar e encontrar informação, usarei como exemplo a busca avançada do Google o *site* de legislação do Senado Federal.

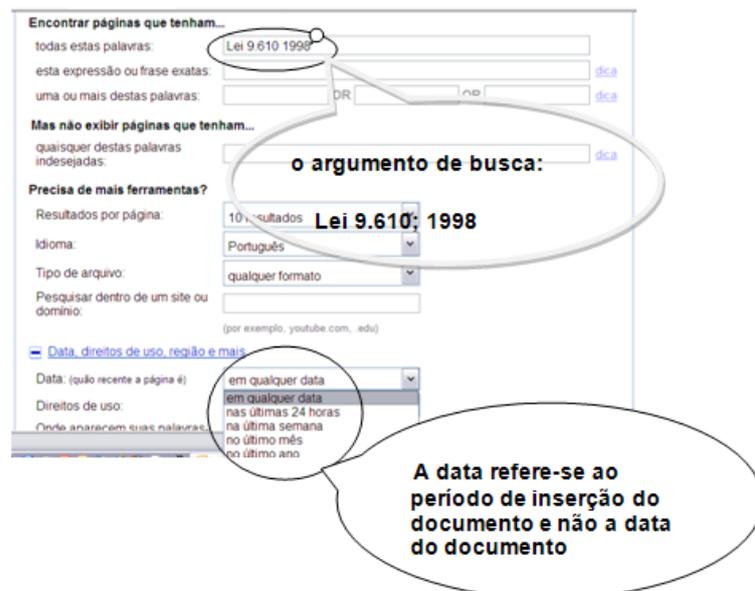
O objeto da pesquisa é a Lei 9610 de 19/02/1998, que dispõe sobre Direitos Autorais.

O primeiro exemplo vem do Google, ambiente de busca aberto na *Web*, cuja pesquisa avançada permite o uso de filtros. Um usuário que não conhece a existência de um *site* especializado em legislação e não sabe que uma fonte de informação primária para essa busca é o site do Senado Federal, provavelmente irá recorrer ao Google.

a) Google – pesquisa avançada

- Pesquisando os metadados do objeto da pesquisa: Lei; 9.610; 1998.

Figura 1 – Busca avançada no Google²



Fonte: GOOGLE, 2011d.

² Disponível em: <http://www.google.com.br/advanced_search?q=google&hl=pt-BR&prmd=ivns>.

b) Site do Prodasen – O Sistema de Processamento de Dados do Senado Federal (SENADO FEDERAL, 2011) é uma biblioteca digital que reúne toda a produção de legislação brasileira com suas respectivas alterações. É uma fonte de informação primária, pois a informação nasce em sua origem.

- Pesquisando os metadados: Lei; 9.610; 1998. Observa-se que existem campos específicos para cada metadado, pois estamos em uma interface de busca especializada em legislação.

Figura 3 – Pesquisa no Portal Senado Federal - Prodasen⁴

The image shows a screenshot of the Senado Federal Portal Legislation search interface. The header features the logo of the Senado Federal and the text 'PORTAL LEGISLAÇÃO'. Below the header, there is a search section titled 'Pesquisa Legislação : Rápida'. This section includes three input fields: 'Tipo' (set to 'LEI'), 'Número' (set to '9610'), and 'Ano' (set to '1998'). A 'Pesquisar' button is located below these fields. On the left side, there is a sidebar menu with the following options: 'Básica', 'Avançada' (with sub-options 'Padronizada' and 'Tradicional'), 'Legislação' (with sub-options 'Rápida' and 'Textual'), 'Biblioteca', and 'Cesta de Ítems'.

Fonte: SENADO FEDERAL, 2011a.

⁴ Disponível em: <<http://www6.senado.gov.br/sicon/index.jsp?action=LegislacaoTextual#>>.

b) Representação do resultado de busca

Por ser uma biblioteca digital especializada, o resultado é um único registro, a própria lei.

Figura 4 – Resultado de busca no Portal Senado Federal - Prodasen⁵

The screenshot shows the 'Portal Legislação' interface. On the left is a 'PESQUISA' sidebar with options like 'Básica', 'Avançada', 'Legislação', 'Biblioteca', and 'Cesta de Itens'. The main content area displays 'Resultado da Pesquisa Legislação : Rápida' and '1 documento'. The document details include: Código 148.431, Origem PODER LEGISLATIVO, Título LEI 9610 de 19/02/1998 - LEI ORDINÁRIA, Apellido LEI DO DIREITO AUTORAL, Data 19/02/1998, Ementa ALTERA, ATUALIZA E CONSOLIDA A LEGISLAÇÃO SOBRE DIREITOS AUTORAIS E DA OUTRAS PROVIDENCIAS, and Publicação DOFC PUB 01 20/02/1998 000003 2 Diário Oficial da União. There are also sections for 'Observação', 'Vide', 'Indexação', and 'Catálogo'. A callout bubble highlights the text 'Um único resultado'.

Fonte: SENADO FEDERAL, 2011b.

O exemplo comparado acima, todavia, não garante que seja mais certo encontrar a informação em uma biblioteca digital, embora, pela lógica, devesse ser mais eficaz. O fato de encontrar a informação não depende apenas de ser ou não um repositório especializado em um tema. Depende também de como esses dados foram tratados e que pontos de acesso à informação foram privilegiados.

Assim, destaco a seguir a importância de critérios na entrada de dados, mas torna-se importante enfatizar, desde já, o papel das bibliotecas digitais na *Web*. Além de serem centrais para as formas de encontro da informação, permitem o acesso a fontes mais fidedignas, numa *Web* democrática e sem parâmetros de confiabilidade.

⁵ Disponível em: <<http://www6.senado.gov.br/sicon/index.jsp?action=LegislacaoTextual#>>.

2.2 A importância dos metadados no tratamento e recuperação da informação

O metadado qualifica a informação, permite a interoperabilidade entre bibliotecas a partir de protocolos de comunicação e corresponde a um conjunto de dados que serão reconhecidos como um registro de informação que descreve um objeto digital. São os atributos de um recurso informacional. A representação descritiva de um objeto digital deve ser representada por metadados, que são os dados sobre dados que ainda podem ser definidos como sendo:

[...] conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode estar em meio eletrônico ou não. Já os formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos (metadados). (ALVES, 2005, p. 115).

As questões críticas que se colocam no momento de definição dos atributos, é “O que será de fato relevante para o usuário?” e “Como identificar atributos relevantes que possam ser acessados através de interfaces de busca?”

Para que se tenha um resultado de busca satisfatório, consistente e pertinente com a busca, é preciso que haja um tratamento da informação na entrada do dado no sistema.

A preocupação em se estabelecer um padrão bibliográfico para a representação de dados em bibliotecas teve início em 1961, com a Conferência de Paris (1961), que, em conjunto com representantes de 53 países que se reuniram para discutir as formas da descrição bibliográfica, estabeleceu a normalização internacional de alguns elementos bibliográficos.

Também nesse período, nos Estados Unidos, a *Library of Congress* (LC) desenvolveu o formato MARC (*Machine-Readable Cataloging*), possibilitando o intercâmbio de dados entre sistemas, formato que evoluiu até os dias de hoje e é amplamente utilizado em bibliotecas para promover a interoperabilidade entre seus registros.

Para efeito deste estudo, não vamos descrever a evolução do histórico de padrões de metadados, e sim nos deter à sua importância nas etapas de alimentação e busca da informação.

Independente de ambiente, físico ou digital, a intenção do usuário é sempre encontrar, identificar, selecionar e obter informação, e é a partir dessas tarefas básicas que estão calcadas as propostas de metadados, na intenção de tornar essas ações mais rápidas e eficazes.

Os metadados representam todos os elementos possíveis de descrição que devem conter na entrada de dados e são essenciais para efetuar a busca e gerar relatórios. Devem, ainda, ser os mais criteriosos e ricos possíveis, descrevendo os recursos informacionais sempre do ponto de vista das necessidades do usuário.

Além disso, seus formatos-padrão auxiliam na tarefa de não esquecer, além de identificar quais são os dados essenciais a serem descritos.

Existem tipos de metadados a serem descritos: Administrativos, Descritivos, De Preservação, Técnico e de Uso. A seguir um quadro com definição de cada tipo e exemplos.

Quadro 1– Tipos de metadados quanto à sua função

TIPO	DEFINIÇÃO	EXEMPLOS
Administrativo	Metadado utilizado na administração de recursos de informação	<ul style="list-style-type: none"> • Aquisição de informação • Direitos de reprodução • Critérios de seleção para digitalização etc.
Descritivo	Metadado para descrição de recursos de informação	<ul style="list-style-type: none"> • Catalogação de registros • Índices especializados etc.
De Preservação	Metadado utilizado para preservação de recursos de informação	<ul style="list-style-type: none"> • Documentação das condições físicas dos recursos etc.
Técnico	Metadado utilizado para conhecer as funções de um sistema ou o comportamento dos metadados	<ul style="list-style-type: none"> • <i>Hardware e software</i> • Dados de segurança • Documentação etc.
De uso	Metadado relativo ao nível e tipo de uso de um recurso de informação	<ul style="list-style-type: none"> • Registros de exibição • Sumário de reuso e de versões etc.

Fonte: Swetlan, 2002, p. 5.

Embora existam inúmeros formatos-padrão de metadados para recursos informacionais digitais, é difícil encontrar na Web uma biblioteca digital que se aproprie desses padrões no momento de tratar a informação.

A Biblioteca Digital da Universidade Estadual de Campinas – Unicamp é um exemplo de serviço que busca ampliar a utilização de metadados para enriquecer o resultado de busca. Podemos identificar, na imagem a seguir, alguns tipos de metadados mencionados no quadro número 1.

Figura 5 – Biblioteca Digital Unicamp, resultado de busca⁶

The screenshot shows a search result for a document titled "Capacitação de professores leigos a distância" by Maria Luiza Soares Inocencio. The page is annotated with callouts identifying different types of metadata:

- Metadados descritivos:** Title [PT], Author(es), and Palavras-chave [PT].
- Metadados de segurança:** Data da defesa: 1999.
- Metadados administrativos:** Idioma, Data de Publicação, Local de Publicação, Código de Acesso, and Orientador.
- Metadados de uso:** Criado, Visitas, and Downloads.

At the bottom, a table provides technical details for the document:

Arquivo	Formato	Tamanho	Tempo estimado para download
Inocencio.M.L.S.pdf	Documento PDF	1281 Kb (1311673 bytes)	1 minuto(s) (Velocidade de conexão de 56 kb/s)

Fonte: UNIVERSIDADE ESTADUAL DE CAMPINAS, 2011.

É especialmente importante a exaustividade de metadados na entrada de dados para permitir que o usuário se beneficie da interface de busca no momento de buscar a informação. Se não há dado, não há informação.

Vivemos uma época em que há um paradoxo de que muita informação é menos informação e não necessariamente algo que nos leva a adquirir conhecimento.

Em plena era da informação, quando o ciberespaço se apresenta a nós como uma grande biblioteca, uma biblioteca com todos os livros e todos os documentos, sistemas de mapeamento lógico tornam-se indispensáveis. Como não nos perderemos e não ficamos exaustos diante de tanta profusão de dados? (LEÃO, 2003, p. 86).

⁶ Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000224332>>.

Para fins deste estudo, explorarei um pouco mais sobre a importância dos metadados e seus formatos-padrão.

2.2.1 FRBR - Requisitos Funcionais para os Registros Bibliográficos

Apresentado em uma Conferência Internacional da IFLA (International Federation of Library Associations and Institutions – Federação Internacional de Associações e Instituições Bibliotecárias), em 1997, o padrão FRBR (*Functional Requirements for Bibliographic Records* – Requisitos Funcionais para os Registros Bibliográficos) foi publicado posteriormente em 1998 com o título de FRBR: final report (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 1998).

O estabelecimento dos requisitos necessários aos registros bibliográficos teve como base as tarefas genéricas realizadas pelos usuários: encontrar, identificar, selecionar e obter informação. Sendo, **encontrar** registros que correspondam ao critério de busca estabelecido pelo usuário, ou seja, localizar registros ou um conjunto de registros, através de um atributo ou de um relacionamento. **Identificar** registros permite ao usuário confirmar se encontrou aquilo que procurava, distinguindo entre dois ou mais registros com características similares. **Selecionar** registros adequados às necessidades do usuário, isto é, escolher registros que atendam as especificações do usuário quanto ao conteúdo, formato físico, etc. ou à rejeição de um registro que não atenda suas necessidades. **Obter** um registro permite ao usuário adquirir o registro, seja através de uma compra, um empréstimo ou através do acesso eletrônico remoto (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 1998).

O RDA - *Resource Description and Access* - é apenas uma extensão do AACR2 – Código Anglo Americano, digamos, uma nova edição baseando-se nos FRBR - *Functional Requirements for Bibliographic Records*, em português, Requisitos Funcionais para Registro Bibliográfico. (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011).

RDA – Em português, **R**ecursos: **D**escrição e **A**cesso.

O RDA está provocando mudanças na catalogação internacional, desde a publicação da segunda edição do *Código de Catalogação Anglo-Americano (AACR2)*.

Esta nova norma de catalogação veio substituir o *Código de Catalogação Anglo-Americano*, ou seja, as *Anglo-American Cataloguing Rules (AACR2)*, que não serão mais atualizadas. Mas, importante mencionar que o RDA não é uma ruptura brusca em relação ao AACR; na verdade, é uma continuação, porém adaptada para um mundo em rede e conectado na *Web*.

Essa informação só é importante para demonstrar o quanto a Biblioteconomia e Ciência da Informação estão sensíveis à busca de novos padrões que auxiliem na descrição de recursos informacionais na *Web*.

2.2.2 Dublin Core Metadata Element Set (DCMES)

Em se tratando de bibliotecas digitais, muitas adotaram o Dublin Core como ponto de partida, mas incluíram metadados específicos, adaptando suas especificidades conforme sua organização, não se valendo totalmente das recomendações destacadas no quadro 1.

O Dublin Core foi adotado como padrão pelo Open Archives Initiative e pela plataforma para repositórios D-Space (GREENBERG; KLAS, 2008).

Os metadados facilitam o compartilhamento da informação, facilitam a gestão da informação e tornam mais efetiva as ferramentas de busca.

São considerados elementos essenciais na descrição de recursos informacionais 15 elementos: Título, Colaborador, Fonte, Criador, Data, Idioma, Assunto, Tipo, Relação, Descrição, Formato, Cobertura, Publicador, Identificador, Direitos. A partir desses elementos, a biblioteca cria/adapta conforme suas necessidades.

O que é muito comum acontecer, e será demonstrado mais adiante, é uma utilização literal desses metadados, sem buscar a compreensão do usuário em relação ao conteúdo do campo. Ex.: Criador é um campo de autoria e poderia ser substituído por "Autoria", mas existe um apego por parte dos profissionais

bibliotecários em relação aos padrões, e isso muitas vezes prejudica o encontro da informação.

Quando falamos sobre a busca na *Web*, o problema aumenta, pois a maioria das páginas HTML não possuem metadados, no máximo uma breve descrição e poucas palavras-chave.

2.2.3 MTD-Br

A Biblioteconomia e Ciência da Informação têm buscado formatos-padrão de metadados para tratar os recursos informacionais, qualificar o dado e proporcionar a recuperação desses metadados. No caso de Teses e Dissertações, existe um padrão internacional, o ETD-MS (*Electronic Theses and Dissertations*), que tem seu correspondente em português, MTD-Br - Padrão Brasileiro de Metadados de Teses e Dissertações.

Alguns elementos do ETD-MS são comuns ao Dublin Core Metadata Element Set (DCMES), como veremos adiante. Seus elementos são: Título, Título Alternativo, Autor, Assunto, Descrição, Descrição/Resumo, Descrição/Nota, Descrição/Versão, Editor, Contribuidor, Contribuidor/Papel, Data, Tipo, Formato, Identificador, Língua, Cobertura, Direitos, Tese/Grau/Nome, Tese/Grau/Nível, Tese/Grau/Disciplina, Tese/Grau/Outorgante.

Tanto o ETD-MS quanto o MTD-Br atendem aos requisitos do DCMES, mas esse cuidado não garante que todos os dados sejam contemplados. Os formatos-padrão são sugestivos e auxiliam na construção de uma tela de entrada de dados para que os dados essenciais não sejam esquecidos, mas o conjunto de dados necessários é uma decisão do profissional bibliotecário, que define a relevância da informação. Muitas vezes os recursos informacionais não são bibliográficos, são corporativos, e nesses casos há que se criar metadados que capturem sua essência e ir além dos padrões recomendáveis.

A título de exemplo, foi selecionada a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) do IBICT – Instituto Brasileiro de Informação Científica e Tecnológica do Ministério da Ciência e Tecnologia, que adotou o MTD-Br, modelo

brasileiro padrão de metadados para a coleta e disseminação das dissertações e teses das universidades conveniadas.

O projeto consiste em coletar a produção intelectual universitária a partir de um convênio com o IBICT. Nesse momento, é aplicado um protocolo da comunicação e os dados que entram no sistema das universidades são replicados para a base do IBICT, que passa a ser um repositório da produção intelectual do conjunto das universidades.

Foi feita uma simulação de busca comparativa na BDTD e no Google a partir dos metadados:

Título: Web Semântica

Assunto: Web Semântica

País: Brasil

Idioma: Português

Grau: Mestrado

Período: Posterior a 2005

a) Biblioteca Digital de Teses e Dissertações – BDTD / IBICT

- Busca na BDTD

Figura 6 – Interface de busca da BDTD⁷

The image shows a web browser window with the BDTD search interface. At the top, there are two tabs: 'Procura Básica' and 'Procura Avançada'. Below the tabs, there are several search criteria, each with a text input field and a dropdown menu labeled 'Todas as Palavras':

- Autor: [Empty text box]
- Resumo: [Empty text box]
- Título: [web semântica]
- Assunto: [web semântica]
- Contribuidor: [Empty text box]
- Instituição de Defesa: [Empty text box]

Below these fields, there is a section for filters with the heading 'Procurar (utilizar um espaço para separar palavras)'. It includes:

- País: [Brasil]
- Grau: [Dissertações]
- Idioma: [Português]
- Ano de Defesa: A partir de: [2005] e/ou até: [Empty text box]

At the bottom of the search area, there are two buttons: 'Procurar' and 'Limpar Procura'.

Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

⁷ Disponível em: <<http://btdt.ibict.br/pt/inicio.html>>.

- Resultado de busca na BDTD

Figura 7 – Resultado de busca da BDTD⁸

Estratégia de busca: grau:Mestre

Classificar por: Relevância

5 documentos encontrados | mostrando os 5 primeiros resultados

Mostrar Detalhes 1-5 de 5

74,60%

Título: [pt] Adaptive-Moodle: Adaptatividade e Interoperabilidade em Ambientes de e-learning utilizando Tecnologias da Web Semântica

Autor: José dos Reis Mota

Contribuidores: Márcia Aparecida Fernandes (Orientador), Alexandre Cardoso (Membro da Banca), Rosa Maria Esteves Moreira da Costa (Membro da Banca)

Instituição de Defesa: <http://www.ufu.br> Universidade Federal de Uberlândia

Assunto(s): Adaptatividade, Interoperabilidade, Web semântica, Inteligência artificial, CIENCIA DA COMPUTACAO, Inteligência artificial, Web semântica, Adaptivity, E-learning, Interoperabiliby, Semantic web, Artificial intelligence

Arquivos: Dissertação

Fonte: Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

b) Google Acadêmico

- Busca no Google Acadêmico

Utilizando-se a pesquisa avançada do Google Acadêmico, verifica-se que não é possível delimitar a pesquisa exatamente como no exemplo da BDTD. Falta campo para selecionar por dissertação e não é possível parametrizar a expressão *Web Semântica* no campo assunto, apenas no campo título, como pode ser visto a seguir. Além disso, há ainda o fato de essa opção reunir a pesquisa acadêmica de produção intelectual em teses, dissertações e artigos de periódicos.

⁸ Disponível em: <<http://bdtd.ibict.br/pt/inicio.html>>.

Figura 8 – Busca avançada no Google Acadêmico⁹

Fonte: GOOGLE ACADÊMICO, 2011a.

- Resultado de busca no Google Acadêmico

Figura 9 – Resultado de busca no Google Acadêmico¹⁰

Fonte: GOOGLE ACADÊMICO, 2011b.

A busca retornou 1 registro, e, mesmo assim, trata-se de um artigo de periódico.

Na pesquisa avançada acadêmica, não é possível parametrizar por tipo de publicação (mestrado/doutorado), e sim por publicação acadêmica geral.

Não é objeto desse tema no momento comparar os resultados, mas é possível perceber que a primeira tela de resultados também não indica o tipo de publicação deste registro. Tratarei desse tema no item 2.5.

⁹ Disponível em: <http://scholar.google.com.br/advanced_scholar_search?hl=pt-BR&as_sdt=0,5>.

¹⁰ Disponível em: <http://scholar.google.com.br/scholar?hl=pt-BR&q=tudonot%C3%ADtulo%3A+%22web+semantica%22&lr=lang_pt&as_ylo=2005&as_vis=0>.

Esses exemplos são apenas ilustrativos, não fazem parte do estudo de caso, mas servem para demonstrar a importância dos metadados no tratamento e busca da informação.

2.3 Vocabulários controlados

Os modelos organizacionais da representação da informação são construídos a partir de uma linguagem documentária que visa uniformizar a linguagem natural representada através da experiência de cada indivíduo. Nas bibliotecas, temos pelo menos dois sistemas de classificação universalmente adotados: a CDD (Classificação Decimal de Dewey) e a CDU (Classificação Decimal Universal). Os livros são, geralmente, organizados a partir desses códigos que garantem, em tese, o agrupamento de livros que tratam de assuntos semelhantes. Esse processo chama-se Classificação.

Em um ambiente digital, a classificação perde importância, pois a recuperação não é mais física. O usuário agrupa os livros por assunto de forma eletrônica. Weinberger (2007) menciona esse novo paradigma: Para que ordenar se vamos desordenar através de um sistema de busca e ordenação variável?

Pois bem, a classificação deixa de ser um elemento essencial e a indexação, atribuição de palavras-chave, ganha maior relevância nesse processo, pois é através dela que o usuário irá recuperar a informação.

Para controlar esse processo de indexação, as bibliotecas adotam vocabulários controlados, que irão contribuir para uma maior consistência e padronização de palavras-chave.

Um vocabulário controlado pode ser desde uma lista de palavras-chave adotadas, quando não se tem uma ferramenta apropriada de gestão, até um Tesouro, que permite estabelecer relações entre os termos e aumentar a oferta de informação.

Um exemplo de aplicação de um Tesouro seria buscar o termo *Comércio Exterior* e como resultado obter registros sobre “importação” e “exportação”, pois são termos que foram relacionados no momento da construção da linguagem com apoio

de ferramenta apropriada para permitir o relacionamento entre os termos no momento da alimentação de dados.

As bibliotecas valem-se desse recurso na intenção de tornar a indexação consistente e padronizada. Ocorre que o usuário tornou-se um internauta e está habituado a digitar termos livres em sistemas de busca; além do que, as interfaces de busca não oferecem índices para que os termos sejam encontrados. A pesquisa geralmente é cega, levando o usuário a um resultado negativo.

Manovich (2011), crítica o design de interfaces de busca simples, quando menciona

Infelizmente, o padrão atual de acesso às mídias - **busca no computador - não nos tira desse paradigma**. A interface da busca é um quadro em branco esperando você digitar algo. Antes de clicar no botão de pesquisar você tem que decidir que palavras-chave ou frases quer procurar. Dessa forma, enquanto a busca traz um aumento dramático na velocidade de acesso, ela assume que você saiba de antemão que algo deve valer a pena na coleção que você vai explorar.

Essa constatação é exemplificada a partir da Biblioteca Digital de Teses e Dissertações do IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia.

a) BDTD

- Busca pela sigla da instituição de ensino “USP”

Figura 10 – Busca na BDTD do IBICT¹¹

The screenshot shows the search interface of the BDTD (Brazilian Digital Library of Theses and Dissertations). The search mode is set to 'Procura Avançada'. The search criteria are:

- Autor: [Empty]
- Resumo: [Empty]
- Título: [Empty]
- Assunto: [Empty]
- Contribuidor: [Empty]
- Instituição de Defesa: **USP**

 Below the search criteria, there are options for 'País', 'Grau', 'Idioma', and 'Ano de Defesa'. The search button is labeled 'Procurar'. A callout bubble points to the 'USP' entry in the 'Instituição de Defesa' field with the text 'Buscando estudos pela sigla USP'.

Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

- Resultado de busca

Figura 11 – Resultado na BDTD do IBICT¹²

The screenshot shows the search results page of the BDTD. The search criteria are:

- Procurar: **instituicao_defesa:"USP"**
- Filtro personalizado: **Sem filtro** - Teses - Dissertações
- Estratégia de busca: **instituicao_defesa:"USP"**
- Classificar por: **Relevância**

 The search button is labeled 'Procurar'. The results section shows the message: 'Nenhum resultado foi localizado para esta consulta'. A callout bubble points to this message with the text 'Resultado Zero'.

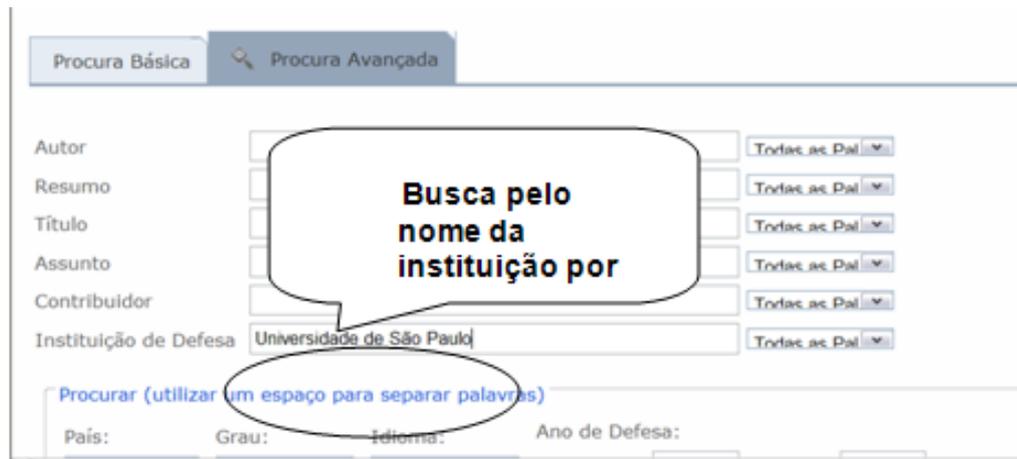
Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

- Busca pela instituição de ensino “Universidade de São Paulo”

¹¹ Disponível em: <<http://bdtb.ibict.br/pt/inicio.html>>.

¹² Disponível em: <<http://bdtb.ibict.br/pt/inicio.html>>.

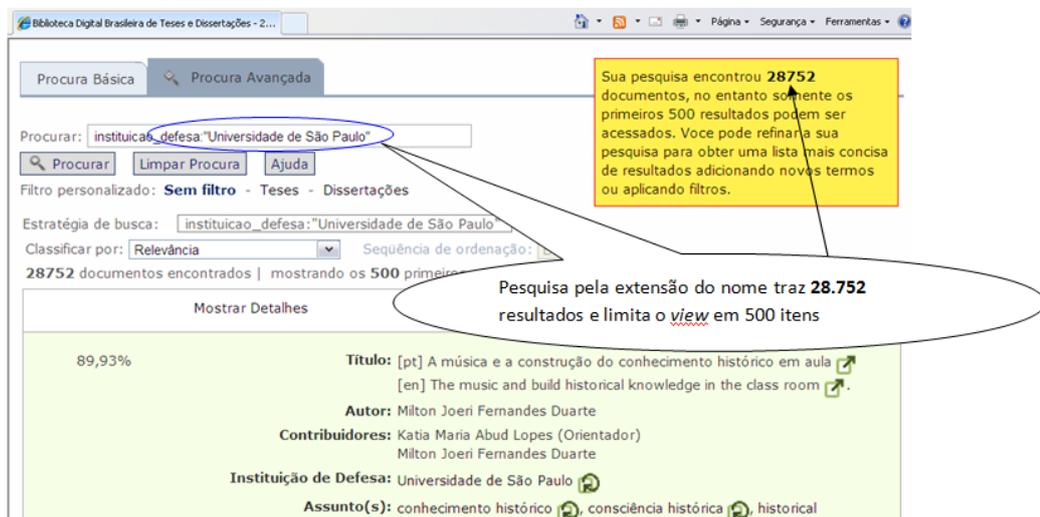
Figura 12 – Busca na BDTD pela instituição de ensino¹³



Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

- Resultado de busca

Figura 13 – Resultado de busca na BDTD pelo nome completo da instituição de ensino¹⁴



Fonte: MINISTÉRIO DA CIÊNCIA E TECNOLOGIA; INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2011.

¹³ Disponível em: <http://bdtb.ibict.br/pt/inicio.html>.

¹⁴ Disponível em: <http://bdtb.ibict.br/pt/inicio.html>.

Este exemplo, além de apontar a importância de utilizar sinônimos em processos de indexação de registros informacionais, exemplifica uma interface de busca que não considerou sinônimos em sua busca e não oferece um índice com as instituições cadastradas.

Como esse serviço denomina-se Biblioteca Digital de Teses e Dissertação e conta com instituições que repassam suas produções intelectuais, seria muito importante ter nesse momento uma visão de quais são as instituições colaboradoras para facilitar a busca. Bastaria a aplicação de um controle de sinônimos e/ou um campo que registrasse as variações de nomenclatura utilizadas para as diversas instituições de ensino para aumentar significativamente o encontro da informação.

A Unicamp é a maior das universidades brasileiras presentes nos repositórios da BDTD com mais de 35.000 itens de dissertações e teses. Porém, ao buscar por “Unicamp” o resultado é zero, igualmente à pesquisa por “USP”.

Um caso interessante, é o Trabalho de Conclusão de Curso (TCC) sobre a análise da produção científica da Universidade Estadual Paulista “Julio de Mesquita”, UNESP, no qual a autora comenta a dificuldade na identificação da produção científica da instituição, por falta de padronização no nome UNESP. Segundo a autora,

[...] houve algumas dificuldades durante as análises, pois a base de dados do ISI não é padronizada para bibliometria. O maior problema foi com os institutos da Unesp, pois grande parte dos artigos estava classificada como Unesp (Universidade) e São Paulo (cidade), mas pertenciam a outros institutos de outras cidades. Portanto, foi necessário padronizar manualmente os institutos e as cidades que possuem campus da Unesp para realizar este trabalho, e mesmo assim, alguns institutos não puderam ser identificados, ficando classificados como “Outros” [...]. (MEDEIROS, 2005, p. 46-47).

Portanto, ter um vocabulário controlado baseado em um Tesauro é muito importante para enriquecer as oportunidades de busca em um ambiente complexo que reúne muitas linguagens e temas diversos, mas o que se vê como tendência é o caminho para um tratamento híbrido da informação, em que o usuário, como colaborador, participa do processo de criação, incluindo suas próprias *tags*.

2.3.1 Taxonomia

Taxonomia é o estudo da classificação. A origem da palavra vem do grego *táksis*=ordem e *ónoma*=nome e originou-se nas Ciências Biológicas, pois refere-se à classificação lógica e científica dos seres vivos, estudo desenvolvido pelo médico e botânico sueco Karl von Linné.

No ambiente da *Web*, os conceitos sobre Taxonomia vêm sendo empregados nas várias formas automatizadas através das quais podemos organizar e representar a informação digital; não apenas na representação do conhecimento de uma determinada biblioteca, mas também no universo corporativo para facilitar a busca e navegação através de uma Intranet ou Portal. Dessa forma, a Taxonomia passou a ser alvo de estudos da Ciência da Informação.

Em uma corporação (daí a expressão *taxonomia corporativa*), os departamentos e os indivíduos possuem as suas linguagens próprias (jargões) que servem no dia a dia para a sua comunicação interna, mas que deixam de ser eficazes em um ambiente digital único, quando a empresa passa a querer se utilizar de uma busca unificada.

Nesse momento, a Taxonomia passa a ser um instrumento unificador dessa linguagem, organizando, representando e buscando conteúdos em ambientes *Web*.

A Taxonomia permite o acesso unificado aos conteúdos e agiliza a tomada de decisão com base na localização da informação. Por ser uma estrutura corporativa, deve refletir o universo da empresa, e os seus termos devem representar todo o conteúdo abrangido em seu escopo.

Em uma Taxonomia já existe uma hierarquia definida, mas não é possível relacionar seus termos, pois por definição, não é esse o seu objetivo.

Um exemplo de Taxonomia em um *site* é o mapa do *site*, onde vemos de forma hierárquica os principais temas contidos no ambiente digital de forma “hiperlinkada”.

Figura 14 – Mapa do site da Comissão de Valores Mobiliários¹⁵

Mapa do Site	
Principal	Página Principal A CVM Links (Sites Relacionados)
Audiência Pública	Audiência Pública
Concurso Público / Estágio	Concurso público Estágios
Contas Públicas da CVM	Contas Públicas da CVM
Dados e Publicações CVM	Cadernos CVM Cartilha Convênios Comparativo Internacional Informativos CVM Pareceres da Procuradoria Jurídica Publicações e Artigos Relatório anual Plano estratégico Informes mensais Revista da CVM
Decisões do Colegiado	Decisões do Colegiado em Processos de Recurso
Legislação e Regulamentação	Atos CVM Convênios Decisões conjuntas Leis Mailing List Modos de Acesso

Fonte: MINISTÉRIO DA FAZENDA, 2011.

Para facilitar as formas de encontrar a informação, a navegação de forma livre enriquece a busca e oferece ao usuário opções que vão além dos menus projetados na interface do ambiente.

A literatura aponta para três tipos de taxonomia:

- a) **Taxonomia descritiva** – construção de um vocabulário controlado que represente, em suas mais variadas relações, a linguagem da empresa;
- b) **Taxonomia de navegação** – representação do modelo mental do usuário, permitindo chegar à informação a partir da navegação. Não envolve a construção do relacionamento entre os termos;
- c) **Taxonomia para gerenciamento de dados** – aplicação em uso mais restrito, funcional. Um único departamento pode melhorar sua produtividade utilizando-se de uma árvore taxonômica para unificar sua linguagem (AQUINO; CARLAN; BRASCHER, 2009).

A taxonomia descritiva seria equivalente à taxonomia corporativa. Envolve um trabalho grande de levantamento, uma equipe de especialistas e tecnologia para suportar o seu desenvolvimento. Entre a taxonomia corporativa e a taxonomia mais aplicada a um departamento, existem alguns graus de separação, e a decisão sobre qual o tipo a ser adotado irá depender das necessidades da empresa.

¹⁵ Disponível em: <<http://www.cvm.gov.br/>>.

No caso de uma Biblioteca Digital, a Taxonomia deverá representar o universo de conhecimento da biblioteca em questão, servindo como opção de busca através da navegação.

O processo de construção de uma taxonomia envolve pessoas-chave da organização e pode ser rápido ou mais longo, dependendo de seu escopo. O importante é entender que seu principal benefício é alavancar o uso de conteúdos em ambiente *Web* a partir de níveis hierárquicos que vão sendo construídos na árvore taxonômica, representando e universalizando a linguagem corporativa e, conseqüentemente, contribuindo no momento da busca.

2.3.2 *Tesouro*

O Tesouro é um recurso sofisticado, mas também rígido, pois, dependendo da área, os termos novos surgem rapidamente, e faz-se necessário alimentar o sistema de forma dinâmica e contínua.

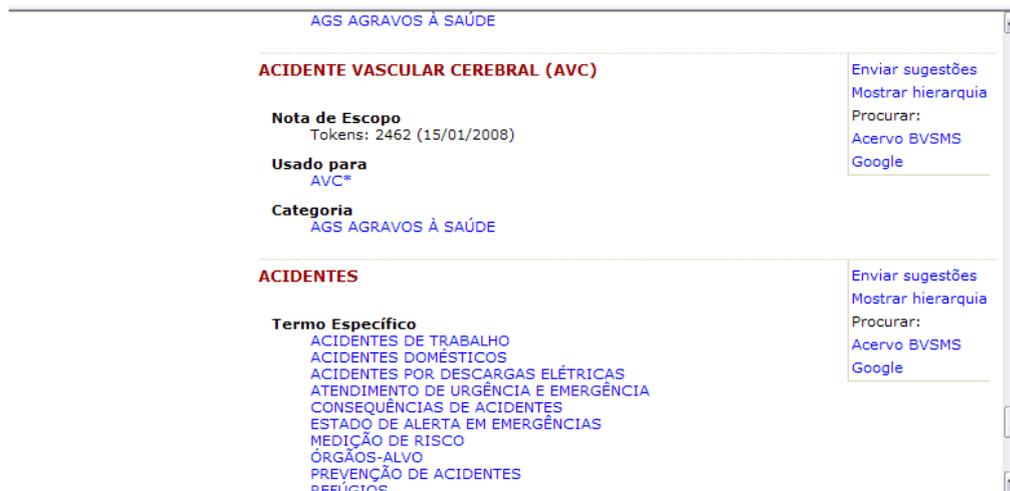
Por estabelecer relacionamento entre termos, no momento da busca, esses relacionamentos são acionados e oferecidos ao usuário. Permite um resultado de busca além do termo específico buscado. Busca-se pelos sinônimos, termos relacionados, termos específicos, termos genéricos e adotados.

O que é pouco praticado, mas recomendável, é acompanhar os *log* de busca dos usuários, para avaliar o quanto o sistema está respondendo ou não aos termos solicitados. Os *logs* de busca são os registros de termos utilizados pelo usuário na busca. Com base neles, é possível reavaliar os termos utilizados no processo de indexação e otimizar a “encontrabilidade” da informação. Exemplo: Se o usuário busca por Dolly, a primeira ovelha clonada, não basta indexar o registro com o termo *clonagem*.

Com essa informação, é possível ajustar os termos, criar novos termos, enriquecer o dicionário com palavras buscadas para aumentar a assertividade.

Um exemplo de uma biblioteca digital que se utiliza de um Tesouro para facilitar a navegação e recuperação da informação é a Biblioteca Virtual em Saúde (BVS). A figura 15 a seguir, ilustra o tratamento e relacionamento entre termos que são representados a partir do termo AVC.

Figura 15 – Apresentação do Tesouro da BVS¹⁶



Fonte: BIBLIOTECA VIRTUAL EM SAÚDE, 2011.

Esse recurso enriquece a busca, pois recupera termos relacionados à palavra buscada.

2.3.3 Folksonomia

Vocabulário controlado ou descontrolado?

A *folksonomia* é um neologismo cunhado por Thomas Vander Wal ao juntar as palavras “folks” (pessoas, em tradução livre) e taxonomia, embora alguns autores ainda contestem o termo *taxonomia*, uma vez que é geralmente estruturado hierarquicamente, e a *folksonomia* não tem princípio hierárquico, nem associativo. A informação pode ser organizada de diferentes formas no mundo digital, mas o que também se questiona é se a *folksonomia* pode atuar como um tipo de estratégia para recuperação colaborativa da informação, por ser uma linguagem natural livre de qualquer padrão ou associação semântica.

A *Folksonomia* é um elemento da Web 2.0 que auxilia o processo de recuperação da informação, uma vez que o próprio usuário associa uma *Tag* ao registro e passa a buscá-lo com a palavra-chave que ele mesmo atribuiu. A livre

¹⁶ Disponível em: <<http://bvsmms.saude.gov.br/cgi-bin/multites/mtwdk.exe?k=default&l=60&s=1&t=2&w=&w2=ABSTIN%CANCIA%20SEXUAL&cp=2&x=&tt=&h=&n=25>>.

rotulação introduz o conceito de indexação colaborativa e por ser livre, trata-se de um vocabulário não controlado.

Essa funcionalidade não só permite que o usuário crie o seu conjunto de *tags*, como também auxilia o profissional bibliotecário a enriquecer o vocabulário controlado da biblioteca digital em questão, uma vez que essas *tags* podem ser incorporadas na linguagem adotada.

O site www.delicious.com é um exemplo de aplicação de *folksonomia*. O usuário pode criar as suas *tags*, compartilhá-las e pesquisá-las pelas *tags* de outros usuários (*social tags*).

O fato de que não há nenhum controle sobre a *tag*, faz com que o mesmo assunto seja encontrado de várias formas, diminuindo a chance de encontro da informação no momento da busca. Ex.: Biblioteca 2.0; Library 2.0; Biblioteca2.0. Neste exemplo, temos 3 palavras-chave com o mesmo significado e cada usuário poderá optar de forma livre, podendo inclusive o mesmo usuário se utilizar das 3 formas sem controle sobre suas próprias *tags*. Essa ocorrência pode ser facilmente observada em uma nuvem de *tag*, no site www.delicious.com no exemplo a seguir.

Por ser um site que permite o registro de sites favoritos na *Web*, existe uma funcionalidade que permite visualizar as *tags* através de uma nuvem de *tags*. As *tags* podem estar em qualquer língua, no singular ou plural, agrupadas sem espaço ou com espaço, estando cada palavra em uma determinada ordem (ex: biblioteca em “B” e digital em “D”).

Figura 16 – Delicious¹⁷

Fonte: DELICIOUS, 2011.

O fato de o espaço entre as palavras representar outra palavra na sintaxe da *tags*, também é uma questão que compromete o resultado, pois usuários menos acostumados com a *Web* estarão, na verdade, criando duas *tags*: biblioteca e 2.0. O correto seria biblioteca2.0 (sem espaço), mas poucos percebem/conhecem esses detalhes de estratégia de busca.

De qualquer forma, o uso da *folksonomia* em bibliotecas digitais é bastante válido, pois é possível criar “My tags” para os registros de interesse, facilitando o reuso da mesma informação. Assim, o próprio usuário pode criar sua taxonomia e controlar suas palavras-chave, dando um sentido a essa prática um tanto quanto anárquica na *Web*.

Uma grande vantagem da aplicação da *folksonomia* é a possibilidade de enriquecer o vocabulário controlado. Com acompanhamento e análise dessas *tags*, o profissional bibliotecário poderá incorporar novos termos e aumentar a chance de encontrar a informação para outros usuários.

A *folksonomia*, utilizada por si só, não representa uma solução para encontrar a informação, mas serve de recurso para as bibliotecas que necessitam controlar seus vocabulários e, ao mesmo tempo, lidar com a diversidade de usuários. A lentidão de processos institucionais também é um fenômeno que estimula essa

¹⁷ Disponível em: <<http://www.delicious.com/tag/>>.

decisão. A análise e aprovação de um termo em um vocabulário controlado, dependendo da instituição, pode demorar meses, sendo que jargões e palavras novas surgem a todo o momento.

2.4 A interface de busca como fator crítico para o encontro da informação

Uma biblioteca digital é estruturada a partir de seus recursos informacionais e o uso que se fará deles. A partir da definição dos metadados que compõem sua representação descritiva, ela está preparada para receber a quantidade de dados necessária para representar um registro, conjunto de metadados que tem em comum o mesmo ID, identificador de registro, em um banco de dados.

A lógica desse banco de dados e seus metadados não é conhecida dos usuários. Estes dependem de uma interface de busca para poder encontrar a informação desejada. Com a profusão de informação disponível na *Web*, com a dependência crescente que temos cada vez mais dos mecanismos de busca para de fato encontrar o que necessitamos, justifica-se a preocupação central em desenvolver interfaces de busca que permitam a recuperação efetiva da informação.

É comum nos depararmos com bibliotecas digitais bem desenvolvidas e amparadas em bancos de dados muito bem estruturados, mas com uma interface de busca que não representa o universo de seu conteúdo e disponibiliza recursos suficientes, impedindo o usuário de encontrar a informação.

As interfaces de busca em um ambiente digital são o meio por onde transita a linguagem comum entre o bibliotecário e usuário. Devem ser consistentes nas mensagens que buscam transmitir. São fatores essenciais que determinam a usabilidade de um ambiente digital.

De acordo com Jakob Nielsen e Molich (apud SANTOS, 2011), são dez os princípios heurísticos de usabilidade:

- 1) Feedback;
- 2) Falar a linguagem do usuário;
- 3) Saídas claramente demarcadas;
- 4) Consistência;
- 5) Prevenir Erros;

- 6) Minimizar a sobrecarga de memória do usuário;
- 7) Atalhos;
- 8) Diálogos simples e naturais;
- 9) Boas mensagens de erro;
- 10) Ajuda e documentação.

As heurísticas são recomendações sem um embasamento científico comprovado, baseados na intuição, bom senso ou na experiência de determinado indivíduo, utilizadas para minimizar ou encontrar erros e falhas durante o desenvolvimento e avaliação de uma determinada interface. (PINHEIRO, 2011).

Heurísticas não são regras, variam de acordo com a necessidade e dinâmica de cada projeto. Novos pontos heurísticos podem ser definidos sempre que surgirem novas necessidades. As necessidades aparecem conforme as interações se tornam mais complexas, levando ao surgimento de novas tecnologias, plataformas e formas de interação.

De acordo com Maculan, Lima e Penido (2011), é notória a evolução das interfaces de bibliotecas digitais, que iniciaram com os catálogos eletrônicos *on-line*, conhecidos na Internet como OPACs (*On-line Public Access Catalogs*), nos quais se faziam consultas padrões por palavra-chave, nome do autor ou título da obra e evoluíram tendo hoje praticamente o acervo integral em ambiente digital.

Atualmente, com o avanço das tecnologias de busca, a biblioteca digital pode oferecer diversificadas opções, que vão além das buscas padrão. Dentre as opções encontradas, têm-se:

Boolean Logic: pesquisa combinações lógicas de múltiplos termos;

Fuzzy Expansion: expande a busca por similaridade da escrita ou da pronúncia;

Wildcard: expande a busca de uma palavra, partindo de uma origem comum, sem considerar se ela pertence ou não a mesma origem linguística;

Exactly like/stem expansion: expande a busca para variações derivadas da palavra originária. A variação inclui nome, pronome, adjetivo, porém todas as palavras derivadas têm a mesma origem linguística;

Case Sensitivity: serve para ignorar ou não o uso da palavra escrita em maiúscula, minúscula ou de ambas as formas;

Stopwords: há um conjunto de palavras que podem ser ignoradas nas buscas;

Proximidade: localiza documentos em que os termos especificados ocorram pelo menos uma vez, e quanto mais próximos estiverem os termos dentro de um mesmo documento maior será sua relevância.

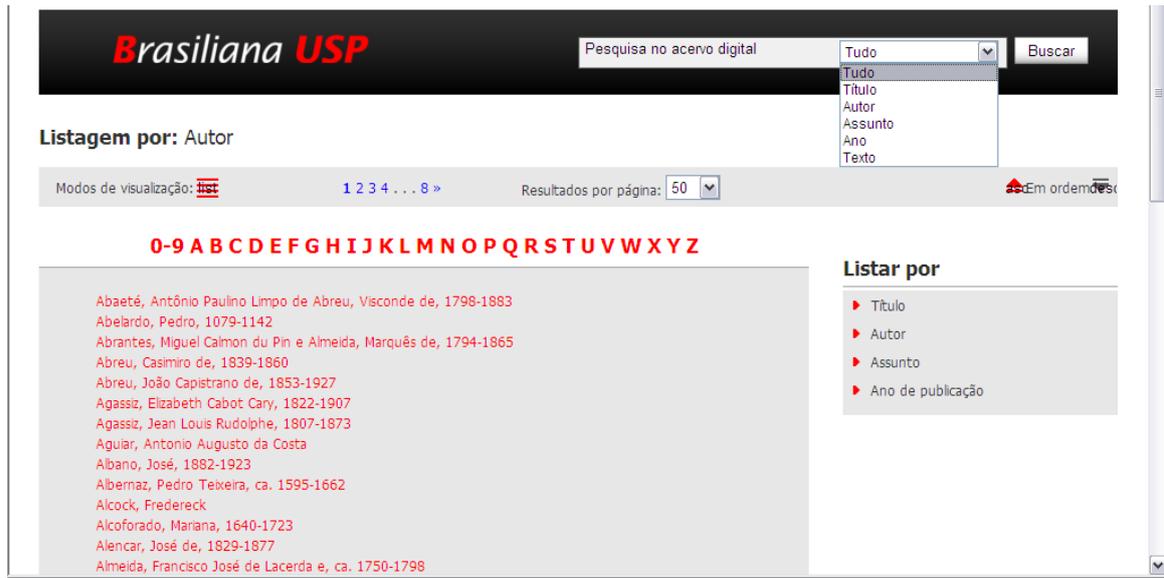
Normalmente, encontra-se em bibliotecas digitais uma interface com duas formas de busca: 1) busca simples; e 2) busca avançada. Na interface simples, estão disponibilizadas poucas opções de busca e na maioria das vezes, apenas um campo em branco; na interface avançada, há um número maior de opções, podendo o usuário limitar e especificar melhor a busca, mas a falta de orientação muitas vezes impede o uso dessa opção.

Um ponto importante a ser considerado no desenvolvimento de interfaces de busca é o que Chang e Rice (1993) chamam de dimensão contextual, na qual o usuário pode construir significado a partir do contexto. Isso ocorre quando é possível refinar a busca, como visto na figura 15.

Em termos de estado da arte em bibliotecas digitais, o projeto da Biblioteca Mindlin, responsável pela implantação da Brasileira USP – um projeto da Reitoria da Universidade de São Paulo, inaugurado em 2009 –, é uma referência.

Mesmo sendo a excelência em termos de concepção, idealização de uma biblioteca digital e uso de tecnologia de ponta, o projeto ainda não dispõe de recursos suficientes para a busca e nem de elementos 2.0 que caracterizam o grau de colaboração entre os usuários.

Figura 17 – Tela principal de busca da Brasileira USP¹⁸



Fonte: UNIVERSIDADE DE SÃO PAULO, 2011.

A sua interface de busca traz um índice agrupado por letra e não possui pesquisa avançada para cruzamento de campos, além de determinar apenas 5 pontos de acesso ao acervo: autor, título, assunto, data e texto.

Esse aspecto será explorado no capítulo 5, em que os estudos de casos serão analisados.

As teorias que envolvem a Interação Humano-Computador (IHC) dão respaldo aos estudos da usabilidade nas interfaces de busca. As teorias da Ciência da Informação (CI) dão respaldo aos estudos de necessidades de informação e comportamento de busca e uso de informação. Assim, é altamente recomendável a aproximação de ambas as áreas para que surjam projetos maduros, inteligentes e eficazes.

Talvez o problema maior da biblioteca não seja o excesso de livros, mas sim a necessidade de novas cartografias. Cartografias mutantes, cartografias reveladoras. Cartografias tão flexíveis e coerentes com as nossas necessidades subjetivas e objetivas que, ao contrário do império borgiano, não as julgaremos inúteis, mas sim vitais para o processo de interação no ciberespaço. (LEÃO, 2003).

Encontrar *links* para as redes sociais em bibliotecas digitais que estejam utilizando a Web 2.0 como canal de comunicação e divulgação tem sido muito comum, e aumentam as iniciativas a cada dia. O que não se encontra com facilidade

¹⁸ Disponível em: <<http://www.brasiliana.usp.br>>.

é uma interface de busca com *design*, estética e funcionalidades que privilegiem os perfis cognitivos de usuários, facilitando o encontro da informação.

2.5 A representação dos resultados de busca

Estudos apontam que no ano de 2007 existiam aproximadamente 550 bilhões de documentos *on-line*, com aproximadamente 7,5 petabytes entre *websites* e base de dados *on-line* (JANSSENS, 2007 apud MAIA; SOUZA, 2008). Para armazenar 7,5 petabytes, em uma pilha de páginas de papel, em que cada página conteria 2500 caracteres, sendo que um byte equivale a 1 caractere, teríamos uma pilha de 300.000 km (1 cm para 100 páginas) o que daria para alcançar a lua ou dar a volta na terra 7,5 vezes. Uma pessoa lendo uma página por minuto gastaria 5,7 bilhões de anos para ler tudo. (JANSSENS, 2007 apud MAIA; SOUZA, 2008). Com todo este volume informacional, algum sistema de recuperação de informação deverá ser necessário, para que uma pessoa recupere rapidamente a informação que deseja em tempo satisfatório. Observa-se pouca mobilização e sensibilização em relação a essa questão até o presente momento. As tecnologias estão muito mais focadas para buscar que para encontrar. Empresas estão comprando sistemas de busca, acreditando que seus conteúdos serão indexados e mais facilmente recuperados. Mas se não são tratados na entrada, não deverão oferecer resultados satisfatórios na saída.

Baeza-Yates e Ribeiro-Neto (1999) definem as funções básicas de um sistema de recuperação de informação, como a representação, armazenamento, organização e acesso à informação, e compreendem que as pesquisas são voltadas à modelagem, classificação dos documentos e sua categorização, arquitetura de sistemas, interface de usuários, sistemas de visualização de dados e filtragem de informação. A noção de relevância, ou seja, recuperar mais documentos que alcancem a necessidade informacional dos usuários e recuperar menos documentos irrelevantes é o objetivo central da recuperação da informação. No entanto, a recuperação efetiva da informação relevante depende tanto das tarefas dos usuários, quanto da visão lógica dos documentos adotados pelos sistemas de recuperação de informação.

Lancaster (1986, p. 131) entende que, para avaliar a efetividade de um sistema de recuperação, devemos determinar o quão bem ele encontra as necessidades dos usuários, levando em consideração também critérios de performance, como: qualidade (medida pela revocação e precisão), esforço (usabilidade do sistema ou até mesmo seu custo de uso) e tempo de resposta.

Para o autor, todos os usuários de um sistema de recuperação de informação têm um pré-requisito em comum: esperam que o sistema seja capaz de recuperar documentos relevantes que contribuam para a satisfação de suas necessidades informacionais. (LANCASTER, 1986).

Um dos maiores desafios da busca é oferecer uma tecnologia e interfaces inteligentes, para que o usuário consiga obter um alto grau de revocação e precisão em seus resultados de busca.

Tanto a precisão como a revocação podem ser expressas na forma de uma equação.

A precisão é indicada pela relevância dos registros obtidos no resultado de uma busca específica, e a revocação é a proporção de registros relevantes recuperados, comparada com a totalidade de registros relevantes existentes na base de um sistema.

O índice de precisão e o índice de revocação são obtidos a partir de numeradores comuns, porém denominadores diferentes.

$$\text{Índice de precisão} = \frac{\text{Número de registros relevantes recuperados}}{\text{Número total de registros recuperados}} \quad (1)$$

$$\text{Índice de revocação} = \frac{\text{Número de registros relevantes recuperados}}{\text{Número total de registros relevantes no sistema}} \quad (2)$$

Essa fórmula se aplica quando conhecemos o total de registros em uma base de dados, o que é o caso em exemplos de bibliotecas digitais que têm o seu universo definido.

A aplicação de métricas ajuda no controle de qualidade. Uma análise detalhada focando os elementos de entrada de dados, tais como metadados e vocabulário controlado, e comparando-os com a qualidade da interface de busca e

representação de resultados irá auxiliar na interpretação dos resultados obtidos nessa avaliação.

A Ciência da Informação estuda o tratamento e organização da informação com o objetivo de possibilitar resultados de busca satisfatórios conforme a demanda do usuário, sem a interferência do “lixo informacional” e a Bibliometria, desde a década de 70, é o campo de estudo que engloba métricas informacionais na Ciência da Informação.

“Existem diversas formas de medição voltadas para avaliar a ciência e os fluxos da informação. Dentre estas, cabe citar a bibliometria, a cienciometria, a informetria e a mais nova delas, a webometria”, um subcampo da Bibliometria, que nos dá algumas formas interessantes de medirmos a qualidade da informação a partir de estudo de *links* e rastreamentos. (VANTI, 2002, p. 152).

No caso de bancos de busca abrangentes como o Google, é muito difícil avaliar a qualidade, pois não se tem ideia do total de registros buscados e muito menos da qualidade desses registros. O que podemos observar é uma tentativa de unificar a representação de dados trazendo o maior número de resultados de busca, muitos dos quais repetidos.

A forma de encontro da informação em projetos de bibliotecas digitais visa romper um paradigma em relação à busca da informação na *Web*. Não há tecnologia de busca que encontre informação, se esta não for disponibilizada adequadamente, planejada para o perfil de usuário desejado e se o resultado de busca não for representado de forma inteligente o suficiente para permitir sua localização.

Ao se formular uma busca em um ambiente *Web*, podemos observar infinitas formas de se representar um resultado de busca. As formas estão relacionadas à maneira como os dados são programados para serem visualizados, o critério de relevância na ordenação de resultados e à tecnologia de busca utilizada pela interface em questão.

A escolha do formato e da ordenação irá interferir na tomada de decisão do usuário, pois a tendência é sempre clicar nos primeiros resultados e não passar das primeiras páginas.

Tomando como exemplo a Lei 9610 pesquisada acima, observei os 4 primeiros resultados de busca do Google.

1) **LEI Nº 9.610, DE 19 DE FEVEREIRO DE 1998.** ... Art. 1º Esta **Lei** regula os direitos autorais, entendendo-se sob esta denominação os direitos de autor e os que ... www.planalto.gov.br/ccivil/leis/L9610.htm - Em cache - Similares

2) Ministério da Cultura - MinC » Lei de Direito Autoral (nº 9610/98)

Lei 9610 de 19 de fevereiro de 1998 Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências ... www.cultura.gov.br/.../lei-de-direito-autoral-nº-961098/ - Em cache - Similares

3) LEI No 9.610, DE 19 DE FEVEREIRO DE 1998

Formato do arquivo: PDF/Adobe Acrobat - Visualização rápida Lei no 9.610/98 – consolidada com proposta de revisão em consulta pública. 1 ... www.cultura.gov.br/.../Lei9610_Consolidada_Consulta_Publica.pdf - Similares

4) LEI No 9.610, DE 19 DE FEVEREIRO DE 1998

Formato do arquivo: PDF/Adobe Acrobat - Visualização rápida
20 fev. 1998 ... LEI No 9.610, DE 19 DE FEVEREIRO DE. 1998. Altera, atualiza ... www.cultura.gov.br/site/wp-content/.../lei_9610-direito-autoral.pdf - Similares

Pode-se observar que existe uma repetição de resultados da mesma fonte (itens 2 a 4) e um emaranhamento de dados não estruturados, dificultando a identificação e seleção por parte do usuário.

Já uma pesquisa efetuada em um ambiente estruturado de dados, como uma biblioteca digital, permite visualizar os resultados de forma melhor organizada..

O exemplo mostrado a seguir foi extraído da Scielo – Scientific Electronic Library Online – [HTTP://search.scielo.org](http://search.scielo.org).

Figura 18 – Possibilidade de ordenação de resultados na SciELO¹⁹



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011e.

Pode-se observar, neste caso, após selecionar assunto, revista, ano de publicação, idioma, coleção, palavra no contexto Resumo e apenas no Brasil, ainda é possível selecionar por ordem de resultado, optando por ano (crescente/decrecente), relevância e Índice SJR.

Essa opção de resultado de busca tem por objetivo atender a necessidade de diferentes perfis de internautas e agrupar os resultados da forma mais rápida para a pesquisa dos resultados obtidos.

Esse aspecto relacionado ao resultado da busca é muito pouco observado, e há pouco ou quase nada de investimento nessa tela. Normalmente, encontra-se uma única forma de saída da informação e, se a pesquisa for na *Web*, via Google, por exemplo, além de ser uma única forma, é também desestruturada e com critérios de relevância desconhecidos.

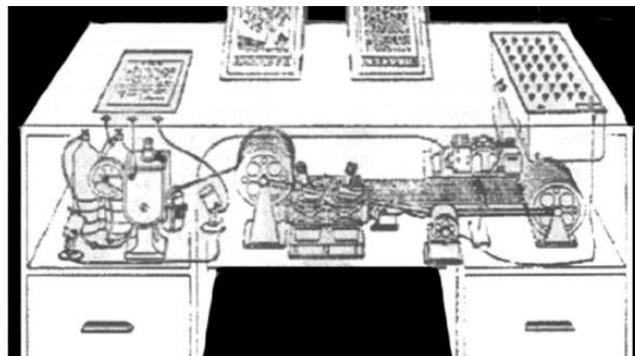
¹⁹ Disponível em: <<http://search.scielo.org/>>.

3 BIBLIOTECAS DIGITAIS

Em 1945, Vannevar Bush, assessor científico do governo norte-americano, chamava a atenção para um sistema similar ao hipertexto, estabelecendo a comunicação entre os cientistas como hoje conhecemos através da Internet. Era uma espécie de arquivo e biblioteca privada mecanizada, que foi chamada de Memex. Consistia em uma mesa com telas translúcidas, teclado, botões, alavancas e mecanismos de armazenamento, gravação e projeção baseados no uso de microfilmes. O processo de ligação entre dois itens seria feito por indexação associativa, em que um item poderia selecionar imediata e automaticamente um ou outro item para ser lido. Bush acrescenta ainda que, ao ligar vários itens, é exatamente como se os itens físicos tivessem sido reunidos para formar um livro. É mais que isso, já que qualquer item pode ser ligado a inúmeras trilhas. (BUSH, 1945 apud VILAN FILHO, 1994).

Bush percebeu que a memória humana operava por associações, e verificou que os sistemas de armazenamento não levavam em consideração os processos cognitivos pessoais. Na realidade, o sistema foi conceitualmente concebido por BUSH (1945). A forma cognitiva que dominava era aquela determinada pelo sistema de armazenamento de informações. Embora suas idéias fossem bastante avançadas, a tecnologia computacional da época não era suficientemente avançada para permitir o desenvolvimento de uma solução mais elaborada. (SILVA, 2005, p. 123).

Figura 19 – A máquina Memex, de Vannevar Bush



Fonte: HISTÓRIA, 2011.

Sem imaginar, Bush já estava idealizando o hipertexto.

O desenvolvimento das bibliotecas digitais está totalmente vinculado aos avanços tecnológicos e à forma de transmissão de dados. Desde o surgimento do primeiro computador conhecido por ENIAC¹, em 1946, até o surgimento da *Web 2.0*, em 2004, muitas transformações e paradigmas foram surgindo.

Em 1971, Michael Hart criou a Biblioteca de Alexandria, primeiro projeto similar a uma biblioteca digital, conhecido como Projeto Gutenberg.

Nesse período, as bibliotecas iniciaram um processo de organização de seus catálogos, transformando-os em catálogos *on-line* e bases de dados organizadas e disponíveis eletronicamente.

A expressão *biblioteca digital* surgiu no início da década de 90, quando deu-se início ao amadurecimento de processos de automação de bibliotecas.

Uma das definições mais utilizadas em literatura especializada para definir biblioteca digital foi dada por um grupo de pesquisadores e utilizada no Workshop on Distributed Knowledge Work Environments, em Santa Fe (EUA), em 1997:

[...] o conceito de “biblioteca digital” não é simplesmente equivalente ao de uma coleção digitalizada dotada de instrumentos de gestão da informação. É, antes, um ambiente que reúne coleções, serviços e pessoas para apoiar todo o ciclo vital de criação, disseminação, uso e preservação de dados, informação e conhecimento. (ATKINS, 1997, tradução nossa).

A definição acima parte do pressuposto que a biblioteca está no ciberespaço e que necessita basicamente de três elementos para se compor: a coleção, os serviços de acesso e os usuários.

Além dessa definição, é importante destacar outra definição considerada relevante na área e uma das mais difundidas, da Digital Library Federation (DLF):

Bibliotecas digitais são organizações que fornecem os recursos, inclusive o pessoal especializado, para selecionar, estruturar, oferecer acesso intelectual, interpretar, distribuir, preservar a integridade e garantir a permanência no tempo de coleções de obras digitais, de modo que estejam acessíveis, pronta e economicamente, para serem usadas por uma comunidade determinada ou por um conjunto de comunidades. (VANTI, 2002).

Uma particularidade importante que caracteriza uma biblioteca é a existência de metadados pré-definidos, ou seja, estruturados e descritos, conforme

¹ Eletronic Numerical Integrator and Calculator. (BREVE, 2011).

mencionado no item 2.2. Essa particularidade diferencia uma biblioteca de um simples repositório de documentos, conteúdos ou livros.

O primeiro a utilizar o termo *biblioteca virtual* foi Tim Berners-Lee e ainda, segundo Tammaro (2008), o conceito empregado nessa definição seria uma coleção de documentos ligados em rede, constituídos por objetos digitais e páginas *Web* produzidos por milhares de autores. O adjetivo *virtual* significa que a biblioteca não existe fisicamente, ela apenas se constitui de uma coleção de *links* que remetem o usuário para outro endereço virtual.

Segundo o Conzorcio Interuniversitario Lombardo per l'Elaborazione Automatica – CILEA (2011), a biblioteca virtual é a *World Wide Web*, que pode ser imaginada como um poderoso organismo que se nutre de informações e cresce de forma caótica.

Um exemplo de experiência recente de biblioteca virtual é o projeto do Google Print (Google Book Search), que apenas disponibiliza *links* para livros digitalizados de algumas coleções do acervo de suas bibliotecas, mas sem qualquer organização.

A página principal do Google livros, similar em inglês ao Google Books, mostra já a partir do menu principal de navegação que existe um conjunto desordenado de assuntos, sem lógica. O objetivo do Google não é ser uma biblioteca digital de livros, e sim uma biblioteca virtual com links para o conteúdo do maior número de livros possíveis a serem digitalizados na íntegra.

A página principal do Google Books, que representa de forma desordenada as áreas de conhecimento representadas no menu à esquerda da figura a seguir, dá a ideia de que o Google não está indo na direção de uma biblioteca digital organizada por temas.

Figura 20 – Página principal do Google Books²



Fonte: GOOGLE BOOKS, 2011.

É possível constatar que já não existem barreiras tecnológicas para que esses sonhos e missões se concretizem. As tecnologias digitais, os mecanismos de busca e a possibilidade rápida de recuperação da informação nos apontam para um futuro já possível de se prever. O conhecimento estará cada vez mais disponível, da mesma forma que a educação poderá ser cada vez mais a distância, em função da qualidade de conteúdos digitais acessíveis.

A diferença está em como cada projeto se apropriará desses recursos, como o desenvolvimento das bibliotecas digitais irá levá-las para além do tratamento e disponibilização da informação. A ruptura de paradigmas e a percepção de que não basta ser digital é preciso ser 2.0, colaborativa é muito importante para a construção de novas interfaces que evoluam nessa direção.

A missão das bibliotecas digitais é apropriar-se do estado da arte em tecnologia e ciência para o tratamento e disponibilização da informação, através de interfaces inteligentes que dialoguem e conduzam o usuário ao seu objetivo.

² Disponível em: <<http://books.google.com.br/bkshp?hl=pt-BR&tab=wp>>.

Já a missão do Google é “organizar as informações do mundo todo e torná-las acessíveis e úteis em caráter universal”. (GOOGLE, 2010). Embora a missão do Google seja organizar a informação, o que o Google faz é disponibilizar a informação para ser buscada, o que é bem diferente de organizar a informação para ser encontrada, como visto no item 2.1.

3.1 A evolução das bibliotecas digitais para a Web 2.0

O termo Web 2.0 foi cunhado por Tim O’Reilly e Dale Dougherty, da O’Reilly Media, em 2004, para estabelecer um marco em relação às mudanças de paradigma e descrever as tendências e os modelos de negócios que sobreviveram à “bolha” do setor de tecnologia nos anos 90 (O’REILLY, 2005).

A colaboração passou a ser o mote da nova *Web*, representando uma nova forma de se comunicar e de compartilhar conhecimento.

O conceito de prosumidor³ surgiu com o desenvolvimento das Intranets, permitindo que sua produção de conteúdo fosse descentralizada e os consumidores de informação passassem a ser também produtores, conteudistas e colaboradores. Essa interação representou um rompimento de paradigma e criou uma grande distância entre ambientes colaborativos e bibliotecas na *Web*.

Os usuários de bibliotecas foram tornando-se mais exigentes, buscando em bibliotecas digitais maior interatividade no ambiente.

Neste cenário de interatividade e colaboração, surgem as bibliotecas digitais 2.0.

A Web 2.0 representa uma mudança de paradigma, uma ruptura de modelos tradicionais na oferta de produtos e serviços e no *design* e tecnologia de interfaces de busca de bibliotecas digitais. O usuário passa a não aceitar mais os catálogos herméticos de bibliotecas tradicionais, pois tem o Google como modelo, supostamente eficiente, com infinitas possibilidades de entrar em contato com informações que nem sequer cogitava existir.

³ Termo que resulta da junção das palavras, produtor e consumidor, que salienta o crescente papel dos consumidores no processo produtivo. (CARDOSO, 2009).

As suas necessidades informacionais evoluem de uma condição passiva, em que ele apenas recebia informação, para uma condição ativa, em que ele passa a comentar, a incluir suas palavras-chave e navegar por itens previamente selecionados por ele a partir de recursos como “My tags” (minhas *tags*) e “My library” (minha biblioteca).

Não basta mais incluir os catálogos na *Web*, os usuários passam a sentir a necessidade de colaborar e compartilhar informações disponíveis nas bibliotecas.

A imagem que hoje representa a relação entre as bibliotecas e seus usuários não é mais um balcão de atendimento que os separa, e sim uma roda da qual ele faz parte.

Figura 21 – Modelo tradicional de atendimento ao usuário



Fonte: Alvim, 2011.

Figura 22 – Modelo de atendimento na web



Fonte: Alvim, 2011.

Segundo Cunha (2008), biblioteca digital é uma coleção de informação digitalizada e organizada, com recursos de busca para texto, imagem e som, que pode ser disponibilizada integralmente ao internauta através de seu computador.

De acordo com Miller (2005), *Biblioteca 2.0* é um termo cunhado por Michael Casey em seu *blog* LibraryCrunch. Embora seus escritos sobre Biblioteca 2.0 sejam baseados com autoridade, Casey (2006) define o termo muito amplamente, arguindo que isso se aplica além dos serviços e inovações tecnológicas. Concordando com Casey (2006), outros bibliotecários blogueiros têm iniciado uma exploração conceitual do que Biblioteca 2.0 pode significar, e, por conta dessa discussão, há muita controvérsia sobre a definição e importância relativa do termo.

O que pode se concluir de fato é a ideia de mesclar o conceito de Biblioteca digital com um repositório de recursos informacionais estruturados em um ambiente digital e com funcionalidades da Web 2.0, em que percebe-se a apropriação de elementos presentes no ciberespaço, conforme representado na figura a seguir.

Figura 23 – Figura representando o conceito de Biblioteca 2.0



<http://dospuntocero.dmaweb.info>

Dídac Margaix



Fonte: Alvim, 2011.

Segundo Maness (2007), o modelo de uma Biblioteca 2.0 poderia ter quatro elementos:

- 1) **É centrada no usuário.** Usuários participam na criação de conteúdos e serviços na presença da biblioteca na *web* etc. O consumo e a criação do conteúdo são dinâmicos, e por isso as funções do bibliotecário e do usuário se misturam.
- 2) **Oferece uma experiência multimídia.** Ambos, coleções e serviços de Biblioteca 2.0, contêm componentes de áudio e vídeo. Embora isso nem sempre seja citado como uma função de Biblioteca 2.0, é aqui sugerido que deveria ser.
- 3) **É socialmente rica.** A presença da biblioteca na *Web* inclui a presença dos usuários. Há tanto formas síncronas (e.g. MI) e assíncronas (e.g. wikis) para os usuários se comunicarem entre si e com os bibliotecários.
- 4) **É comunitariamente inovadora.** Este é talvez o aspecto mais importante e singular da Biblioteca 2.0. Baseia-se no fundamento das bibliotecas como serviço comunitário, mas entende que as comunidades mudam, e as bibliotecas não devem apenas mudar com elas, elas devem permitir que os usuários mudem a biblioteca. Ela busca continuamente mudar seus serviços, achar novas formas de permitir que as comunidades, não somente indivíduos, busquem, achem e utilizem informação.

Ainda segundo Maness (2007), Biblioteca 2.0 é uma comunidade virtual centrada no usuário. Ela é socialmente rica, quase sempre um espaço eletrônico igualitário. Enquanto o Bibliotecário 2.0 deveria atuar como um facilitador e prover suporte, ele ou ela não é necessariamente o primeiro responsável pela criação do conteúdo. Os usuários interagem e criam recursos (conteúdo) uns com os outros e com os bibliotecários. De várias formas, é uma realidade virtual para bibliotecas, uma manifestação *Web* da biblioteca como lugar. Uma presença da biblioteca na *Web* em Biblioteca 2.0 inclui a presença da constituição daquela biblioteca e utiliza tanto as mesmas aplicações e tecnologias como sua comunidade.

O conceito de que a biblioteca deve mudar com o usuário e este deve poder modificar a biblioteca é um conceito essencialmente novo para os profissionais da área, que estão habituados a atuar em uma via de mão única.

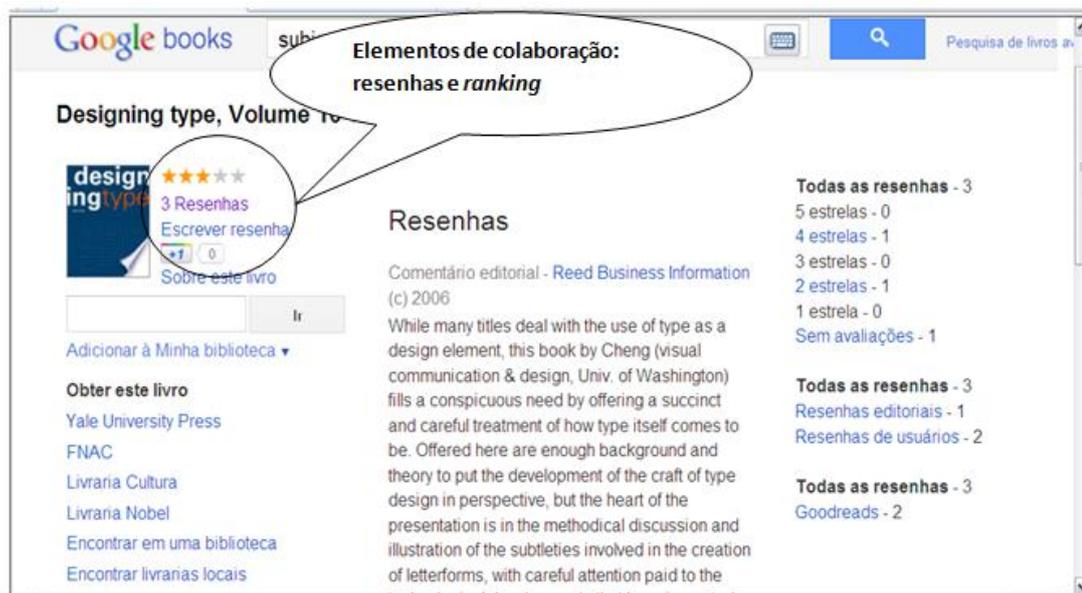
Uma vez na *Web*, não é mais possível querer ditar as regras. Trata-se de uma questão de sobrevivência e de respeito, poder ouvir o usuário e ir transformando-se com ele. O sistema é dinâmico e a transformação agora é contínua.

Em uma época quando não se conhece mais o usuário, pois ele migrou do espaço físico para o virtual, a interação é praticamente zero, com exceção de algumas possibilidades permitidas por ferramentas de comunicação, há de se criar novos indicadores e monitorar o acesso à biblioteca para que sua missão seja sustentada por parâmetros pré-definidos e aceitáveis, de acordo com as possibilidades que a tecnologia nos oferece.

Se já é um desafio tratar e disponibilizar a informação em um ambiente virtual com base nos elementos vistos no capítulo anterior, torna-se um desafio maior ainda incorporar os elementos da Web 2.0, permitindo que o usuário intervenha em seu processo de criação comentando e rotulando registros, por exemplo.

O Google é um exemplo de como será o futuro das bibliotecas digitais. Ao criar o projeto Google Books, ele introduz os elementos de colaboração e enriquece o registro, conforme demonstrado na tela a seguir.

Figura 24 – Google Books⁴



Fonte: GOOGLE BOOKS, 2011.

Com as tecnologias que estão disponíveis nos dias de hoje, falta às bibliotecas digitais evoluir e ultrapassar a barreira do tratamento da informação, para atingir a excelência na recuperação da informação.

⁴ Disponível em: <<http://books.google.com.br/bkshp?hl=pt-BR&tab=wp>>.

3.2 As bibliotecas nas redes sociais

As redes sociais, segundo Marteleto (2001, p. 72), representam “[...] um conjunto de participantes autônomos, unindo idéias e recursos em torno de valores e interesses compartilhados”. A autora ressalta, ainda, que só nas últimas décadas o trabalho pessoal em redes de conexões passou a ser percebido como um instrumento organizacional, apesar de o envolvimento das pessoas em redes existir desde a história da humanidade.

As bibliotecas estão descobrindo as redes sociais. A cada dia veem-se novos *blogs*, páginas no Facebook, YouTube, novos endereços no Twitter e agora no Google+.

As redes sociais criaram para as bibliotecas a possibilidade do *marketing* independente, autônomo, desvinculando suas demandas tecnológicas da área de Informática e permitindo que os próprios profissionais bibliotecários criassem seus canais de comunicação e divulgação, oferecendo novos produtos e serviços. Os usuários que haviam se afastado das bibliotecas físicas por conta da possibilidade de acesso aos catálogos *on-line*, voltaram a interagir com as bibliotecas a partir de comentários, *chat*, Skype e tantos outros meios de comunicação que estão disponíveis no ciberespaço.

São espaços importantíssimos de colaboração e interação, que complementam os serviços prestados na disponibilização de seus catálogos.

Não são apenas as bibliotecas digitais que estão aderindo às redes sociais. Esses espaços virtuais vêm sendo considerados uma alternativa para bibliotecas ainda sem recursos e com pouca infraestrutura, como uma forma de estabelecer um canal de comunicação e divulgação entre seus usuários.

A seguir, um exemplo de como as bibliotecas digitais estão utilizando as redes sociais de forma integrada.

Ao mesmo tempo em que a Biblioteca Florestan Fernandes da FFLCH – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo –

USP, criou um *blog*, destacou um *plug-in*⁵ para o Twitter e disponibilizou o *feed*⁶ para ser seguida no agregador de notícias.

Figura 25 – Página principal do blog da FFLCH⁷



Fonte: BLOG DA BIBLIOTECA FLORESTAN FERNANDES, 2011.

Da mesma forma que o usuário colabora e participa de redes sociais ele deveria participar da criação da inteligência das bibliotecas digitais, comentando, rotulando, criando a sua própria biblioteca, marcando seus favoritos, votando nos registros mais significativos. Esse aspecto será explorado no capítulo 6.

O fato de as bibliotecas digitais estarem mais presentes nas redes sociais, mas não envolvidas em desenvolver funcionalidades 2.0 em seus catálogos on-line, leva a uma possível conclusão de que existe uma preocupação com a socialização e compartilhamento do conhecimento, mas não existe ainda um amadurecimento voltado para a construção da inteligência coletiva de seus catálogos na Web, a partir de recursos da Web 2.0, conforme será visto no capítulo 5.

⁵ É um programa de computador usado para adicionar funções a outros programas maiores, provendo alguma funcionalidade especial ou muito específica. Geralmente pequeno e leve, é usado somente sob demanda. (WIKIPÉDIA, 2011a).

⁶ É um formato de dados usado em formas de comunicação com conteúdo atualizado frequentemente, como sites (sítios) de notícias ou blogs. (WIKIPÉDIA, 2011b).

⁷ Disponível em: <<http://bibliofflch.wordpress.com/2011/04/>>.

4 OS PERFIS COGNITIVOS DO USUÁRIO NA WEB

A internet, a Web 1.0 e a Web 2.0 foram transformando os perfis cognitivos dos usuários. O hipertexto, a linguagem, a quantidade de informação e a busca são alguns dos elementos que permitiram a descoberta de novas habilidades, novos raciocínios, novos hábitos e outras necessidades informacionais.

Pode-se comparar o fenômeno Web 2.0 ao celular. Ambos criaram necessidades que não existiam. Não se tinha necessidade de comunicação como se tem nos dias de hoje e nem de compartilhamento da informação e/ou conhecimento como se pratica hoje.

As redes sociais são outro fenômeno que surgiram criando novas necessidades. Novos vínculos afetivos e novas formas de linguagem surgem, e códigos vão sendo criados para tornar cada vez mais rápida a comunicação.

A evolução das Tecnologias da Informação e Comunicação superaram as tecnologias de *software* para disponibilizar conteúdos em bibliotecas digitais; e com isso, o usuário que estava habituado a apoiar-se em uma biblioteca para suas pesquisas, migrou para o ciberespaço, onde as possibilidades de descobrir novos conhecimentos são praticamente ilimitadas.

No caso das bibliotecas digitais, observa-se uma quantidade expressiva de conteúdo digital disponível, mas seus catálogos continuam herméticos e acessíveis por intermédio de uma linguagem que o usuário não domina e a interface de busca, em geral, não oferece possibilidades de novas descobertas, conforme mencionado no item 2.4.

Um dos principais desafios de profissionais de bibliotecas digitais, além de construir interfaces de busca inteligentes, consiste em educar o usuário da *Web*, ajudando-o a adquirir novas competências informacionais.

Segundo Dudziak (2003), a verdadeira mediação educacional ocorre quando o bibliotecário convence o aprendiz de sua própria competência, inculcando-lhe autoconfiança para continuar o aprendizado, transformando-o em um aprendiz autônomo e independente. A natureza da experiência do aprendizado mediado no âmbito da biblioteca pode variar de acordo com cada situação e o grau de interação humana.

O fato de o usuário estar diante de uma máquina não transfere, e sim aumenta a responsabilidade do bibliotecário em tornar a interface de busca o mais amigável possível, considerando-se os diversos perfis de usuários/educandos. As competências informacionais devem ser adquiridas durante o processo de busca, na comunicação entre o usuário e a máquina.

4.1 A Linguagem hipermidiática

Vannevar Bush, já em 1945, publicou na revista americana *Atlantic Monthly*, o artigo “As we may think”, em que descreve uma máquina, chamada Memex, capaz de propiciar leitura e escrita não lineares e armazenar uma biblioteca multimídia de documentos. Segundo Bush (1945), a mente humana não funciona de forma linear, e sim por associação. Sem poder imaginar, Bush já antevia o hipertexto.

Em 1965, Theodore Nelson cunhou o termo *hipertexto* em seu projeto Xanadu²⁷, cuja proposta era implementar uma rede de publicações eletrônicas, instantânea e universal – um verdadeiro sistema hipertexto, um universo documental. O termo *hipertexto*, no conceito de Nelson, estava relacionado à ideia de leitura/escrita não linear em sistemas informatizados, reforçando a visão de Bush em 1945.

Durante suas pesquisas, Nelson descreveu muitas das ideias implantadas nos sistemas hipertextuais atuais.

Embora o conceito de hipertexto tenha sido esboçado com o projeto Memex de Vannevar Bush, foi apenas na década de 90 que se desenvolveu a linguagem HTML²⁸ e o protocolo de comunicação HTTP^{*29}, os quais possibilitaram a produção e a disseminação de documentos hipertextos pela rede mundial de computadores – a Internet.

²⁷ Theodor Holm Nelson, idealizador do hipertexto, propôs a criação de uma biblioteca eletrônica (World Publishing Repository _ Project Xanadu), onde seriam armazenados os textos integrais de documentos. Xanadu seria um " lugar mágico da memória literária onde nada seria esquecido". (BORGES, 2003).

²⁸ HTML – Abreviatura de HyperText Markup Language – linguagem composta por um conjunto de comandos de formatação e utilizada na criação de documentos.

²⁹ Tradução de *Hypertext Transfer Protocol*, faz a comunicação entre o cliente e o servidor por meio de mensagem.

Com o surgimento da WWW³⁰, o usuário migra definitivamente para o *ciberespaço*, termo cunhado pelo escritor americano William Gibson no início da década de 1980, para conceituar o ambiente virtual criado pelas redes de computadores.

Para Santaella (2007a, p. 156), o espaço é uma extensão tridimensional, sem fronteiras, na qual objetos e eventos ocorrem e têm posição e direção relativas, desde a explosão do universo digital que trouxe consigo a emergência do ciberespaço.

Para a autora, o conceito de espaço vem se tornando uma moeda cada vez mais corrente, especialmente a partir do desenvolvimento da tecnologia móvel, que traz em si um conceito amplo de espaço e desterritorialidade, conceito este representado hoje por todas as tecnologias móveis que nos permitem acessar a Internet de qualquer lugar do planeta.

Uma definição que se assemelha ao que se pretende apresentar é a definição de Ciberespaço dada por Lévy (1999, p. 17):

O ciberespaço é o novo meio de comunicação que surge da interconexão mundial de computadores. O termo especifica não apenas a infraestrutura material de comunicação digital, mas também o universo oceânico de informações que ele abriga, assim como os seres humanos que navegam e alimentam este serviço.

Leão (2004) ressalta seis características do ciberespaço, identificadas por Cloninger:

- a) constitui-se como uma rede de telecomunicações de vários para vários;
- b) possui caráter multimidiático;
- c) engloba o conceito de banco de dados;
- d) está sujeita a automação e programabilidade;
- e) produz uma mudança no paradigma do tempo e no conceito de “ao vivo”;
- f) pode ser acessível por equipamentos móveis.

A partir desta definição, inserem-se os elementos fundamentais deste estudo: o internauta e seus perfis cognitivos, suas linguagens, o universo de informações, seus fluxos informacionais e como é possível encontrar informação nesse meio. A esses elementos acrescenta-se a Web 2.0, ambiente propício à colaboração e ao desenvolvimento da inteligência coletiva. Portanto, o ciberespaço passa a ser um

³⁰ Sigla de World Wide Web, a Web é um Sistema hipertextual que funciona sobre a Internet.

ambiente de troca, interação, exposição e muita colaboração, que estimula o coletivo inteligente, pois é ausente de controle centralizado e é o ambiente onde todos os elementos desenvolvidos neste trabalho se encontram. A informação é o que todos buscam na *Web*, seja em *sites*, em Portais, em bibliotecas digitais ou tantos outros repositórios existentes neste ambiente.

Assim, surge a linguagem hipermidiática, que, segundo Santaella, é

[...] uma linguagem em que convergem o texto escrito (livros, periódicos científicos, jornais, revistas), o audiovisual (televisão, vídeo, cinema) e informática (computadores e programas informáticos). Aliada às telecomunicações (telefones, satélites, cabo) das redes eletrônicas e a tecnologia das informações digital conduziu a disseminações da internet que resultou da associação de dois conceitos básicos, o de servidores de informações com o hipertexto. (SANTAELLA, 2005).

Trata-se, de fato, de uma nova linguagem, em um novo tipo de meio ou ambiente de informação, no qual ler, perceber, escrever, pensar e sentir adquirem características inéditas.

Não se trata apenas de transferir o universo de conteúdo textual para o ambiente hipertextual. Significa reconstruir significados e criar nós entre conteúdos, com o objetivo de agregar conhecimento ao conteúdo pré-existente, criando o efeito hipertextual.

Sem sombra de dúvida trata-se de um novo paradigma. Os textos deixam de ser lineares e passam a representar relações entre si. Essa característica nodal, que permite desconstruir e reconstruir outro texto a partir do original, só é possível pelo surgimento de protocolo de comunicação, que permite que cada página tenha um endereço como forma de construção de nós no ciberespaço.

Essa nova forma de pensar, de escrever, representar ideias e conhecimentos não acompanhou a evolução das bibliotecas, que, por mais digitais que tenham se tornado, ainda preservam uma linguagem textual, priorizando a forma linear de busca e navegação. Por mais multimídia que elas sejam em seu conteúdo, o conceito hipermidiático se aplica muito timidamente no momento de encontrar uma informação.

Segundo Santaella (2004, p. 38),

as linguagens do ciberespaço são linguagens hipermidiáticas. Inteirar-se da natureza palinódica da semiose na hipermídia, que se expressa na sua estrutura reticular, nodal, constitui porta de entrada para o conhecimento do ciberespaço.

Todas as manifestações de linguagem que surgem no ciberespaço fazem parte de uma cibercultura.

Por necessidade de estabelecer um senso comum e buscar um padrão que facilite reunir o conhecimento, as bibliotecas valem-se da utilização de uma linguagem documentária para representar o conhecimento. Esse recurso acaba limitando a biblioteca a uma linguagem própria, destituída de significado para o usuário. Um exemplo claro dessa problemática ocorreu em uma biblioteca infanto-juvenil, quando um aluno, ao fazer uma pesquisa sobre clonagem, tentou localizar informação a partir do nome da ovelha Dolly, não obtendo sucesso em sua busca.

Essa distância entre o tratamento da informação e a linguagem do usuário, só tendem a diminuir o uso das bibliotecas como fonte de pesquisa na *Web*.

A linguagem sempre exerceu papel fundamental no processo comunicacional e na transação de conhecimento, seja por meio do sujeito contemplativo, seja por meio do sujeito imersivo, cada qual com suas características. Lévy (2003) ressalta que

Assim que penetramos no universo da *Web*, descobrimos que ele constitui não apenas um imenso “território” em expansão acelerada, mas que também oferece inúmeros “mapas”, filtros, seleções para ajudar o navegante a orientar-se” (LÉVY, 2003, p. 85).

Além da linguagem, a Internet oferece diversos gêneros textuais distribuídos em distintos ambientes virtuais, como o ambiente *Web*, *chat*, *e-mail*, fórum de discussão, jogos, criação personagens, áudio e vídeo (videoconferências), além de outros. Desta forma, os gêneros surgem dentro de ambientes como locais que permitem culturas variadas e estão sujeitos a inúmeras linguagens.

Se você tem um parente em idade adolescente, é muito fácil observar essa característica. Procure ler uma mensagem em seu Messenger (MSN) ou um conversa qualquer por torpedo (SMS) entre amigos. São linguagens quase indecifráveis e estabelecidas entre eles como uma espécie de códigos.

Transportando esse exemplo para o ambiente de uma biblioteca digital, imagine tentar representar todas as maneiras pelas quais o usuário irá fazer uma busca? Esse desafio será abordado e aprofundado no item referente a interfaces de busca inteligentes.

Outro conceito importante nos aspectos da linguagem é a Interatividade. O saber se constrói pela interação entre seus participantes. Uma biblioteca digital é um ambiente vivo que pode alimentar-se do saber de seus usuários, bastando para isso se abrir para a interatividade. A interatividade ultrapassa barreiras e atrai usuários desejosos pelos mesmos conhecimentos. A interatividade passa a ser quase uma linguagem, pois se sobrepõe aos limites cognitivos e se compreende a partir de um interesse comum.

Santaella (2004) define a interatividade na comunicação mediada por computador como sendo:

Nesse novo contexto, o emissor não emite mais mensagens, mas constrói um sistema com rotas de navegação e conexões. A mensagem passa a ser um programa interativo que se define pela maneira como é consultado, de modo que a mensagem se modifica na medida em que atende às solicitações daquele que manipula o programa. (SANTAELLA, 2004, p. 163).

Assim deveriam ser os ambientes virtuais. Mas muitos ainda estão distantes de permitir essa desconstrução e reconstrução.

4.2 Os usuários no papel de internautas

Dentre os vários desafios em transportar um acervo de biblioteca para o ciberespaço, está a dificuldade de acompanhar o pensamento e a lógica do usuário.

Quando tínhamos apenas bibliotecas físicas, os chamados usuários, na época, abordavam um bibliotecário em busca de uma informação; existia uma técnica, denominada “estratégia de busca”, que nos permitia entender o que o usuário buscava e, assim, o ajudávamos a encontrar. O contato físico, a comunicação e a convivência, também eram elementos importantes para se conhecer cada usuário.

Constata-se que a forma linear da leitura de livros e periódicos que marcou o acesso ao saber até o início do século XXI já não se assemelha ao acesso labiríntico e hipertextual do conhecimento, necessário nos dias de hoje para produzir significados, como vimos no capítulo acima.

Alguns autores, no entanto, sugerem que podemos classificar os internautas de alguma maneira, e isso pode ajudar a projetar melhores interfaces e soluções para as bibliotecas que estão migrando para o ciberespaço ou sendo criadas nele.

Esta problemática tem sido objeto das reflexões de muitos teóricos, tais como Lucia Santaella (2007b), que, ao investigar sobre o perfil cognitivo das novas gerações, estabeleceu uma análise progressiva das transformações intelectuais experimentadas pela nossa espécie a partir do contato com as tecnologias desde a Idade Média.

Em relação ao período medieval, Santaella (2004) apresenta a existência de um leitor “contemplativo” de objetos e signos duráveis, imóveis e manuseáveis, tais como os livros, as pinturas, os mapas. Com base na autora, “esse leitor não sofre, não é açoitado pelas urgências do tempo” (SANTAELLA, 2004, p. 24), tem tempo para meditar, para a contemplação, tendo o sentido da visão se sobressaindo aos demais e a imaginação como seu complemento.

Com o advento da Idade Moderna e o processo de urbanização desencadeado, principalmente a partir do século XIX, profundas mudanças foram se operando no modo de vida das pessoas: a velocidade não mais se pautava pela tração animal, mas sim pelo ritmo das locomotivas e, logo em seguida, dos automóveis. Neste período, “viver passou a significar adaptar-se à congestão de imagens na retina” (SANTAELLA, 2004, p. 29).

Em seu livro, *Navegar no ciberespaço*, a autora apresenta o leitor movente, “aquele que nasce com o advento do jornal e das multidões nos centros urbanos habitados de signos [...], o leitor apressado de linguagens efêmeras, híbridas, misturadas.” (SANTAELLA, 2004, p. 29).

Para Santaella, o jornal é o primeiro rival do livro e se configura numa linguagem híbrida a partir da impressão mecânica aliada ao telégrafo e à fotografia, possuindo um caráter efêmero, feito para durar o curto tempo das suas notícias, criando assim um leitor “fugaz, novidadeiro, de memória curta, mas ágil. Um leitor que precisa esquecer, pelo excesso de estímulos, e na falta do tempo para retê-los” (SANTAELLA, 2004, p. 29).

No entanto, esse leitor movente, tendo sofrido tão significativas mudanças cognitivas, ainda se configura como muito distante do perfil cognitivo do leitor “imersivo” que surge com o advento da informática e “que navega entre nós e conexões alineares pelas arquiteturas líquidas dos espaços virtuais” (SANTAELLA, 2007b, p. 31).

O leitor imersivo, em razão de suas experiências no ciberespaço, desenvolve outras sensibilidades, percepções, enfim, um novo tipo de cognição que o torna “um leitor em estado de prontidão, conectando-se entre nós e nexos, num roteiro multilinear e labiríntico que ele próprio ajudou a construir” (SANTAELLA, 2007b, p. 33). Assim, pode-se constatar o protagonismo desse novo sujeito, que não só deixou de ser um simples receptor passivo, mas tem liberdade de escolha e de produção dos conteúdos que deseja acessar.

É também um leitor marcado pela interatividade, aprendendo e ensinando velozmente por meio da interação constante que, também, implica em constante inovação. Acostumado com o estilo Google de buscar, esse leitor não quer tentar e tentar e confunde buscar com encontrar. Quando o resultado lhe aparece na casa dos dois zeros ou mais, ele se contenta com a busca e aceita os resultados como sendo os melhores possíveis. Esse comportamento oferece uma ameaça às bibliotecas digitais, que necessitam se valer de vocabulários controlados e muitos padrões para gerir seu acervo digital, pois constantemente os resultados de busca são zero ou próximos de zero, pois o leitor desconhece o conteúdo oferecido, e o ponto forte de bibliotecas são os catálogos padronizados e os vocabulários controlados e não as tecnologias e interfaces de busca.

Aprofundando o perfil cognitivo do leitor imersivo, a autora observa três níveis de acesso às tecnologias digitais, demonstrando diferenças na maneira de navegação partindo do “internauta errante”, passando pelo “detetive” e tornando-se um “navegador previdente”. (SANTAELLA, 2004).

É esse “leitor imersivo” com suas sutilezas que nos interessa explorar para propor uma maior eficácia no desenvolvimento das interfaces de busca. Não é possível criar uma maneira única de buscar se existe mais de um perfil de leitor a navegar.

O internauta errante vale-se do instinto para adivinhar as rotas que deve seguir rumo aos objetivos que deseja atingir. Tal comportamento garante ao internauta errante um componente marcadamente lúdico no contato com o

ciberespaço. Ele navega em territórios ainda desconhecidos, portanto, não pode se guiar pelo suporte da memória, recebendo assim, uma surpresa a cada movimento.

Podemos deduzir desse perfil, que uma interface complexa, com busca avançada e muitas escolhas para parametrizar sua busca, não será de grande efeito. Esse é o usuário que prefere o Google, onde ele digita uma palavra e passa a clicar aleatoriamente pelos inúmeros resultados que surgem, descobrindo um mundo novo a cada clique.

O internauta detetive conta com recursos de sua memória, podendo seguir “as trilhas dos índices de que os ambientes hipermidiáticos estão povoados” (SANTAELLA, 2007b, p. 178), conseguindo desenvolver estratégias de busca baseadas em seus erros e acertos. Seu percurso é caracterizado como um processo auto-organizativo, particular daquele que aprende com a experiência.

Esse é um perfil que se frustra com caminhos mal formulados, que o impedem de estabelecer uma lógica para poder retornar ao ponto de partida. Estabelecendo uma metáfora entre a História de João e Maria, se esse usuário não consegue retornar às suas migalhas de pão e não reconhece a estratégia, ficando impedido de se organizar, será possivelmente um usuário crítico a essa interface e se perderá nessa imensa biblioteca.

O leitor internauta previdente passou por um processo de aprendizagem. Por isso, mostra familiaridade com os ambientes informacionais e movimenta-se usando a lógica da previsibilidade, podendo antecipar as consequências de suas escolhas.

Esse leitor só chegará a esse ponto caso se depare com uma interface madura e inteligente, conforme abordado no item 2.4.

Os tipos, errante, detetive e previdente, configuram também três níveis de leitura imersiva, os quais estão presentes e se alternam no cotidiano de aprendizagens do leitor.

Segundo a autora, o leitor imersivo não deve perder de vista a “sua posição de explorador, cúmplice e co-criador” do ciberespaço (SANTAELLA, 2007b, p. 180).

Vale ponderar que os tipos de leitores podem se fundir e um tipo pode vir a ser os três tipos dependendo do momento e do cenário, mas o que importa realmente é adquirir essa consciência de que o leitor internauta não é um só e pode se transformar dependendo de suas necessidades básicas de informação ou conexão.

O conceito de liquidez, abordado também por Lucia Santaella (2007a) é mais um aspecto importante ao se construir uma interface de busca. A *Web*, com suas redes sociais e a comunicação sendo estabelecida de inúmeras formas diferentes, impôs uma linguagem líquida, que se desfaz, não possui protocolos e permite seu entendimento a partir de inúmeros instrumentos. Às vezes uma palavra, às vezes um *emoticom* e às vezes um ícone. Qualquer manifestação escrita torna-se uma linguagem e fica difícil querer impor uma maneira de se buscar a informação.

Essas transformações, provocadas por um novo ambiente de interação que estimulou a existência de uma linguagem líquida e é dinamicamente transformador, consistem em um desafio para o desenvolvimento de novas interfaces de busca em bibliotecas digitais. Como criar uma solução única baseada em tantas possibilidades de perfis de usuários é a questão que se coloca. Quais são os elementos essenciais para promover o encontro da informação em uma biblioteca digital? Essa questão será abordada no capítulo 6.

4.3 Os usuários de bibliotecas digitais e a mudança de paradigma

No universo da Biblioteconomia, convencionou-se chamar o solicitante, de usuário da biblioteca. A literatura relacionada aos processos de atendimento do bibliotecário faz referência ao usuário.

A partir da década de 90, as bibliotecas começam a romper esse paradigma e começam a perceber seus usuários como clientes, incorporando em si mesmas os conceitos da qualidade, chegando inclusive a desejar e obter a certificação ISO 9000, como conseguiu, por exemplo, a Biblioteca da Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo - ESALQ/USP de Piracicaba.

Com a chegada da Internet e da *Web*, ocorre mais uma ruptura de paradigma, pois o usuário que transformou-se em cliente migrou para o ambiente *Web* e passou a acessar a biblioteca através de um computador. O que mudou? Mudou a forma de relacionamento, dificultou o trabalho do profissional bibliotecário, pois esse deixou de ter instrumentos para levantamento de perfil de usuário e criou a figura do internauta.

A *Web* era muito recente e o Google ainda não existia. Como auxiliar o usuário a encontrar a informação? O único instrumento disponível passou a ser a interface de busca. A interface, como sua própria definição estabelece, passa a ser o canal, o instrumento, a linguagem e a única possibilidade de o usuário chegar até a informação.

Esse novo leitor marca o último paradigma até o momento vivido pelas bibliotecas digitais em relação ao seu usuário. Além de estar na *Web*, ele é muito mais informado, questionador e busca aprofundar suas questões, não se contentando mais com visões e soluções lineares. Há muita necessidade de criar textos e registros “hiperlinkados” para permitir que o usuário tenha experiências semelhantes às aquelas encontradas a partir do Google.

Um recurso desconhecido e quase nunca adotado é a análise dos *logs* de busca para acompanhar, analisar e incorporar termos que são buscados, mas não encontrados. Esse relatório deve ser extraído do administrador do sistema, que é o profissional responsável pela segurança e manutenção do ambiente e pode ajudar na manutenção e evolução de um vocabulário controlado.

5 ESTUDOS DE CASOS DE INTERFACES DE BUSCA EM BIBLIOTECAS DIGITAIS

Para apontar os elementos essenciais necessários ao encontro da informação, escolhi dois exemplos que revelam as soluções encontradas por suas bibliotecas digitais no desenvolvimento de suas interfaces de busca.

Na análise dos casos, foram consideradas as práticas relacionadas ao tratamento da informação na entrada de dados e os recursos oferecidos na interface de busca.

Os critérios utilizados na seleção dos casos levaram em consideração:

- a) o perfil dos serviços prestados;
- b) a comunidade alcançada;
- c) a utilização de tecnologias, considerando-se os padrões de bibliotecas digitais na *Web*.

A metodologia adotada neste estudo de caso foi a pesquisa qualitativa não comparativa, de duas bibliotecas digitais, tomando como base as heurísticas de Nielsen e Molich (apud SANTOS, 2011), mencionadas no item 2.4, adaptadas e aplicadas ao universo das bibliotecas digitais, com base nos seguintes elementos:

- a) Metadados** – Tratamento da informação;
- b) Comunicação** – Como os elementos da interface se comunicam conduzindo o internauta no processo da busca;
- c) Visão de conjunto** – Existe um taxonomia ou recursos de navegação? Como é possível conhecer o universo informacional no ambiente digital?
- d) Busca simples** – Qual o processo na condução da busca?
- e) Busca avançada** – A utilização de conectores é explorada? Existem exemplos da sintaxe a ser utilizada? É possível combinar ou selecionar campos na busca? É possível refinar a busca?
- f) Mecanismos de ajuda** – Existem recursos que conduzem o usuário no processo de busca?
- g) Representação de resultados de busca** – Os resultados aparecem em uma única ordem ou oferecem possibilidade de ordenação de registros? Os metadados representados no resultado de busca são suficientes para que o usuário selecione a informação encontrada?

h) Funcionalidades da Web 2.0 – Quais são as funcionalidades da Web 2.0 utilizadas pela biblioteca digital? Qual o grau de colaboração do ambiente digital?

Com esses elementos definidos, passo à análise dos seguintes casos:

- a) SciELO – Biblioteca da Scientific Electronic Library Online – www.scielo.br
- b) Sapiencia – Biblioteca Digital de Teses e Dissertação da PUC-SP – www.sapiencia.pucsp.br

5.1 SciELO - Scientific Electronic Library Online - www.scielo.br

De acordo com o site oficial, SciELO é uma biblioteca digital de revistas científicas brasileiras em formato eletrônico, que organiza e publica textos completos de revistas na *Web*, assim como produz e publica indicadores de seu uso e impacto. (SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011).

O projeto iniciado em meados de 1997 é resultado de uma parceria entre a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), o Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (Bireme) e editores de revistas científicas; durante o seu desenvolvimento, recebeu o nome de Biblioteca Científica Eletrônica *On-line*, cuja sigla SciELO corresponde à sua versão em inglês.

Por ser uma biblioteca digital que reúne a produção científica selecionada de periódicos ibero-americanos, seus indicadores refletem não apenas a produção intelectual desses países, mas também são altamente relevantes para a avaliação de pesquisa, desenvolvimento e educação no Brasil.

O conjunto de acervo no Brasil abrange 261 títulos de periódicos.

A metodologia utilizada no desenvolvimento da biblioteca é da própria SciELO. Ainda conforme o *site*, a metodologia SciELO é um conjunto de normas, guias, manuais, programas de computador e procedimentos operacionais dirigidos à preparação de textos de periódicos científicos em formato eletrônico, incluindo, entre outras, as seguintes funções: armazenamento de textos estruturados em bases de dados, publicação dos periódicos na Internet ou em outros meios, recuperação de artigos e outros textos por seu conteúdo, produção regular de relatórios de uso e

indicadores bibliométricos, aprimoramento de critérios para a avaliação da qualidade de periódicos e o desenvolvimento de procedimentos e políticas para a preservação de publicações eletrônicas. (SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011).

Seu público alvo são pesquisadores, docentes e demais profissionais da área acadêmica e científica.

5.1.1 A interface de busca e seus elementos

Apresento a seguir a página principal da interface de busca da SciELO envolvendo todos os países abrangidos pelo projeto.

Figura 26 – Tela principal da SciELO³¹

The screenshot displays the SciELO main interface with the following sections:

- Sobre SciELO**: Includes links for 'Indicadores Bibliométricos', 'Acesse por OAI e RSS', and 'Portal SciELO.org antigo'.
- coleções**: Lists countries with flags: Argentina, Brasil, Chile, Colômbia, Costa Rica, Cuba, and Espanha.
- Lista periódicos**: Features a search box with the text 'Pesquisa periódicos' and 'Entre com uma ou mais palavras', and a 'pesquisar' button.
- Por ordem alfabética - todos**: Lists letters A through Z.
- Por assunto - todos**: Lists 'Ciências' and 'Agrárias'.
- SciELO em números**: Displays statistics: 867 Periódicos, 22.783 Fascículos, 335.871 Artigos, and 7.032.048 Citações.
- Twitter**: Includes the Twitter logo and the text 'SciELO no Tw Siga e infor'.
- Notícias**: Lists 'Fadiga e erro médico', 'Sistema de informação estatística', and 'Foto de índio'.
- Press Releases**: A section for news releases.

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011d.

A interface propicia acesso aos textos completos de artigos através de um índice de autor e um índice de assuntos, utilizando-se de um formulário de pesquisa de artigos, que busca os elementos que o compõem, tais como autor, palavras do título, assunto, palavras do texto e ano de publicação.

³¹ Disponível em: <<http://www.scielo.org/php/index.php?lang=pt>>.

A seguir analiso os aspectos positivos e negativos da interface de busca com base nos elementos mencionados no início do capítulo 5.

a) Metadados

Aspectos positivos: A biblioteca é transparente. Por possuir índices de seus principais pontos de acesso, é praticamente impossível obter-se um resultado zero de pesquisa, pois o fato de um termo estar contemplado no índice sinaliza que existe ao menos um registro a ser encontrado.

Aspectos negativos: Os índices são poluídos. Como o exemplo mostrado a seguir no item d), observa-se em todos os índices uma poluição de dados que mostra um descuido com o controle da qualidade na entrada da informação, podendo levar o usuário a um erro em sua estratégia de busca.

A seguir, um exemplo de tela mostrando o aspecto positivo e ao mesmo tempo negativo dos metadados na interface.

Figura 27 – Resultado de busca da SciELO pelo índice “Assunto” na letra M³²

The screenshot displays the SciELO search interface. At the top, there is a search bar with the SciELO Brasil logo on the left. The search bar contains the text 'alfa assunto pesquisa' and 'autor assunto pesquisa'. Below the search bar, the text 'Coleção da biblioteca' is displayed. The interface shows the search results for the 'Assunto' index under the letter 'M'. The search results are displayed in a list box with the following terms: MβL, M&E, M+A'AND D, M-1 AND M-2 NEURONS, M-ALKOXY PHENOLS, M-CHLOROPHENYLPIPERAZINE, M-CPBA OXIDATION, M-CRESOL, M-LEARNING, and M-METHOXY-P-HYDROXY-BENZOIC ACID. Below the list box, there are three buttons: 'próximo', 'adicionar', and 'pesquisa'. The interface also shows 'Base de dados : article' and 'Índice Assunto'.

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

³² Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Percebe-se claramente que existe um erro no critério de entrada ou seleção de termos. Por mais que sejam termos específicos de uma área de conhecimento, certamente um usuário não buscaria expressões como algumas dessas representadas no índice.

A iniciativa de índice é muito bem vinda, mas há que se tomar muito cuidado, pois essa é uma das funcionalidades que mais transparência dá aos catálogos. Se houver inconsistência nos índices o projeto passará uma insegurança ao usuário.

b) Comunicação

Aspectos positivos - Em sua tela principal, a SciELO dá demonstrações de tentativas de comunicação com usuários oferecendo opções de navegação e índices para consulta.

Figura 28 – Pesquisa na SciELO de artigo no índice “Autor”³³

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011b.

³³ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/?IsisScript=iah/iah.xis&base=article^dlibrary&index=AU&lang=p&fmt=iso.pft&form=B>>.

O site possui 3 versões de línguas: português, espanhol e inglês.

Aspectos negativos - A abreviação, a falta de explicação de termos, tais como *config*, as expressões *formulário básico* e *formulário livre*, além de termos soltos como *alfa*, não auxiliam o usuário na compreensão do que deve ser feito. São abreviaturas e expressões sem significado universal.

As notas com orientação de busca não oferecem exemplos de fácil compreensão.

c) Visão de conjunto

Aspectos positivos – Existe a informação referente ao universo a ser pesquisado.

Aspectos negativos – O único canal que informa o universo a ser pesquisado é a busca por navegação por periódico em ordem alfabética. Esse é o único momento em que é possível saber que existem 261 periódicos na rede SciELO no Brasil.

Falta uma mensagem que apresente o ambiente e permita conhecer seu conteúdo no primeiro contato com a biblioteca.

d) Busca simples

Aspectos positivos - Possui conectores para combinação de termos e permite uma configuração básica. Nessa opção de pesquisa, por autor e assunto, há a possibilidade de se guiar através de índice. Conforme tela já demonstrada no item b).

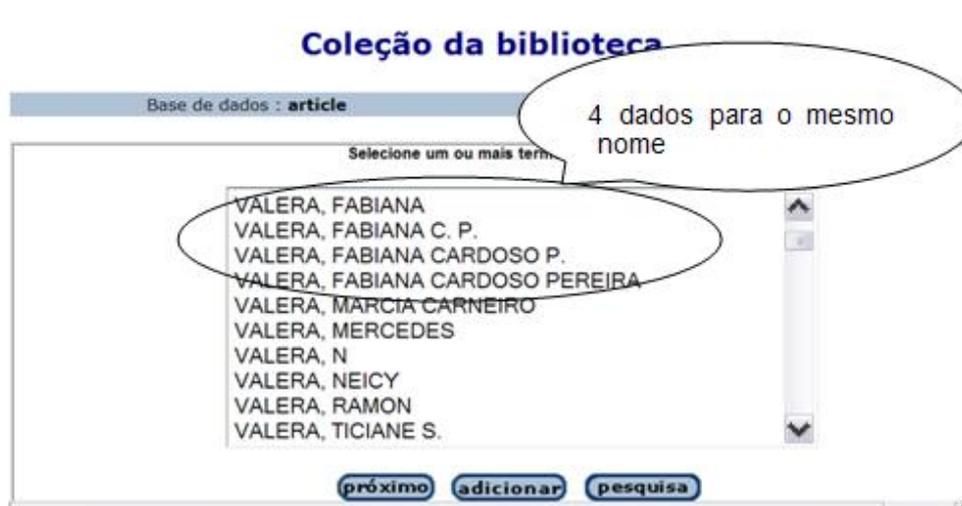
Esse recurso é um avanço em interfaces de busca, e é difícil de se encontrar exemplos com essas características na busca simples.

Aspectos negativos - A busca simples é feita a partir do *formulário básico*. Como é mais comum encontrarmos as expressões de interface de busca como busca simples e busca avançada, esse formulário não transmite sua função e confunde o usuário. É difícil perceber em qual opção de busca o usuário se encontra, pois a mudança de cores é quase imperceptível.

Os índices revelam a inexistência de critérios na entrada de dados, pois as origens são muitas e não há uma padronização de termos. Isso resulta em redundância de dados e poluição visual para o usuário, conforme pode ser visto na figura a seguir.

Ex.: Pesquisando pelo primeiro nome, Fabiana Valera, apareceriam as 4 opções de nome, ao invés de uma só.

Figura 29 – Exemplo na SciELO de redundância no índice autor³⁴



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

e) Busca avançada

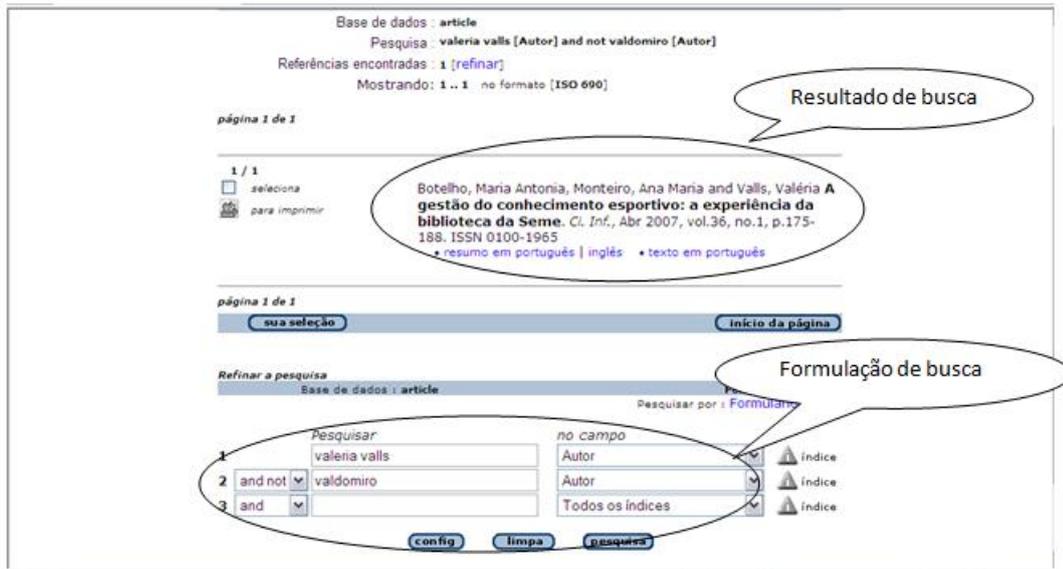
Aspectos positivos – Permite combinar até três termos utilizando-se de conectores e selecionando cada termo em todos os índices da base de dados. Existem filtros que permitem o refinamento da busca. Existe também uma quantidade significativa de campos a serem utilizados na busca, recurso pouco explorado por bibliotecas digitais.

Aspectos negativos – A busca avançada é feita a partir do *formulário livre*, e, como foi dito na busca simples, a barreira é o próprio índice que está presente para facilitar, mas poluído acaba prejudicando a busca.

No exemplo a seguir, selecionei o autor *Valéria Valls* e restringi o autor *Valdomiro*, pois sei que ambos escrevem juntos e não era um artigo com ambos que estava buscando.

³⁴ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Figura 30 – Busca avançada na SciELO a partir de formulário livre³⁵



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

A estratégia de busca é um recurso de pouco domínio do usuário e a interface não demonstra estar sensibilizada com essa questão. Os conectores estão em inglês e não há exemplos de como utilizá-los.

f) Mecanismos de ajuda

Aspectos positivos – Uma interface bastante intuitiva, com recursos amigáveis como o índice e notas, como pode ser visto na tela a seguir, contendo uma breve explicação de como construir uma estratégia de busca. As notas encontram-se no *formulário livre*.

³⁵ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Figura 31 – Orientação na SciELO para formulação de busca³⁶

Entre uma ou mais palavras

Todas as palavras **(AND)** Qualquer palavra **(OR)**

config **pesquisa**

Notas :

- Não digite operadores lógicos (AND, OR ou AND NOT) entre as palavras. Separe as palavras por espaço.
- Use o símbolo de truncagem \$(cifrão) para pesquisar palavras com mesma raiz. Exemplo: educ\$ recupera educación, education, educação, etc.
- Selecione a opção **Todas as palavras (AND)** para relacionar as palavras (restringe o escopo da pesquisa), ou a opção **Qualquer palavra (OR)** para somar as palavras (amplia o escopo da pesquisa).
- Para pesquisar em outros campos ou para especificar o campo de pesquisa use o [Formulário Básico](#).

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

Aspectos negativos – Na tela *formulário básico*, onde se dá a busca avançada, não há ajuda para compreender a utilização de conectores. A área de Notas não foi explorada.

g) Representação de resultados de busca

Aspectos positivos – A possibilidade de refinamento de busca, busca a partir da busca, é um recurso sofisticado e pouco presente em uma interface de busca.

³⁶ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Figura 32 – Resultado na SciELO de pesquisa com possibilidade de refinamento da busca³⁷

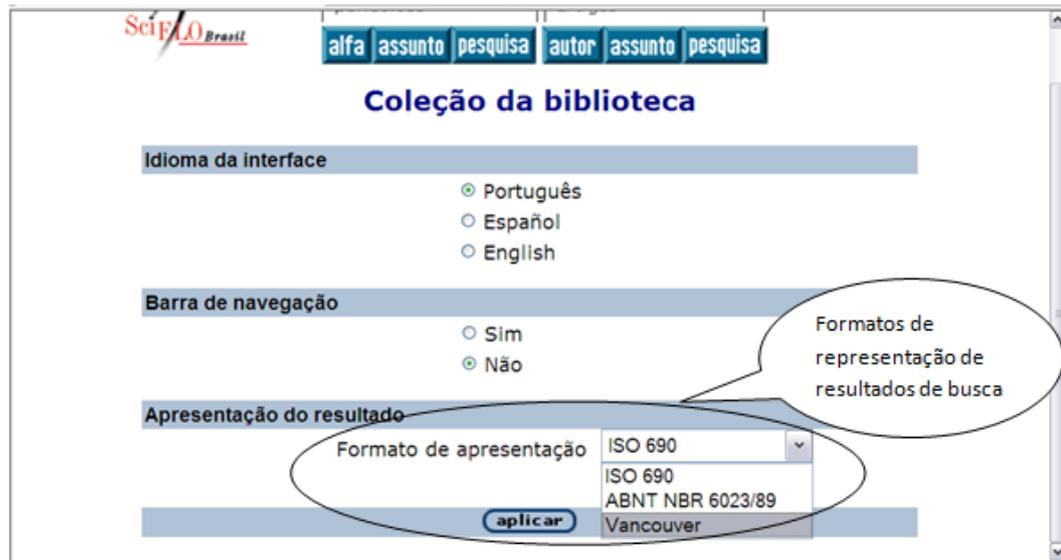


Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011c.

Outra funcionalidade positiva são os formatos de saída na forma de representação de resultados de busca. A tela a seguir mostra que há 3 formatos padrão para citar os dados de resultados. Vale destacar que esse também é um recurso pouco explorado e contribui para a citação do artigo em um trabalho acadêmico.

³⁷ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/#refine>>.

Figura 33 – Formatos de citação disponíveis³⁸



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

Aspectos negativos – Não é possível determinar o critério de ordenação de resultados. O formato facilita a identificação de metadados, mas não contribui na seleção de itens relevantes por não possuir critério de ordenação. Uma breve explicação sobre a origem das normas de referências bibliográficas também seria bem vindo.

h) Funcionalidades Web 2.0

Aspectos positivos – Os serviços *Meu Scielo* e *Serviços personalizados* são bastante sofisticados e remetem o usuário às principais redes sociais. A interação, nesse sentido, é ampla, permitindo também personalizar e criar sua coleção de biblioteca.

³⁸ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Figura 34 – Acesso às redes sociais a partir de login no ambiente da SciELO³⁹



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011f.

Aspectos negativos – Não foram observadas funcionalidades de comentários em registros e nem a possibilidade de inclusão de *tags*.

5.1.2 Considerações sobre a Biblioteca on-line da SciELO

Ao analisar o conjunto de elementos, e os aspectos positivos e negativos, observo que existem funcionalidades sofisticadas como o refinamento de busca e opções de formato de saída das citações, mas por outro lado o conjunto da análise dá a impressão de faltar um controle de qualidade na entrada de dados e mais mecanismos de ajuda no processo de busca.

A coleção não é hermética, pois possui índices e visão de conjunto, mas não dá para afirmar que sua interface é intuitiva, pois existe uma economia no uso de sinais símbolos, como abreviações de palavras não usuais.

Como os usuários são especializados e utilizam a biblioteca com freqüência, as barreiras que seriam intransponíveis para usuários leigos, tendem a ser

³⁹ Disponível em: <http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0103-18132010000200007&lng=pt&nrm=iso&tlng=pt>.

compreendidas e internalizadas por usuários assíduos, que assimilam e se apropriam de suas funcionalidades.

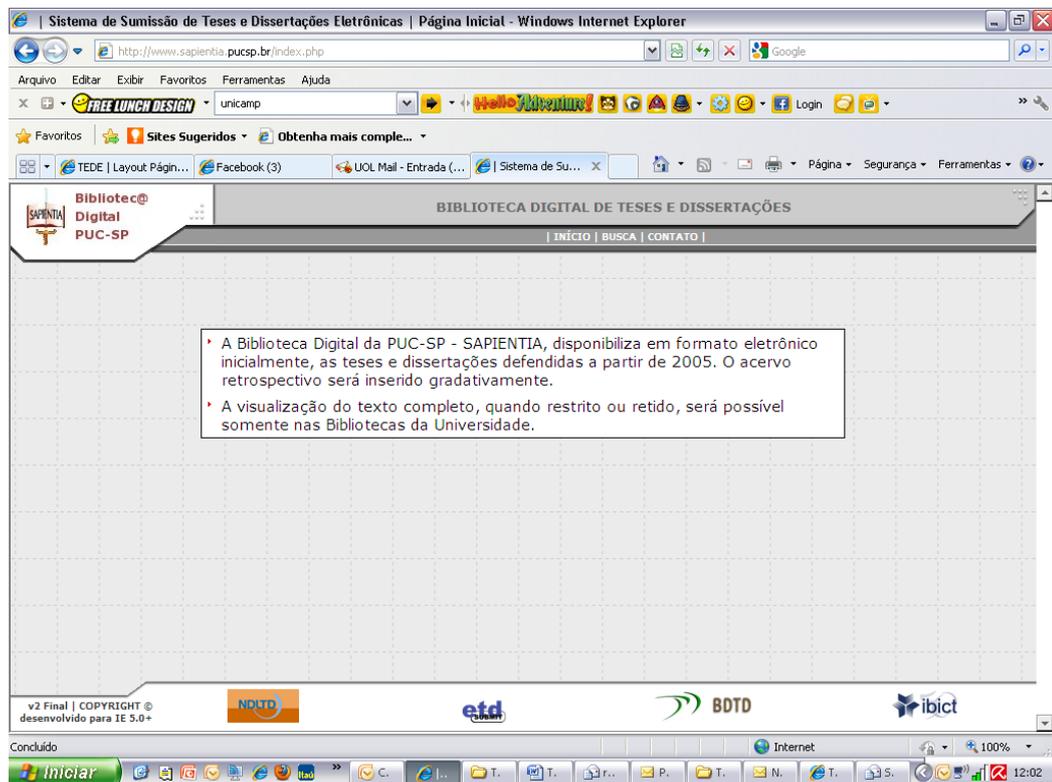
5.2 Sapientia – Biblioteca Digital de Teses e Dissertações – PUC-SP

A Biblioteca Digital da PUC-SP – SAPIENTIA disponibiliza em formato eletrônico as teses e dissertações defendidas a partir de 2005.

5.2.1 A interface de busca e seus elementos

Apresento a seguir a tela principal da Biblioteca Digital de Teses e Dissertações da PUC – SP.

Figura 35 – Página principal da BDTD Sapientia – PUC-SP



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

a) Metadados

Aspectos positivos – Pode-se observar, pelo resultado de busca, que há um vocabulário controlado multilíngue.

Aspectos Negativos – Não é possível analisar aspectos de normalização e qualidade da informação, pois o catálogo é hermético. Não há índices e nem informações sobre o uso eventual de formatos-padrão para os metadados.

Figura 36 – Resultado de busca da BDTD Sapientia⁴⁰

communication area, including language as reasoning, language as cognition, hypermedia language, 11 the associative mind and the digital influence in non-linear thinking are also used. Finally, the communication method RCPTIA proposed, presents a dynamic taxonomy that associates personal contexts, institutional contexts and technological contexts with any digital file or internet address. In this way, the method enables users to use their own cognitive systems integrated with document repositories indexing systems in order to retrieve information the same way that it is retrieved by their own mind

Titulação: Doutor em Comunicação e Semiótica: Signo e significação nas Mídias

Contribuidor (es): [Orientador] Sergio Bairon Blanco Sant'anna

Assunto(s): [pt] COMUNICACAO
 [pt] contexto, recuperação de informação, busca, cognição, hipermedia, redes
 [en] knowledge management
 [en] context
 [en] information retrieval
 [en] search
 [en] cognition
 [en] hypermedia
 [en] network
 [pt] Gestão do conhecimento
 [pt] Sistemas de recuperação da informação
 [pt] Redes de informação

Data da defesa: 17/10/2005

Arquivo(s): PDF - Sem legenda

Termos em português e inglês

Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

b) Comunicação

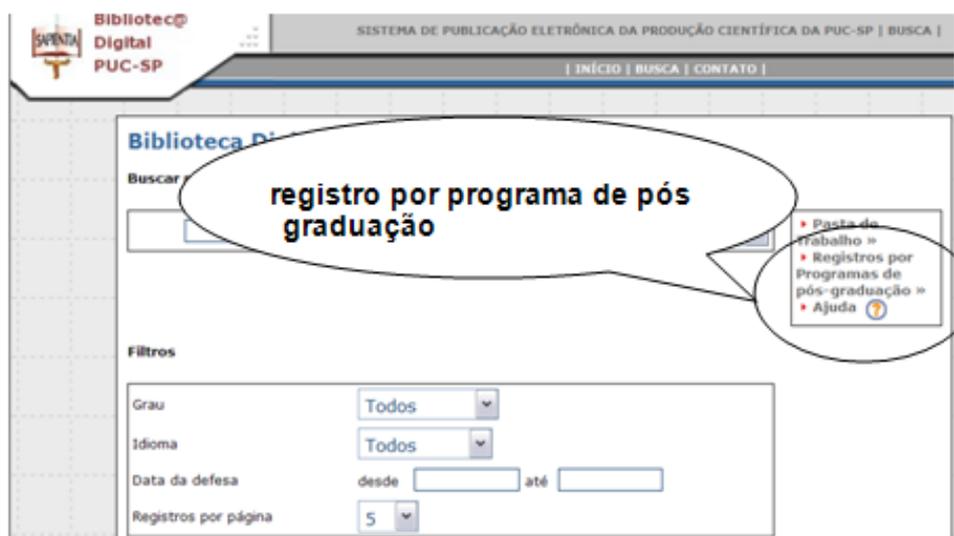
Aspectos positivos – A interface é simples, intuitiva e possui um canal para contato visível e acessível ao usuário.

Aspectos negativos – A opção de busca por programa, no lado direito da tela, poderia ser mais bem representada. Não é intuitivo perceber que há uma segunda opção de busca, conforme demonstrado na tela abaixo. A comunicação está mais voltada para dentro. Como a biblioteca encontra-se na Web, a utilização de usuários que não conhecem a estrutura da instituição depende de uma

⁴⁰ Disponível em: <www.sapientia.pucsp.br>.

comunicação mais eficaz. O uso de expressão como por exemplo, *Pasta de Trabalho*, não é intuitivo para o usuário externo e talvez também não seja para o usuário interno.

Figura 37 – Página Secundária da BDTD Sapiencia⁴¹



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

c) Visão de conjunto

Aspectos positivos – A Biblioteca informa na tela inicial que o conjunto de teses e dissertações, corresponde ao período posterior a 2005. Na tela de pesquisa por Programa de Pós graduação é possível ter uma visão de conjunto.

Aspectos negativos Como não há índices e nem exemplos, o conjunto representando ou os termos indexados são desconhecidos do usuário.

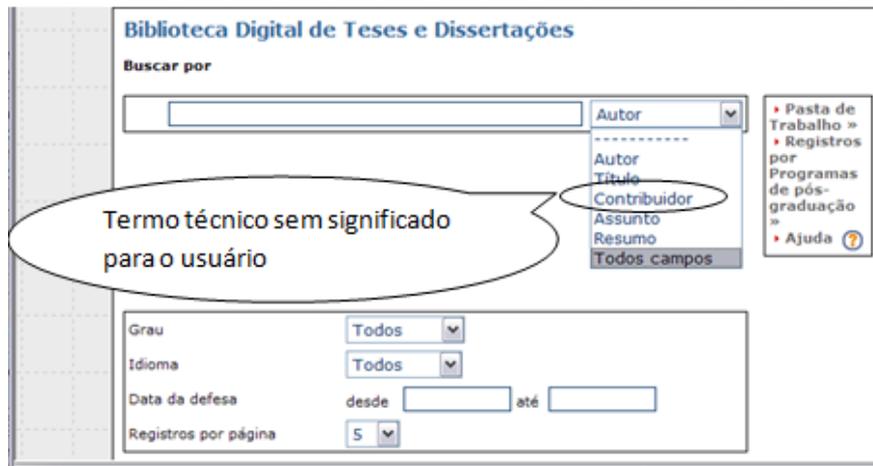
d) Busca simples

Aspectos positivos – A busca simples pode ser feita pelos principais pontos de acesso ao recurso informacional.

Aspectos negativos – A utilização de campos com nomenclatura desconhecida pela maioria de usuários. Exemplo: campo *contribuidor* significa o Orientador. *Contribuidor* é um termo utilizado pelo formato-padrão Dublin Core, mencionado no item 2.2.2 e poderia ter sido adaptado para “orientador”. Com isso perde-se uma oportunidade de oferecer uma busca mais ampla.

⁴¹ Disponível em: <www.sapiencia.pucsp.br>.

Figura 38 – Busca simples da BDTD Sapiencia⁴²



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

A Biblioteca tampouco se utiliza de recursos de navegação por taxonomia, conforme demonstrado no item 2.3.1, ou uma visão de conjunto de temas para facilitar a busca pelo universo de conhecimento existente no ambiente digital.

e) Busca avançada

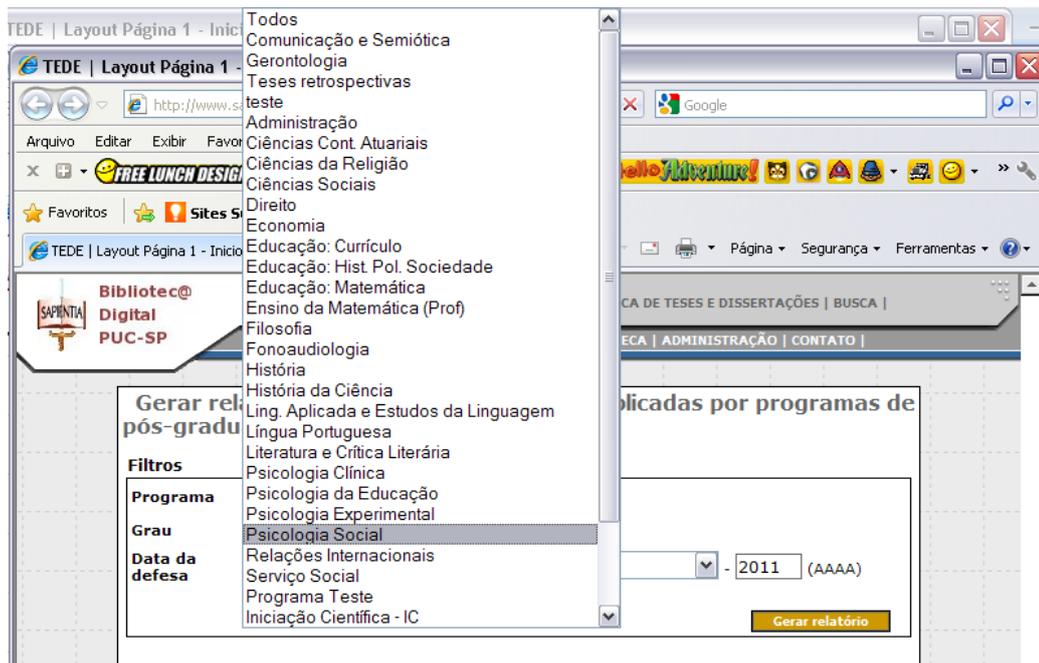
Aspectos positivos – Não há possibilidade de aspecto positivo, visto não haver busca avançada.

Aspectos negativos – Não há menção de busca avançada no ambiente e nem a possibilidade de cruzamentos por campos através de conectores.

Ao clicar em *Programa de Pós graduação*, chega-se à opção de busca por área. Essa busca corresponderia à busca avançada, mas a Biblioteca optou por considerá-la uma segunda tela de busca, conforme a figura a seguir.

⁴² Disponível em: <www.sapiencia.pucsp.br>.

Figura 39 – Página de busca por Programa de Ensino da BDTD Sapientia⁴³



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

Essa opção de busca poderia ser mais um campo e poderia ficar na primeira tela para permitir um refinamento logo na primeira busca.

f) Mecanismos de ajuda

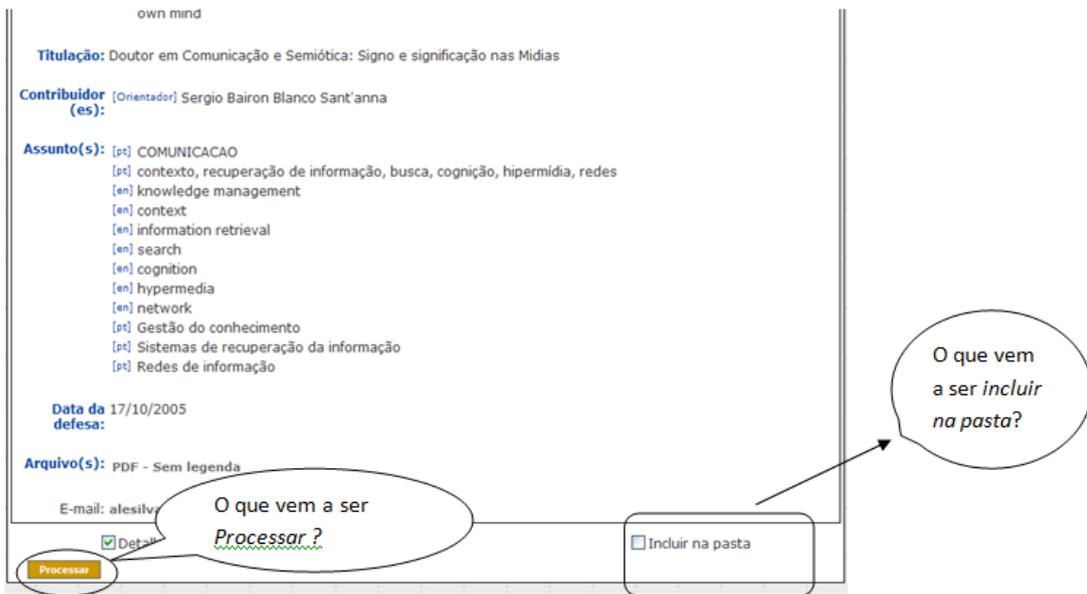
Aspectos positivos Não há possibilidade de aspecto positivo, visto não haver nenhum mecanismo de ajuda.

Aspectos negativos –Não foi identificado nenhum *pop up*, nenhum exemplo ou dica de como pesquisar.

Processar a informação também não é uma palavra usual para representar a busca. Normalmente utiliza-se expressões como *pesquisar*, *buscar*, *seguir* ou mesmo o símbolo de uma lupa para representar a busca.

⁴³ Disponível em: <www.sapientia.pucsp.br>.

Figura 40 – Página de resultado de pesquisa da BDTD Sapiencia⁴⁴



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

g) Representação de resultados de busca;

Aspectos positivos – Não há aspecto positivo devido à restrição de recursos nessa etapa do processo de busca.

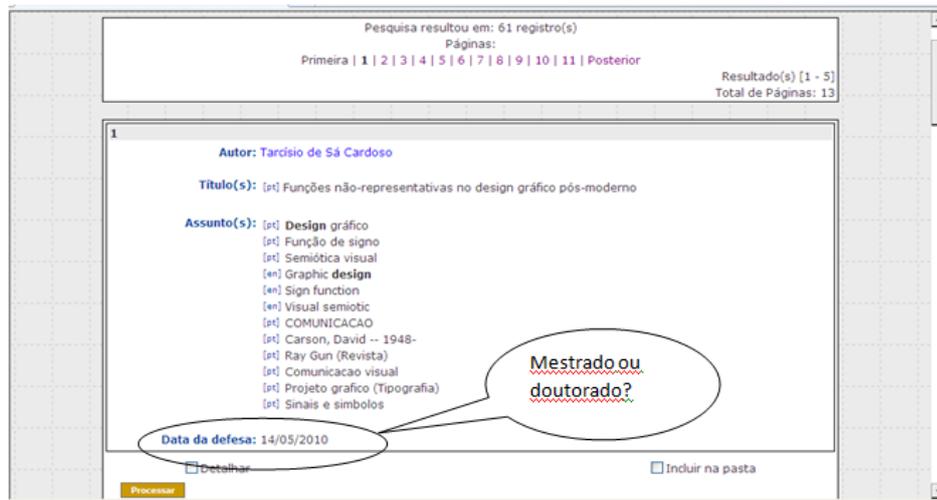
Aspectos negativos – Sem opção de ordenação e sem definição de lógica.

Ao fazer uma pesquisa, não se consegue estabelecer a ordem em seu resultado de busca e nem entender qual foi o critério estabelecido para a representação de telas.

Ao observar a primeira tela do resultado de busca (sem detalhamento), nota-se que os campos selecionados para representar a informação não são suficientes. Exemplo: Informa a data da defesa, mas não diz se é mestrado ou doutorado.

⁴⁴ Disponível em: <www.sapiencia.pucsp.br>.

Figura 41 – Página de resultado da pesquisa da BDTD Sapiientia⁴⁵



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO, 2011.

h) Funcionalidades da Web 2.0

Aspectos positivos – Não há possibilidade de aspecto positivo.

Aspectos negativos – Não há nenhum recurso que remeta o usuário às funcionalidades da Web 2.0.

5.2.2 Considerações sobre a BDTD Sapiientia da PUC-SP

A interface de busca idealizada pela Biblioteca Digital da PUC-SP poderia explorar as possibilidades de desenvolvimento de estratégia de busca na busca simples e avançada, reforçando a comunicação com o seu usuário e os recursos existentes, oferecendo recursos de navegação por taxonomia ou índice, a exemplo da biblioteca SciELO no estudo de caso do capítulo 5.1.

Elementos da Web 2.0 que permitam “taguear”, comentar, compartilhar, acompanhar ou atendimento *on line*, por exemplo, também seriam bem vindos em tempos de redes sociais e maior interatividade.

Como seu universo de conhecimento é bem delimitado e existem diversos exemplos de bibliotecas com a mesma característica na Web, seus recursos

⁴⁵ Disponível em: <www.sapiientia.pucsp.br>.

poderiam ser ampliados e enriquecidos, a exemplo da Biblioteca Digital de Teses e Dissertação da Unicamp na figura 5.

5.3 Comentário sobre os estudos de casos

Os aspectos positivos e negativos sobre cada elemento estabelecido para estudo em ambos os casos analisados, revela a possibilidade de extremos opostos que são possíveis de se encontrar em exemplos de bibliotecas digitais na Web. Embora seu universo de conhecimento seja distinto, seu usuário possui uma semelhança e seu conteúdo instiga à pesquisa em ambos os casos.

Embora o foco desse estudo não seja tecnológico, é visível a diferença existente entre os dois casos em relação a esse aspecto. A SciELO possui recursos mais sofisticados de busca e tratamento da informação. É mais aberta, mais amigável, mais intuitiva e possui em seu conjunto possibilidades maiores de alcançar bons índices de “encontrabilidade” da informação. Talvez por ser projetada exclusivamente para o ambiente Web e não ter um público interno como a Sapienia.

A Sapienia talvez tenha bons índices de retorno com usuários internos, mas a passagem da tela 1 para a tela 2 de busca é um problema que deve ser sanado, pois uma boa parte de usuários deve ter dificuldade em chegar nessa opção de busca.

Como o objetivo desse estudo não foi comparar as bibliotecas e sim analisar dois exemplos nos quais pudessem ser aplicados os elementos de análise, passo agora a recomendar o conjunto de elementos essenciais que poderiam contribuir para melhorar a qualidade das interfaces de bibliotecas digitais, diminuindo o tempo de busca e aumentando a encontrabilidade da informação.

6 ELEMENTOS ESSENCIAIS PARA O ENCONTRO DA INFORMAÇÃO EM BIBLIOTECAS DIGITAIS: RECOMENDAÇÕES

Uma biblioteca digital, que tem por objetivo reunir o conjunto de informações de uma área do conhecimento, em tese, teria de ser mais eficaz que uma busca livre na Web no momento da recuperação da informação, pois supõe-se que seu universo seja conhecido pelo usuário.

Baseando-me neste trabalho, em minha experiência e nos estudos de casos apresentados, faço a partir de agora algumas recomendações para projetos de bibliotecas digitais do ponto de vista de seus elementos essenciais.

6.1 Metadados

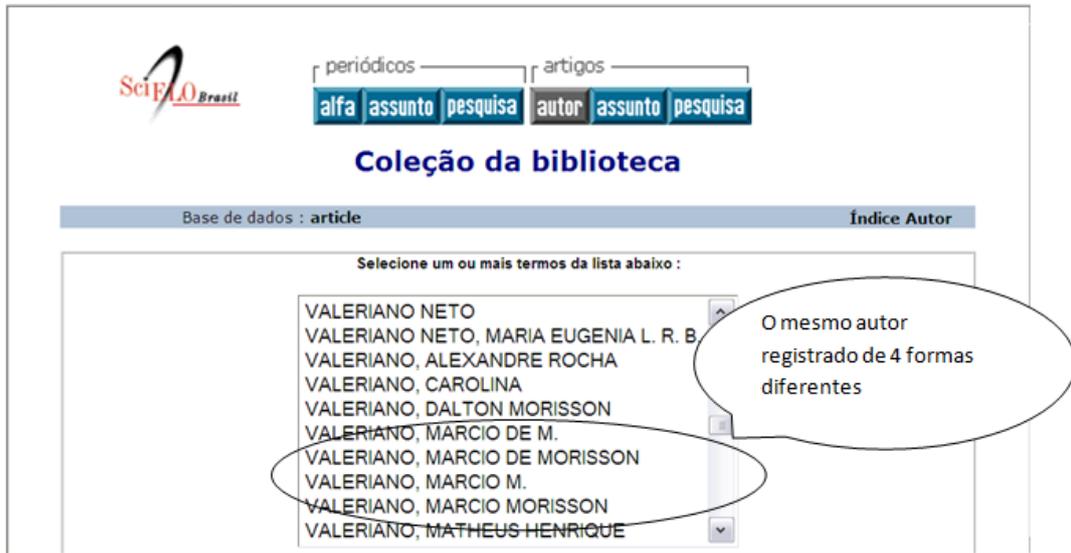
Como visto no item 3.2, a entrada de dados é um fator crítico de sucesso para a recuperação da informação. A qualidade da representação descritiva dos recursos informacionais, a definição do formato-padrão de metadados, o vocabulário controlado adotado e/ou a linguagem natural de um processo de indexação, farão parte de um registro em um sistema de informação que representará um recurso informacional.

Para ilustrar essa afirmação, segue uma pesquisa em 3 sequências de telas.

Passo 1: Pesquisa pelo índice de autor

Neste exemplo, observa-se que o fato de o autor estar mencionado com 4 nomes diferentes, comprometerá a busca, conforme provado nas telas seguintes.

Figura 42 – Resultado na SciELO de pesquisa por índice de autor⁴⁶



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

Passo 2: Pesquisa por *Artigo* pelo primeiro e último nome do autor. Como no índice acima, percebe-se que duas das quatro entradas não trazem o último nome, já deduz-se que o resultado de busca será incompleto. Essa constatação está comprovada no passo 3.

Nessa pesquisa foram encontrados 5 artigos.

Figura 43 – Resultado na SciELO de pesquisa por índice de autor⁴⁷



Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

⁴⁶ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

⁴⁷ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Passo 3: Pesquisa por *Artigo* pelo primeiro e segundo nome do autor. Como no índice acima, as 4 entradas possuem os dois nomes pesquisados, deduz-se que o resultado de busca trará 100% do resultado. Observa-se que a pesquisa aumenta de 5 para 11 resultados. Nessa pesquisa, foram encontrados 11 artigos, 6 a mais que na pesquisa anterior.

Figura 44 – Refinamento da busca na SciELO⁴⁸

Pesquisa pelos dois nomes que encontram-se nas 4 entradas de autor identificadas no PASSO 1

Resultado: 11

SciELO Brasil

periódicos artigos

alfa assunto pesquisa autor assunto pesquisa

Coleção da biblioteca

sua seleção voltar resultado nova pesquisa config fim da página

Base de dados: article

Pesquisa: valeriano [Autor] and marcio [Autor]

Referências encontradas: 11 [refinar]

Mostrando: 1 .. 10 no formato: [ISO 690]

página 1 de 2 ir para página 1 2

1 / 11

seleciona

para imprimir

Canavesi, Vanessa, Ponzoni, Flávio Jorge and Valeriano, Márcio Morisson **Estimativa de volume de madeira em plantios de *Eucalyptus* spp. utilizando dados hiperespectrais e**

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

Exemplos como o acima, demonstrando inconsistência nas bases de dados com impacto no resultado de busca, são facilmente encontrados em interfaces de busca na Web.

Tem aumentado o número de estudos sobre a importância de metadados para a descrição de recursos informacionais. Se formos pensar na Web Semântica como uma realidade em breve, fica difícil conceber uma arquitetura que não estructure os dados para poder agregar inteligência à busca semântica. Há necessidade de se atribuir uma carga semântica aos termos, e isso será desenvolvido a partir de atributos aplicados aos termos na busca a partir de metadados. Ex: Ao buscar manga, será necessário atribuir carga semântica ao termo: manga de camisa, manga fruta, cavalo manga-larga e assim por diante.

⁴⁸ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Os metadados tendem a ser cada vez mais importantes em uma Web cada vez mais universal, pois são os atributos que farão a diferença na criação do significado no momento da busca.

6.2 Comunicação

O aspecto comunicacional em um ambiente digital é muito importante, pois torna-se necessário utilizar-se de diversos recursos que possam substituir a interação homem-homem e preservar a comunicação em uma interação homem-máquina. Esse aspecto é usualmente negligenciado, pois se pressupõe que o usuário, ao acessar um ambiente digital, já está alfabetizado digitalmente. Vivemos essa experiência todos os dias ao tentar nos cadastrar em algum *site*, ao tentar preencher um formulário ou fazer um busca mais específica na Web. Não é tão intuitivo e simples como deveria ser.

É muito importante que o usuário sinta-se guiado em todo o seu processo de busca. Evitar termos técnicos ou desconhecidos, jargões da área, siglas, abreviações e orientações vagas são algumas das recomendações.

Utilizar atendimento *on-line* via *chat*, interação por voz, *pop-ups* ou exemplos, pode ajudar o usuário a compreender melhor a interface.

No exemplo da Biblioteca da PUC-RJ ilustrada no item 6.6, observam-se vários pontos de interrogação na interface, que, ao serem clicados, abrem *pop-ups* explicativos.. Este recurso contribui para tornar claros os passos seguintes.

No caso do Portal do Sebrae, www.sebrae.com.br, a utilização da comunicação na primeira pessoa aproxima o usuário da interface. No exemplo ilustrado no item 6.5, a comunicação foi invertida e o *site* torna-se um espelho do usuário.

Ex.:“Como faço para...?” Testes e provas de conceito também são bem-vindos no momento de aprovação de um projeto. Convidar o usuário a testar o ambiente, filmar o usuário tentando pesquisar uma informação pode ajudar a identificar problemas na interface.

Os estudos de Usabilidade e Arquitetura da Informação contribuem com metodologias para provas de conceito.

Difícilmente encontramos um ambiente com muitos recursos comunicacionais, mas pesquisar muito e extrair os melhores exemplos de bibliotecas digitais já maduras podem ajudar muito para elevar um projeto para um patamar mais profissional.

É muito comum um projeto acabar no momento que deveria começar. É a partir do momento em que ele está no ar, sendo testado e utilizado, que deve-se monitorar para corrigir e desenvolver soluções comunicacionais. Quando um usuário comete um erro de compreensão ou de intuição, fica claro que nesse ponto deve haver uma intervenção. Monitore, dialogue, pesquise, registre e implante as melhores práticas.

6.3 Visão de conjunto

Ao pesquisar em uma biblioteca digital, é muito importante ser informado sobre o conjunto de recursos informacionais existentes e disponíveis para consulta. Isso por si só já contribui para auxiliar na formulação de uma busca, mesmo que a interface seja básica.

Oferecer recursos de navegação por temas, pelos itens mais buscados, mais votados e mais baixados, também ajuda. Muitas vezes, as interfaces não possuem índices de autor, título ou assunto e oferecem apenas um campo em branco para digitar uma palavra. Essa é a solução mais pobre de interface que se pode encontrar. Nada há a fazer, além de adivinhar palavras aleatoriamente. (MANOVICH, 2007).

Informações quantitativas também são recomendáveis, pois dizem muito sobre as características do acervo.

O Portal de Curta-Metragens, www.portacurtas.com.br, pode ser considerado uma biblioteca digital contendo o repositório de filmes curta-metragem premiados e informa em sua tela principal, a quantidade de vídeos disponíveis em seu acervo.

Figura 45 – Página principal do site⁴⁹

Fonte: PORTAL DE CURTAS, 2011.

O exemplo a seguir da Biblioteca Digital Ann Arbor, dá uma ideia de opções de busca por navegação com visão de conjunto. Ao sugerir opções, a interface já está oferecendo uma noção de conjunto através dos itens mais recentes e/ou mais acessados/consultados.

Figura 46 – Página principal: Anne Arbor Digital Library⁵⁰

Fonte: ANNE ARBOR DIGITAL LIBRARY, 2011.

⁴⁹ Disponível em: <<http://www.portacurtas.com.br/index.asp>>.

⁵⁰ Disponível em: <www.aadl.org>.

Outra forma seria antecipar a busca e sugerir uma navegação pela estrutura hierárquica do *site* através de uma taxonomia ou de um Tesouro. Um exemplo desse recurso é o *site* Flora Brasiliensis. Voltada para a comunidade de pesquisadores e cientistas, a biblioteca não tem a preocupação de atender usuários que desconhecem o tema. Permite apenas a navegação pelo nome científico do conjunto da flora brasileira. Para tal, foi desenvolvida uma taxonomia sobre a Flora Brasileira que permite inclusive um pré-conhecimento sobre os grupos/famílias de plantas.

Assim, o recurso de busca é muito simples, mas a estrutura hierárquica e a busca por navegação podem compensar a falta de outros recursos de pesquisa. Clicando no ícone *Informação*, tem-se o acesso ao registro.

A visão de conjunto é essencial para sugerir formas de busca. Segundo Manovich (2011),

[...] Interfaces padronizadas para coleções enormes de mídias tais como lista, galeria, grade (grid) e slide não nos permitem ver os conteúdos de uma coleção inteira. Essas interfaces normalmente só disponibilizam alguns itens por vez (independentemente de você estar em modo de navegação ou procura). Esse método de acesso não nos permite compreender a “forma” de toda uma coleção e observar padrões interessantes.

Como podemos descobrir coisas interessantes em coleções massivas? Ou seja, como podemos examiná-las de forma eficiente e eficaz, sem o conhecimento do que queremos encontrar?”

Figura 47 – Taxonomia do site Flora Brasiliensis⁵¹

OPINÃO

FLORA BRASILIENSIS

A OBRA

O PROJETO

DIGITE O NOME CIENTÍFICO

 buscar

CLASSIFICAÇÃO SEGUNDO A FLORA BRASILIENSIS

fam. *Urticinae*                       

6.4 Busca simples

Mesmo que o projeto seja simples e a única opção seja a busca por um único campo, a experiência tem de ser única. Não deve ser uma busca cega em um campo em branco na base da tentativa e erro.

A busca pode ser guiada, se oferecer, por exemplo, um índice de palavras por cada ponto de acesso ao recurso informacional. Ex.: Autor: índice de autor, título, índice de título e assim por diante.

Dos exemplos analisados e selecionados, um caso de como é possível ir muito além do básico é a Biblioteca Digital da Unicamp.

A pesquisa simples dá opções de campos para pesquisa, explica que as aspas são utilizadas para se buscar expressões e dá uma breve explicação de como utilizar os conectores. Além disso, mostra o universo de possibilidades na busca pelas diversas bibliotecas, destaca as teses mais baixadas e mais visitadas, sugerindo uma busca por navegação.

Figura 48 – Interface de busca da Biblioteca Digital da Unicamp⁵²

Bases Disponíveis	
Dissertações e Teses	
Trabalhos de Conclusão de Curso	
Eventos	
Hemeroteca - CMU	
Revistas Eletrônicas	
Produção Técnico-Científica Digital	
Coleções Raras e Especiais	

Pesquisa

Entrar com uma ou mais palavras:

Pesquisar: Palavra-Chave

Use as aspas para expressões. Exemplo: "bibliotecas digitais"

Todas as palavras **(AND)** Quaisquer palavras **(OR)**

Pesquisar em:

Todos
Dissertações e Teses
Trabalhos de Conclusão de C
Eventos
Hemeroteca - CMU

Mostrar: resultado(s) por página

Formato de Apresentação:

Documentos	
Total:	44.698
Mestrado:	22.833
Doutorado:	12.682
Livre Docência:	230
Outros:	8.953

Teses mais baixadas

- [O conhecimento matemático e o uso de jogos na sala...](#)
- [\(In\)disciplina na escola : cenas da complexidade d...](#)
- [As bases fisiológicas da estrutura triádica da sem...](#)
- [As dificuldades de aprendizagem em leitura e escri...](#)
- [O professor de alunos com deficiência mental : co...](#)

Teses mais visitadas

Pesquisa por Índices

Pesquisar:

Selecionar Índice:

Palavra-Chave Autor Orientador Título

Área de Concentração Ass...

Fonte: UNIVERSIDADE ESTADUAL DE CAMPINAS, 2011.

⁵² Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000224332>>.

Outro exemplo interessante é a Biblioteca do Planalto, que permite a visão geral dos assuntos legislativos e, com isso, a busca por navegação se torna possível.

Figura 49 – Visão de conjunto e busca por navegação⁵³



Fonte: PRESIDÊNCIA DA REPÚBLICA, 2011.

Embora a busca denomine-se simples, ela acaba se tornando a mais complicada, pois não costuma oferecer nenhuma pista. Em geral, é uma busca que depende unicamente das competências informacionais que o usuário tem, de pesquisar em bibliotecas digitais.

A linguagem do usuário normalmente é natural, conflitando com a linguagem padronizada das bibliotecas e dificultando muito a obtenção de sucesso na experiência de busca. Nesse aspecto, o Google é valorizado, pois dificilmente uma palavra registrará zero de resultado e essa “sensação” de ter encontrado a informação dá ao usuário a impressão de que o Google é sempre o melhor caminho.

As bibliotecas digitais necessitam investir na busca simples. Esse passo do projeto não é visto como crítico. Vê-se como crítico o tratamento da informação e, mesmo assim, perde-se na qualidade por falta de indicadores e rigor na definição de metadados.

⁵³ Disponível em: <<http://www4.planalto.gov.br/legislacao>>.

Introduzir o hábito de analisar as buscas efetuadas pelos usuários através dos *logs* de busca é altamente recomendável. Esse é um indicador a ser acompanhado para avaliar a possibilidade de encontrar a informação. Quantos resultados de busca foram negativos? Quais as palavras utilizadas que retornaram como negativas? Como é a linguagem natural do usuário? Qual o perfil de usuários que pesquisam a biblioteca na *Web*? Esses questionamentos são necessários na avaliação da interface de busca e no tratamento da informação.

6.5 Busca avançada

A busca avançada é outro aspecto complexo em uma construção de interface de busca, pois os usuários, em sua maioria, desconhecem regras simples de sintaxe e combinação de conectores a partir de lógica booleana⁵⁴.

A maioria dos usuários digita uma palavra ou mais seguida de espaço, o que não indica uma expressão e sim um conector **E**, em geral, invisível no campo de busca.

Ex1: Se busco meu nome separado apenas por espaço: Renate Landshoff, o sistema compreende que estou buscando Renate **E** Landshoff; portanto, o resultado será todos os registros com o nome **Renate** seguido por todos com o nome **Landshoff** e, por fim, todos os registros com ambos os nomes.

Ex2: Se busco “Renate Landshoff”, entre aspas como uma expressão, o sistema entende que ambas as palavras tem de estar juntas.

Essa sintaxe e muitas outras relacionadas à busca, são desconhecidas de um usuário básico, portanto a interface de uma busca avançada requer um cuidado especial.

Exemplos, orientações claras e curtas, *pop-ups*, informações adicionais quando o resultado for nulo, recursos de cruzamento entre campos e refinamento de busca, são algumas recomendações importantes.

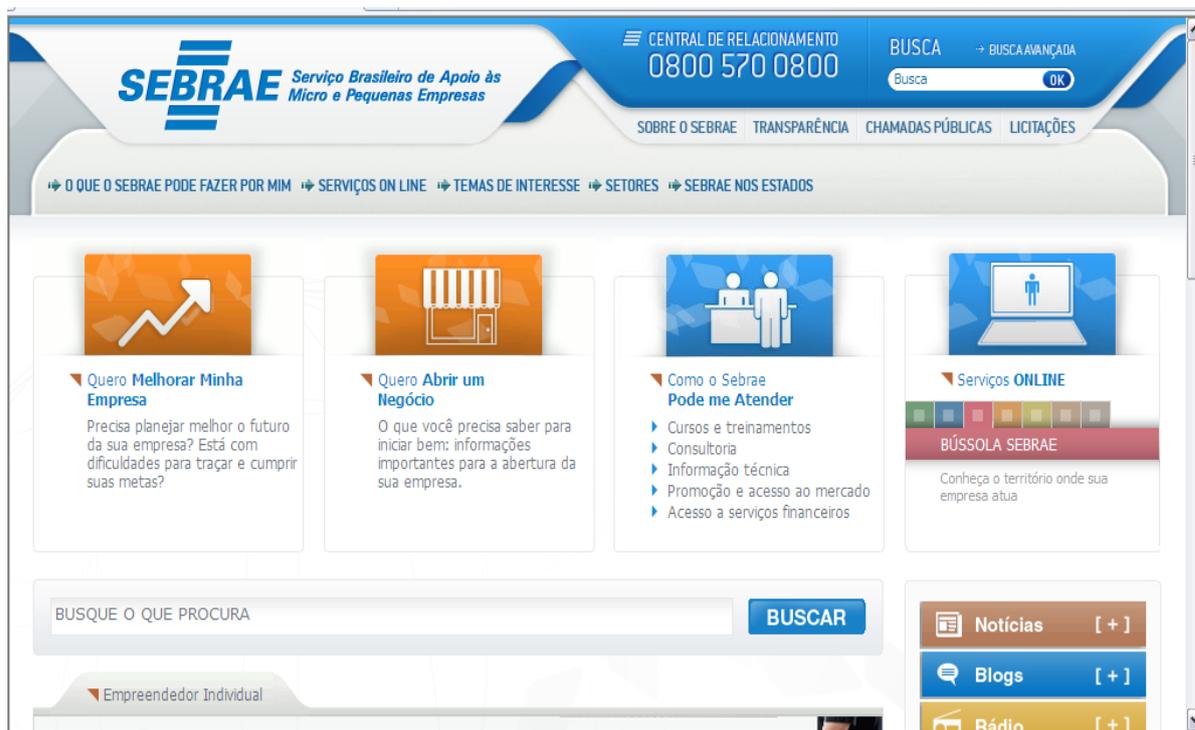
⁵⁴ A lógica booleana se refere ao sistema de lógica matemática denominada Álgebra booleana, cujo nome teve origem no matemático inglês George Boole. É utilizada para criar regras lógicas ou expressões. (INTRODUÇÃO, 2011).

Não encontrei durante a pesquisa nenhuma interface de busca que ofereça opções a mais de um perfil de usuário, mas, dependendo da abrangência do *site*, seria recomendável optar por um modelo de busca com base nos perfis cognitivos de usuários, conforme explanado no item 4. Recomendo, no mínimo, uma interface voltada para o usuário básico e outra voltada para o usuário avançado.

O Portal Sebrae é um caso típico de ambiente voltado para múltiplos perfis cognitivos de usuários, pois tem como cliente usuários completamente distintos .

No caso de *sites*, como o Sebrae, que têm um grande número de usuários novos a cada dia, torna-se especialmente importante categorizá-lo na hora da busca. Pensando nisso, o Sebrae desenvolveu uma seleção natural, permitindo a busca por tipo de negócio e o tempo de vida do empreendedor. Poderia haver uma interface diferenciada para cada um desses tipos de usuários. O pipoqueiro que quer regularizar o seu negócio é completamente distinto do empreendedor que quer aumentar seu capital e ampliar sua empresa. De certa forma, o Portal distingue seus clientes pelos “portlets”⁵⁵ e promove uma busca diferenciada a partir da navegação. Esse recurso é muito bem vindo e comunica-se facilmente com o usuário, pois ele se identifica com o seu momento.

⁵⁵ É um componente visual independente que pode ser utilizado para disponibilizar informações dentro de uma página.

Figura 50 – Portal SEBRAE⁵⁶

Fonte: SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS, 2011.

Já na SciELO, onde os acadêmicos, pesquisadores e cientistas vão rapidamente diminuindo a curva de aprendizagem por consultar o site com frequência, a preocupação com o desenvolvimento da interface muda de foco.

A busca avançada passa a ter combinação de campos com refinamento de busca e é possível pesquisar combinando um assunto com um autor e ainda selecionando um autor que não se deseja.

A possibilidade de campos que existe na formulação da busca, também contribui para aumentar as chances de encontro da informação e tornar os resultados mais precisos.

⁵⁶ Disponível em: <<http://www.sebrae.com.br/>>.

Figura 51 – Busca avançada da SciELO⁵⁷

Um autor e não o outro (somente o primeiro autor)

períódicos artigos
alfa assunto pesquisa autor assunto pesquisa

Coleção da biblioteca

Formulário básico
Pesquisar por : Formulário livre

pesquisar
qualidade

1
2 and valls
3 and not waldomiro

no campo
Assunto
Todos os índices
Palavras do título
Autor
Assunto
Resumo
Ano de publicação
Registro de Ensaio Clínicos
Tipo de artigo
Afiliação - Organização
Afiliação - País, País
Áreas geográficas

Search engine: IAH

BIREME/OPAS/OMS - Centro Latino-Americano e do

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

A tela a seguir possibilita o refinamento da busca. Esse é um recurso muito pouco explorado em interfaces de busca: permitir que o usuário recomece uma pesquisa a partir do resultado de uma primeira pesquisa.

Figura 52 – Refinamento de busca SciELO⁵⁸

períódicos artigos
alfa assunto pesquisa autor assunto pesquisa

Coleção da biblioteca

sua seleção enviar resultado nova pesquisa **config** fim da página

Base de dados : article
Pesquisa : qualidade [Assunto] and valls [Autor] and not waldomiro [Autor]

Referências encontradas : 1 [refinar]

Mostrando: 1 .. 1 no formato [ISO 690]

página 1 de 1

1 / 1
seleciona
para imprimir

Valls, Valéria Martin. **O enfoque por processos da NBR ISO 9001 e sua aplicação nos serviços de informação.** Ci. Inf., Ago 2004, vol.33, no.2, p.172-178. ISSN 0100-1965
• resumo em português | inglês • texto em português

página 1 de 1
sua seleção início da página

Fonte: SCIENTIFIC ELETRONIC LIBRARY ONLINE, 2011a.

⁵⁷ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

⁵⁸ Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>.

Tratando-se de um projeto de Bibliotecas Digitais na *Web* e por ser sempre um conteúdo especializado, a recomendação é que sempre haja a opção de busca avançada, pois usuários de bibliotecas tendem a ter mais conhecimento no uso de sintaxes para elaboração de estratégias de busca e existe uma curva de aprendizagem rápida por ser de uso frequente.

6.6 Mecanismos de ajuda

Mecanismos de ajuda devem ser usados sem economia. O usuário não deve ter perguntas sem respostas. Devem ser incluídos o máximo de *pop-ups* e pontos de interrogação, como faz a Biblioteca Digital da PUC-RJ.

Figura 53 – Página principal da biblioteca⁵⁹



Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO, 2011.

⁵⁹ Disponível em: <<http://www.dbd.puc-rio.br/>>.

Incluir exemplos e implantar ajuda em todas as telas, não apenas na tela principal, é outra recomendação importante. É muito comum observar que há uma preocupação na apresentação, mas ela desaparece no decorrer das telas e das demais interfaces.

São recursos já existentes em grande parte das bibliotecas digitais estrangeiras e começam a despontar nas bibliotecas brasileiras, como visto no decorrer deste estudo.

O *chat*, por exemplo, pode ser encontrado na biblioteca digital da PUC-RJ, conforme tela a seguir.

Figura 54 – Chat⁶⁰

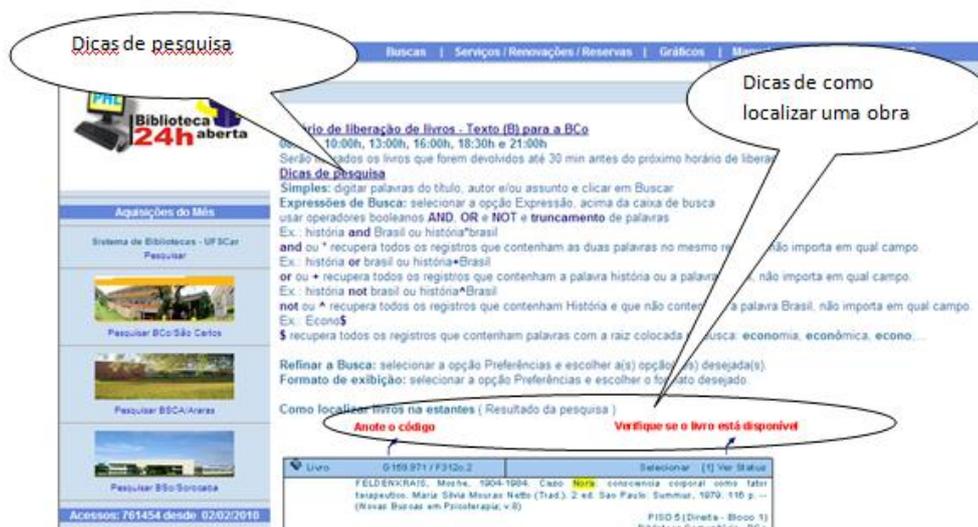


Fonte: PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO, 2011.

⁶⁰ Disponível em: <<http://www.dbd.puc-rio.br/>>.

Um exemplo de mecanismo de ajuda detalhado é a Biblioteca Digital da Universidade Federal de São Carlos.

Figura 55 – Mecanismo de ajuda detalhado⁶¹



Fonte: UNIVERSIDADE FEDERAL DE SÃO CARLOS, 2011a.

Só é possível abrir mão de recursos de ajuda quando a usabilidade do ambiente for extremamente intuitiva; essa característica deve ser testada e aprovada pelos usuários.

No caso do Google, a interface é bastante intuitiva, mas mesmo assim alguns passos ainda são obscuros.

Um exemplo é a página principal do Google onde vê-se a expressão “Estou com sorte”, demonstrada na figura a seguir. Não é possível saber o significado dessa expressão.

⁶¹ Disponível em: <<http://200.9.84.221/cgi-bin/wxis.exe?IsisScript=phl82.xis&cipar=phl82.cip&lang=por>>.

Figura 56 – Ambiente inicial da interface de busca do Google⁶²

Fonte: GOOGLE, 2011c.

A recomendação mais importante no caso desse recurso é identificar grupos de usuários e aplicar testes distintos para validar conceitos e critérios selecionados.

Já começam a despontar, ainda fora do Brasil, tecnologias para atendimento de usuário *on-line*, em que o responsável é avisado pelo sistema quando um usuário encontra-se supostamente confuso no ambiente. Essa situação é detectada a partir de alguns indicadores, como resultados negativos, várias tentativas de busca ou palavras vazias (por exemplo, preposições).

Desenvolver tutoriais *on-line* também é um recurso recomendável, mas eles têm de ser curtos e muito eficazes na comunicação e demonstração. Apresentações longas são ineficazes. O importante é ter o foco no perfil do usuário a ser comunicado, explorando inclusive níveis diferentes de tutoriais. Por exemplo: um para jovens e outro para adultos, um para nível básico de usuário e outro para um nível avançado.

⁶² Disponível em: <www.google.com.br>.

6.7 Representação de resultados de busca

Esse é um momento crítico do processo de busca. No momento de representar um resultado de busca, a interface tem de se preocupar com critérios de ordenação dos resultados e com a representação de metadados no registro recuperado.

A forma como os metadados são dispostos e a flexibilidade em sua visualização, influenciarão na 'encontrabilidade' da informação.

Conforme mencionado no item 2.2.1, o usuário deve encontrar, identificar, selecionar e obter a informação.

É comum encontrar soluções estáticas que não nos permitem modificar a ordem e nem identificar o conjunto de metadados de cada registro. O Google é um exemplo. Na pesquisa a seguir, buscando a sigla CMS (Content Management System), o resultado é uma miscelânea de dados combinados e repetidos, sem metadados identificados e sem referências básicas, como título e data do conteúdo. Tampouco é possível ordenar os resultados. O critério da lista é selecionado pelo Google e desconhecido pelo usuário, conforme figura a seguir, onde é possível observar que três resultados levam o usuário à mesma *URL*.

Figura 58 – Ordenação de resultado de busca da Biblioteca da UFSC⁶⁴



Fonte: UNIVERSIDADE FEDERAL DE SANTA CATARINA, 2011.

Fazendo uma analogia com um cardápio de um restaurante, o ideal seria oferecer os resultados de busca como o ritual de um restaurante: primeiro uma degustação de um primeiro resultado, em seguida um detalhamento maior da busca oferecendo mais informações que auxiliem no processo de seleção e em seguida a possibilidade de visualização da tela com todos os metadados. O usuário poderia ter a opção de já fazer uma seleção prévia a partir da tela de degustação, pois o conjunto de informações já deveria ser relevante. A recomendação principal na representação da busca é a priorização de metadados relevantes que criem significado para o usuário finalizar seu processo de busca.

Como recomendação especial, trago uma importante contribuição para a área de Ciência da Informação ao apresentar Lev Manovich, pesquisador de destaque na área de novas mídias, mídias digitais e design. Manovich, dentre outros focos, estuda as dificuldades da busca no ciberespaço. Suas críticas são respeitadas e sua expressiva produção intelectual traz inúmeras contribuições para o projeto de novas bibliotecas digitais, sendo uma delas a proposta de tornar a informação mais visual.

Os artistas da visualização de dados transformam o caos informacional de pacotes de dados que se locomovem através da rede de formas claras e ordenadas (...) A visualização de dados nos permite enxergar padrões e estruturas por detrás do vasto e aparente fortuito conjunto de dados (...) Os dados quantitativos são reduzidos a seus padrões e estruturas, os quais, a

⁶⁴ Disponível em:

<http://aspro02.npd.ufsc.br/pergamum/biblioteca/index.php?resolution2=1024_1&tipo_pesquisa>.

seguir explodem em inúmeras imagens visuais ricas e concretas. (MANOVICH, 2004, p. 149).

Outro destaque importante, alertando sobre as construções de interfaces de busca é igualmente importante para este estudo

novas relações de percepção sensível com a tecnologia são incorporadas à nossa realidade: passamos a exigir interfaces que sejam amigáveis, divertidas e expressivas, recompensadoras, satisfatórias. (MANOVICH, 2007).

São inúmeras as citações de Manovich que caberiam neste estudo, mas o importante neste momento de finalização, é recomendar a ampliação de horizontes e fronteiras como objeto de estudo para otimizar a encontrabilidade de informação em bibliotecas digitais, além de não se distanciar dos elementos essenciais acima destacados.

7 CONSIDERAÇÕES FINAIS

A quantidade de informação disponível na *Web* não deve prejudicar a qualidade da informação encontrada. Ao projetar uma biblioteca digital, faz-se necessária a utilização de interfaces de busca mais amigáveis, novas cartografias e novos mapas capazes de revelar onde e como encontrar a informação desejada.

O desenvolvimento de um projeto de biblioteca digital, além de ser multidisciplinar como já foi dito, deve envolver também usuários, pois eles representam a necessidade e o olhar do cliente.

A seguir destaco três aspectos que devem ser observados como recomendação sobre a conclusão deste estudo: a entrada de dados, o usuário e as interfaces de busca na *Web*.

Do ponto de vista de entrada de dados, todo cuidado é pouco. É a entrada que garantirá a qualidade da busca e saída da informação. Investir em metadados para descrever o tipo de dado, com foco nas necessidades do usuário, são princípios básicos de uma boa representação descritiva. A qualidade da informação está totalmente vinculada ao tratamento que se dá ao dado e à forma como se descreve um recurso informacional. Ao tomar a decisão sobre a metodologia a ser adotada na construção de uma biblioteca digital, faz-se necessário pensar no todo. Cada dado que representa e descreve um recurso informacional deve ter um significado e deve ser uma possibilidade de busca no momento de projetar a interface de busca.

Do ponto de vista do usuário, sabe-se que são inúmeros os perfis cognitivos e que é praticamente impossível desenvolver uma interface única que atenda a perfis tão distintos. O usuário necessita adquirir habilidades informacionais para reconhecer quando a informação é necessária e ser capaz de localizá-la e utilizá-la de forma efetiva. O profissional bibliotecário deve participar ativamente desse processo, criando tutoriais, momentos de capacitação, campanhas e recursos de usabilidade nas várias interfaces do ambiente visando munir o usuário de competências informacionais.

Criar níveis de interface para tipos variados de usuários também é recomendável. Se é possível identificar os vários perfis de usuário em uma biblioteca

digital, certamente é possível identificar seus perfis cognitivos na *Web*, facilitando, assim, a sua navegação pelo ambiente.

Não faltam bons exemplos e não falta tecnologia. É possível ir além; portanto, é preciso de muita cautela e estudo antes de implantar uma solução definitiva.

A busca por padrões, por recursos da *Web 2.0*, deve abrir o horizonte de quem alimenta os dados, pois são possibilidades hoje que enriquecem a informação e contribuem para um resultado de busca mais eficaz. O usuário deve, cada vez mais, ser inserido nesse processo de construção de inteligência coletiva.

Do ponto de vista das interfaces de busca está tudo por fazer. Embora esse estudo não tenha sido exaustivo, minha experiência profissional permite afirmar que são poucas as bibliotecas digitais que se utilizam de interfaces inteligentes e funcionalidades 2.0 em sua essência. A *Web 2.0* traz uma rica oportunidade ainda não explorada em sua plenitude. Vê-se um exemplo aqui e outro acolá, mas é mais comum encontrarmos *links* para as redes sociais, onde as bibliotecas estão presentes menos timidamente, que encontrar bibliotecas permitindo o uso de agregadores de notícia, *tags*, comentários, votação e demais tipos de intervenções colaborativas que enriqueceriam seu conteúdo. Muitas iniciativas vêm sendo realizadas para modificar completamente a busca na *Web*. A mais conhecida é a *Web Semântica*, conhecida como a *Web 3.0*. A *Web Semântica* busca a construção de uma *Web "inteligente"*, dotada de significado e com uma carga semântica relevante que permita categorizar e filtrar a informação.

A era da informação mal começou. Nós transformamos o meio ambiente, mas a nossa própria transformação é muito mais lenta. A tecnologia e a evolução humana não correm juntas na mesma velocidade.

A informação está disponível para um grupo seletivo de profissionais, que sabe encontrar, compreender a linguagem de bancos de dados e entender a forma como a informação foi representada no ambiente digital. O desafio torna-se enorme quando um usuário básico, que não possui de muitos critérios de seleção, tenta encontrar uma informação na *Web*.

Finalizando, este estudo não pretende ser conclusivo, pois o sujeito é um ambiente dinâmico exposto a inúmeras alterações de configurações em um curto espaço de tempo. A próxima geração da *Web*, a *Web Semântica*, deverá mudar completamente a forma de buscar informação, aumentar muito sua capacidade de ser encontrada e talvez nem existam mais interfaces. Elas são camadas que serão

dispensadas a cada inteligência que for sendo agregada às tecnologias de busca, como por exemplo, a busca por voz, que tornou desnecessário hoje saber escrever a palavra, bastando apenas pronunciá-la para encontrá-la.

As teorias que envolvem a Interação Humano-Computador (IHC) darão respaldo aos estudos da usabilidade e aproximarão cada vez mais o homem da máquina. As teorias da Ciência da Informação (CI) darão respaldo aos estudos de necessidades de informação e comportamento de busca e uso de informação, e encontrarão soluções, que se tornarão cada vez mais intuitivas.

Percebe-se que o Google caminha na direção de ser uma imensa biblioteca virtual, com *links* para conteúdos cada vez mais ricos e interfaces cada vez mais lúdicas e inteligentes. Quanto às bibliotecas digitais, certamente cumprirão sua missão de representar uma parte do universo de conhecimento com recursos de navegação, busca, tratamento da informação e interação cada vez mais sofisticados.

A partir desse estudo, concluo que o estado da arte deixa de ser um fim para ser um caminho eternamente perseguido pelos avanços tecnológicos e possibilidades de soluções inesgotáveis.

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.
- ALVIM, Luísa. Viva biblioteca viva. Disponível em: <<http://vivabibliotecaviva.blogspot.com/>>. Acesso em: 03 jun. 2011.
- ANNE ARBOR DIGITAL LIBRARY. Disponível em: <www.aadl.org>. Acesso em: 04 jun. 2011.
- AQUINO, Idalécio J.; CARLAN, Eliana; BRASCHER, Marisa B. Princípios classificatórios para a construção de taxonomias. **PontodeAcesso**, Salvador, v. 3, n. 3, p. 196-215, dez. 2009.
- ATKINS, Daniel E. (Org.). **Report of the Santa Fe planning workshop on distributed knowledge work environments**: digital libraries. 1997. Disponível em: <<http://www.sis.pitt.edu/~repwkshop/papers/dl1997.pdf>>. Acesso em: 02 jun. 2011.
- BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern information retrieval**. New York: ACM Press, 1999.
- BIBLIOTECA VIRTUAL EM SAÚDE. Disponível em: <<http://bvsmis.saude.gov.br/cgi-bin/multites/mtwdk.exe?k=default&l=60&s=1&t=2&w=&w2=ABSTIN%CANCIA%20SEXUAL&cp=2&x=&tt=&h=&n=25>>. Acesso em: 03 jun. 2011.
- BLOG DA BIBLIOTECA FLORESTAN FERNANDES. Disponível em: <<http://biblioofflch.wordpress.com/2011/04/>>. Acesso em: 20 jun. 2011.
- BORGES, Maria Manuel Marques. Biblioteca digital: materialização e utopia. **Revista da Faculdade de Letras**, Porto, v. 2, p. 653-664, 2003. Disponível em: <<http://ler.letras.up.pt/uploads/ficheiros/2936.pdf>>. Acesso em: 07 jun. 2011.
- BRAGA, G. M. Informação, ciência da informação: breves reflexões em três tempos. **Ciência da Informação**, Brasília, v. 24, n. 1, 1995. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewArticle/534>>. Acesso em: 20 jun. 2011.
- BREVE história do computador parte II. Disponível em: <<http://cadinformatica.br.tripod.com/>>. Acesso em: 01 jun. 2011.

BUSH, Vannevar. As we may think. **The Atlantic Monthly**, jul. 1945. Disponível em: <<http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/>>. Acesso em: 02 maio 2011.

CARDOSO, Vasco. Prosumidores: o papel activo dos consumidores do século XXI. **Creativeblink**, 2009. Disponível em: <<http://www.creativeblink.com/2009/02/prosumidores-o-papel-activo-dos-consumidores-do-seculo-xxi/>>. Acesso em: 04 jun. 2011.

CASEY, Michael. Born in the biblioblogosphere. **LibraryCrunch**, 3 jan. 2006. Disponível em: <http://www.librarycrunch.com/2006/01/post_1.html>. Acesso em: 25 jul. 2011.

CHANG, Shan-Ju; RICE, Ronald E. Browsing: a multidimensional framework. **Annual Review of Information Science and Technology**, v. 28, p. 231-276, 1993.

CONZORCIO INTERUNIVERSITARIO LOMBARDO PER L' ELABORAZIONE AUTOMATICA. Disponível em: <<http://www.cilea.it/>>. Acesso em: 02 jun. 2011.

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

DELICIOUS. Explore everyone's tags. Disponível em: <<http://www.delicious.com/tag/>>. Acesso em: 02 jun. 2011.

DUDZIAK, Elizabeth Adriana. Information Literacy: princípios, filosofia e prática. **Ciência da Informação**, Brasília. v. 32, n. 1, p. 23-35 jan./abr. 2003. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/123/104>>. Acesso em: 20 jun. 2011.

FLORA BRASILIENSIS. Disponível em: <<http://florabrasiliensis.cria.org.br>>. Acesso em: 04 jun. 2011.

GIBSON, William. **Neuromancer**. New York: Ace Books, 1984.

GOOGLE ACADÊMICO. Disponível em: <http://scholar.google.com.br/advanced_scholar_search?hl=pt-BR&as_sdt=0,5>. Acesso em: 01 maio 2011a.

GOOGLE ACADÊMICO. Resultado. Disponível em: <http://scholar.google.com.br/scholar?hl=pt-BR&q=tudonot%C3%ADtulo%3A+%22web+semantica%22&lr=lang_pt&as_ylo=2005&as_vis=0>. Acesso em: 01 maio 2011b.

INTRODUÇÃO à lógica booleana. Disponível em:

<http://help.sap.com/saphelp_46c/helpdata/pt/5b/d2330e43c611d182b30000e829bfe/content.htm>. Acesso em: 20 jun. 2011.

LANCASTER, F. W. **Vocabulary control for information retrieval**. 2. ed. Arlington: Information Resources, 1986.

LEÃO, Lucia. **Cibercultura 2.0**. São Paulo: U.N. Nojosa, 2003.

LEÃO, Lucia. **Derivas**: cartografias do ciberespaço. São Paulo: Annablume, 2004.

LÉVY, P. **Cibercultura**. São Paulo: Editora 34, 1999.

LÉVY, P. **Cibercultura**. São Paulo: Editora 34, 2003.

MACULAN, Benildes Coura Moreira dos Santos; LIMA, Gercina Angela Borém de Oliveira; PENIDO, Patrícia. Taxonomia facetada como interface para facilitar o acesso à informação em bibliotecas digitais. **Revista ACB**, Florianópolis, v. 16, n. 1, p. 234-249, jan./jun. 2011. Disponível em: <http://www.revista.acbsc.org.br/index.php/racb/article/viewFile/746/pdf_48>. Acesso em: 10 jun. 2011.

MAIA, Luiz Cláudio Gomes; SOUZA, Renato Rocha. Medidas de similaridade em documentos eletrônicos. In: Encontro Nacional de Pesquisa em Ciência da Informação, 9., 2008, São Paulo. **Anais...**São Paulo: USP, 2008.

MANESS, Jack M. Library 2.0 theory: web 2.0 and its implications for libraries. **Webology**, v. 3, n. 2, jun. 2006. Disponível em: <<http://www.webology.ir/2006/v3n2/a25.html>>. Acesso em: 28 jul. 2011.

MANOVICH, Lev. Contra a busca. Disponível em: <<http://lab.softwarestudies.com/2011/07/contra-busca.html>>. Acesso em: 30 jul. 2011.

MANOVICH, Lev. **Information as an Aesthetic Event**. 2007. Disponível em <http://www.festivalmiden.gr/textslinks/Lev%20Manovich_Information%20as%20an%20Aesthetic%20Event_2007.htm>. Acesso em: 07 maio 2011.

MANOVICH, Lev. Visualização de dados como uma nova abstração e anti-sublime. In: LEÃO, Lucia (Org). **Derivas**: cartografia do ciberespaço. São Paulo: Annablume, 2004.

MARTELETO, Regina Maria. Análise de redes sociais: aplicação nos estudos de transferência da informação. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 71-81, jan./abr. 2001.

MEDEIROS, Ana Paula Santulo Custódio de. **Análise bibliométrica da produção científica da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP**. 2005. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia e Ciência da Informação) –Universidade Federal de São Carlos, São Carlos, 2005.

MILLER, Paul. Web 2.0: building the new library. **Ariadne**, n. 45, out. 2005. Disponível em: <<http://www.ariadne.ac.uk/issue45/miller>>. Acesso em: 06 jul. 2011.

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA (Brasil); INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Biblioteca Digital Brasileira de Teses e Dissertações. Disponível em: <<http://bdtd.ibict.br/pt/inicio.html>>. Acesso em: 28 jul. 2011.

MINISTÉRIO DA FAZENDA (Brasil). Comissão de Valores Mobiliários. Disponível em: <<http://www.cvm.gov.br/>>. Acesso em: 18 jun. 2011.

MORVILLE, P. **Ambient Findability**: what we find changes who we become. Sebastopol: O'Reilly, 2005.

O'REILLY, Tim. What is web 2.0? design patterns and business models for the next generation of software. **O'REILLY**, 30 set. 2005. Disponível em: <<http://oreilly.com/web2/archive/what-is-web-20.html>> Acesso em: 31 jul. 2011.

OLIVEIRA, Marlene de (Coord.). **Ciência da Informação e Biblioteconomia**: novos conteúdos e espaços de atuação. Belo Horizonte: Editora da UFMG, 2005.

PINHEIRO, Gabriel. A importância da consistência em interfaces móveis. **Webinsider**, 13 abr. 2011. Disponível em: <<http://webinsider.uol.com.br/2011/04/13/a-importancia-da-consistencia-em-interfaces-moveis/?>>. Acesso em: 20 jun. 2011.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO. Sapiencia. Disponível em: <www.sapiencia.pucsp.br>. Acesso em: 03 jun. 2011.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO. Divisão de bibliotecas e documentação. Disponível em: <<http://www.dbd.puc-rio.br/>>. Acesso em: 07 jun. 2011.

PORTAL DE CURTAS. Disponível em: <<http://www.portacurtas.com.br/index.asp>>. Acesso em: 04 jun. 2011.

PRESIDÊNCIA DA REPÚBLICA (Brasil). Disponível em:
<<http://www4.planalto.gov.br/legislacao>>. Acesso em: 02 jul. 2011.

SABBATINI, Renato. A explosão da informação. **Informática**, 23 dez. 1997.
Disponível em: <<http://www.sabbatini.com/renato/correio/cp971223.htm>>. Acesso em: 20 jun. 2011.

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

SANTAELLA, Lucia. **Matrizes da linguagem e pensamento: sonora, visual e verbal**. 3. ed. São Paulo: Iluminuras; Fapesp, 2005.

SANTAELLA, Lucia. **Navegar no ciberespaço: o perfil cognitivo do leitor imersivo**. 2 ed. São Paulo: Paulus, 2007b.

SANTAELLA, Lúcia. **Navegar no ciberespaço: o perfil cognitivo do leitor imersivo**. São Paulo: Paulus, 2004.

SANTOS, Robson L. G. dos. **Usabilidade e métodos de avaliação de usabilidade de interfaces web**. Disponível em:
<http://xa.yimg.com/kq/groups/20750039/2055696000/name/usab_metodos.pdf>. Acesso em: 20 jun. 2011.

SARACEVIC, Tefko, Ciência da informação: origem, evolução e relações. **Perspec. Ci. Inf.**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em:
<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>>. Acesso em: 10 jun. 2011.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Coleção da biblioteca. Disponível em:
<<http://www.scielo.br/cgi-bin/wxis.exe/iah/>>. Acesso em: 02 jun. 2011a.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Coleção da biblioteca. Autor.
Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/?IsisScript=iah/iah.xis&base=article^dlibrary&index=AU&lang=p&fmt=iso.pft&form=B>>. Acesso em: 02 jun. 2011b.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Coleção da biblioteca. Refinar pesquisa. Disponível em: <<http://www.scielo.br/cgi-bin/wxis.exe/iah/#refine>>. Acesso em: 02 jun. 2011c.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Disponível em:
<<http://www.scielo.org/php/index.php?lang=pt>>. Acesso em: 02 jun. 2011d.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Pesquisa. Disponível em: <<http://search.scielo.org/>>. Acesso em: 02 jun. 2011e.

SCIENTIFIC ELETRONIC LIBRARY ONLINE. Trabalhos em Linguística Aplicada. Disponível em: <http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0103-18132010000200007&lng=pt&nrm=iso&tlng=pt>. Acesso em: 02 jun. 2011f.

SENADO FEDERAL (Brasil). Portal legislação. **Pesquisa legislação rápida.** Disponível em: <<http://www6.senado.gov.br/sicon/index.jsp?action=LegislacaoTextual#>>. Acesso em: 22 jun. 2011a.

SENADO FEDERAL (Brasil). Portal legislação. **Resultado:** pesquisa legislação rápida. Disponível em: <<http://www6.senado.gov.br/sicon/index.jsp?action=LegislacaoTextual#>>. Acesso em: 22 jun. 2011b.

SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS. Disponível em: <<http://www.sebrae.com.br/>>. Acesso em: 05 jun. 2011.

SILVA, Alexandre Campos. **Gestão de Conhecimento:** linguagem, forma e impacto na comunicação em redes de informação. 2005. 250 f. Tese (Doutorado em Comunicação e Semiótica)—Pontifícia Universidade Católica de São Paulo, 2005.

SWETLAN, Anne J. Gilliland. **Introduction to metadata:** setting the stage. maio 2000. Disponível em: <<http://www.slis.kent.edu/~mzeng/metadata/Gilland.pdf>>. Acesso em: 07 jun. 2011.

TAMMARO, Anna Maria; SALARELLI, Alberto. **A Biblioteca digital.** Brasília: Brique de Lemos, 2008.

UNIVERSIDADE DE SÃO PAULO. Brasileira USP. Disponível em: <<http://www.brasiliana.usp.br/>>. Acesso em: 10 abr. 2011.

UNIVERSIDADE ESTADUAL DE CAMPINAS. Biblioteca Digital da UNICAMP. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000224332>>. Acesso em: 10 jun. 2011.

UNIVERSIDADE FEDERAL DE SÃO CARLOS. Biblioteca aberta 24h. Disponível em: <<http://200.9.84.221/cgi-bin/wxis.exe?IscScript=phl82.xis&cipar=phl82.cip&lang=por>>. Acesso em: 07 jun. 2011a.

UNIVERSIDADE FEDERAL DE SANTA CATARINA. Biblioteca universitária.

Disponível em:

<http://aspro02.npd.ufsc.br/pergamum/biblioteca/index.php?resolution2=1024_1&tipo_pesquisa>. Acesso em: 07 jun. 2011.

VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002.

VILAN FILHO, Jayme Leiro. Hipertexto: visão geral de uma nova tecnologia de informação. **Ciência da Informação**, Brasília, v. 23, n. 3, p. 295-308, set./dez. 1994.

WEINBERGER, David. **A nova desordem digital**. São Paulo: Campus, 2007.

WIKIPÉDIA. Feed. Disponível em: <<http://pt.wikipedia.org/wiki/Feed>>. Acesso em: 03 jun. 2011b.

WIKIPÉDIA. Plugin. Disponível em: <<http://pt.wikipedia.org/wiki/Plugin>>. Acesso em: 03 jun. 2011a.

REFERÊNCIAS CONSULTADAS

ALVARENGA, Lidia; SILVA, Daniela Lucas da. Organização e representação do conhecimento na Ciência da Informação: revisão da literatura. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 3, n. 1, 2010. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/27/57>>. Acesso em: 20 jun. 2011.

ARANTES, Priscila; SANTAELLA, Lucia (Orgs.). **Estéticas tecnológicas: novos modos de sentir**. São Paulo: Educ, 2008.

CAMPELLO, B.; ABREU, V. F. G. Competência informacional e a formação do bibliotecário. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 10, n. 2, p. 178-193, jul./dez. 2005.

CASTELLS, Manuel. **A galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade**. Rio de Janeiro: Jorge Zahar, 2003.

CATARINO, Maria Elisabete; BAPTISTA, Ana Alice. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DataGramZero**, Rio de Janeiro, v. 8, n. 3, jun. 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_04.htm>. Acesso em: 20 jul. 2011.

CENDÓN, Beatriz Valadares. Ferramentas de busca na Web. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 39-49, jan./abr. 2001. Disponível em: <<http://www.scielo.br/pdf/ci/v30n1/a06v30n1.pdf>>. Acesso em: 18 jun. 2011.

JOHNSON, Steven. **Cultura da Interface: como o computador transforma nossa maneira de criar e comunicar**. Rio de Janeiro: Zahar, 2001.

LEÃO, Lucia (Org.). **InterLab: labirintos do pensamento contemporâneo**. São Paulo: Editora Iluminuras; FAPESP, 2002.

LEÃO, Lucia (Org.). **O chip e o caleidoscópio: reflexões sobre as novas mídias**. São Paulo: Ed. SENAC, 2005.

LEÃO, Lucia. **A estética do labirinto**. São Paulo: Ed. Universidade Anhembi Morumbi, 2002.

LOURENÇO, Cíntia Azevedo. Metadados: o grande desafio na organização da web. **Informação & Sociedade**, João Pessoa, v. 17, n. 1, p. 71-80, jan./abr. 2007. Disponível em: <<http://cintialourenco.eci.ufmg.br/downloads/466-986-1-PB.pdf>>. Acesso em: 20 maio 2011.

MARCONDES, Carlos H. (Org.) et al. **Bibliotecas digitais: saberes e práticas**. 2. ed. Salvador: EDUFBA; Brasília: IBICT, 2006.

MORIN, Edgard. **O método: a natureza da natureza**. Local: Portugal Editora Publicações Europa-América Ltda, Ano. 1997

PASSARELLI, B. Do Mundaneum à WEB Semântica: discussão sobre a revolução nos conceitos de autor e autoridade das fontes de informação. **DataGramZero**, Rio de Janeiro, v. 9, n. 5, out. 2008. Disponível em: <http://www.dgz.org.br/out08/Art_04.htm>. Acesso em: 14 mar. 2011.

PRADO, Renato Silva de Almeida. **Arquitetura de interface: análise de formas de organização da informação na interação entre pessoas e códigos**. 2006. 107 f. Dissertação (Mestrado em Comunicação e Semiótica)–Pontifícia Universidade Católica de São Paulo, São Paulo, 2006.

RUSSO, Mariza. **Fundamentos de Biblioteconomia e Ciência da Informação**. Rio de Janeiro: Editora e-papers, 2010.

SANTAELLA, Lucia. A estética das linguagens líquidas. In: ARANTES, Priscila; SANTAELLA, Lucia (Orgs.). **Estéticas tecnológicas: novos modos de sentir**. São Paulo: Educ, 2008.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa; ALVES, Rachel Cristina Vesú. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. **DataGramZero**, Rio de Janeiro, v. 10, n. 6, dez. 2009. Disponível em: <http://www.dgz.org.br/dez09/Art_04.htm>. Acesso em: 18 jun. 2011.

TAPSCOTT, Don; WILIAMS, Anthony D. **Wikinomics: como a colaboração em massa pode mudar o seu negócio**. São Paulo: Nova Fronteira, 2007.

VIEIRA, Angel Freddy Godoy; GARRIDO, Isadora dos Santos. Folksonomia como uma estratégia para recuperação colaborativa da informação. **DataGramZero**, Rio de Janeiro, v. 12, n. 2, abr. 2011. Disponível em: <http://www.dgz.org.br/abr11/Art_02.htm>. Acesso em: 20 jun. 2011.

YIN, Robert K. **Estudo de caso**. Porto Alegre: Bookman, 2005.