

Serviços de descoberta em rede: a experiência do modelo Google para os usuários de bibliotecas universitárias

Caterina Marta Groposo Pavão

Doutora; Universidade Federal do Rio Grande do Sul (UFGRS), Porto Alegre, RS, Brasil;
caterina@cpd.ufrgs.br

Sônia Elisa Caregnato

Doutora; Universidade Federal do Rio Grande do Sul (UFGRS), Porto Alegre, RS, Brasil;
sonia.caregnato@ufrgs.br

Resumo: Contextualiza o uso dos serviços de descoberta em rede nas bibliotecas universitárias, que necessitam atender a demanda de usuários cada vez mais exigentes no que se refere à performance durante a busca e a recuperação da informação. Pretende colaborar para o entendimento sobre essas ferramentas, salientando a formação do seu índice central, a estruturação da busca facetada, os critérios de relevância utilizados e a apresentação dos resultados numa lista única. Utiliza o método bibliográfico para selecionar a literatura nacional e internacional relevante acerca do tema. Conclui que as bibliotecas devem revisitar o alcance de seus catálogos para manter-se conectadas com os modelos mentais dos usuários e suas necessidades de informação. Serviços de descoberta em rede, integrado com os demais serviços de informação, têm o potencial de melhorar a experiência de busca. Estas ferramentas devem ser projetadas baseadas na experiência de busca e comportamento dos usuários para equilibrar as necessidades e expectativas dos mesmos com as capacidades do sistema.

Palavras-chave: Serviços de descoberta. Bibliotecas universitárias. Busca e recuperação da informação.

1 Introdução

Os profissionais da informação que atuam nas bibliotecas de instituições de ensino superior, desde muito têm se preocupado em oferecer aos seus usuários mecanismos de busca que consigam recuperar de forma rápida e eficiente a informação contida no acervo da biblioteca. Atualmente, a preocupação é permitir a recuperação, não só

do conteúdo desses acervos, mas de toda a informação contida em repositórios, catálogos e bases de dados, em qualquer suporte, esteja ela dentro ou fora da biblioteca.

Como uma possível solução para o problema da busca em acervos distribuídos e na tentativa de atender às expectativas dos usuários, surgiram na década passada os metabuscadores, ou ferramentas de busca federada, que se caracterizam pela pesquisa simultânea em diversas fontes, sejam elas bancos de dados de bibliotecas, bases de dados comerciais, repositórios institucionais, revistas eletrônicas, a internet, entre outras, e, também, pela apresentação dos resultados em uma lista única. Se, por um lado, esta tecnologia veio agilizar a busca de informação, por outro lado o tempo de resposta para apresentação dos resultados não tem sido satisfatório, visto que os metabuscadores utilizam os mecanismos de busca de cada uma das fontes pesquisadas.

Para minimizar os problemas relacionados com o tempo de resposta e apresentação dos resultados, surgem, no início desta década, os *Web Scale Discovery Services*, denominados neste trabalho, em tradução livre, de serviços de descoberta em rede ou apenas serviços de descoberta. O princípio básico dos serviços de descoberta é fornecer um índice único de metadados¹ previamente coletados, que permitem uma busca unificada, substituindo a busca federada, ou seja, a busca em cada uma das bases provedoras de conteúdos, utilizada até então. Hospedados local ou remotamente, disponibilizam uma variedade de informações que podem incluir os diferentes tipos de conteúdo licenciado, conteúdo local e o próprio catálogo da biblioteca, tudo combinado num índice único.

As tecnologias de descoberta em rede a partir de índices unificados oferecem grande potencial para simplificar a busca e a recuperação da informação no meio acadêmico e torná-la mais efetiva. Entretanto, como em todas as soluções tecnológicas, muitos detalhes precisam ser entendidos e ajustados. As formas como esses recursos estão sendo implementados têm implicações tanto para as bibliotecas como para os usuários.

As bibliotecas universitárias, segundo Luther e Kelly (2011), estão frente ao desafio de oferecer aos usuários recursos com a simplicidade do Google, que é o que a nova geração de usuários espera. Ao mesmo tempo em que oferecem mecanismos

de busca na rica coleção de material impresso e digital que possui, que é o que o usuário precisa para atender às demandas acadêmicas.

Os usuários da “Geração Google”, “como são chamados os adolescentes nascidos a partir de 1993, depois da popularização do computador” (OLIVEIRA, 2011), e que hoje representam a grande maioria dos estudantes universitários, são muito dependentes das tecnologias, capazes de criar, manipular ou interagir com o conteúdo digital em diferentes plataformas, meios, linguagens e formatos digitais. A informação pode estar disponível em fontes que podem ser acessadas por uma variedade de canais de informação e utilizadas para satisfazer às necessidades do dia a dia, do trabalho, dos estudos ou das horas de lazer. No entanto, mesmo para esta geração, é por vezes muito difícil decidir qual a fonte apropriada para buscar informação de maneira eficiente e, sobretudo, avaliar criticamente os resultados recuperados.

Para Vaughan (2011a), a combinação de grandes repositórios de conteúdo com interfaces acessíveis e intuitivas tem o potencial de facilitar muito o processo de pesquisa, fornecendo identificação rápida da informação, acesso ao conteúdo e capacidade de ranquear a relevância dos documentos recuperados de um grande repositório de informações.

Os usuários demandam novos produtos e serviços no âmbito biblioteconômico, desde o acesso em rede à biblioteca, até a possibilidade de reunir toda a informação disponível mediante ferramentas distintas dos clássicos catálogos. Assim, as tecnologias de comunicação e informação se adaptam às características desses usuários proporcionando técnicas de recuperação de informação que se moldem a novos comportamentos.

Uma mudança no atual conceito de busca e recuperação da informação no contexto acadêmico implicará em mudança de atitude dos profissionais da informação para atender às demandas dos usuários, seja fornecendo recursos de aprendizagem, auxiliando na seleção da informação mais adequada, ou incentivando o espírito crítico, essenciais para que o usuário seja bem sucedido no processo de seleção da informação recuperada e que corresponda a suas expectativas. Entretanto, atender a todas essas demandas só será possível mediante o entendimento não só dos serviços de descoberta em si, mas da forma como os usuários se comportam ao utilizá-los.

Assim, o propósito deste trabalho é o de explorar as funcionalidades dos sistemas de descoberta em rede, salientando a formação do seu índice central, a estruturação da busca facetada, os critérios de relevância utilizados e a apresentação dos resultados numa lista única, a fim de atender às demandas de informação da “Geração Google”.

Para este fim, utilizou-se o método bibliográfico para levantar e selecionar a literatura nacional e internacional acerca do tema no Portal de Periódicos da CAPES, principalmente nas bases de dados *Lisa* e *Web of Science* e no *Scientific Electronic Library Online* (SciELO), incluindo todos os países da América Latina e Caribe. Como o tema é novo, não foi imposta limitação de data de publicação. Os dados foram reunidos, sistematizados e analisados no período de 2010 a 2013.

2 Os serviços de descoberta em rede

Possibilitar e facilitar aos usuários o acesso à informação constitui-se na função central dos profissionais da informação. O desafio para bibliotecas universitárias, apañadas na mudança sísmica dos recursos impressos para os digitais, é oferecer uma experiência com a simplicidade do Google e de amplo acesso às coleções de materiais impressos e digitais da biblioteca.

Howard e Wiebrands (2011) argumentam que durante séculos o catálogo tem sido o esteio da biblioteca e uma das principais ferramentas para acesso às coleções. Os *Online Public Access Catalog* (OPACs) baseados na *web* têm sido amplamente implementados e usados na maioria das bibliotecas, tornando-se uma ferramenta importante para bibliotecários e usuários. Suas limitações, no entanto, ficaram cada vez mais óbvias, particularmente quando comparadas a outras ferramentas de pesquisa na *web* que ganharam proeminência e popularidade.

Os catálogos de bibliotecas refletem o método de organização de informações de uma época pré-internet e, como tal, "fazem sentido" para bibliotecários e usuários experientes. No entanto, não são simples de usar ou navegar, principalmente quando comparadas ao Google que, como diversos estudos mostram, é a primeira ferramenta de busca para muitos estudantes e professores.

Os serviços de descoberta baseiam-se na abordagem do Google no que diz respeito à disponibilização de uma caixa única de busca, tempo de resposta, quantidade e ordenamento de resultados. Segundo Siqueira (2013), o Google nasceu trazendo uma promessa extremamente difícil, qual seja, indexar a quase totalidade das páginas da *web* sem negligenciar metodologias de seleção e armazenamento de termos e *links* que possibilitassem melhorias constantes nos resultados de busca. Utilizando o *PageRank* como algoritmo de ordenamento de qualificação das páginas da *web*, almejava verificar a probabilidade de satisfação de uma dada busca pela identificação dos *links* mais pertinentes em uma página, conforme critérios de citação externa de outras páginas.

Os serviços de descoberta se diferenciam do Google pelo seu conteúdo. Os desenvolvedores dessas ferramentas firmam acordos com os mais diversos fornecedores de conteúdo e editoras para obter acesso aos seus metadados e, de preferência, ao texto completo dos documentos. Assim, cada desenvolvedor pode disponibilizar milhares, senão milhões, de novos itens incluídos no índice unificado e proporcionar aos usuários uma ampla experiência de descoberta de conteúdos de relevância acadêmica.

Uma pesquisa na internet pode ser considerada suficiente e conveniente em situações do cotidiano, porém, no meio acadêmico, a precisão de um sistema que recupera somente documentos relevantes e a revocação, que permite recuperar todos os documentos relevantes, são vitais. Os motores de busca disponíveis na internet podem ser o ponto de partida para as pesquisas acadêmicas, mas as bibliotecas podem fornecer muito mais. Elas contêm ricas coleções e recursos cuidadosamente selecionados que não são indexados de forma consistente por esses motores.

2.1 As diferentes gerações de recursos de busca em bibliotecas

Segundo Codina, Abadal e Rovira (2010), historicamente, três grandes tipos de soluções tecnológicas têm sido utilizados no momento de implementar sistemas de informação capazes de dar conta dos conteúdos da *web*: indexação, coleta e busca federada. Essas três soluções são as que as bibliotecas acadêmicas buscam para solucionar o problema da grande quantidade de informação que devem disponibilizar

aos usuários. A indexação é a solução mais antiga e é utilizada na maioria dos motores de busca e todos os serviços de recuperação da informação para bibliotecas. Para substituir os complexos sistemas de análise e indexação de documentos, a coleta de metadados (*harvesting*) cria um índice comum a partir dos dados codificados contidos em diversas fontes e que seguem normas e protocolos comuns. A busca federada é a solução mais recente no contexto de bibliotecas, mas já vinha sendo usada pelos buscadores da *web* e consiste em enviar, simultaneamente, a mesma pergunta a diversos motores de busca.

A busca federada, ou por meio de metabuscadores, é largamente utilizada, pois facilita o acesso às diversas coleções da biblioteca. Porém, quando a metabusca é realizada em diversos provedores de serviços de bases de dados, o entusiasmo com a ferramenta arrefece e percebe-se que o usuário ainda não alcança o resultado que gostaria. Segundo Way (2010), isso acontece devido a fatores como lentidão, incapacidade de acessar uma ou outra base no momento da busca, irrelevância de informações recuperadas e duplicação nos resultados, entre outros. A constatação dessas limitações levou à busca de novas soluções.

As bibliotecas acadêmicas, principalmente, estão voltando sua atenção para uma nova geração de recursos de busca denominada *web service discovery* ou *web scale discovery services*. Na língua portuguesa ainda não há uma terminologia que identifique estes serviços. Em trabalho apresentado no 24º. Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação (CBBDD), Maranhão (2011) utiliza a denominação “serviços de descoberta em escala na Internet” e “serviço de descoberta na Internet”. Neste estudo, optou-se por denominá-los serviços de descoberta em rede ou simplesmente serviços de descoberta, por considerar que a internet é apenas o meio para estabelecer comunicação utilizando as facilidades e infraestrutura da rede global de computadores.

O termo *web-scale* foi introduzido no vocabulário das bibliotecas por Lorcan Dempsey, vice-presidente da *Online Computer Library Center* (OCLC) que, em 5 de janeiro de 2007, publicou no seu blog: “Eu gosto da expressão *web-scale*. Ela é utilizada largamente pela Amazon e outros na discussão das suas ‘plataformas’ de serviços como S3 e EC2.” (DEMPSEY, 2007, tradução nossa)². Dempsey observou como o termo vinha sendo utilizado por organizações como a Amazon para descre-

ver seus serviços de infraestrutura infinitamente escalável, como o *Elastic Compute Cloud* (EC2) ou *Simple Storage Service* (S3). A OCLC também vem adotando esse conceito para explicar a maneira como oferece seus produtos e serviços através de sua plataforma global, o WorldCat e, mais recentemente, o WorldCat Local.

Para Breeding (2012), *web-scale* representa um novo paradigma alternativo para as bibliotecas, tanto em termos de gerência de suas atividades internas, como na maneira como elas proporcionam acesso às coleções e serviços para seus usuários. Segundo o autor, quatro características devem estar presentes quando pensamos em *web-scale*: larga escala de plataformas tecnológicas, aplicações disponibilizadas por meio do *software* como um serviço de hospedagem múltipla, abordagens de dados solidamente agregados e acordos cooperativos entre as bibliotecas participantes.

O conceito de *web-scale* nos serviços de descoberta remete a sistemas de larga escala, que carregam uma conotação de âmbito maciçamente grande, em tamanho ou extensão. Entendendo o enorme alcance da *web*, este ponto de referência é útil para entender como estes serviços operam em diferentes domínios, tais como aqueles relativos às bibliotecas. Nelas, soluções *web* de descoberta em escala tentam fornecer acesso instantâneo a mais ampla gama de conteúdo, incluindo não apenas as coleções locais, mas também os materiais fornecidos como recursos externos.

2.2 Características dos serviços de descoberta em rede

O princípio de funcionamento do serviço de descoberta parte da coleta prévia de metadados de catálogos, bases de dados, repositórios, entre outros, por meio do protocolo *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH)³, *File Transfer Protocol* (FTP) ou outro mecanismo semelhante. Com isso, cria-se uma base de dados própria que reúne todas as informações recuperadas a partir de uma pesquisa única, cujos resultados retornam para os usuários numa única resposta.

A vantagem primeira e inicial em relação à busca federada ou metabuscadores é que a existência de uma base de dados com todos os metadados já reunidos assegura que sempre será oferecida uma resposta ao usuário final, retornando resultados mais rapidamente, pois não depende de tempo de conexão, nem de possibilidade de acesso às diversas fontes. (MARANHÃO, 2011, p. 8).

As principais características dos serviços de descoberta em rede são apontadas resumidamente por Vaughan (2011a). Primeiramente, o autor cita o índice profundamente detalhado criado a partir dos metadados coletados remotamente do catálogo da biblioteca local ou qualquer outro recurso de informação contratado ou gratuito que a biblioteca queira disponibilizar aos usuários. Em seguida, a busca disponível a partir de uma única caixa de pesquisa inspirada no Google. A terceira característica é a capacidade de resposta que fornece os resultados rapidamente e classificados por relevância, numa interface moderna que oferece as funcionalidades e o *design* intuitivo esperados pelos usuários de hoje. Por último, o autor destaca a flexibilidade que permite que o sistema seja hospedado localmente ou remotamente pelos fornecedores de conteúdos, assim como a personalização para que a biblioteca o adapte ou altere com o intuito de atender a especificações individuais.

Os serviços da descoberta em rede, presumivelmente, apresentam um rol de informações mais exato, relevante e apropriado às necessidades dos usuários do que, por exemplo, uma página *web* que utiliza motores de busca e costuma mostrar os resultados das primeiras páginas baseado em índices de popularidade. Nos serviços de descoberta, o índice é centralizado, controlado e detalhado, baseado em esquemas normalizados para todos os tipos de conteúdos, permitindo buscas rápidas e recuperação de resultados classificados por relevância, o que lhe proporciona maior confiabilidade. Também reduzem a apresentação de itens duplicados e apresentam registros enriquecidos com capas, citações (WOS, JCR, DOI), alertas, RSS⁴ e o texto completo, quando permitido pelo editor.

2.3 Problemas associados aos serviços de descoberta em rede

A centralização dos registros, embora possa fornecer um método eficaz para os serviços de descoberta, na opinião de Patil e Gopal (2011), pode também apresentar problemas, como um ponto único de falha e pontos de estrangulamento. Além disso, as questões relacionadas com a escalabilidade⁵, replicação de dados e manipulação de versões de serviços dos fornecedores devem ser levadas em conta.

Além de fornecer aos usuários os dados solicitados em listas de resultados, há que se ter em mente a necessidade de proporcionar serviços de valor agregado,

num contexto de interação completa entre a ferramenta e o usuário que possibilite a recuperação da informação relevante. Segundo Palmonari e outros (2011), nas buscas em serviços de descoberta em rede os usuários necessitam de certa experiência que lhes permita identificar os dados pertinentes e, em seguida, localizar os serviços relevantes. Para os autores, os dados são as informações detalhadas sobre uma necessidade específica, enquanto os serviços executam processos envolvendo os dados e retornando um resultado informativo.

Neste sentido, pode-se inferir que os autores acreditam que é uma tarefa difícil para o usuário selecionar a informação relevante dentre a vasta gama de informação recuperada, se o serviço de descoberta não propiciar mecanismos que facilitem a seleção. Entre estes mecanismos, pode-se citar a deduplicação de registros⁶; a ferberização⁷; a filtragem dos resultados por autor, título, assunto, entre outros; o agrupamento por similaridade; a visualização prévia dos conteúdos e a possibilidade de classificação por relevância.

Mesmo os relatos mais otimistas sobre os serviços de descoberta apontam deficiências. Lown, Sierra e Boyer (2013) apontam falhas em todos os serviços disponíveis, entre elas, as restrições na customização da interface, a falta de uma cobertura completa de todos os recursos da biblioteca, o alto custo e alguns problemas de usabilidade. Também, a incerteza de qual o espaço que estas ferramentas de descoberta devem ocupar no amplo conjunto de recursos de informação que as bibliotecas oferecem.

2.4 Determinação da relevância dos resultados

Por padrão, uma busca num serviço de descoberta é uma busca por palavras nos metadados e no texto completo dos documentos, a qual retorna as informações classificadas por relevância. Segundo Vaughan (2011b), estes serviços utilizam algoritmos de relevância proprietários, com diferentes pesos atribuídos aos vários campos de metadados, determinados por parâmetros tais como proximidade e frequência. Porém, diferentes parâmetros podem ser aplicados a diferentes tipos de conteúdo. Por exemplo, os cálculos de relevância para artigo de periódicos podem incluir ou não artigos *peer-review* e o número de vezes que o artigo foi citado. Atualidade é um

fator para a determinação da relevância para quase todos os tipos de conteúdo incluídos no índice dos serviços de descoberta. O algoritmo de relevância deve ser continuamente ajustado para ampliar a eficiência da recuperação.

Como um único item pode ser oriundo de diversos fornecedores de conteúdos, os serviços de descoberta criam um registro que mescla e acomoda os metadados mais significativos de cada provedor. Independentemente de saber se o conteúdo de um determinado documento vem de um fornecedor ou de vários fornecedores, o sistema deve corrigir erros e normalizar os dados. Registros mesclados ajudam na eliminação de duplicidade de resultados, que é realizada por rotinas específicas para este fim. Além da melhoria constante do algoritmo de relevância, vários componentes do recurso de recomendação estão avançando com o objetivo de retornar as informações relevantes para atender uma determinada busca.

É essencial que os resultados sejam apresentados de maneira que os itens mais importantes estejam no topo da lista. Para Breeding (2012), os serviços de descoberta disponíveis no mercado já progrediram significativamente na forma de ordenação dos resultados de busca. Mesmo assim, o autor vê a necessidade de continuar melhorando os algoritmos de relevância, para que mesmo os pesquisadores mais sofisticados tenham confiança de que serão brindados com a melhor representação dos recursos disponíveis para atender suas necessidades de informação.

Na *web*, especialmente nos motores de busca do tipo Google, a relevância é baseada no conteúdo completo de páginas e a frequência em que os *links* são clicados pelos usuários, entre outros elementos. Nos serviços de descoberta, em muitos casos, os índices são construídos apenas com escassos metadados da descrição de um item, proporcionando um mínimo de dados para a classificação de relevância. Em um mundo ideal, os serviços de descoberta teriam acesso a metadados robustos, incluindo a indexação por pessoal especializado, além dos cabeçalhos de assuntos fornecidos pelos autores, assim como do resumo e o texto integral de todos os materiais que compõem o acervo da biblioteca. Mas as melhorias mais importantes para a relevância no contexto das bibliotecas, segundo Breeding (2012), envolvem a incorporação de outros fatores que permitem medir a importância de qualquer resultado em relação à consulta do usuário. Os motores de busca da *web* têm utilizado uma abordagem cada vez mais personalizada em relação às buscas, alterando os *rankings*

de relevância de acordo com qualidades inferidas pela localização física, pesquisas anteriores e pelos *clicks* do usuário.

2.5 A indexação facetada

Outro aspecto a considerar ao avaliar o sucesso ou insucesso dos usuários de serviços de descoberta em rede, diz respeito à eficaz utilização da navegação facetada.

Nos serviços de descoberta, as facetas são um conjunto de etiquetas significativas e organizadas de tal forma a refletir os conceitos relevantes para um determinado conjunto de dados. São a representação de categorias, propriedades, atributos, características, relações, funções ou conceitos, que são fundamentais para um conjunto de documentos ou entidades e que podem representar um interesse especial para um grupo de usuários. Segundo Fagan (2010), a navegação facetada oferece ao usuário subcategorias relevantes pelas quais ele pode ter uma visão geral dos resultados e, em seguida, restringi-los. As facetas geralmente incluem autores, assuntos e formatos, mas podem incluir qualquer campo logicamente criado a partir de um registro *Machine Readable Cataloging* (MARC).

As facetas desempenham várias funções em uma interface e geralmente aplicam ontologias simples. De acordo com McGuinness (2003), o impacto das ontologias pode ser notado nos seguintes aspectos: controle de vocabulário, organização e apoio à navegação, estruturas de guarda-chuva para expandir conteúdos, suporte de navegação, suporte à busca e suporte à desambiguação. Essas funções oferecem várias vantagens potenciais para o usuário, pois utilizam sistemas da categoria coerentes, são previsíveis, expõem visualizações prévias para orientar o usuário, mostram como voltar ao estado anterior, sugerem alternativas lógicas e ajudam a evitar conjuntos de resultados vazios quando as buscas são refinadas. Como desvantagem pode-se citar o fato de que as categorias de interesse do usuário nem sempre são conhecidas com antecedência e, portanto, aspectos importantes podem não ser mostrados. Além disso, a construção automática de facetas garante parcialmente o sucesso de uma busca.

Fagan (2010) comprova, a partir de estudos realizados em diversas interfaces, que a navegação facetada é uma técnica conveniente para apoiar a exploração e

a descoberta. O autor acrescenta que as facetas são úteis para a criação de estruturas de navegação e que a categorização facetada auxilia na recuperação eficiente em banco de dados e ajuda a evitar “becos sem saída”. Os usuários são mais rápidos nas buscas quando utilizam um sistema facetado, encontram mais resultados e o sucesso na recuperação de resultados relevantes é maior. Por conseguinte, a satisfação dos usuários é maior num sistema facetado, o que os tornam mais confiantes, pois permite que interajam facilmente com as características específicas das interfaces facetadas e vão além de apenas um clique nas facetas quando é permitido.

2.6 As soluções comerciais

Os serviços de descoberta em rede podem ser tanto de código aberto como proprietários. A luta para conquistar a credibilidade entre as bibliotecas tem se acirrado nos últimos anos, quando surgiu uma variedade de novas soluções, muitas vezes representando diferentes modelos conceituais. As opções de código aberto são preferidas pelos bibliotecários pelo espírito de desenvolvimento colaborativo. Isto tem feito com que as empresas que fornecem sistemas proprietários ofereçam um acesso mais aberto aos seus produtos, com o objetivo de aumentar a interoperabilidade, a extração de dados, a personalização dos recursos e a possibilidade de criar novas funcionalidades.

Nos artigos publicados por Breeding (2010, 2011), Joc e Chang (2010) e Burke (2010) são citados e descritos alguns serviços de descoberta disponíveis no mercado, tanto de código aberto como proprietários. Dentre os principais estão o *Aquabrowser*, *Encore*, *Endeca*, *EBSCO Discovery Service*, *FAST*, *Meresco*, *Primo Central*, *PurpleSearch*, *Summa*, *Summon*, *TouchPoint*, *VuFind* e *Worldcat Local*. O Quadro 1 mostra um comparativo dessas ferramentas levando em conta plataforma, instalação, licença, suporte, treinamento e principais características.

Quadro 1 – Comparação entre sistemas de descoberta em rede

	Aqua browser	Encor	Endeca	EBSCO Discovery	FAST	Meresco	Primo	Purple Search	Summa	Summon	Touch Point	VuFind	Worldcat Local
Plataforma suportada													
Solaris							X				X	X	
Linux						X	X				X	X	
Windows												X	
Opções de hospedagem/ instalação													
Instalação local	X					X	X	X				X	
Hospedagem	X			X			X	X		X			X
Licença													
Proprietária	X	X	X	X			X			X			X
Fonte aberta						X			X			X	
Suporte													
Fornecedor				X			X						X
Terceiros						X							
Comunidade						X						X	
Treinamento													
Fornecedor							X						
Terceiros													
Comunidade												X	
Características													
Indexação centralizada				X			X			X	X	X	X
Busca federada				X			X				X		X
Harvesting / Indexação	X	X	X	X	?	X	X	?	X	X	X	X	X
Resultado por ranking de relevância	X			X		X	X			X	X	X	X
Marcação do usuário	X			X		X	X				X	X	X
Comentários do usuário	X			X		X	X				X	X	X
Guardar itens da busca											X	X	
Invidação de disponibilidade de item (no catálogo)				X		X	X			X	X	X	X
Reserva						X	X				X	X	

	Aqua browser	Encor	Endeca	EBSCO Discovery	FAST	Meresco	Primo	Purple Search	Summa	Summon	Touch Point	VuFind	Worldcat Local
Exibição de capas dos livros	X			X		X	X			X	X	X	X
Refinar resultados por diferentes categorias	X			X		X	X			X	X		
Recomendação de serviços	X		7			X	X			X	X		
Mostrar resultados facetados	X			X		X	X			X	X	X	X
API						X	X			X			
Suporte a dispositivos móveis				X		X	X			X	X	X	X
Sugestões "Você quis dizer"						X	X			X	X	X	X
RSS				X			X			X			

Fonte: Adaptado de Unified Resource Discovery Comparison ([2015?]).

Segundo Burke (2010), os serviços de descoberta levaram algum tempo para chegar ao mercado porque eles exigem um modelo de tecnologia e mão-de-obra inteiramente nova. O acervo das bibliotecas é constituído dos seus próprios materiais e de assinatura de bases de dados de editores comerciais. Por conseguinte, os metadados devem ser coletados tanto do catálogo da biblioteca como dos editores, demandando acordos com milhares deles. Por sua vez, as informações devem ser indexadas uma única vez e posteriormente atualizadas à medida que incrementos e revisões acontecem localmente ou por parte dos editores, o que significa que elas devem ser constantes, a cada minuto. Os metadados que estão sendo inseridos variam em qualidade e tamanho e devem ser mapeados em um esquema rico e consistente para permitir a pesquisa e classificação por relevância. O modelo é realmente o do Google, ou seja, um grande índice que trata cada item de informação igualmente e faz com que cada um seja localizado, rápida e facilmente. A inserção deste volume de dados, não apenas inicialmente, mas de forma contínua, significa que são necessários programas especiais que permitam a atualização constante e imediata dos índices de busca. Assim como o Google indexa milhões de itens a cada hora, um serviço de descoberta bem sucedido deve fazer o mesmo.

2.7 Os usos dos sistemas de descoberta

Way (2010) constata em seu estudo que os serviços de descoberta em rede representam uma mudança dramática na forma como as bibliotecas oferecem acesso às coleções. O armazenamento da informação que existia em bases de assunto, editor ou provedor de conteúdo em muitos aspectos não existe mais ou não são mais importantes. O que importa é que uma fonte indexa a informação e a entrega ao usuário pelo serviço de descoberta da sua biblioteca. Por outro lado, não quer dizer que as ferramentas de descoberta são perfeitas ou que os usuários deixarão de usar o Google para suas pesquisas acadêmicas. Ainda assim, os resultados do estudo sugerem que esta nova categoria de recurso tem o potencial de mudar radicalmente a maneira como os usuários interagem e descobrem a riqueza das informações disponíveis nas coleções da biblioteca.

Entretanto, o serviço de descoberta pode passar aos usuários a visão de que tudo está disponível on-line e em texto completo e isto pode contribuir para a decepção dos mesmos com o serviço. No estudo de Howard e Wiebrands (2011), os bibliotecários entrevistados salientam: “Estamos criando uma expectativa de que todo material está disponível eletronicamente”. Em alguns casos mais extremos, o sistema provocou queixas por parte dos estudantes, por exemplo, quando o resultado da busca indicava a existência de um livro fisicamente localizado na prateleira da biblioteca, mas que não estava disponível na forma digital e *on-line* integral, fato que era visto como um fracasso aos olhos dos estudantes.

O comportamento de busca e recuperação da informação dos usuários que utilizam serviços de descoberta nas bibliotecas acadêmicas poucas vezes tem sido objeto de estudo. As pesquisas com frequência abordam aspectos como a quantidade de acessos ao serviço e a bases de dados específicas, porém aspectos relacionados à interação do usuário com o serviço, o que ele espera encontrar, de que forma, como se comportam perante os resultados recuperados e às possibilidades apresentadas pelo serviço ainda são uma incógnita. Da mesma forma, não são abordados aspectos sobre as motivações ou necessidade de informação que o fazem procurar o serviço de descoberta e não o catálogo tradicional da biblioteca ou um recurso de busca fora da área acadêmica.

3 Conclusão

O impacto do Google nas bibliotecas acadêmicas tem sido implacável na última década, modificando a experiência e expectativas dos usuários em relação à busca e recuperação da informação e por isso é amplamente discutido nestes espaços. O usuário espera uma experiência de busca simplificada, rápida, que dê acesso ao conteúdo dos documentos on-line, principalmente que espelhe o uso do Google e outros motores de busca. Evidentemente, os serviços de descoberta em rede são uma tentativa de atender a essa necessidade.

Um dos recursos mais poderosos dos serviços de descoberta em rede é a sua capacidade para atender às expectativas dos usuários a partir de um ponto único de acesso, apoiados por uma robusta e abrangente ferramenta de busca que recupera conteúdo de relevância científica e acadêmica.

O objetivo deste artigo foi sistematizar a compreensão sobre os serviços de descoberta em rede e mostrar a necessidade de estudos mais aprofundados para determinar se a combinação de resultados em recursos heterogêneos, em uma única lista, ajuda ou atrapalha os usuários das bibliotecas. Devem-se buscar maneiras para auxiliá-lo a remover a ambiguidade dos vários tipos de recursos e a navegar de forma eficiente no vasto ambiente de informação que o sistema oferece.

Uma questão importante a ser investigada é a percepção dos usuários quando são apresentados a uma plataforma de descoberta como principal mecanismo de pesquisa no *site* de uma biblioteca. Uma caixa de busca única pode provocar expectativas significativas nos usuários, sendo importante considerar os problemas que podem ocorrer quando a capacidade da caixa de busca não atende às suas expectativas, pois ele pode assumir falsamente que as funções de pesquisa de biblioteca são semelhantes às de um motor de busca da *web* e que todo o conteúdo da instituição, acadêmico ou não, será acessado digitalmente.

Embora as bibliotecas estejam inclinadas a projetar suas *homepages* em torno de uma caixa única de busca, para dar uma impressão de simplicidade e funcionalidade, a cobertura inadequada de recursos pode frustrar os usuários e os recursos informacionais significativos da biblioteca podem ficar mascarados. Portanto, torna-se

imprescindível investigar como os usuários buscam, recuperam e o que esperam ao pesquisar o conteúdo da biblioteca em um único ambiente de busca

As bibliotecas devem revisitar o alcance de seus catálogos para manter-se conectadas com os modelos mentais dos usuários e suas necessidades de informação. As bibliotecas devem entender como situar a ferramenta de descoberta e as ferramentas especializadas na web de uma maneira que faça sentido para os seus usuários.

Novos recursos de informação, tais como os serviços de descoberta em rede, têm o potencial de melhorar a experiência de busca dos usuários das bibliotecas universitárias, mas apenas se for possível integrá-los com outros serviços de informação e recursos, para que eles sejam úteis e compreensíveis. Estas ferramentas devem ser projetadas cuidadosamente, baseadas na experiência de busca e comportamento dos usuários para equilibrar as necessidades e expectativas dos mesmos com as capacidades do sistema.

Referências

BREEDING, M. The state of the art in library discovery 2010. **Computer in Libraries**, v. 30, n. 1, p. 31-34, 2010. Disponível em:

<<http://www.librarytechnology.org/ltg-displaytext.pl?RC=14574>>. Acesso em: 8 set. 2015.

BREEDING, M. The new frontier: automation marketplace 2011. **Library Journal**, v. 136, n. 6, p. 24-6, 28-34, 2011. Disponível em: <

<http://lj.libraryjournal.com/2011/03/library-services/automation-marketplace-2011-the-new-frontier/>>. Acesso em: 8 set. 2015.

BREEDING, M. Library web-scale. **Computers in Libraries**, v. 32, n. 1, p. 19-22, Jan. 2012. Disponível em: <[http://www.librarytechnology.org/ltg-](http://www.librarytechnology.org/ltg-displaytext.pl?RC=16577)

[displaytext.pl?RC=16577](http://www.librarytechnology.org/ltg-displaytext.pl?RC=16577)>. Acesso em: 8 set. 2015.

BURKE, J. Discovery versus disintermediation: the new reality driven by today's end-user. IN: BIENNIAL CONFERENCE AND EXHIBITION, 15., 2010, Melbourne, Australia. **VALA2010 Proceedings**. Melbourne: VALA - Libraries, Technology and the Future Inc., 2010. Disponível em:

<http://www.vala.org.au/vala2010/papers2010/VALA2010_57_Burke_Final.pdf>.

Acesso em: 8 set. 2015.

CODINA, L.; ABADAL, E.; ROVIRA, C. Búsqueda federada en el ecosistema de la e-ciencia: el caso Science Research. **El Profesional de la Información**, v. 19, n. 1, p. 77-85, enero/feb. 2010. Disponível em:

<<http://www.elprofesionaldeinformacion.com/contenidos/2010/enero/11.pdf>>.

Acesso em: 8 set. 2015.

DEMPSEY L. **Web scale**. Lorcan Dempsey's Weblog: on libraries, services and networks [blog]. 05 jan. 2007. Disponível em:

<<http://orweblog.oclc.org/archives/001238.html>>. Acesso em: 8 set. 2015.

FAGAN, J. D. Usability studies of faceted browsing: a literature review. **Information Technology and Libraries**, p. 58-66, June 2010. Disponível em:

<<http://ejournals.bc.edu/ojs/index.php/ital/article/view/3144>>. Acesso em: 08 set.

2015.

HOWARD, D.; WIEBRANDS, C. Culture shock: librarians' response to web scale search. In: ALIA INFORMATION ONLINE CONFERENCE AND EXHIBITION, 15., 2011, Sydney. **Conference Papers...** Sydney: Australian Library and Information Association, 2011. Disponível em:

<<http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=7208&context=ecuworks>>. Acesso em: 8 set. 2015.

JOC, K.; CHANG, K. The impact of discovery platforms on the information-seeking behaviour of EFL undergraduate students. In: BIENNIAL CONFERENCE AND EXHIBITION, 15., 2010, Melbourne. **VALA2010 Proceedings**. Melbourne: VALA - Libraries, Technology and the Future Inc., 2010. Disponível em:

<http://www.vala.org.au/vala2010/papers2010/VALA2010_122_Joc_Final.pdf>.

Acesso em: 8 set. 2015.

LOWN, C.; SIERRA, T.; BOYER, J. How Users Search the Library from a Single Search Box. **College & Research Libraries**, may 2013, p. 227-241.

LUTHER, J.; KELLY, M. C. The next generation of discovery. **Library Journal**, v. 136, n. 5, p. 66-71, 2011. Disponível em:

<<http://lj.libraryjournal.com/2011/03/technology/the-next-generation-of-discovery/>>. Acesso em: 8 set. 2015.

MARANHÃO, A. M. N. Dos catálogos aos metabuscadores e serviços de descoberta na internet: uma visão geral. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 24., 2011, Maceió. **Anais...** Maceió: UFAL, 2011. 13 p. Disponível em:

<http://www.academia.edu/1512956/Dos_Cat%C3%A1logos_aos_Metabuscadorese_Servi%C3%A7os_de_Descoberta_na_Internet>. Acesso em: 8 set. 2015.

McGUINNESS, D. L. Ontologies come of age. In: FENSEL, D. et al. (Ed.). **Spinning the semantic web: bringing the World Wide Web to its full potential**. Cambridge, MIT Press, 2003, p. 179-184. Disponível em: <[http://www-ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-\(with-citation\).htm](http://www-ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-(with-citation).htm)>. Acesso em: 8 set. 2015.

OLIVEIRA, José Palazzo Moreira de. O mito da Geração Google. 2011. Disponível em: <http://palazzo.pro.br/Joomla/index.php?option=com_content&view=article&id=102:cronica-28&catid=78&Itemid=549&lang=pt-BR>. Acesso em: 29 set 2015.

PALMONARI, M. et al. Aggregated search of data and services. **Information Systems**. v. 36, n. 2, p. 134-150, Apr. 2011. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0306437910000979>>. Acesso em: 8 set. 2015.

PATIL, N.; GOPAL, A. Comparative study of mechanisms for web service discovery based on centralized approach focusing on UDDI. **International Journal of Computer Applications**. v. 14, n. 1, p. 28-31, Jan. 2011. Disponível em: <<http://www.ijcaonline.org/volume14/number1/pxc3872328.pdf>>. Acesso em: 8 set. 2015.

SIQUEIRA I. C. P. Mecanismos de busca na *web*: passado, presente e futuro. **Ponto de Acesso**, Salvador, v.7, n.2, p. 47-67, ago. 2013. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/6355/6136>>. Acesso em: 8 set. 2015.

UNIFIED RESOURCE Discovery Comparison. [2015?]. Disponível em: <<https://sites.google.com/site/urd2comparison/home/comparison>>. Acesso em: 8 set. 2015.

VAUGHAN, J. Web scale discovery: what and why? **Library Technology Reports**, v. 47, n. 1, p. 5-11, Jan. 2011a.

VAUGHAN, J. Serials solutions Summon. **Library Technology Reports**, v. 47, n. 1, p. 22-29, Jan. 2011b.

WAY, D. The impact of web-scale discovery on the use of a library collection. **Serials Review**, v. 36, n. 4, p. 214-220, Dec. 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0098791310000882>>. Acesso em: 8 set. 2015.

Agradecemos à CAPES pela concessão de bolsa de doutorado sanduíche na Universidade Complutense de Madrid para a primeira autora e ao Prof. Dr Juan Antonio Martínez Comeche, pelo auxílio na condução da pesquisa de campo.

Web Services Discovery: a Google experience for users in University Libraries

Abstract: The article contextualizes the use of web services discovery in university libraries, which intend to meet the need of increasingly demanding users when it comes to performance during information seeking and retrieval. The research work aims to contribute to the understanding of these tools, stressing the building of its central index, the structure of faceted search, the relevance criteria used and the presentation of results in a single list. It employed the bibliographical research method for selecting the relevant national and international literature on the subject. It concludes that libraries should revisit the scope of their catalogs to stay connected with the mental models of users and their information needs. Web scale discovery services, integrated with other information services, have the potential to improve the searching experience. These tools should be designed based on users' information seeking behavior and experience to balance their needs and expectations with the capabilities of the system.

Keywords: Web Services Discovery. University Library. Information Seeking and Retrieval Behavior.

Recebido em 10/09/2015
Aceito em 27/10/2015



¹ Metadados são informações estruturadas que possuem semântica padronizada, utilizados para representar as informações digitais de maneira bibliográfica. Disponível em:

<<http://hdl.handle.net/10183/12730>>. Acesso em: 12 ago. 2013.

² “I like the expression *web-scale*. It is used heavily by Amazon and others in discussion of their 'platform' services like S3 and EC2.” (DEMPSEY, 2007).

³ O Protocolo OAI-PMH é uma especificação aberta de coleta e transferência de registros de metadados, é utilizado para garantir a interoperabilidade entre sistemas. (*THE OPEN...*, 2008).

⁴ *Really Simple Syndication* é um padrão desenvolvido em linguagem XML que permite aos responsáveis por sites e blogs divulgarem notícias ou novidades destes. Disponível em: <http://www.infowester.com/rss.php>. Acesso em: 06 abr. 2014.

⁵ Escalabilidade é a característica de um sistema ou uma função que descreve a sua capacidade para lidar e atuar sob uma carga de trabalho aumentada ou expandida. Um sistema que se adapta bem será capaz de manter ou mesmo aumentar o seu nível de desempenho ou de eficiência quando submetido a grandes demandas operacionais. Disponível em: <<http://www.investopedia.com/terms/s/scalability.asp#axzz1abmeRsR2>>. Acesso em: 12 out. 2011.

⁶ Deduplicação de registros é a eliminação de registros duplicados via software.

⁷ Ferberização é a integração das diferentes expressões e manifestações de uma obra em um mesmo resultado.