

DOS SISTEMAS DE INFORMAÇÃO FEDERADOS À FEDERAÇÃO DE BIBLIOTECAS DIGITAIS

Sueli Mara S. P. Ferreira
Leonardo Fernandes Souto

RESUMO

As bibliotecas digitais foram importantes recursos informacionais surgidos no final do século passado. Porém, logo se mostraram incapazes de satisfazer a totalidade das necessidades e demandas de informação dos usuários por vários motivos, tais como, dispersão das fontes, divergências nas interfaces de busca, falta de integração dos conteúdos dentre outros. Como decorrência surgem os projetos de Federação de Bibliotecas Digitais. O objetivo desse *paper* é apresentar o conceito de Sistemas de Informação Federados e tecnologias de recuperação de informação correlatas, buscando caracterizar tanto as atuais propostas de gestão como as tecnologias utilizadas em projetos de Federação de Bibliotecas Digitais. Inclui uma discussão sobre sistemas de informação federados, busca federada e sistema de coleta de metadados/*harverting*, identifica alguns aplicativos de *harvesting* disponíveis para construção da infraestrutura tecnológica para sistemas dessa natureza e, finalmente, descreve algumas iniciativas nacionais e internacionais de Federação de Bibliotecas Digitais.

PALAVRAS-CHAVE

Sistemas de informação federados; Federação de bibliotecas digitais; Busca federada; Coleta de metadados/*harvesting*; Provedores de serviço

FROM FEDERATED INFORMATION SYSTEMS TO A DIGITAL LIBRARIES FEDERATION

ABSTRACT

Digital libraries have been important informational resources which appeared in the end of the last century. However, they have not been effective in satisfying the whole set of users' information needs and demands. This ineffectiveness is drive by technical factos which include the distribution of sources, distinction in the search interfaces, lack of integration of contents. As one of the solutions for these problems, the federation of digital libraries approach has been suggested. The goal of this paper is to present the concept of Information System Federation and to characterize the management proposal and technologies involved in projects that consider the approach of Federation of Digital Libraries today. Hence, the paper includes a discussion about federated systems of information, federated search and about harvesting, and it also identifies some of the available software in the market that enable the implementation of a harvesting system. Finally, it describes some national and international initiatives of Federation of Digital Libraries.

KEYWORDS

Information system federation; Federation of digital libraries; Federated search; Harvesting; Service providers

1 INTRODUÇÃO

Com o exponencial crescimento de informações registradas em diferentes fontes tais como, *sites*, bases de dados e sistemas na Web, emerge uma demanda crescente por novas tecnologias e pesquisas que suportem, com maior efetividade, as funcionalidades de pesquisa e indexação de conteúdos, bem como as necessidades e especificidades dos usuários.

As bibliotecas digitais foram importantes recursos informacionais surgidos no final do século passado. Porém, logo se mostraram incapazes de satisfazer a totalidade das necessidades e demandas de informação dos usuários por vários motivos.

Para Marcondes e Sayão (2001, p. 26) as bibliotecas digitais “estão sendo criadas em número crescente, armazenadas em diferentes servidores isolados, operadas por interfaces de busca diferentes, o que obriga um usuário a uma dispendiosa busca, *site a site*, para encontrar informações relevantes” e “do ponto de vista de um usuário acadêmico ou pesquisador, o interessante e confortável seria poder submeter sua necessidade de informação e interagir com uma *única interface*, recebendo as informações de diferentes fontes, de forma consolidada”.

Visando atender a essa demanda, os projetos de bibliotecas digitais deixam de ser idealizados somente a partir da concepção de sistemas de informação simples e isolados e se voltam ao desenvolvimento de sistemas de informação estruturados a partir da integração de dados de diferentes sistemas. Nascem, assim, os projetos de Federação de Bibliotecas Digitais.

O objetivo deste *paper* é apresentar o conceito de Sistemas de Informação Federados e tecnologias de recuperação de informação correlatas, buscando caracterizar tanto as atuais propostas de gestão como as tecnologias utilizadas em projetos de Federação de Bibliotecas Digitais.

2 SOBRE SISTEMAS DE INFORMAÇÃO FEDERADOS

Iniciar uma discussão sobre Sistemas de Informação Federados, ou mesmo sobre Federação de Bibliotecas Digitais, significa repensar tais sistemas sob dois aspectos: como uma nova forma de gestão ou como um novo suporte tecnológico. Tais aspectos estão relacionados ao uso atual que se faz do termo “federação” por profissionais de diferentes áreas.

No que se refere aos aspectos de gestão, uma “federação” é concebida como um grupo de instituições independentes, trabalhando juntas de modo formal ou informalmente, que concordam em dar suporte a um conjunto de serviços comuns e padronizados.

Federações, portanto, atuam fortemente embasadas no compartilhamento de esforços, na cooperação de equipes, na convergência e integração das atividades, dos serviços, das práticas, dos procedimentos, dos conteúdos, das tecnologias e dos atores envolvidos em cada um dos processos de gestão da informação.

Trata-se, portanto, de um novo modelo de gestão com valores agregados tanto aos produtores e geradores de conteúdos como aos usuários finais a procura desses conteúdos. Tais valores estão relacionados, dentre outros, à maior escalabilidade nos resultados de busca, integração de conteúdos dispersos, maior agilidade no atendimento às necessidades da comunidade, maior visibilidade da produção local a nível global, garantia de interoperabilidade do conteúdo e, ainda, enorme economia em termos de recursos humanos, financeiros e tecnológicos.

Do ponto de vista tecnológico, o conceito de Federação deriva de antigos estudos sobre integração de dados estocados em distintos sistemas. Segundo Busse et al (1999) na década de 80 o foco principal estava em coletar dados dispersos em muitas aplicações proprietárias e incompatíveis por meio de um sistema de gestão de bases de dados centralizadas (DBMS).

Já na década de 90, a necessidade era combinar dados estocados em diferentes sistemas de gerenciamento de bases de dados. Como resultado surgem os *multi-database systems* e, mais tarde, emerge o termo *bases de dados federadas* que, para Sheth e Larson (1990) apud Medeiros e Alencar (1999), pode ser definido como uma coleção de sistemas de bancos de dados

independentes, cooperativos, possivelmente heterogêneos, que são autônomos e que permitem o compartilhamento de todos ou alguns de seus dados, sem afetar as suas aplicações locais.

Como resultado do *Workshop Engineering Federated Database Systems (EFDBS 1997)*¹, em Barcelona na Espanha, a comunidade científica - observando a necessidade de integrar também diversificados sistemas digitais às bases de dados federadas - propõe uma nova terminologia: sistema de informação federado (SIF).

Segundo Wehr (1992) sistema de informação federado significa uma nova geração de softwares, cuja principal tarefa é operar de maneira global com capacidade de envolver as diversificadas fontes de dados existentes. Sendo essas fontes revestidas das mesmas características já mencionadas de heterogeneidade, autonomia e distribuição, mas agora possibilitando a inclusão daquelas com informação semi-estruturada e até mesmo não estruturada. Isso, na visão de Conrad et al. (1999, p.1) representa uma ampliação da noção clássica de sistemas de bases de dados federadas e *multi-database systems*.

Atualmente, no que se refere à recuperação de informação em sistemas de informação distribuídos, diversos autores propõem diferentes tipologias. A proposta de maior aceitabilidade na literatura (LIU et al., 2002; SULEMAN, 2002; DOWNES, 2003 e outros)², até o momento, identifica dois tipos de tecnologia: a busca federada e o sistema de coleta de metadados/*harvesting*.

A figura 1, a seguir, proposta por Suleman (2002, p.15), ilustra essas duas tecnologias.

¹ Outros eventos ocorreram, sob o nome de Engenharia de Sistemas de Informação Federados/*Engineering Federated Information Systems*- EFIS: 2003 em Coventry/UK, 2001 em Berlin/Germany; 2000 em Dublin/Ireland e em 1999 em Kühlungsborn, Germany.

² Tanto LIU (2002) como Downes (2003) mencionam o princípio de uma terceira forma de recuperação de informação que denominam de "*peer-to-peer search systems*".

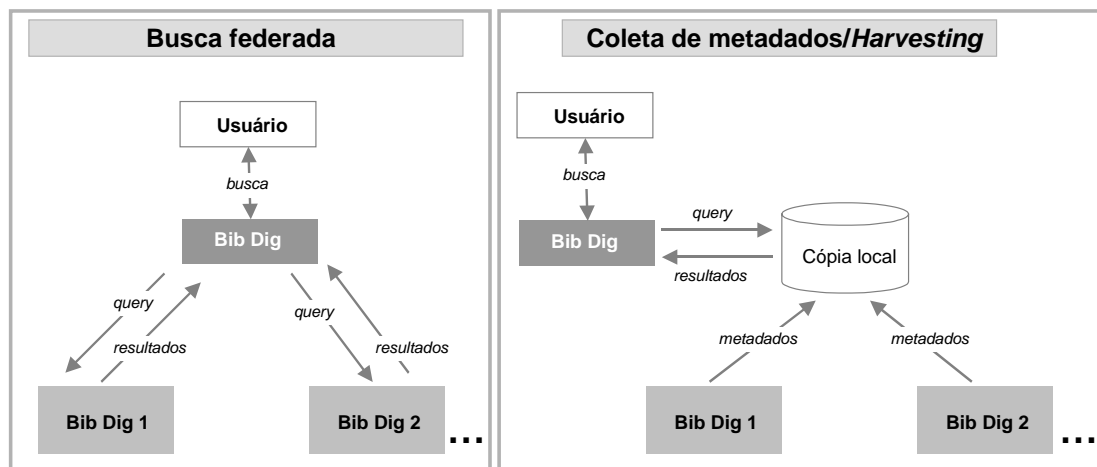


Figura 1: Modelo de busca federada e de coleta de metadados/harvesting (SULEMAN, 2002, p. 15)

2.1 BUSCA FEDERADA

O termo “busca federada” possui vários sinônimos, apresentados por diferentes autores: metabusca, busca distribuída, busca paralela, portal de busca, *broadcast search*, *cross-database search*.

Essa variação terminológica é possível de ser encontrada inclusive nos estudos da NISO-MI – NISO *Metasearch Initiative*³. Segundo Fryer (2004, p. 17) isso ocorre devido à divergência de pontos de vista entre os vendedores de soluções especializadas (como os conhecidos metabuscadores Google, All the Web e Alta Vista) e aquelas oferecidas por sistemas como Dogpile, Vivisimo ou Metacrawler, por exemplo.

Adotando a expressão mais frequentemente utilizada – busca federada - entende-se que essa busca consiste na transmissão de uma pergunta, com a sintaxe apropriada, a um grupo de bases de dados dispersas, agrupando os resultados coletados das bases de dados, e apresentando-os em um formato sucinto e unificado, com duplicação mínima e permitindo que o usuário da biblioteca classifique o resultado agrupado, ajustando-o por meio de vários critérios (JACSO, 2004, p. 17).

Detalhando mais essa definição, Wadham (2004, p. 20) define a busca federada como uma técnica pela qual as fontes de informação múltiplas podem ser pesquisadas ao mesmo tempo,

incluindo catálogos de bibliotecas, bases de dados comerciais, mecanismos de busca *Web* e outros metadados ou bases de dados digitais, sendo que o mecanismo de busca federado, usando um protocolo de busca tal como o Z39.50 por exemplo, recolhe a informação a partir de várias fontes e fornece um resultado de busca integrado ao usuário.

Por conduzir busca separada em cada uma das coleções (ou repositórios) integrantes, a busca federada é mais adequada para Federação com poucos integrantes, mas detentores de grandes coleções, pois segundo Liu et al. (2002) o contrário acarretaria problemas de escalabilidade. Segundo Downes, (2003) tal busca seria praticamente impossível em ambientes grandes e distribuídos como a web, por exemplo, pois a *query* de busca deveria ser replicada dezena, centenas e até milhares de vezes nos vários sites individuais disponíveis.

Em particular, esse tipo de busca tem sido muito utilizado em OPAC's (*Online Public Access Catalog*). Um exemplo brasileiro de busca federada é o CRUESP/Bibliotecas, que é um consórcio que tem por missão a consolidação do trabalho participativo e integrado dos Sistemas das Bibliotecas das Universidades Estaduais Paulistas (USP, UNESP e UNICAMP), com foco na cooperação, compartilhamento e a racionalização dos recursos. Iniciou suas atividades em 1999 e conta atualmente com 89 bibliotecas. Disponibiliza seus produtos por meio de um Portal, com a integração de diferentes recursos, utilizando o protocolo Z39.50. Dentre os produtos destacam-se: *Unibibli WEB* (um sistema de busca federada que permite a consulta simultânea aos catálogos bibliográficos das três universidades) e *ERL* – Electronic Reference Library – (por meio do Portal da Pesquisa - <http://www.portaldapesquisa.com.br/> - oferece um recurso de busca federada que permite a consulta simultânea a diferentes bases de dados referenciais). O acesso pode ser feito pelo site <http://bibliotecas-cruesp.usp.br> .

Dentre os sistemas de busca federada internacional existentes, dois merecem destaque por sua completude. O projeto MERLOT - *Multimedia Educational Resource for Learning and Online Teaching* (<http://www.merlot.org/Home.po>) - que se refere ao desenvolvimento cooperativo e gratuito de recursos baseados na web para que professores, alunos e profissionais em geral possam facilmente encontrar e disponibilizar materiais digitais de aprendizagem com suas respectivas avaliações e indicações de usos mais apropriados. Sua busca federada

³ Maiores informações: http://www.niso.org/committees/MS_initiative.html

(<http://fedsearch.merlot.org/main/search.jsp>) é realizada entre os bancos de dados de dezenas de universidades, faculdades isoladas, associações de pesquisadores, bibliotecas digitais e empresas americanas, podendo ser selecionadas as áreas de *Multimedia Educational Resource for Learning and Online Teaching* (Merlot Physics) ou *Digital Resource Collections for Physics and Astronomy Education* (Compadre).

Outro sistema é o sistema EduSource (<http://209.87.57.50:8080/axis/search/index.html>), sob coordenação do Grupo Canarie do Canadá, cuja proposta é a criação de uma robusta infraestrutura tecnológica para o compartilhamento nacional de repositórios de objetos de aprendizagem. Esse projeto é considerado um dos mais importantes projetos para a nação no século 21. Sua fase inicial foi o inventário das ferramentas, sistemas, protocolos e práticas disponíveis no país, de modo a definir componentes para um framework que garantisse a interoperabilidade entre as várias instituições. Hoje já existem cerca de 11 instituições participando desse projeto e disponibilizando seus conteúdos para a busca federada.

2.2 SISTEMA DE COLETA DE METADADOS/HARVESTING

Em Federações que utilizem o sistema de coleta de metadados/harvesting existe uma facilidade de busca centralizada intitulada “*harvester*” ou “*agregador*” que visita cada um dos repositórios parceiros e autônomos coletando os metadados disponíveis, abastecendo um repositório global com eles e oferecendo uma interface única de busca. A partir daí, os usuários são direcionados diretamente ao registro original e/ou documento completo localizado nas diversas coleções locais dos parceiros.

A vantagem dessa coleta em relação à busca federada é que a existência de um banco dos metadados assegura que sempre seja oferecida uma resposta ao usuário final, além de garantir a preservação do próprio metadado ao longo do tempo. Enquanto na busca federada, problemas no servidor ou alteração de endereço levam a perda integral dos dados (ANAN et.al., s.d.).

A coleta de metadados/*harvesting* ganhou maior ênfase a partir do surgimento da OAI-*Open Archive Initiative* - Iniciativa dos Arquivos Abertos⁴ em 1999, que propõe um *framework* técnico e organizacional para coleta de metadados padronizados de modo a facilitar a recuperação do conteúdo estocado em repositórios digitais distribuídos. Segundo Feng, Jeusfeld e Hoppenbrouwers (2001), esse *framework* é o mais importante esforço da comunidade científica para garantir a interoperabilidade entre coleções digitais, pois oferece o protocolo OAI/PMH – *Protocol Metadata Harvesting* para uso e aplicação em larga escala.

Além do protocolo OAI/PMH, essa Iniciativa preconiza o uso de softwares livres, arquitetura aberta e pautada no conceito de “provedores de dados” e “provedores de serviços”. Os “provedores de dados” são entidades que implementam e gerenciam repositórios digitais utilizando o Protocolo OAI-PMH e expõem assim, seus metadados para serem coletados por provedores de serviço ou agregadores. O “provedor de serviço” implementa e gerencia a ferramenta que coleta automaticamente os metadados expostos pelos provedores de dados, via o Protocolo OAI-PMH, os organiza e oferece produtos e serviços de valor agregado ao usuário final via interface única de acesso.

Com base nessa Iniciativa, vários softwares vêm sendo desenvolvidos tanto para a implementação e criação de provedores de dados⁵ como para provedores de serviços. Os softwares já existentes e disponíveis para a consolidação de um sistema de recuperação para uma Federação pautada em coleta de metadados/*harvesting* estão apresentados no quadro a seguir.

⁴ Por “arquivos” entende-se todo e qualquer sistema, banco de dados, repositório ou biblioteca digital que disponibilize conteúdo digital via web; e “abertos” entende-se a característica desse arquivo no que se refere a ser público, livre, gratuito e de fonte aberta.

⁵ Exemplos de softwares para implementar provedores de dados já em uso no Brasil são: o SEER- Sistema de Editoração Eletrônica da Revistas customização feita pelo IBICT do OJS – *Open Journal System*; o DICI – Diálogo Científico para repositórios temáticos, customização feita pelo IBICT do software Eprints e o Dspace – customização feita pela Portcom/Intercom do software desenvolvido pelo MIT e HP.

Quadro 1 – Provedores de Serviços

PROVEDORES DE SERVIÇO	CARACTERÍSTICAS
<p>ARC – A Cross Archive Search Service</p> <p>http://arc.cs.odu.edu</p>	<p>É um produto de uso livre desenvolvido pela <i>Old Dominion University Digital Library Reserach Group</i> (ODU) a qual tem se dedicado a projetos de softwares que provêm busca através de repositórios que utilizam a funcionalidade de metadados coletados via OAI. Foi o primeiro aplicativo desenvolvido para cumprir o papel de provedor de serviço, oferecendo assim uma interface para pesquisa de metadados coletados em arquivos abertos.</p>
<p>METALIS</p> <p>http://metalisp.cilea.it</p>	<p>É um provedor de serviço que permite a busca em metadados sobre Biblioteconomia e Ciência da Informação, coletados a partir de arquivos abertos por meio do protocolo OAI-PMH. É possível utilizar a mesma arquitetura aberta em projetos semelhantes voltados para áreas específicas ou instituições. O código de programação da arquitetura pode ser copiado a partir de http://metalisp.cilea.it/credits.html.</p>
<p>OAIster</p> <p>http://oaister.umdl.umich.edu/o/oaister</p>	<p>É um projeto da <i>University of Michigan Digital Library Production Service</i>. Tem por objetivo criar uma coleção digital de livre acesso e faz a coleta de metadados a partir de repositórios institucionais estruturados com base na OAI.</p>

Fonte: Elaboração própria

3 FEDERAÇÃO DE BIBLIOTECAS DIGITAIS

Conforme já mencionado, muitos paradigmas relacionados ao desenvolvimento de softwares têm mudado, pois os sistemas modernos estão sendo estruturados observando-se a interoperabilidade entre eles, mas, principalmente, respeitando-se a autonomia entre seus componentes. Essa idéia é fundamental para a Federação de Bibliotecas Digitais, pois diferentemente das áreas de engenharia de software em que os componentes interoperáveis podem representar as partes de um software, aqui os componentes correspondem aos vários provedores de dados partícipes da Federação. Isto significa que se pode estar falando de várias instituições distintas técnica, geográfica e administrativamente (BUSSE ET AL., 1999, p. 3).

Segundo Shreeves et al. (2005, p. 223), “a federação de recursos digitais tem se tornado muito importante na realização do potencial total das bibliotecas digitais” e “é freqüentemente alcançada através da agregação de metadados descritivos” sendo assim, “as decisões que os desenvolvedores de recursos tomam quanto à criação, manutenção e garantia da qualidade de seus metadados podem ter um significativo impacto sobre os agregadores e provedores de serviços”.

Revedo a evolução das tecnologias de bibliotecas digitais⁶, sobretudo as americanas iniciadas com recursos federais provenientes da *National Science Foundation* (NSF), Mischo (2005) sugere que ocorreram duas fases de projetos: a primeira em 1994, com a fundação de seis projetos no programa Digital Libraries Initiative, intitulado de DLI-1 e a segunda fase em 1998, com o DLI-2 que emergem dois projetos protótipos.

A DLI-1 tinha como proposta desenvolver e implementar tecnologia de rede e computação para fazer com que coleções eletrônicas de larga escala fossem interoperáveis e acessíveis entre si. Uma série de produtos foi derivada dessa fase, entre eles: a importante contribuição do projeto da *Illinois University* sobre tecnologias para a publicação on-line de coleções em pequena escala (por exemplo, dos periódicos científicos possibilitando o acesso integral a seus fascículos) dentre elas: apresentação de texto completo usando HTML e *Cascading Style Sheets* (CSS), links internos entre citações e notas de rodapé; links para artigos citados usando o padrão DOI (Digital

Objetc Identifier) e ainda as OpenURLs dentre outras⁷. Essa fase foi fundamental para o desenvolvimento de normas, padrões e guias de boas práticas disponíveis hoje para bibliotecas digitais.

A DLI-2 é marcada por uma colaboração internacional feita entre os projetos da *Cornell University* e o Projeto *ePrint* do Reino Unido visando o desenvolvimento e adaptação do protocolo e especificações do OAI-PMH. Embora, os projetos do DLI-1 sirvam para identificar e definir importantes documentos e metadados padrões, é na fase 2 que surgem os estudos relativos a sua Federação, pois os projetos da DLI-1, ainda segundo Mischo (2005), não tinham os requisitos tecnológicos para federar e integrar o que estava disperso em um universo de recursos de informação distribuídos. Como decorrência, surgem as pesquisas para o desenvolvimento de mecanismos de busca distribuídos com o *NSDL Core Integration Project*⁸, e os projetos de federação de bibliotecas digitais respaldados no desenvolvimento tecnológico proveniente dos sistemas de informação federados conforme já mencionado.

Arms (2000) menciona que o termo biblioteca digital federada “descreve um grupo de organizações, trabalhando juntas de maneira formal ou informalmente, que concordam em assumir uma série de serviços e padrões comuns, buscando a interoperabilidade entre elas. Seus parceiros podem ter diferentes plataformas, desde que implementem serviços compatíveis com os padrões determinados. Necessitam concordar tanto com os padrões tecnológicos, quanto com as políticas e diretrizes (incluem, portanto normas de direitos autorais, segurança, privacidade e acordos financeiros)”.

Autores como Feng, Jeusfeld e Hoppenbrowsers (2001) acreditam que o futuro das bibliotecas digitais vai além do armazenamento e do registro de conteúdos e sim, que passam a ser visualizadas como um ambiente de criação, aquisição, compartilhamento e propagação do conhecimento, deixando entrever assim que seu foco está cada vez mais centrado no desenvolvimento das competências informacionais dos indivíduos.

⁶ Vale lembrar que embora não intitulado como biblioteca digital o projeto de Vannevar Bush em 1945 já traduzia tal idéia.

⁷ Pode-se mencionar ainda que no Brasil, a BIREME desenvolve, nessa mesma época, a tecnologia SCIELO utilizando o conceito de marcação das referências e citações de modo a permitir indicadores úteis para estudos bibliométricos e análise dos fatores de impacto dos periódicos incluídos em sua coleção.

⁸ Projeto iniciado em janeiro de 2003, com previsão de término em dezembro de 2006. Maiores informações: http://nsdl.org/resources_for/library_builders/projects.php?this_sort=pi&keyword=&project_id=227656

Nesse sentido, Coleman e Oxnam (2002) corroboram quando descrevem que as bibliotecas digitais cada vez mais devem ser visualizadas como sendo ambientes integrados pela composição de *espaços de informação*, de *aprendizado* e de *interação*. Por *espaço de informação* se refere às inúmeras coleções (conteúdos cada vez mais complexos, heterogêneos, distribuídos, estruturados ou semi e não estruturados) que necessitam ser organizadas e gerenciadas de maneira integrada. Por *espaço de aprendizagem*, a biblioteca deve se apresentar como uma possibilidade de oferecer novas estruturas e combinações dos e entre os conteúdos focando-se o aprendizado do usuário e seus respectivos estilos de aprendizagem. Finalmente, enquanto *espaço de interação* a Federação deve proporcionar situações e recursos que libertem o usuário das limitações físicas para a interação, valendo-se das tecnologias de interação virtual corroborando, assim, com a construção coletiva de conhecimento.

Apesar dos estudos sobre Federações de Bibliotecas Digitais serem considerados recentes, já é possível encontrar projetos bem estruturados e com bons resultados, como as experiências nacionais e internacionais citadas, a seguir, a título de ilustração.

3.1 Iniciativas brasileiras de federações de bibliotecas digitais

- **Biblioteca Digital de Teses e Dissertações (BDTD):** é um projeto que visa à integração dos sistemas de informação de teses e dissertações das Instituições de Ensino Superior (IES) brasileiras, bem como estimular o registro e a publicação de teses e dissertações em meio eletrônico. Conta com um modelo distribuído e faz uso das tecnologias de arquivos abertos, sendo as IES provedores de dados e o IBICT um agregador, o qual coleta metadados de teses e dissertações destes provedores de dados e fornece serviços de informação sobre esses metadados, expondo-os para serem coletados por outros provedores de serviços. No caso de IES que ainda não implementaram suas bibliotecas digitais de teses e dissertações, o IBICT oferece gratuitamente o TEDE (Sistema de Publicação Eletrônica de Teses e Dissertações), quanto às instituições que já implantaram suas bibliotecas digitais de teses e dissertações a partir de tecnologias diferentes do sistema TEDE, a integração à BDTD pode ser feita através da implementação da camada do Protocolo OAI-PMH em seus repositórios de dados, expondo metadados no padrão MTD-BR. O coletador usado pelo IBICT é o *DBUnion*. A busca é feita a partir de uma interface única, a qual executa a consulta em um repositório local que agrega os metadados originários de diferentes instituições. O repositório de metadados da BDTD expõe

seus dados para o coletador da NDLTD da *Virginia Tech* descrita a seguir. O acesso pode ser feito em URL: <http://www.ibict.br>.

- **Biblioteca Virtual em Saúde (BVS):** é gerenciada pela BIREME (Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde) e é “visualizada como a base distribuída do conhecimento científico e técnico em saúde registrado, organizado e armazenado em formato eletrônico nos países da Região, acessível de forma universal na Internet de modo compatível com as bases internacionais” (BIREME). A BVS “representa uma expansão do modelo atual de cooperação técnica ao promover a produção e operação descentralizada de fontes de informação multimídias, conectadas em rede, com acesso direto e universal, sem limitações geográficas e de horário” (BIREME, 1998). Oferece um recurso de busca federada que pesquisa simultaneamente as diversas bases de dados da BVS. Pode ser acessada em URL: <http://www.bireme.br>.
- **Federação Lusófona de Bibliotecas Digitais em Ciências da Comunicação:** gerenciada pela Portcom/Rede de Informação em Comunicação dos Países de Língua Portuguesa vinculada a Intercom/Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, tem como proposta a integração da produção técnica, científica e acadêmica de associações e instituições de ensino e pesquisa em ciências da comunicação de países de língua portuguesa. Para tanto, atua fortemente no desenvolvimento de metodologias de repositórios digitais para uso compartilhado pela comunidade científica, no fomento do registro e disseminação dessa produção em formato eletrônico seguindo padrões internacionais, na preservação da memória da área e na capacitação da comunidade científica lusófona em comunicação. Essa Federação é fruto das atividades de inserção da Portcom no mundo virtual iniciadas no ano de 2000. Oferece, ao usuário final, o serviço de busca via interface única, a qual recorre a metadados coletados (*harvesting*) pelo software *ARC/A Cross Archive Search Service*. Tais metadados são provenientes de provedores de dados utilizando o protocolo OAI, em especial dos repositórios implementados pelos atuais parceiros da Portcom, quais sejam revistas científicas utilizando o software SEER (Sistema de Editoração Eletrônica de Revistas) – versão brasileira do OJS/*Open Journal System* – e repositórios utilizando o software *Dspace* e o DICÍ (Diálogo Científico) - versão brasileira do *Eprints* (FERREIRA, SOUTO, 2005). O acesso pode ser feito em URL: <http://www.portcom.intercom.org.br>.

3.2 Iniciativas estrangeiras de federações de bibliotecas digitais

São inúmeras as iniciativas de projetos de federação de bibliotecas digitais em outros países, podendo ser citados: a *SMETE Open Federation* (<http://smete.org>), a *ARCHON* (<http://archon.cs.odu.edu/>) e *ETRDL* (<http://etrdl.isti.cnr.it/>), dentre outras. A seguir, são descritas duas outras iniciativas, consideradas de maior reconhecimento por parte da comunidade científica, diante da relevância dos estudos envolvidos em seu desenvolvimento, a tal ponto de servirem como modelo para a criação de outros projetos de Federação de Bibliotecas Digitais:

- ***Networked Digital Library of Theses and Dissertations (NDLTD)***: é uma federação de bibliotecas digitais que integra várias coleções de teses e dissertações provenientes de vários países. Teve início em 1996. Inicialmente, fez uso de um sistema de busca federada, mas como alternativa desenvolveu um sistema de coleta de metadados, a partir do protocolo OAI-PMH, tendo como padrão de metadados o formato *ETDMS (Electronic Theses and Dissertations Metadata Standard)*. Em seu *site* oficial, a NDLTD disponibiliza vários recursos de busca, dentre eles o Catálogo de União dos metadados das dissertações e teses eletrônicas, disponível para consulta em um portal mantido pela VTLS. Quanto à interface busca, esse projeto faz uso do sistema Virtua. O acesso pode ser feito em URL: <http://www.ndltd.org/browse.en.html>.
- ***Networked Computer Science Technical Reference Library (NCSTRL)***: é uma federação de bibliotecas digitais que provém material informacional para área de Ciência da Computação. Resulta da fusão de dois projetos anteriores CSTR e NSF WATERS que atuaram isoladamente até 1994. A NCSTRL tornou-se unificada em 1995 e teve a adesão dos primeiros membros internacionais (DAVIS; LAGOZE, 2000 ; MALY ET AL., 1994). Inicialmente, fez uso da busca federada por meio de um protocolo chamado Dienst. Sua primeira versão enviava as solicitações de consulta a vários servidores. Atualmente, oferece um serviço de coleta de metadados baseado na arquitetura de OAI e mantém um índice mestre. Pode ser acessada em URL: <http://www.ncstrl.org/>.

4 CONSIDERAÇÕES FINAIS

Baseado nos conceitos de bases de dados federadas chegou-se ao conceito de sistemas de informação federados e seus componentes de recuperação baseados na busca federada e na coleta de metadados/*harvesting*.

Colocam-se como evidentes as tendência e necessidade atuais quanto ao compartilhamento de recursos informacionais e uma das soluções disponíveis está vinculada ao conceito embutido no bojo de sistemas de informação federados e, mais especificamente, na federação das bibliotecas digitais.

Percebe-se que os avanços das tecnologias de constituição, interação e interoperabilidade e outras tantas têm e terão um considerável impacto no futuro dessas Federações, inclusive já se evidenciando na literatura uma grande convergência de pesquisas advindas de várias áreas como Interação Humano Computador (ICH), Recuperação de Informação, Tecnologia de Educação e Design Instrucional.

Pensar, portanto a implementação de Federação de Bibliotecas Digitais, conforme discutido nesse *paper*, representa uma revisão quanto à gestão e as tecnologias envolvidas em tais sistemas dessa natureza, mas também se refere à revisão da atual concepção de nossas bibliotecas digitais, seus propósitos, objetivos e, finalmente, uso das tecnologias e adoção integral do processo de interatividade em sua potencialidade.

REFERÊNCIAS

ANAN, H. et al. **Preservation and Transition of NCSTRL Using an OAI-Based Architecture**. Disponível em: <http://www.cs.odu.edu/~liu_x/paper/ncstrl/ncstrl.pdf>. Acesso em: 15 maio 2006.

ARMS, W.Y. *Digital Libraries*. [s.l.]: M.I.T. Press, 2000.

BIREME. **A Biblioteca Virtual em Saúde**. Disponível em: <http://www.bireme.br/bvs/P/sobre_bvs.htm>. Acesso em: 08 maio 2006.

BIREME. **Resumo executivo**. In.: REUNIÓN DEL SISTEMA LATINOAMERICANO Y DEL CARIBE DE INFORMACIÓN EN CIÊNCIAS DE LA SALUD, 1998, Costa Rica. Disponível em: <<http://www.bireme.br/bvs/P/pdoc.htm>>. Acesso em: 08 maio 2006.

BUSSE, S. et al. **Federated Information Systems: concepts, terminology and architectures**. Technical Report, " TU Berlin 99-9, 1999. Disponível em: <<http://citeseer.ist.psu.edu/busse99federated.html>>. Acesso em: 04. abr. 2006.

COLEMAN, A.; OXNAM, M. Interactional digital libraries: introduction to a special issue on interactivity in digital libraries. **Journal of Digital Information**, v. 2, n. 4, 2002. Disponível em: <<http://jodi.tamu.edu/Articles/v02/i04/editorial/>>. Acesso em: 04. Abr. 2006.

CONRAD, S. et al. Engineering Federated Information Systems - Report of EFIS '99 Workshop. **SIGMOD Record**, v. 28, n. 3, Sep. 1999. Disponível em: <<http://www.sigmod.org/sigmod/record/issues/9909/index.html>>. Acesso em: 04 abr. 2006.

DAVIS, J. R.; LAGOZE, C. NCSTRL: Design and deployment of a globally distributed digital library. **Journal of the American Society for Information Science**, v. 51, n. 3, p. 273-280, 2000. Disponível em: <<http://www3.interscience.wiley.com/cgi-bin/fulltext/69501282/PDFSTART>>. Acesso em: 15 maio 2006.

DOWNES, S. Distributing Learning Objects. **Best of Australian Flexible Learning Community: 2001-2004**. Publicado em 5 Oct. 2003. Disponível em: <http://community.flexiblelearning.net.au/GlobalPerspectives/content/article_4715.htm>. Acesso em: 11.May.2006.

FENG, L.; JEUSFELD, M.; HOPPENBROUWERS, J. Towards knowledge-based digital libraries. *SIGMOD Record*, v. 30, n. 1, p. 41-46, Mar. 2001. Disponível em: <<http://infolab.uvt.nl/pub/fengl-2001-27.pdf>>. Acesso em: 16. May 2006.

FERREIRA, S. M. S. P.; SOUTO, P. C. N. Federação de bibliotecas digitais lusófonas em ciências da comunicação. In: MARCONDES, C. H. et al. **Bibliotecas digitais: saberes e práticas**. Salvador/Brasília: UFBA/IBICT, 2005. p. 325-342.

FRYER, D. Federated Search Engines. **Online**, v. 28, n. 2, p. 16-19, Mar./Apr. 2004.

IBICT. **Sobre a BDTD**. Disponível em: <<http://bdttd.ibict.br/utilitarios/sobre.jsp>>. Acesso em: 08 maio 2006.

JACSO, P. Thoughts About Federated Searching. **Information Today**, v. 21, n. 9, Oct. 2004, p. 17-20.

LIU, X. et al. Federated Searching Interface Techniques for Heterogeneous OAI Repositories. **Journal of Digital Information**, v. 2, n. 4, article no. 106, May 2002. Disponível em: <<http://jodi.ecs.soton.ac.uk/Articles/v02/i04/Liu/>>. Acesso em: 15 maio 2006.

MALY, K. et al. **Wide area technical report service: Final report**. Technical Report, Department of Computer Science, Old Dominion University. 1994, June 6.

MARCONDES, C. H.; SAYAO, L. F. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. **Ciência da Informação**, v. 30, n. 3, p. 24-33, set./dez. 2001.

MEDEIROS, C. B.; ALENCAR, A. C. **Qualidade dos dados e interoperabilidade em SIG** In.: WORKSHOP BRASILEIRO DE GEOINFORMÁTICA, 1., Campinas, 1999. Disponível em: <<http://www.geoinfo.info/portuguese/geoinfo1999/papers/CBauzer.pdf>>. Acesso em: 14 maio 2006.

MISCHO, W. L. Digital Libraries: challenges and influential work. **D-Lib Magazine**, v.11, n.7/8, July/Aug 2005. Disponível em: <<http://dlib.ejournal.ascc.net/dlib/july05/mischo/07mischo.html>>. Acesso em: 14. maio. 2006.

SHETH, A.; LARSON, J. Federated Database Systems for Managing Distributed, Heterogeneous and Autonomous Databases. **ACM Computing Surveys**, v. 22, n. 3, p.183–236, 1990.

SHREEVES, S. L. et al. **Is ‘quality’ metadata ‘shareable’ metadata? The implications of local metadata practices for federated collections.** In.: NATIONAL CONFERENCE OF ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES, 12., Minneapolis, 2005. p. 223-237. Disponível em: <<http://www.ala.org/ala/acrl/acrl/events/shreeves05.pdf>>. Acesso em: 25 abr. 2006.

SULEMAN, H. **Open digital libraries.** 2002. 190 f. Dissertation (Doctor of Philosophy in Computer Science and Applications) - Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, 2002. Disponível em: <<http://scholar.lib.vt.edu/theses/available/etd-11222002-155624/>>. Acesso em: 28 abr. 2006.

WADHAM, R. L. Federated Searching. **Library Mosaics**, v. 15, n. 1, p. 20, Jan./Feb. 2004.

WEHR, H. Integrating Heterogeneous Data Sources into Federated Information Systems. In.: WORKSHOP ON GENERATIVE AND COMPONENT-BASED SOFTWARE ENGINEERING, 2002. **Anais eletrônicos...** Disponível em: <http://www.old.netobjectdays.org/node02/de/Conf/publish/slides.html#Integrating_Heteroge.t>. Acesso em: 10 maio 2005.

SUELI MARA S. P. FERREIRA

Professora Doutora da Escola de Comunicações e Artes da USP
Coordenadora da Rede de Informação em Comunicação dos
Países de Língua Portuguesa/Portcom
Email: smferrei@usp.br

LEONARDO FERNANDES SOUTO

Doutorando do Programa de Ciência da Informação da Escola de
Comunicações e Artes da USP.
Bibliotecário da Biblioteca Central da UNICAMP
Email: lfsouto@unicamp.br

Recebido em: 30 de maio de 2006
Aceito para publicação em: 20 de junho de 2006