Plano de Estudo

1. Motivação

Devido ao crescente processo de digitalização de documentos tais como artigos científicos, jornais, revistas, teses, entre outros, surgiu a necessidade de armazená-los de forma organizada para se obter acesso rápido e eficiente a esses documentos. Como uma forma de suprir tal necessidade, as bibliotecas digitais foram então desenvolvidas.

Essas bibliotecas têm se multiplicado na Internet nos últimos anos e tem sido pauta de muitos debates no meio científico. Numa comparação entre a quantidade de bibliotecas virtuais e de bibliotecas como as conhecemos hoje, o número de bibliotecas digitais ainda é muito pequeno. Suas vantagens e desvantagens têm sido discutidas e ainda não se sabe se elas substituirão ou não as bibliotecas localizadas nas universidades.

Juntamente com os repositórios de dados virtuais, muitos serviços também têm sido implementados. São bastante comuns provedores de serviços de buscas numa certa federação de repositórios virtuais. Como esses tipos de serviços, as bibliotecas digitais são um recurso muito interessante para pesquisadores, estudantes e para as próprias universidades.

2. Objetivo

Temos como objetivo nesse estudo a discussão de aspectos polêmicos dessas bibliotecas virtuais como direitos autorais dos documentos depositados bem como a abordagem de aspectos técnicos: a implementação de repositórios e serviços e o levantamento dos serviços existentes e idéias para novos serviços.

3. Tópicos a serem desenvolvidos

a) O que são as bibliotecas digitais. Exemplos populares na Internet.

As bibliotecas digitais podem ser repositórios de documentos digitalizados em formato PDF, TXT, GIF, JPG, DOC, etc, ou repositórios de informações descritivas desses documentos como título, assunto, autores, número de páginas, formato, onde se encontra disponível, etc. Esse tipo de informação é comumente chamado de metadado.

b) Discussão de vantagens e desvantagens das bibliotecas virtuais sobre as bibliotecas nas universidades.

O custo financeiro de se manter um documento numa biblioteca digital é muito menor do que mantê-lo nas prateleiras das bibliotecas. Porém, as bibliotecas nas universidades envolvem um contexto que facilita o entendimento da literatura nela contida.

c) O Protocolo OAI.

Como uma proposta da *Open Archives Iniative*, surgiu o protocolo OAI. Este provê um conjunto de ferramentas de inter-operabilidade independente da aplicação.

d) Provedores de Dados e Provedores de Serviços. Exemplos.

Os Provedores de Dados administram sistemas que dão suporte ao protocolo OAI como um meio de expor metadados. Já os Provedores de Serviços usam metadados via protocolo OAI como uma base para construção de serviços úteis.

e) *Harvester*.

Um *harvester* é uma aplicação cliente que atende às requisições do protocolo OAI. É operado por um provedor de serviço como um meio de coletar dados de um ou mais repositórios.

f) Repositório.

Um repositório é um servidor acessível pela Rede que é capaz de processar as seis requisições constituintes do protocolo OAI. Um repositório é gerenciado por um provedor de dados para expor metadados aos *harvesters*.

g) Técnicas de implementação de *archives* e *harvesters*.

Existem alguns *frameworks* de interfaces que podem ser implementados para trabalhar com repositórios arbitrários. Alguns exemplos são OAICat e OAIHarvester.

Entre algumas técnicas interessantes de implementação, destaca-se o modelo SODA (Smart Objects, Dumb Archives) criado pelo comitê técnico de desenvolvimento de bibliotecas digitais da *Old Dominium University*. Este modelo transfere a funcionalidade tradicionalmente associada com os *archives* aos próprios objetos arquivados.

h) *Harvesters* famosos.

Entre os harvesters mais populares destacam-se:

· DP9 portal que expõe os metadados dos repositórios para motores de busca como o Google, por exemplo.

· ARC faz buscas numa federação de repositórios virtuais.

· NDLTD banco de teses e dissertações digitalizadas.

Ainda serão objeto de estudos harvesters como my.OAI, OAIster, etc.

i) *Harvesters* em desenvolvimento.

Há dois grandes projetos em andamento no presente momento. São eles: FEDORA da Virginia Tech e CHIPS N´ SALSA da ISTEC juntamente com a Universidade do Novo México.

j) Propostas de novos *harvesters*.

No Congresso sobre Política de Informação em Bibliotecas Digitais na UNICAMP entre os dias 20 e 21 de março de 2003, o Prof. Dr. Johan van Reenen da ISTEC e Universidade do Novo México mencionou a necessidade de um serviço para se poder identificar os repositórios escondidos na Internet.

Também há idéias dor Prof. Dr. Imre Simon do IME-USP a respeito de um serviço de gráficos para medição da evolução da quantidade de documentos depositados nos repositórios virtuais.

k) *Archives* para *peer-revision* eletrônica.

Existe na Rede uma discussão polêmica a respeito de *peer-revision* via Internet, e já podemos encontrar repositórios *pre-print* na Internet, que armazenam artigos que devem passar pelo processo de *peer-revision*.

l) Protocolo OAI versus direitos autorais.

É muito polêmica também a questão dos direitos autorais dos metadados e documentos encontrados via protocolo OAI. Uma calorosa discussão deu-se na lista *OAI-general*, mantida pela *Open Archives Iniative,* que será explorada também na monografia.

4. Bibliografia

§ Peter Lyman. *Scholarly Communication and Technology.* Conferência sobre Documentos Digitais e o Futuro da Comunidade Acadêmica. Emory University. 1997.

§ http://www.dtc.umn.edu/~odlyzko/doc/peer.review.txt

§ Ian Parberry. *A Guide for New Referees in Theorical Computer Science*. Universidade de Norte do Texas. 1994.

§ http://www.openarchives.org/OAI/2.0/guide-repository.htm

§ http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm

§ http://www.openarchives.org/tools/tools.html

§ Michael L. Nelson, Kurt Maly, Mohammead Zubair, Stewart N. T. Shen, *SODA: Smart Objects, Dumb Archives*. Conferência sobre Pesquisa e Tecnologia Avançada para Bibliotecas Digitais. Paris. 1999.

§ Lista de discussão OAI-general. Subject: *OAI and intellectual property issues*.

§ H. N. Jerez, R. Jordan, J. Van Reenen. *Distribuited Learning Objects and Metadata Harvesting Implementation on the Chips n´ Salsa Project*. Universidade do Novo México.

§ Xiaoming Liu, Kurt Maly, Mohammad Zubair, Michael L. Nelson. *Arc An OAI Service Provider For Cross-Archive Searching*. Old Dominium University.

§ Herbert Van de Sompel, Carl Lagoze. *Notes From the Interoperability Front: A Progress Report On The Open Archives Iniative*. Research Library, Los Alamos Laboratoy.

§ Xiaoming Liu, Kurt Maly, Mohammad Zubair, Michael L. Nelson. *DP9: An OAI Gateway Service For Web Crawlers*. Old Dominium University.

§ Yochai Benkler. *Coases´ Penguim, or, Linux And The Nature of The Firm*.