

EIXO TEMÁTICO: Os Open Archives e os repositórios institucionais

Sistema de acesso à informação baseado em *Open Archives*: a experiência do Holmes

Nanci Oddone

Prof. Dra. em Ciência da Informação,
ICI/UFBA, neoddone@ufba.br

Ricardo Sodré Andrade

Graduando em Arquivologia e bolsista de
iniciação científica (PIBIC/CNPq), ICI/UFBA.
ricardo@feudo.org

Resumo

O Holmes (<http://www.holmes.feudo.org>) é um recurso eletrônico de acesso à informação idealizado e executado no âmbito do Grupo de Estudos em Políticas de Documentação e Informação – GEDINFO. Sua concepção resultou de discussões que visavam encontrar meios de superar uma das maiores dificuldades dos usuários do Portal de Periódicos da CAPES: a inexistência de busca por assunto. O projeto inicial tinha como meta criar uma ferramenta que facilitasse a indexação e a recuperação dos artigos publicados nos vários periódicos nacionais de Arquivologia, Biblioteconomia e áreas correlatas. Em pouco tempo, contudo, verificou-se que seria possível agregar outras fontes dessas áreas e oferecer um mecanismo confiável de integração e consolidação da documentação científica produzida em Ciência da Informação. Aproveitando a infra-estrutura do protocolo OAI-PMH que já existia em periódicos nacionais e estrangeiros, desenvolveu-se o aplicativo com o software livre *PKP OAI Harvester*. No momento, o Holmes reúne 26 provedores de dados, que se dividem entre anais de eventos, repositórios textuais e revistas científicas. Mais de 13 mil documentos já foram indexados e o número tende a aumentar, a partir da inclusão de novos provedores e da regularidade na alimentação dos atuais provedores. Recentemente, após solicitação, o Holmes foi indexado pelo Unesco Libraries Portal.

Palavras-chave: *open archives*; sistema de recuperação da informação; documentação científica; metadados

Introdução

A explosão na produção do conhecimento científico a partir da segunda metade do século passado acarretou na identificação e reconhecimento da importância dos aspectos referentes à disseminação e acesso da informação gerada.

Publicar de forma a alcançar o maior número de interessados e obter/dispor de informações da maior relevância possível e em quantidade satisfatória são objetivos dos pesquisadores de todas as áreas, enquanto produtores de conhecimento, assim como objeto de estudo dos profissionais da ciência e tecnologia da informação, enquanto gestores de informação e do conhecimento.

A alta disponibilidade de informações no atual contexto globalizado (onde as tecnologias da informação e comunicação transformaram e aproximaram também o meio acadêmico) aliada à necessidade de apropriação para produção de mais conhecimento fez com que novas soluções

fossem discutidas, no intuito de contribuir com o aperfeiçoamento do fluxo de conhecimento da comunidade acadêmica.

O conceito de *open archives* surge, neste cenário, pela necessidade do desenvolvimento e promoção de padrões de interoperabilidade, isto é, operação colaborativa de diversos atores diferentes em um mesmo contexto, em vistas à facilitação e disseminação de conteúdos científicos. A *Open Archives Initiative* denomina o grupo que mantêm o protocolo OAI-PMH, uma “linguagem de comunicação”, ou protocolo, que torna possível a interoperabilidade entre sistemas, como os que serão descritos ao longo do texto.

A infra-estrutura OAI envolve os produtores de trabalhos acadêmicos (pesquisadores), provedores de dados OAI (repositórios, periódicos, etc.), provedores de serviços OAI e usuários, todos utilizando o OAI-PMH como protocolo para a interação. Os pesquisadores são os produtores de informação, estes publicam seus trabalhos nos provedores de dados, que por sua vez armazenam e preservam as informações utilizando a arquitetura OAI, os provedores de serviço são aplicativos que traduzem as potencialidades da arquitetura OAI em sistemas funcionais de busca e recuperação de informação para os usuários.

No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), têm fomentado o uso de sistemas compatíveis com o protocolo OAI-PMH, como o Sistema Eletrônico de Editoração de Revistas (SEER), a versão traduzida do *Open Journal System*, do *Public Knowledge Project* (OJS – PKP).¹ Márdero Arellano, Santos e Fonseca expõem algumas características do SEER:

Esta ferramenta contempla ações primordiais à automação das atividades de editoração de periódicos científicos, permitindo completa autonomia na tomada de decisões sobre o fluxo editorial, a publicação e o acesso por parte do editor; ele define as etapas do processo editorial, de acordo com a política definida pela revista, mas dispende de assistência e registro on-line em todas as fases do sistema de gerenciamento. Na etapa de submissão, o sistema disponibiliza um espaço para comunicação com o editor e permite também o acompanhamento da avaliação e editoração do trabalho. (Márdero Arellano et al., 2005)

Também podemos citar o *Open Conference System*, o sistema utilizado pela Comissão Científica deste XIV Seminário Nacional de Bibliotecas Universitárias, que auxilia na gestão de eventos científicos, facilitando a criação dos anais de evento e disponibilizando-os através de um sistema compatível com o protocolo OAI-PMH e o DSpace, que permite a criação de repositórios institucionais, como as denominadas Bibliotecas Digitais.

Holmes – o “detetive da Ciência da Informação”

Os provedores de serviços OAI são sistemas que disponibilizam funcionalidades utilizando o potencial do protocolo OAI-PMH, como a busca em diversos provedores de dados OAI

¹ Mais informações acerca desses programas podem ser encontradas em <<http://www.pkp.ubc.ca/>>.

simultaneamente. Um exemplo é o Holmes - o “detetive digital” da Ciência da Informação (disponível na internet desde 10/04/2006), primeiro provedor de serviços OAI na área de CI no Brasil. Ao dia de escrever esse trabalho, o Holmes permitia uma busca integrada em 26 (vinte e seis) repositórios ou periódicos de CI aderentes ao OAI-PMH.

Um dos mais conhecidos repositórios compatíveis com a OAI em CI disponível, em quantidade de registros, é o *E-LIS – E-prints in Library and Information Science* (<http://eprints.rclis.org>) – que possui cerca de 4.160 trabalhos em sua base de dados. No entanto, o Holmes, enquanto ferramenta que possibilita acesso a textos acadêmicos (não é um repositório), consegue dispor à busca aproximadamente 13.500 (treze mil e quinhentos) trabalhos em CI, o que demonstra sua funcionalidade e utilidade, pois evita a necessidade de se percorrer e efetuar buscas isoladas nos sistemas de cada um dos provedores de dados (repositórios, periódicos etc.) cadastrados em seu sistema, atentando para o fato de que boa parte dos trabalhos apresentados na busca passou pelo processo de avaliação pelos pares (*peer-review*).

É importante salientar que o Holmes registra os números acima declarados com cerca de um mês de existência. O *software* no qual o Holmes é baseado, o *PKP OAI Harvester*,² facilita a inclusão de novos provedores de dados para o processo de coleta dos metadados através da interface de gerenciamento do sistema. O processo de coleta (*harvesting*) dos metadados do provedor de dados incluído no sistema é o primeiro passo para a formação da base do Holmes, que armazena os metadados de todos os provedores de dados (mas não os conteúdos acadêmicos) e possibilitam o funcionamento deste sistema.

O levantamento bibliográfico se torna uma tarefa menos complicada, pois, em um mesmo ambiente, diversos veículos de informação científica são consultados.

A quantidade de artigos em CI da base do Holmes aumenta em dois casos, com a adição de novos provedores de dados no sistema ou quando esses repositórios disponibilizam mais trabalhos. O Holmes possui uma atualização semanal, o que significa que os trabalhos acadêmicos novos em provedores de dados que já integram o Holmes são localizados e seus metadados coletados.

Após solicitação, o Holmes foi incluído no Unesco Libraries Portal.

Próxima versão do Holmes

Em meados do ano 2006, o projeto PKP anunciou a versão 2 do PKP OAI Harvester, o software no qual o Holmes está baseado.

Essa versão deverá proporcionar as seguintes características e funcionalidades ao Holmes:

² Disponível gratuitamente em <<http://pkp.sfu.ca/pkp-harvester/>>.

1. Coletar metadados OAI registrados em diversas formas, como Dublin Core não-qualificado, extensões Dublin Core do Open Journal Systems e Open Conference Systems, MODS e MARCXML;
2. Interface de busca que permitem pesquisas simples e avançadas, consultando múltiplos campos;
3. Coleta seletiva de metadados em repositórios multidisciplinares;
4. Capacidade de extensão de suas capacidades por meio de plugins

Esperamos migrar o Holmes de versão tão logo quanto possível.

Considerações finais

O Holmes é uma demonstração do potencial que a adoção padrões de metadados, sendo uma das vantagens a interoperabilidade apresentada pela arquitetura OAI.

Atividades de pesquisa, como o levantamento de referências, se torna mais eficiente diante da disponibilidade de acesso crescente ao conhecimento científico.

No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) fomenta o uso de sistemas *Open Archives* por meio da tradução de sistemas compatíveis com OAI-PMH, treinamento de recursos humanos e outras atividades relacionadas, com isso, é possível garantir que a adoção da plataforma OAI seja um investimento seguro para as instituições brasileiras, incluindo as bibliotecas universitárias que desejam publicar conteúdos em meio digital.

Referências

Márdero Arellano, Miguel Angel; Regina dos Santos; Ramón da Fonseca.. SEER: Disseminação de um sistema eletrônico para editoração de revistas científicas no Brasil. In: **Arquivística.net**, vol 1, n 2, 2005. Disponível em <<http://www.arquivistica.net/ojs/viewarticle.php?id=33&layout=abstract>>. Acesso em 06 de maio de 2006.

PUBLIC KNOWLEDGE PROJECT. **Open Archives Harvester**. Disponível em <<http://pkp.sfu.ca/?q=harvester>>. Acesso em: 18 ago. 2006.