

DIÁLOGO CIENTÍFICO: *EPrints* como um ambiente virtual aberto da comunicação científica

VIANA, CASSANDRA LÚCIA DE MAYA
Instituto Brasileiro de Informação em Ciência e Tecnologia
cassandra@ibict.br
MÁRDERO ARELLANO, MIGUEL ÁNGEL
Instituto Brasileiro de Informação em Ciência e Tecnologia
miguel@ibict.br

RESUMO

O *EPrints* (versões 2.2.1 e 2.3.7), software livre que favorece a comunicação entre pesquisadores, foi traduzido e customizado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia. A primeira versão do “Diálogo Científico” foi desenvolvida para a área de Ciência da Informação. Destinado à criação de repositórios temáticos, vem sendo compartilhado com instituições de ensino e pesquisa no Brasil. Apresentam-se seus aspectos técnico-operacionais. A versão atual inclui novas funcionalidades, tais como: melhor interface gráfica do usuário, tabela JITA, em português, novas opções de navegação etc. Atualmente, visa o intercâmbio e a integração com outros arquivos abertos em âmbito internacional.

Palavras-chave: Repositório digital, repositório temático, *EPrints*, comunicação científica

1. Introdução

A comunicação científica estruturou-se em meio à sistematização do processo de produção, disseminação e utilização do conhecimento científico. Valendo-se de órgãos editoriais, veículos formais e informais de divulgação e recursos bibliográficos, e principalmente, após o surgimento da imprensa, ela se estabeleceu, através dos séculos, como um importante mecanismo de julgamento do mérito do trabalho dos estudiosos das ciências.

Ainda nos dias de hoje os pesquisadores dependem da disseminação dos resultados das pesquisas, da literatura publicada e da informação científica registrada para manter e estimular o ciclo produtivo do conhecimento e obter o reconhecimento de seu trabalho científico por parte da sociedade.

O surgimento e a expansão da Web e das tecnologias a ela associadas, a diminuição dos custos das tecnologias da informação e da comunicação (TIC), aliados à crise econômica mundial e aos altos custos do sistema de publicação impressa tiveram forte repercussão no meio acadêmico-científico. A reação veio sob a forma de crítica e oposição ao modelo clássico de avaliação do trabalho científico: a “medida crua de impacto da pesquisa, baseada na contagem das publicações em si” (Harnad, 2006), segundo o fatalismo do “*publish or perish*”.

Desencadearam-se, então, sucessivos movimentos em favor do livre acesso à informação científica. No auge desse processo, e sob a liderança de Ginsparg, Harnad, Lagoze, Van de Sompel e Luce, a partir do ano 2000 foi, então, institucionalizada a Open Archives Initiative (OAI). Surge, então, um novo paradigma,

sob o qual a informação científica é considerada insumo básico para o desenvolvimento científico e tecnológico de uma nação, devendo ser de livre circulação, quando financiada com recursos públicos.

Sob esta nova perspectiva, “acredita-se que o modelo tradicional formal de comunicação científica entre pares está sendo revisto de modo muito revelador e promissor pelos *open archives*”. (Moreira, 2005).

Concretamente, através das ações propostas e do esforço colaborativo de grande número de pesquisadores de origem multidisciplinar, foram desenvolvidas soluções tecnológicas e estabelecidos padrões de interoperabilidade para melhorar a disseminação da produção científica. Esses esforços têm-se materializado na geração de sistemas de publicação e preservação de conteúdos digitais temáticos e institucionais, tais como: o Open Journal Systems, o EPrints, o DSpace, entre outros.

2. Implementação do Eprints pelo IBICT

O software EPrints é o que melhor representa os padrões e ideais da *Open Archives Initiative*. (Kuramoto, 2005).

Com base nesse software, foi implementado, pelo IBICT, o Diálogo Científico (DICI), em língua portuguesa, que oferece um ambiente virtual aberto para a comunicação e o debate entre os pares e a publicação de trabalhos científicos, em texto integral, diretamente na Internet. É operacionalizado através de auto-arquivamento, i.e. do depósito pelos próprios autores e oferece uma infra-estrutura que permite o processamento automático de comentários e a produção de versões do documento original.

Além de um maior acesso à produção científica, proporciona a sua preservação digital, tornando o processo de produção do conhecimento científico mais transparente, dinâmico e democrático.

Inicialmente, o Eprints foi traduzido para o português e adaptado para o Brasil tendo sido criada uma versão padrão customizada para repasse a instituições do sistema de informação em C&T do país. Para a representação de assuntos e descrição do conteúdo temático dos documentos a serem depositados, foi adotada a tabela de áreas de conhecimento do CNPq, aplicável a qualquer associação ou instituição científica.

Em 2003, durante o XIII ENDOCOM, o projeto DICI foi divulgado para a comunidade científica de Comunicação e Ciência da Informação (CI).

Em 2004 foram disponibilizadas no site do IBICT informações sobre o software com o objetivo de auxiliar na instalação e implementação desse modelo de repositórios no país. Nesse mesmo ano, durante o II Simpósio de Bibliotecas Digitais, foi então lançado o DICI, como um repositório Eprints para a área de CI.

Também foi feita uma apresentação dessa ferramenta, em 2005, durante o 57º Congresso da SBPC, em Fortaleza, Ceará.

3. Especificações da versão 1.1 do DICI

A versão 1.1 do DICI foi produzida com base na versão 2.2.1 do Eprints, o qual foi desenvolvido em linguagem Perl.

Utiliza os padrões tecnológicos da OAI para manutenção dos objetos digitais bem como para garantir a interoperabilidade com outros repositórios baseados nos princípios do acesso livre à informação científica, a saber:

- a) estrutura de metadados Dublin Core (DC), estando visíveis na Web mesmo para usuários *anonymous*;
- b) protocolo OAI PMH (Protocol for Metadata Harvesting, configurado como provedor de dados);
- c) servidor de banco de dados (MySQL), sistema operacional (ambiente Unix/Linux), servidor web (Apache), servidor de aplicações e demais tecnologias para a sua instalação, baseadas em plataforma web e em softwares de código aberto;
- d) linguagem XML para transmissão do conteúdo dos dados das aplicações;
- e) aceita praticamente todos os formatos de arquivos

Permite o depósito de mais de 20 tipos de documentos, desde os pré-prints, até conteúdo integral de livros.

3.1 Características da interface Web

A interface Web apresenta, em sua página principal, as seguintes opções: Navegação, Pesquisa, Cadastro, Área de autor e Ajuda.

A opção de Navegação permite percorrer os registros por Área de Conhecimento e por Ano; a opção de Pesquisa permite a busca no texto completo, ou ainda Busca Simples ou Avançada; e a Área de Autor oferece as opções de ver e alterar o cadastro, alterar os critérios de recebimento de Informações sobre os documentos depositados e a alteração do endereço de e-mail.

3.2 Repositórios Digitais Temáticos

O DICI permite a criação de repositórios virtuais por área do conhecimento, também denominados de repositórios temáticos.

Para representar as diversas áreas de conhecimento, utiliza a Tabela de Áreas do CNPq, a qual é dividida em 9 grandes grupos e cuja estrutura hierárquica compreende também subáreas, sendo estas gradativamente detalhadas até o nível de especialização alcançado pelo conhecimento em cada uma delas.

Quando é feito o repasse dessa tecnologia, é selecionada e instalada apenas a área de conhecimentos correspondente ao interesse da instituição receptora.

3.3 Gerenciamento dos e-prints

A estrutura gerencial do DICI está dividida em três seções:

- Editorial: dedicada à verificação da qualidade dos metadados, orientação na submissão e publicação.
- Técnica: concentrada na implementação, melhoria e desenvolvimento de funcionalidades e ferramentas que permitam a interoperabilidade e,
- Administrativa: encarregada das políticas de uso, seu impacto na comunidade científica e sua integração com outras comunidades.

Existem 3 funções principais no sistema: a do Autor, a do Editor e a do Administrador.

As ferramentas do Autor lhe permitem: ver/alterar seu cadastro; alerta sobre novos documentos depositados e gerar novas versões de seus documentos depositados.

O papel do Autor no sistema consiste, basicamente, no depósito de documentos de sua autoria, na atualização da versão dos documentos de sua autoria, conforme comentários e sugestões recebidos, e no comentário dos trabalhos de outros autores.

O Fluxograma de depósito de documentos envolve as seguintes etapas:

- a) Seleção do tipo de documento a ser depositado;
- b) Opção de: comentar documento de outro autor, ou de responder a um comentário recebido, ou ainda de efetuar alterações e nova versão de seus próprios trabalhos;
- c) Inclusão dos metadados correspondentes ao documento a ser depositado;
- d) indicação do formato de arquivo em que consiste o documento e do nível de segurança para sua consulta;
- e) seleção da localização (internet, HD), nível de compressão do conteúdo (compactado ou não), e indicação da quantidade e tipos de documentos que serão depositados (*upload*);
- d) digitação do caminho completo para o local onde será obtido cada documento a ser depositado;
- d) verificação final do depósito
- d) apresentação do Termo de Aceite relacionado aos direitos de autoria e distribuição do documento
- e) confirmação do depósito e orientação quanto à sua disponibilidade no sistema

O papel do Editor é exercido através da revisão da adequação do conteúdo ou tema do documento, da validação e correção do conteúdo dos metadados e da autorização final para a publicação do item no repositório.

Além das opções do autor, ao Editor também é permitido: ver o status do sistema (dados estatísticos); pesquisa com refinamento para a correção e exclusão de depósitos; pesquisa no cadastro de autores para visualização e alteração de todos os seus registros.

O sistema envia um e-mail de alerta para o editor da respectiva área temática, assim que novos depósitos são efetuados no repositório.

A função do Administrador é a de gerenciar o desempenho e acompanhar o desenvolvimento da utilização do sistema. Além opções disponíveis aos autores e editores, o administrador também dispõe de recursos para adicionar autores e editores e de gerenciar a tabela de áreas, podendo criar, apagar, editar ou relacionar categorias de assunto.

No que se refere às questões de copyright, por tratar-se de repositórios de acesso livre, todos os trabalhos depositados continuam sendo propriedade dos autores. Portanto, cabe a estes a responsabilidade sobre o que é submetido e a decisão sobre quais conteúdos poderão ser acessados ou não, garantindo assim sua propriedade intelectual.

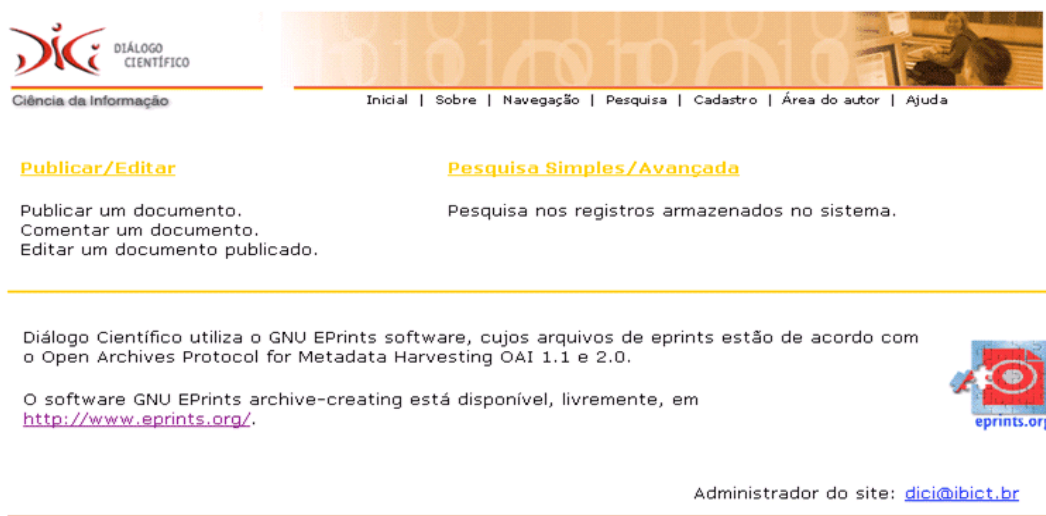
3.4 Repasse para a comunidade científica brasileira

A primeira instituição a adotar o sistema para criação de Diálogos Científicos usando o *Eprints*, em setembro de 2001, foi a Sociedade Brasileira de Genética (<http://sbg.ibict.br>).

Em abril desse mesmo ano o IBICT iniciou a implantação do DICI, em parceria com a PORTCOM/Rede de Informação em Comunicação dos Países de Língua Portuguesa, mantida pela INTERCOM/Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, que criou o repositório Arena Científica (<http://reposcom.portcom.intercom.org.br/handle/1904/18519>) para os pesquisadores lusófonos da área das Ciências da Comunicação. Este repositório, cuja interface foi personalizada segundo as características dessa entidade, está sendo utilizado, atualmente, pelos estudantes de pós-graduação, docentes e pesquisadores desta área como um dos módulos do Portal PORTCOM.

As atividades relacionadas à disseminação e orientação das sociedades científicas brasileiras quanto ao uso do *software Eprints*, têm sido realizadas constantemente pelos técnicos do Laboratório, especialmente através de sua participação em eventos nacionais e internacionais. Além disso, esses técnicos têm acompanhado a sua implantação nas instituições de ensino e pesquisa.

O pacote de software DICI foi instalado e já está sendo adotado nas seguintes instituições: Sociedade Brasileira de Genética; PORTCOM; Fundação Cenecista Presidente Kennedy; IMP da USP; e no próprio Instituto, como repositório de Ciência da Informação (<http://dici.ibict.br>).



DICI DIÁLOGO CIENTÍFICO
Ciência da Informação

Inicial | Sobre | Navegação | Pesquisa | Cadastro | Área do autor | Ajuda

Publicar/Editar

- Publicar um documento.
- Comentar um documento.
- Editar um documento publicado.

Pesquisa Simples/Avançada

Pesquisa nos registros armazenados no sistema.

Diálogo Científico utiliza o GNU EPrints software, cujos arquivos de eprints estão de acordo com o Open Archives Protocol for Metadata Harvesting OAI 1.1 e 2.0.

O software GNU EPrints archive-creating está disponível, livremente, em <http://www.eprints.org/>.

Administrador do site: dici@ibict.br

Figura 1. Página principal do DICI

No caso deste último, foi implementada uma metodologia de arquivamento de artigos publicados em periódicos nacionais da área. Foi solicitada autorização aos editores e inseridos dos metadados e os textos completos de mais de 600 artigos. O crescimento do repositório pode ser observado no gráfico, coletado e disponibilizado pelo sistema Celestial, do projeto Registry of Open Access Repositories (ROAR):

Diálogo Científico

Generated by <http://archives.eprints.org/>

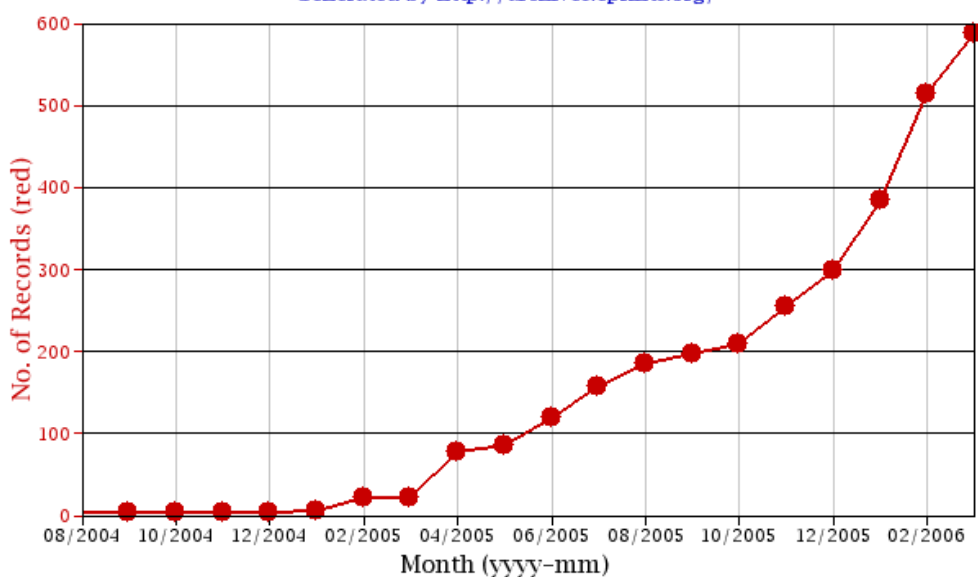


Figura 2. Diálogo Científico – Gráfico de Registro Cronológico de Depósitos

O repasse dessa tecnologia pelo IBICT envolve palestras e treinamento dos profissionais de informação da instituição parceira, para atendimento de sua comunidade de usuários e para gerenciamento da ferramenta objetos digitais nela depositados.

4. Um novo “Diálogo Científico”

A versão 1.2 do DICl será instalada e configurada a partir da versão 2.3.7 do EPrints. A ferramenta será disponibilizada, assim como a anterior, no site do IBICT.

Estão previstas alterações e inclusão de novas funcionalidades para a barra de ferramentas; interface gráfica do usuário: e funções internas do sistema.

4.1 Implementação da nova versão do DICl

Para esta implementação serão realizadas, dentre outras, as seguintes tarefas:

- Tradução da versão 2.3.7 do Eprints para a língua portuguesa;
- revisão e consolidação Tabela JITA, já traduzida para a língua portuguesa;
- migração dos dados para a nova versão;
- migração do sistema para novo servidor; e
- elaboração da documentação do sistema

Estão previstas, ainda, atividades de monitoramento da implantação, nas instituições que venham a adotar esta nova versão, bem como a atualização permanente dos técnicos envolvidos e troca de experiências, através da criação e manutenção de uma comunidade virtual de aprendizagem colaborativa, coordenada pela equipe DICl no IBICT.

Dentre as novas funcionalidades, destacamos:

- dispositivo de segurança para *logoff* do sistema;
- inclusão de novas opções de tipos de documentos para depósito;

- c) novas opções de navegação: por país, autor ou editor, título do livro ou periódico, e conferência;
- d) controle estatístico de uso do sistema; e
- e) FAQ (perguntas freqüentes).

4.2 Integração com E-LIS

Desde 2004 existe uma parceria entre o IBICT e o projeto E-LIS da CILEA (Conorzio Interuniversitario Lombardo per l'Elaborazione Automatica), Itália.

Em ambas as iniciativas existe uma prática de publicação semelhante, na qual os editores locais atuam como intermediários no depósito da produção científica da área.

Neste último aspecto, o engajamento dos autores é fundamental para a concretização do auto-arquivamento, de comentários e da publicação de novas versões de pré-prints depositados pelos participantes cadastrados. Também é grandemente importante o papel dos editores na avaliação dos trabalhos a serem aprovados para publicação nos espaços de comunicação que a ferramenta permite deverá ter maior reconhecimento.

Uma das metas a serem alcançadas na nova versão é a integração do DIC1 de CI com o E-LIS. Foi realizada a tradução do Esquema JITA¹ para classificação de assuntos da área de CI, ele funciona em dois níveis de descrição para cada elemento, procurando auxiliar no processo de submissão.

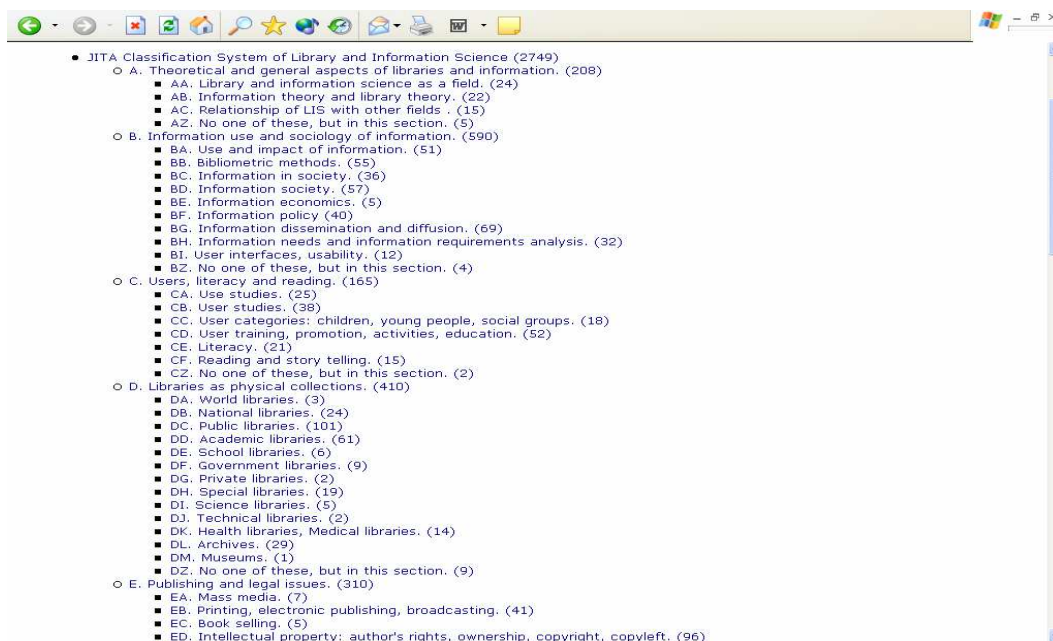


Figura 3. JITA – Sistema de Classificação em Biblioteconomia e Ciência da Informação.

¹ Várias instituições já usam o JITA. JITA é o acrônimo de José (Barrueco), Imma (Subirats Coll), Thomas (Krichel), Antonella (De Robbio) fundadores do E-LIS.

Também pretende-se trabalhar no sentido de estabelecer uma infra-estrutura tecnológica que promova a integração dos repositórios temáticos que usam o EPrints.

A criação desse sistema agregador torna-se viável com a implementação da camada de provedor de serviços dos dispositivos de *harvesting*, que efetuam a coleta dos metadados expostos pelos repositórios (provedores de dados) que adotam os mesmos padrões e protocolos.

No caso do padrão OAI, já existem diversas iniciativas desta natureza, porém, na área de Ciência da Informação, o único provedor de serviços conhecido é o Metalis (<http://metalis.cilea.it/>). Este serviço já realiza a coleta dos metadados dos seguintes repositórios:

- a) Archive Ouverte en Sciences de l'Information et de la Communication;
- b) ArXiv;
- c) Caltech Library System Papers and Publications;
- d) CNR Bologna Research Library;
- e) Digital Library of Information Science and Technology;
- f) E-Prints in Library and Information Science;
- g) Librarians' Digital Library;
- h) Mémoires de 3 cycle en Sciences de l'Information et de la Communication;
- i) Thèse-EN-ligne;

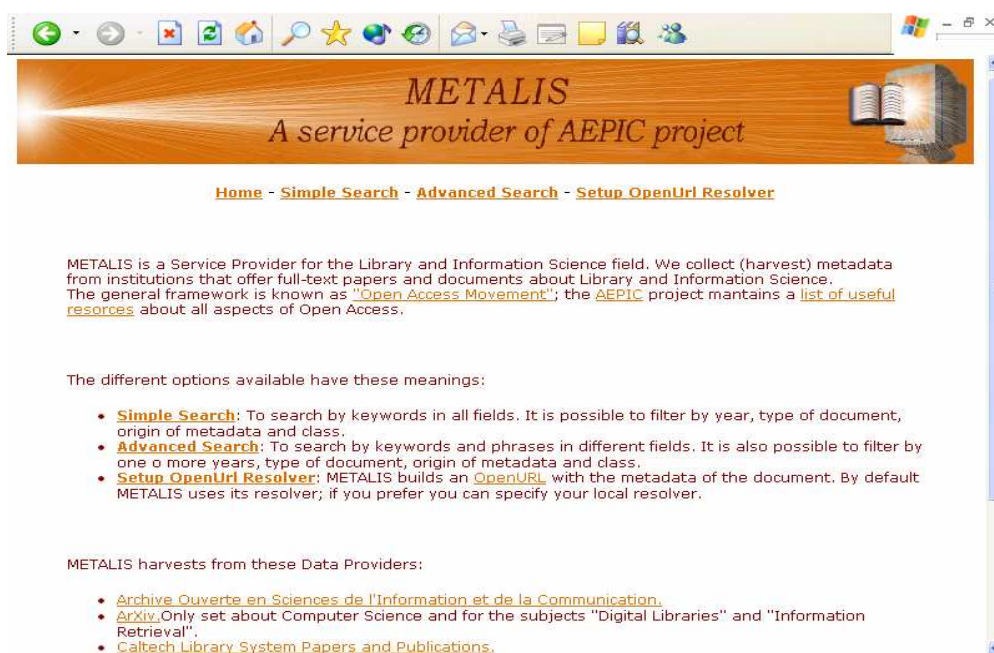


Figura 4. – Homepage do METALIS.

5. Conclusões

As praticas de arquivamento atuais, mesmo dentro de mecanismos como o Eprints continuam dependendo da dinâmica de comunicação de cada área. A adoção dessa ferramenta pelas instituições poderá favorecer o surgimento de novos

colégios invisíveis integrados em rede e a promoção de novas práticas através de políticas que estimulem a democratização do acesso livre à informação científica.

O IBICT publicou, em setembro de 2005, o Manifesto Brasileiro de apoio ao Acesso Livre à Informação Científica (<http://www.ibict.br/openaccess/>). Esta iniciativa contou com o apoio da Academia Brasileira de Ciências (ABC), Sociedade Brasileira para o Progresso da Ciência (SBPC), Rede Nacional de Ensino e Pesquisa (RNP) e universidades estaduais e federais. Na visão do Instituto, “criar repositórios não tem custo alto. A Internet também é uma vitrine onde podemos expor nossas idéias e torna-se importante o registro de conteúdos científicos brasileiros na Web” (Kuramoto, 2005). Também, como resultado deste movimento, a SBPC propôs que as agências de fomento invistam 10 por cento dos recursos destinados à compra de periódicos científicos na implantação de repositórios de acesso livre.

A perspectiva futura do projeto DICI é a de trabalhar colaborativamente com um maior número de instituições, tanto IES como Institutos de Pesquisa, no país, de modo a promover a integração, tanto dos repositórios brasileiros da área de CI, quanto o compartilhamento de práticas de depósito e discussão entre autores.

O pacote de *software* DICI está tecnicamente habilitado para interoperabilidade com outros sistemas cuja infra-estrutura tecnológica seja adequada aos padrões e protocolos da OAI. Esta sua propriedade é a mais importante para permitir sua integração posterior, com outros serviços de informação, tanto em nível nacional quanto internacional.

Integrando os esforços individuais e agregando toda a infra-estrutura tecnológica já implantada, esses sistemas podem disponibilizar o conteúdo integral de todo e qualquer documento que tenha sido depositado em algum repositório Eprints.

Tal como no caso do Metalis, pretende-se não somente aderir a essas iniciativas, mas também, fomentar o desenvolvimento desses provedores de serviços nas diversas áreas em que forem implementados repositórios temáticos de Eprints.

Deste modo o IBICT poderá contribuir para a dinamização do ciclo produtivo do conhecimento dessas comunidades científicas e manter-se engajado no processo histórico da transição dos mecanismos de divulgação e avaliação das ciências que, aparentemente, está levando ao surgimento de um novo modelo de comunicação científica.

5. Referências Bibliográficas

COSTA, Sely M. S. O novo papel das tecnologias digitais na comunicação científica. In: MARCONDES, Carlos H. , KURAMOTO, Hélio, TOUTAIN, Lídia B., SAYÃO, Luís. **Bibliotecas digitais: saberes e práticas**. Salvador: UFBA; Brasília: IBICT, 2005. p. 167-185.

DIÁLOGO Científico: Ciência da Informação. Disponível em: <<http://dici.ibict.br/>>. Acesso em 23 de março de 2006.

DIÁLOGO Científico: gráfico de registro cronológico de depósitos. In: REGISTRY of Open Access Repositories. Disponível em: <<http://archives.eprints.org/?url=http%3A%2F%2Fdici.ibict.br%2F>>. Acesso em 23 de março de 2006.

HARNAD, Stevan R. **Publish or perish — self-archive to flourish**: the green route to open access. Disponível em: <<http://eprints.ecs.soton.ac.uk/11715/01/harnad-ercim.pdf>>. Acesso em: 17 de março de 2006.

JITA classification schema of Library and Information Science. Disponível em: <<http://eprints.rclis.org/jita.html>>. Acesso em: 24 de março de 2006.

KURAMOTO, Hélio: [A iniciativa]. In: IBICT (Portal na Internet). **Notícias**: Ibict lança manifesto pelo acesso livre à informação científica. Disponível em: <<http://www.ibict.br/noticia.php?page=5&id=142>>. Acesso em: 24 de março de 2005a.

KURAMOTO, Hélio. Software livre para bibliotecas digitais. In: MARCONDES, Carlos H., KURAMOTO, Hélio, TOUTAIN, Lídia B., SAYÃO, Luís. **Bibliotecas digitais**: saberes e práticas. Salvador: UFBA; Brasília: IBICT, 2005b. p. 147-164.

LAGOZE, Carl, VAN DE SOMPEL, Herbert . **The open archives initiative**: building a low-barrier interoperability framework. Disponível em: <<http://www.openarchives.org/documents/jcdl2001-oai.pdf>>. Acesso em 17 de março de 2006.

METALIS: a service provider of AEPIC project. Disponível em: <<http://metalis.cilea.it/>>. Acesso em: 24 de março de 2006.

MOREIRA, Walter. Os Colégios virtuais e a nova configuração da comunicação científica. **Ciência da Informação**, Brasília, v.34, n.1, jan./abr.2005. p.57-63.

WEITZEL, Simone da R. E-prints: modelo da comunicação científica em transição. In: FERREIRA, Sueli M. S. P., TARGINO, Maria das G. **Preparação de revistas científicas**: teoria e prática. São Paulo: Reichmann e Autores, 2005. p. 161-193.