

Repositórios Digitais Institucionais: a experiência do Instituto de Pesquisas Energéticas e Nucleares (IPEN)

Mery P. Zamudio Igami

Biblioteca *Terezine Arantes Ferraz*

1 Repositórios Digitais Institucionais – RDI

A tecnologia da informação modificou significativamente todas as atividades rotineiras da sociedade. O surgimento das redes de comunicação transformou o ambiente onde elas atuam mudando o conceito de espaço e tempo no âmbito da comunicação e informação pela instantaneidade da transmissão de dados.

Na área de comunicação científica não foi diferente. Toda a cadeia de produção, disseminação, publicação e armazenamento de documentos científicos experimentou uma grande transformação, impactando de forma positiva a produção e consumo da informação científica na área acadêmica. O surgimento da internet e da WWW, na década de 80, aceleraram as mudanças na forma de comunicar e disseminar os trabalhos acadêmicos que passaram do suporte exclusivamente em papel para o suporte digital.

Na mesma época houve uma significativa crise mundial devido aos altos preços praticados pelas editoras no processo de assinatura das publicações periódicas, tornando inviável a sustentação das coleções pelas bibliotecas e afetando diretamente a consulta pelos usuários. Como reação da própria comunidade científica surgiu então o movimento de livre acesso à informação científica, mais conhecido por *Open Access Initiative* (KURAMOTO, 2012).

Esse movimento defende a disponibilização na internet da literatura acadêmica e científica permitindo que seja lida, descarregada, distribuída, impressa e pesquisada livremente.

É importante o registro histórico desse movimento não somente porque está diretamente relacionado à criação dos RDIs, mas também porque ele possibilitou várias mudanças estruturais no fluxo da comunicação científica. Os editores perderam a exclusividade de distribuição (comercialização) da pesquisa científica, uma vez que na era digital é possível que isso seja realizado de forma descentralizada e por iniciativa de cada autor.

1.1 O movimento do *Open Access Initiative*- OAI

A partir de 1999, os movimentos em favor do acesso livre à literatura científica, principalmente à literatura produzida sob financiamento público, consolidaram-se.

A comunidade acadêmica envolvida nesses movimentos corroborou o seu apoio por meio de declarações formais, dentro das quais se destaca a *Declaração de Budapest – BOAI* ¹ em 2002, onde foram emitidos os conceitos básicos do OAI, e a *Declaração de Berlim* ² em 2003, onde se destacou a natureza institucional dos repositórios digitais de livre acesso como um espaço organizado baseado em regras de utilização que conferem credibilidade aos materiais publicados e em políticas de arquivamento a longo prazo.

O movimento *Open Archives Initiative* desenvolveu padrões de interoperabilidade buscando facilitar a comunicação eficiente de conteúdo. Foram também estabelecidas duas modalidades para viabilizar a disponibilidade dos conteúdos: a *via verde* que diz respeito ao autodepósito nos RDs e a *via dourada* que inclui as revistas que publicam os artigos com livre acesso.

O primeiro repositório digital (ArXiv.org) ³ foi instalado no laboratório de física, em Los Álamos, EU, coordenado pelo físico Paul Ginsparg (1996), motivado pela insatisfação com os rumos na política de aquisição das publicações científicas e a crescente inviabilidade das bibliotecas americanas de manter as coleções atualizadas. Daquela época para a atual aconteceram inúmeras mudanças positivas nos modelos de funcionamento dos repositórios promovidas pela evolução da TI.

¹ Budapest Open Access Initiative - <http://www.budapestopenaccessinitiative.org>

² Open Access - <http://openaccess.mpg.de/Berlin-Declaration>

³ General Information about Arxiv - <http://arxiv.org>

Não existe um modelo único para a criação dos RDIs; cada instituição define a política de depósito adequada ao contexto da sua comunidade. O importante é que essa comunidade valide e utilize a plataforma tornando-a uma vitrine da sua produção científica.

No país, o movimento de Acesso Aberto e construção de Repositórios Digitais Institucionais tiveram um grande incentivo e apoio técnico do Instituto Brasileiro de Informação em Ciência e Tecnologia – (IBICT) (INSTITUTO BRASILEIRO DE CIÊNCIA E TECNOLOGIA, 2009). De acordo com o Diretório internacional de RDIs – o DOAR⁴, até o momento, há no país 153 repositórios digitais em funcionamento, dos quais a maioria é mantida pelas instituições onde a comunidade produz os trabalhos científicos. Desses RDIs, 74% utilizam o *software Dspace*⁵.

1.2 O Projeto de Lei sobre os Repositórios Digitais

Em 2007, foi apresentado o Projeto de Lei no senado PL1120/2007 (BRASIL, 2007), que torna obrigatório nas universidades e institutos de pesquisa o desenvolvimento e implantação de RDIs nas instituições de origem, bem como torna obrigatório o autodepósito da produção científica por parte dos seus pesquisadores. Além disso, o Projeto de Lei determina a formação de uma comissão de alto nível para discutir, propor e estabelecer uma política de livre acesso à informação científica. Além de simplesmente criar mecanismos de disseminação da informação científica, o Projeto propõe a criação de uma comissão para discutir políticas e estabelecer diretrizes a serem adotadas em nível nacional pelas instituições que produzem informação científica.

1.3 O que são os RDIs?

Os RDIs constituem atualmente um avanço na gestão do conhecimento intelectual. Podem receber diversas denominações tais como equipamentos institucionais, observatórios digitais, instrumentos ou ferramentas de gestão de conhecimento e outras.

⁴ The Directory of Open Access Repositories – *Open DOAR* - <http://www.opendoar.org>

⁵ Software livre para construção de Institucionais Digitais (Dspace) - <http://www.ibict.br/pesquisa-desenvolvimento-tecnologico-e-inovacao/Sistema-para-Construcao-de-Repositorio-digital-Institucionais-Digitais>

Um dos aspectos mais importantes dos RDIs é que eles estão inseridos na política institucional de gestão do conhecimento e são essencialmente instrumentos gerenciais de onde se obtém indicadores para auxiliar na gestão da política científica das instituições de pesquisa.

Sob a ótica documentária um RDI constitui um serviço de informação científica em ambiente digital, interoperável e de acesso aberto, dedicado ao gerenciamento da produção científica de uma instituição. Contempla a reunião, armazenamento, organização, preservação, recuperação e, sobretudo, a ampla disseminação e visibilidade da informação científica produzida na instituição.

1.3.1 O que caracteriza um repositório digital?

- a) a organização, armazenamento, acesso e distribuição de documentos digitais;
- b) a utilização de padrões abertos acessibilidade;
- c) as funcionalidades que o *software* selecionado permite;
- d) a utilização de metadados com padrão internacional – interoperabilidade;
- e) o arquivo digital anexado;
- f) o acesso de buscadores para elaboração de *rankings* internacionais - visibilidade;
- g) o autodepósito dos trabalhos;
- h) a monitoração, controle e elaboração de indicadores.

2 O RDI do IPEN

O prestígio de uma instituição de pesquisa está fortemente vinculado aos resultados dos elementos intangíveis, por exemplo, qualidade no ensino da pós-graduação, e tangíveis como os resultados da pesquisa, entre eles a Produção científica. Tradicionalmente, o IPEN, por meio da sua unidade de informação, mantém a sua produção científica organizada e atualizada desde sua criação em 1956. No entanto, devido ao avanço da tecnologia

da informação, tornou-se necessário atualizar a plataforma gerenciadora dessa informação já armazenada, com a consequente modernização no processo de gestão do capital intelectual.

No decorrer de 2013, seguindo outras instituições no país, como por exemplo, a Universidade de São Paulo, a Universidade Estadual Paulista e a Universidade Estadual de Campinas, a Direção do IPEN decidiu criar o seu repositório, uma vez que a matéria prima para concretizar esta ação já se encontrava disponível e organizada de tal forma que se garantiria a sustentabilidade do RDI.

A política de gestão da produção científica é uma prática rotineira no IPEN. Diferentemente da maioria dos repositórios em funcionamento, o RDI do IPEN já nasce com um volume significativo de matéria-prima digital organizada, com políticas definidas de auto arquivamento e de preservação da produção científica.

Porém, o fator determinante de sucesso foi o modelo de indexação e organização dos documentos adotado pela biblioteca desde o início em 1976, quando ainda se constituía em uma base de dados referencial. Este fato facilitou o processo de migração e compatibilização dos metadados para o novo padrão internacional adotado, o *Dublin Core*, garantindo interoperabilidade com outros repositórios digitais, uma das principais características dos RDIs.

Em novembro de 2013, foram iniciados os trabalhos de construção do RDI do IPEN. Inicialmente foram programadas visitas do grupo de informática e da biblioteca a instituições com experiências bem-sucedidas, como, por exemplo, o Instituto de Pesquisas Espaciais – (INPE) em São José dos Campos e a Universidade de São Paulo- USP, com o objetivo de conhecer melhor o projeto de RDIs dessas instituições e selecionar o modelo mais adequado à realidade do IPEN.

As visitas técnicas embasaram a decisão de selecionar o *software Dspace* para confecção do repositório do IPEN. O Dspace é um *software* livre especializado no gerenciamento de repositórios resultante de um projeto colaborativo da *MIT Libraries* e da *Hewlett-Packard Company*, amplamente utilizado no mundo. No Brasil, graças ao incentivo do IBICT em Brasília, o DSpace é utilizado por 75% das instituições que possuem repositórios.

2.1 O processo de migração

Em dezembro de 2013, foi iniciado o processo de criação do novo repositório do IPEN, após um estudo preliminar do *software* e contratação de uma empresa para a capacitação das equipes da biblioteca e da informática.

Nos meses seguintes foram efetuados os testes de implantação do *Dspace*, configuração, desenho da nova arquitetura do banco de dados, compatibilização e criação de novos metadados, desenvolvimento de funcionalidades.

Na época, a produção científica contabilizada era de 19.200 itens. A migração gradual desses itens iniciou-se em julho de 2014.

Internamente, além de todas as funções inerentes e habituais a um repositório, o diferencial do RDI do IPEN diz respeito à integração com os dados registrados no sistema de informação gerencial e planejamento do instituto, o SIGEPI. Trata-se de uma funcionalidade especial, desenvolvida em ambiente externo ao *Dspace*, porém acrescentada no momento do registro do documento no *Dspace* via metadado específico, de tal forma que cada trabalho registrado no RD se correlaciona, automaticamente com a atividade do Plano Diretor à qual o autor do IPEN está vinculado.

Para essa finalidade foi efetuado um exaustivo trabalho de identificação dos nomes completos dos autores do IPEN (lista de autoridades), com as suas respectivas variáveis na forma de citação e agrupamento dos trabalhos. Foi atribuído um número único de identificação para cada autor, o *ID autor IPEN*, formato semelhante ao *Researcher ID* utilizado pela base internacional *Web of Science*. Essa facilidade constituiu um filtro que ao ser acionado reúne e apresenta rapidamente todos os trabalhos de um determinado autor.

Ao RDI está vinculada também uma tabela estatística, atualizada *on-line* a cada registro introduzido. Essa funcionalidade destina-se à obtenção de dados para a elaboração de índices e quadros estatísticos a serem confeccionados fora do ambiente do *Dspace* para subsidiar os pesquisadores e os gestores do IPEN, com informações para a monitoração, administração e elaboração de indicadores científicos fornecidos para o Plano Diretor do IPEN, o relatório anual da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior(CAPES) e outros, bem como para cumprir as exigências dos órgãos financiadores de pesquisa.

O modelo adotado para a coleta dos trabalhos também difere de outros. Nesta primeira etapa foi considerado mais conveniente o serviço de informação (biblioteca) responsabilizar-se pela coleta, depósito e complementação dos dados, utilizando as fontes de informação nacionais e internacionais. Assim, a unidade de informação realiza periodicamente a verificação dos trabalhos publicados pelos pesquisadores nas bases *Web of Science*, *Scopus*, *Scielo*, *INIS* e Currículo Lattes. Após a eliminação das duplicatas e localização dos documentos na íntegra efetua-se o depósito no RDI.

Trata-se de um modelo de depósito que está sendo testado, há vantagens e desvantagens. A principal vantagem diz respeito à economia de tempo do pesquisador uma vez que se elimina a obrigatoriedade do autodeposito, tornando, porém, imprescindível por parte do pesquisador à atualização do seu Currículo Lattes. A maior desvantagem foi absorvida pela Biblioteca, qual seja a coleta periódica dos documentos publicados bem como a localização e obtenção da evidência dos mesmos.

Importante registrar que a construção do RD foi uma iniciativa muito nova na instituição, que demandou um trabalho detalhado, cuidadoso e gradual das equipes de desenvolvimento de sistemas e da biblioteca, dado a pouca familiaridade com a nova plataforma. Alia-se a esse fato a opção de utilizar um software livre, onde certas funcionalidades ainda não foram desenvolvidas.

Em setembro de 2014, após oito meses de trabalho, as equipes conseguiram migrar 23% das coleções, foi uma tarefa que demandou ajustes, adequações, testes e decisões internas, dado o grande nível de detalhamento dos dados, motivo de comemoração das equipes, após cada etapa finalizada e bem sucedida. Contamos também com a assessoria de uma empresa especializada no assunto, a qual guiou a equipe da informática em cada funcionalidade desenvolvida. Nesse ritmo de trabalho, foi possível completar o processo de migração em fevereiro de 2015. Finalmente, em junho desse mesmo ano, o RDI foi disponibilizado na Internet.

2.2 Desempenho do RDI em 2015

Atualmente, o novo RDI:

- ✓ dispõe de uma identidade própria e permanente; (ex: ;
- ✓ utiliza as estatísticas de uso do *Google Analytics*;

- ✓ é reconhecido pelo Google Scholar e outras máquinas de busca da Internet;
- ✓ participa dos *rankings* internacionais de repositórios institucionais (ex.: Webometrics);
- ✓ tem registro nos Diretórios Internacionais de RIs (ex.: DOAR);
- ✓ interoperava com sistemas congêneres.

Conforme foi mencionado, a criação do RDI e a migração dos dados constitui a primeira etapa dentro de um projeto maior, qual seja institucionalizar e operacionalizar a gestão da produção científica do IPEN. Após o processo se consolidar em entrar em rotina de funcionamento, as próximas etapas a serem desenvolvidas pelas equipes, dizem respeito à:

- ✓ automação da coleta dos documentos;
- ✓ otimização das ferramentas de busca;
- ✓ obtenção de relatórios estatísticos customizados;
- ✓ otimização do *lay-out* do RDI;
- ✓ acréscimo de novas coleções;
- ✓ desenvolvimento de novas funcionalidades.

A pesquisa no RDI é muito simples e segue os padrões da maioria das bases de dados. É possível pesquisar por autor, título, assunto, ID de autor IPEN, ano de publicação, combinar palavras-chave de busca e assim por diante. Apresenta também uma tabela estatística funcionando em tempo real, distribuída por nível monográfico; isso permite que a direção e o corpo técnico científico acompanhe a evolução quantitativa dos documentos depositados no RDI.

Uma das últimas funcionalidades acrescentadas ao RD, diz respeito à sinalização dos autores do IPEN. Ao clicar no símbolo do autor é possível acessar a página do pesquisador, onde se apresentam os seguintes dados: as variáveis do nome, uma tabela com todos os trabalhos constantes no RDI, bem como um quadro resumo numérico. Dessa forma o autor pode visualizar, acessar e monitorar à sua produção científica com o texto completo de forma rápida e simples. O quadro resumo se configura ainda como uma linha do tempo para o autor.

Por se constituir em um equipamento totalmente apoiado na tecnologia da informação, o RDI é um recurso institucional dinâmico sujeito a constantes atualizações e melhorias, o qual deve garantir a confiabilidade dos dados providos bem como assegurar a sua sustentabilidade.

3 Referências bibliográficas

BRASIL. Projeto de Lei do Senado 1120/2007 Dispõe sobre o processo de disseminação da produção técnico-científica pelas instituições de ensino superior no Brasil e dá outras providências. ***Diário oficial da Republica Federativa***, Brasília, DF, Disponível em: <
<http://www.camara.gov.br/sileg/integras/461698.pdf> > Acesso em: set. 2014.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT Repositórios digitais. Site institucional. Brasília: 2009. Disponível em:
<http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorio-digitais-digitais> >Acesso em set. 2014.

KURAMOTO, H. ***Acesso livre***: como tudo começou. Brasília, 2012. Blog. Disponível em: <Acesso em: set. 2014.

SP, janeiro de 2016