

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

Mery P. Zamudio Igami

Biblioteca *Terezine Arantes Ferraz*

1 Repositórios Digitais Institucionais

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 financiada por recursos públicos, 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 Repositórios Digitais, 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 destacam : a Declaração de

Budapest – BOAI – ¹ em 2002, onde foram emitidos os conceitos básicos do OAI, e a Declaração de Berlim² 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 auto depósito 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.

Não existe um modelo único para a criação dos RDs; 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 RDs – 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 RDs, 74% utilizam o software Dspace⁵.

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 RDs nas instituições de origem, bem como torna obrigatório o auto depósito 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.

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>

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

⁵ Software livre para construção de Repositórios Institucionais Digitais (Dspace) – <https://antigo.ibict.br/tecnologias-para-informacao/dspace>

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 RDs?

Os RDs se constituem atualmente em um avanço na gestão do conhecimento intelectual. Podem receber diversas denominações tais como equipamentos institucionais, observatórios digitais, repositórios digitais, instrumentos ou ferramentas de gestão de conhecimento e outras. Os RDs também podem ser classificados, dependendo do seu objetivo e contexto onde estão instalados como Repositórios temáticos, disciplinares e institucionais.

Um dos aspectos mais importantes dos RDs é que eles estão inseridos na política institucional de gestão do conhecimento, são essencialmente instrumentos gerenciais de onde podem ser obtidos indicadores para auxiliar na gestão da política científica das instituições onde operam. Sob a ótica documentária um RD se constitui em um serviço de informação científica funcionando em ambiente digital, inter operável 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 auto depósito dos trabalhos;
- h) a monitoração, controle e elaboração de indicadores.

2 O Repositório Digital (RD) 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 desse conhecimento intelectual explícito.

No decorrer do segundo semestre 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 RD.

Diferentemente da maioria dos repositórios em funcionamento, o RD do IPEN já nasce com um volume significativo de matéria-prima digital organizada e indexada via vocabulário controlado especializado na área nuclear, com políticas definidas para o arquivamento e a 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. Todos os documentos registrados na antiga base de dados receberam um número único de controle interno, este fato facilitou o processo de migração e compatibilização dos metadados para o novo padrão internacional adotado, o Dublin Core, que por sua vez garante a interoperabilidade com outros repositórios digitais, uma das principais características dos RDs.

Em novembro de 2013, foram iniciados os trabalhos de construção do RD 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 RDs 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. Trata-se de 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 metadados específicos, 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 se constituiu em um filtro que ao ser acionado reúne e apresenta rapidamente todos os trabalhos de um determinado autor.

2.2. O modelo de depósito do RD

O modelo adotado para a coleta dos trabalhos também difere de outros. Nesta primeira fase foi considerado mais conveniente que o serviço de informação (Biblioteca) se responsabilizasse pela coleta, depósito e complementação dos dados, utilizando as fontes de informação nacionais e internacionais. Ao contrário do modelo de auto depósito adotado pela maioria dos RDs. 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 RD.

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 auto depósito, tornando, porém, imprescindível por parte do pesquisador a atualização do seu Currículo Lattes. A maior desvantagem foi absorvida pela Biblioteca, qual seja o processo da coleta periódica dos documentos publicados bem como a localização e obtenção da evidência dos mesmos. Deve-se considerar que foi possível adotar esse modelo de funcionamento do RD por se tratar de uma única comunidade atuante em um Instituto de Pesquisas, nada comparado as múltiplas comunidades atuantes, por exemplo em uma universidade.

Importante registrar também 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 (TI) 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 RD entrou em operação e foi disponibilizado na Internet.

2.3. A integração do RD com o sistema de gestão do IPEN

O modelo administrativo do IPEN é matricial e está dividido em Centros por área de especialização. Os pesquisadores integrantes de cada Centro geram e publicam seus trabalhos científicos. O RD contempla a mesma divisão e os Centros são codificados e registrados em metadados específicos. O Plano Diretor da Instituição é o instrumento unificador e de gestão. O RD está vinculado a esse Plano de gestão para fornecer os indicadores da Produção científica. Quando um documento é depositado no RD, este é registrado automaticamente ao Centro ao qual os autores do trabalho pertencem, à tabela estatística, à página do pesquisador e ao RD. 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 elaboração de indicadores científicos, registrados no relatório anual para a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) para o MCTI e outros, bem como para cumprir as exigências dos órgãos financiadores de pesquisa.

2.4 Desempenho do RD a partir de 2015

- ✓ - dispõe de uma identidade própria e permanente: <http://repositório.ipen.br>
- ✓ - 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 RDs (ex.: DOAR);
- ✓ - interopera com sistemas congêneres.

Conforme foi mencionado, a criação do RD 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;
- ✓ - acréscimo de novas coleções;
- ✓ - desenvolvimento de novas funcionalidades e atualização de versões do Dspace .

A pesquisa no RD 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 RD.

Uma das últimas funcionalidades acrescentadas ao RD em 2018, diz respeito à marcação dos autores do IPEN com o logotipo do IPEN. Ao clicar nesse logotipo no nome 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 RD, 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.

Em 2019, foi identificada a necessidade de acrescentar mais uma funcionalidade ao RD. Objetivando facilitar o processo de referência para os alunos, foi instalada uma funcionalidade para apresentar a referência bibliográfica, elaborada de acordo com a normas da ABNT (adaptação do IPEN) de cada tipo de documento consultado.

Por se constituir em um equipamento totalmente apoiado na tecnologia da informação, o RD 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

1 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 República Federativa, Brasília, DF, Disponível em: < <http://www.camara.gov.br/sileg/integras/461698.pdf> > Acesso em: set. 2014.

2 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.

3 KURAMOTO, H. Acesso livre: como tudo começou. Brasília, 2012. Blog. Disponível em: <https://kuramoto.wordpress.com/2012/02/27/acesso-livre-como-tudo-comecou/> > Acesso em nov. 2015