

Repositórios digitais da Embrapa: acesso livre a produção técnico-científica

Rosângela Lacerda de Castro (EMBRAPA) - rosangela.lacerda@embrapa.br

Marilaine Schaun Pelufê (EMBRAPA) - maripelufe@gmail.com

Rosângela Galon Arruda (Embrapa) - rosangela.arruda@embrapa.br

Resumo:

O Movimento de Acesso Aberto revolucionou a comunicação científica possibilitando novas formas de publicação, divulgação e, principalmente, maior visibilidade da pesquisa. Os repositórios institucionais são produtos deste movimento e caracterizam-se como ferramentas capazes de organizar e tornar acessíveis conteúdos de universidades e instituições produtoras de conhecimento científico. Este trabalho descreve os repositórios institucionais da Embrapa: o Repositório de Acesso Livre à Informação Científica (ALICE) e o serviço de Informação Tecnológica em Agricultura (Infoteca-e), citando suas características, estatísticas de consultas e downloads, procedimentos para depósito e critérios adotados para preservar os documentos digitais nesses repositórios.

Palavras-chave: *Repositório institucional. Embrapa. Pesquisa agropecuária. Produção científica. Comunicação científica*

Área temática: *Temática I: Tecnologias de informação e comunicação – um passo a frente*

Repositórios digitais da Embrapa: acesso livre a produção técnico-científica

Resumo

O Movimento de Acesso Aberto revolucionou a comunicação científica possibilitando novas formas de publicação, divulgação e, principalmente, maior visibilidade da pesquisa. Os repositórios institucionais são produtos deste movimento e caracterizam-se como ferramentas capazes de organizar e tornar acessíveis conteúdos de universidades e instituições produtoras de conhecimento científico. Este trabalho descreve os repositórios institucionais da Embrapa: o Repositório de Acesso Livre à Informação Científica (ALICE) e o serviço de Informação Tecnológica em Agricultura (Infoteca-e), citando suas características, estatísticas de consultas e downloads, procedimentos para depósito e critérios adotados para preservar os documentos digitais nesses repositórios.

Palavras-chave: Repositório institucional. Embrapa. Pesquisa agropecuária. Produção científica. Comunicação científica.

Área temática: Tecnologias de Informação e Comunicação – um passo a frente

1 INTRODUÇÃO

A internet como instrumento de acesso a informação e ao conhecimento possibilita novas formas de armazenamento e recuperação da informação, dentre elas estão as publicações eletrônicas e os repositórios digitais.

As publicações eletrônicas disponibilizadas na Web têm grande repercussão dentro da comunidade científica, pois permitem o acesso e divulgam a informação em larga escala, revolucionando toda a estrutura de produção, disseminação e acesso à informação que estava em vigor antes do advento da Internet (LOPES, 2004).

Moreno, Leite e Arellano (2006), corroboram esta afirmação ao dizer que:

Os periódicos eletrônicos, os repositórios institucionais e os repositórios temáticos de acesso aberto, aliados à tecnologia de arquivos abertos estão sendo utilizados pelas comunidades científicas para apoiar e tornar mais ampla a divulgação dos resultados das pesquisas bem como maximizar o seu impacto, criando mecanismos para legitimar e estimular a publicação dos trabalhos produzidos. (MORENO; LEITE; ARELLANO, 2006, p. 84).

A expressão *acesso aberto* significa a disponibilização livre na Internet da literatura acadêmica ou científica, permitindo a qualquer pessoa ler, baixar, copiar, imprimir o texto integral dos documentos.

Para Fachim et al. (2009):

O Movimento de Acesso Aberto à Informação Científica vêm propondo que a informação científica seja disponibilizada gratuitamente; o que é favorecido pelos avanços constantes das tecnologias de informação e comunicação (TIC) dos últimos anos, gerando uma demanda do uso da web para a disseminação dos resultados de pesquisas. (FACHIM et al., 2009, p. 220).

Brody e Harnard (2004) vão além ao afirmar que o acesso aberto a produção científica maximiza o acesso e acelera o impacto da pesquisa e consequentemente sua produtividade, progresso e recompensas.

Para a comunicação científica, o acesso aberto possibilitou novas formas de publicação, divulgação e principalmente, maior visibilidade da pesquisa, a qual é matéria prima para o desenvolvimento de um país.

Um dos elementos principais do movimento de acesso aberto é a criação de repositórios institucionais. Mais do que uma ferramenta tecnológica, estes têm sido compreendidos como um modelo alternativo de comunicação científica que complementa as vias tradicionais de disseminação do conhecimento (LEITE, 2009).

Neste contexto, foram criados sistemas de informação digital abertos denominados repositórios digitais, tendo por base metodologias validadas em instituições de reconhecido mérito técnico e científico no Brasil (Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT) e no exterior (Massachusetts Institute Of Technology – MIT e Scholarly Publishing and Academic Resources Coalition – SPARC) que emergiram do movimento de acesso aberto, cuja finalidade é armazenar, preservar, divulgar e possibilitar o acesso à produção científica, tecnológica, cultural e artística da instituição.

2 REPOSITÓRIO DIGITAL

Com o surgimento de iniciativas de comunicação científica no modelo acesso aberto, são criados a partir de 2002 os repositórios com mecanismo de

coleta de metadados, a fim de potencializar não só a organização e o armazenamento dos conteúdos, mas também sua recuperação de forma mais rápida e mais ampla do que o modelo do processo de comunicação científica até então em vigor. Conforme Leite e Costa (2006), isso acontece porque os repositórios devem, necessariamente, utilizar protocolos como o OAI-PMH (Protocol for Metadata Harvesting), que permite a interoperabilidade dos conteúdos armazenados em arquivos abertos.

Os repositórios digitais podem ser considerados como serviços de informação científica - em ambiente digital - dedicados ao gerenciamento de toda a produção intelectual de uma instituição. Eles reúnem um conjunto de documentos coletados, armazenados, organizados e disponibilizados eletronicamente que permitem o livre acesso a seu conteúdo. Heery e Anderson (2005) citam algumas características que diferem os repositórios de outras coleções digitais: o conteúdo disponibilizado nos repositórios é depositado pelo seu criador ou por terceiros; a arquitetura do repositório gerencia tanto seu conteúdo quanto seus metadados; o repositório oferece um conjunto mínimo de serviços básicos, por exemplo: upload, download, pesquisa e controle de acesso.

Os repositórios surgiram inicialmente como repositórios digitais temáticos, que incluíam apenas o conteúdo de determinada área do conhecimento. Em seguida a ideia evoluiu para que esses repositórios se agrupassem e ficassem sob a responsabilidade de uma instituição e voltados para a divulgação da produção científica, passando para a denominação de Repositório Institucional (ROSA, 2011).

Mas existem repositórios digitais com ambas as características, sendo institucional e temático, como é o caso dos Repositórios Digitais da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), que são repositórios institucionais temáticos na área de ciências agrárias e afins.

3 REPOSITÓRIOS INSTITUCIONAIS DA EMBRAPA

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa), lançou, em abril de 2011, dois repositórios institucionais, conforme a natureza da informação (científica ou tecnológica): Acesso Livre à Informação Científica da Embrapa (Alice)

e o serviço Informação Tecnológica em Agricultura (Infoteca-e). Com o objetivo de aumentar a visibilidade e acessibilidade de sua produção técnico-científica, as quais se relacionam às áreas de atuação dos centros de pesquisa da Embrapa, os repositórios foram criados para reunir, armazenar, organizar, recuperar, preservar e permitir o acesso, com texto completo, da produção técnico-científica dos pesquisadores da instituição.

Chan e Costa (2005), citados por Kuramoto (2006), defendem que o acesso ao conhecimento, basicamente na agricultura, medicina e tecnologia, pode ajudar a criar uma forte infraestrutura social, econômica e técnica, essenciais ao processo de desenvolvimento.

Para Alberts (2002), citado por Leite e Costa (2006), a informação científica e técnica é, fundamentalmente, um bem público global, que deve estar livremente disponível para o benefício de todos.

Os repositórios institucionais da Embrapa são produtos de informação desenvolvidos pelo Sistema Embrapa de Bibliotecas (SEB), sob a responsabilidade da Embrapa Informação Tecnológica, e foram construídos utilizando o software DSpace¹, que, segundo Santos Júnior (2010), é recomendado pelo IBICT para a construção de repositórios digitais no Brasil.

Alimentados a partir das bases de dados gerenciadas pelas bibliotecas dos centros de pesquisa da Embrapa, por meio do sistema de gerenciamento de bibliotecas Ainfo², os repositórios intervêm em duas questões estratégicas: contribuem para o aumento da visibilidade da instituição e servem como indicador tangível da sua qualidade (RODRIGUES, 2007 citado por FACHIN et al., 2009). A seguir são descritas características, estatísticas, procedimentos para depósito e os critérios adotados para a preservação dos documentos digitais que estão armazenados nos repositórios da Embrapa.

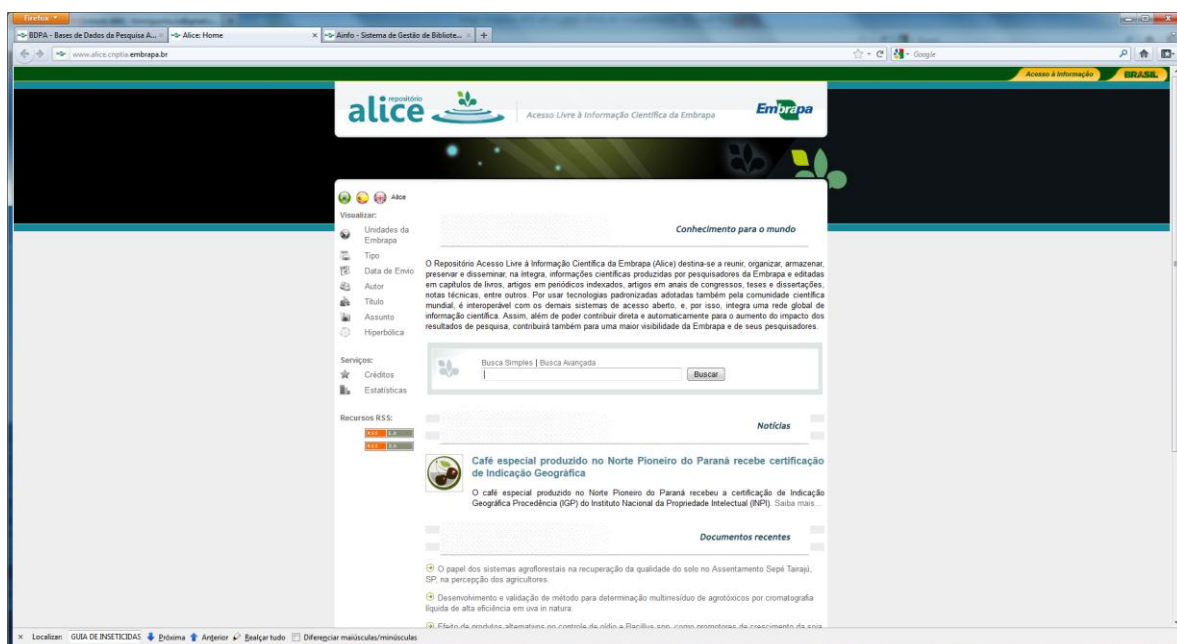
3.1 Repositório Acesso Livre à Informação Científica da Embrapa (Alice)

¹ Dspace é um software de código de fonte aberto que fornece facilidades para o gerenciamento de acervos digitais.

² Ainfo é um sistema de gerenciamento de bibliotecas que compreende a organização, controle e disseminação de informações, tanto em meio físico como digital, desenvolvido pela Embrapa Informática Agropecuária.

O Repositório Acesso Livre à Informação Científica da Embrapa (Alice) (Figura 1) destina-se a reunir, organizar, armazenar, preservar e disseminar, na íntegra, informações científicas produzidas por pesquisadores da Embrapa e editadas em capítulos de livros, artigos em periódicos indexados, artigos em anais de congressos, teses e dissertações, notas técnicas, entre outros. Por usar tecnologias padronizadas adotadas também pela comunidade científica mundial, é interoperável com os demais sistemas de acesso aberto, e, por isso, integra uma rede global de informação científica. Assim, além de poder contribuir direta e automaticamente para o aumento do impacto dos resultados de pesquisa, contribuiu também para uma maior visibilidade da Embrapa e de seus pesquisadores.

Figura 1 – Homepage do Alice



Fonte: <<http://www.alice.cnptia.embrapa.br>>. Acesso em: 27 mar. 2013.

3.2 Informação Tecnológica em Agricultura (Infoteca-e)

O serviço Informação Tecnológica em Agricultura (Infoteca-e) (Figura 2) reúne e permite acesso a informações sobre tecnologias produzidas pela Embrapa relacionadas com as áreas de atuação de seus centros de pesquisa. Suas coleções compreendem conteúdos editados pela própria Embrapa, com linguagem adaptada de modo que produtores rurais, extensionistas, técnicos agrícolas, estudantes e

professores de escolas rurais, cooperativas e outros segmentos da produção agrícola possam assimilá-los com maior facilidade e, assim, apropriarem-se de tecnologias geradas pela Embrapa.

Figura 2 – Homepage do Infoteca-e



Fonte: <<http://www.infoteca.cnptia.embrapa.br/>>. Acesso em: 27 mar. 2013.

Atualmente, os repositórios da Embrapa contam com 63.080 itens sendo 39.347 itens no Alice e 23.733 itens no Infoteca-e. Esses números são constantemente atualizados uma vez que registros e uploads de toda a produção científica, que alimenta os repositórios, são realizados diariamente no sistema Ainfo.

3.3 Dados Estatísticos

Os repositórios digitais da Embrapa foram criados em 2011 e neste curto período de existência já contabilizam números extraordinários de consultas e downloads.

As estatísticas dos repositórios são geradas nas próprias páginas dos repositórios, pois o DSpace permite, através do módulo de estatística, obter informações sobre consultas e downloads inclusive com a identificação do país de origem dessas consultas e downloads, estatísticas de itens e coleções que foram mais consultados ou que tiveram o maior número de downloads, além de algumas

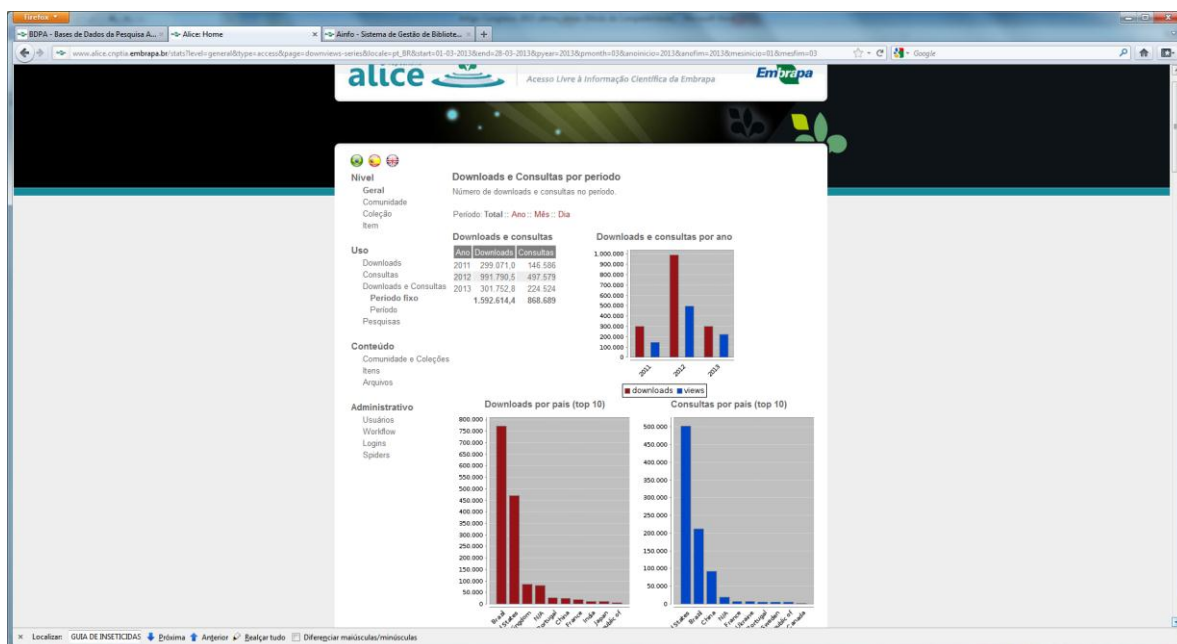
estatísticas administrativas.

3.3.1 Estatísticas de uso do repositório Alice

No ano de 2011 foram realizadas no repositório Alice 146.586 consultas e 299.071 downloads. Em 2012 o Alice contabilizou 497.579 consultas e 991.790 downloads. Em 2013 já conta com 224.524 consultas e 301.752 downloads.

Desde a sua criação em 2011 até março de 2013, o repositório Alice totaliza 868.689 consultas e 1.592.613 downloads³ (Figura 3).

Figura 3 – Número de Consultas e Downloads no Alice



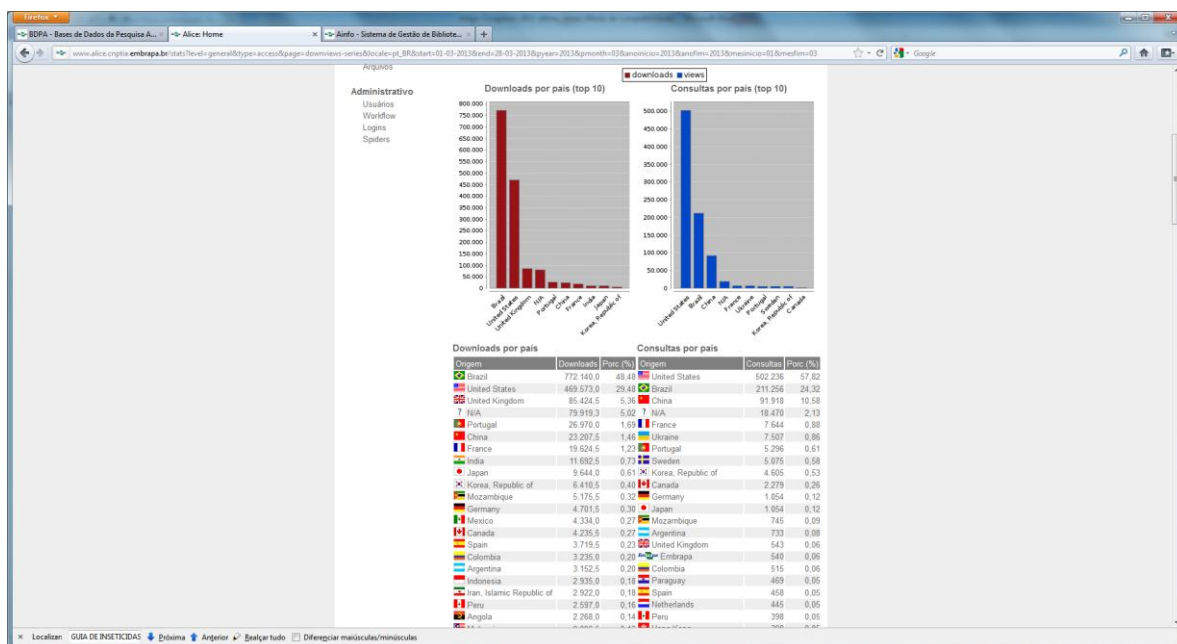
Fonte: <<http://www.alice.cnptia.embrapa.br>>. Acesso em: 28 mar. 2013.

Quanto ao país de origem das consultas, 57% são realizadas pelos Estados Unidos, seguido por Brasil (25%); China (10%) e 9% é pulverizado em vários outros países do mundo (Figura 4).

Os países que mais fizeram downloads foram o Brasil com 48%, os Estados Unidos com 29% e o Reino Unido 5%. 18% de downloads foram feitos por outros países no mundo.

³ Dados coletados em 28 de março de 2013

Figura 4 – Consultas e Downloads por país no Alice



Fonte: <<http://www.alice.cnptia.embrapa.br>>. Acesso em: 27 mar. 2013.

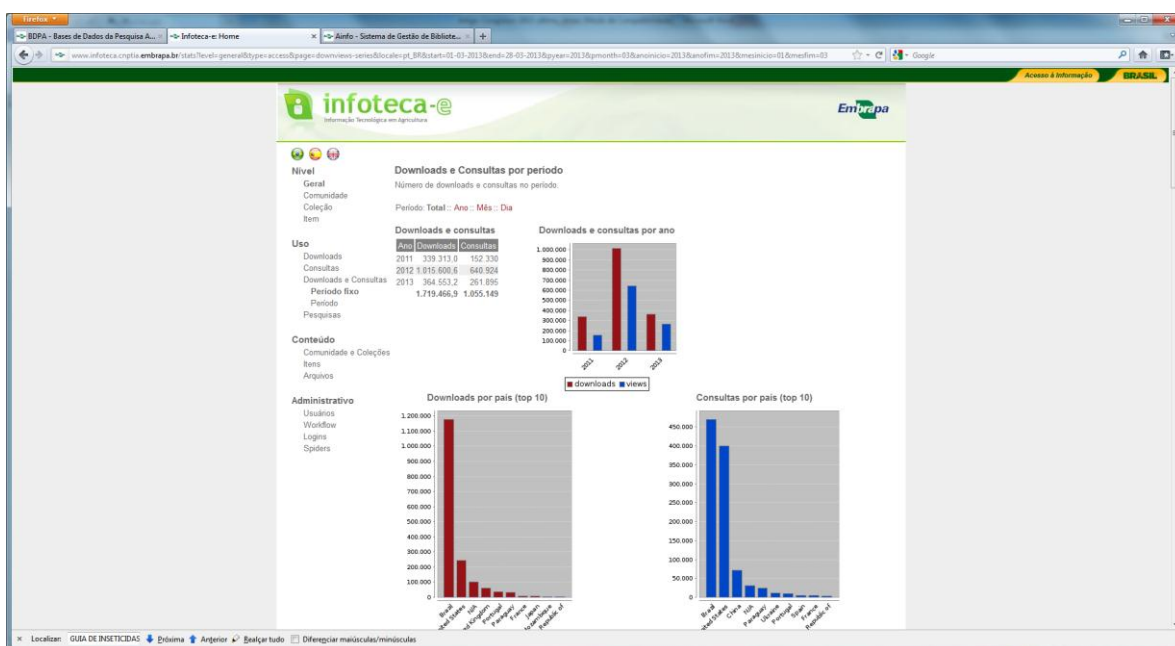
3.3.2 Estatísticas de uso do repositório Infoteca-e

A Infoteca-e contabilizou no ano de 2011 152.330 consultas e 339.313 downloads. Em 2012, foram efetuadas 640.924 consultas e 1.015.600 downloads. No ano de 2013 já contabiliza 261.895 consultas e 364.553 downloads.

Desde a sua criação em 2011 até março de 2013, a Infoteca-e já totalizou 1.055.149 consultas e 1.719.466 downloads⁴ (Figura 5).

⁴ Dados coletados em 27 de março de 2013

Figura 5 – Downloads e Consultas na Infoteca-e

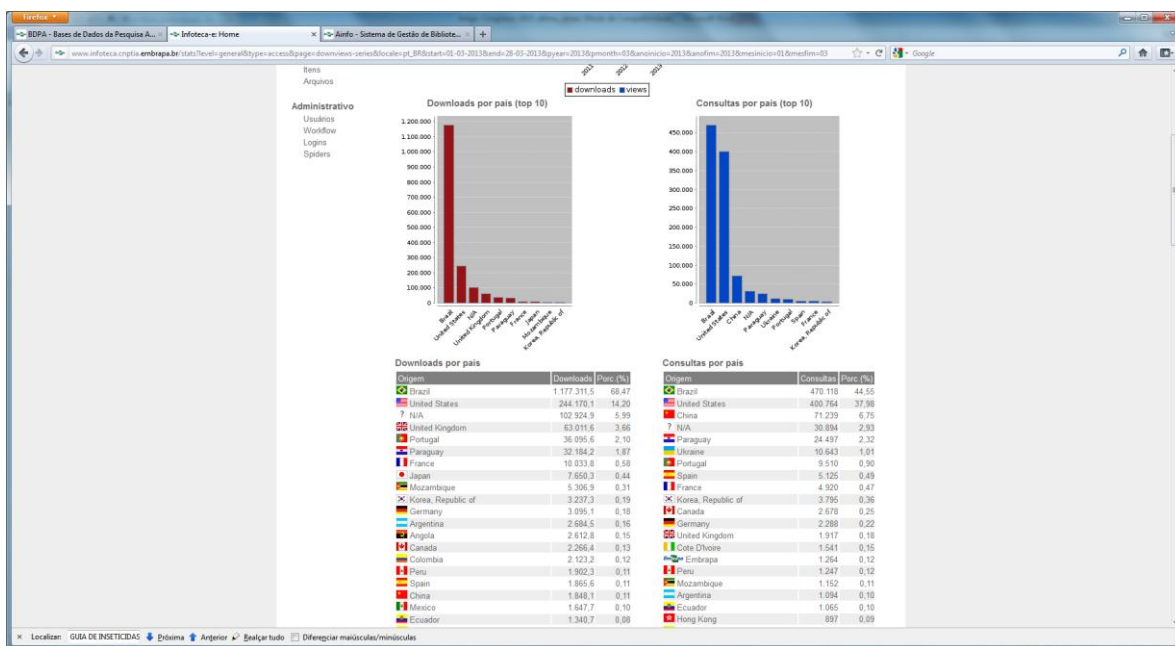


Fonte: <<http://www.infoteca.cnptia.embrapa.br/>>. Acesso em: 27 mar. 2013.

Os países que mais consultaram a Infoteca-e foram o Brasil com 44% e Estados Unidos com 38%.

Brasil e Estados Unidos também foram os países que mais fizeram downloads, sendo 68% no Brasil e 14% nos Estados Unidos (Figura 6).

Figura 6 – Consultas e Downloads na Infoteca-e por país



Fonte: <<http://www.infoteca.cnptia.embrapa.br/>>. Acesso em: 27 mar. 2013.

As estatísticas nos repositórios institucionais tem diversas aplicabilidades e auxiliam nos processos de avaliação dos repositórios. Segundo Leite (2009), também podem oferecer aos autores feedback do uso de suas publicações através das estatísticas de consultas e downloads e origem dos usuários que utilizam os repositórios.

3.4 Depósito

O processo de depósito dos documentos nos repositórios é feito pelos bibliotecários da Embrapa e não diretamente pelos pesquisadores. Os bibliotecários fazem a catalogação e indexação dos arquivos no sistema Ainfo e fazem o carregamento do arquivo digital para os repositórios.

A manutenção dos repositórios é realizada de forma sistemática e está integrada ao sistema Ainfo pelo qual é possível depositar os documentos digitais diretamente nos repositórios.

No momento da catalogação o bibliotecário faz o upload do arquivo no formato pdf e inclui na coleção digital a que este documento se enquadra nos repositórios, conforme sua natureza científica ou tecnológica. Este procedimento simplifica muito o depósito de documentos digitais e também potencializa a tarefa nas bibliotecas da Embrapa.

3.5 Critérios de Publicação

Quanto a disponibilização dos documentos nos repositórios, é importante ressaltar que toda a produção científica dos pesquisadores deve ser registrada no Ainfo mas nem todos os documentos podem ser disponibilizados na íntegra nos repositórios.

Os documentos editados pela Embrapa que sejam para publicação em acesso aberto são disponibilizados na íntegra nos repositórios. Como, por exemplo, os artigos técnico-científicos de autoria de pesquisadores da Embrapa publicados em periódicos que adotam uma licença que permita disponibilização do texto eletronicamente como SciELO ou outros.

Já os artigos publicados em periódicos que não são de acesso aberto ou em livros e capítulos de livros dos quais a Empresa não disponha de termo de cessão

de direitos patrimoniais assinado pelos autores, são depositados nos repositórios com restrição para disponibilização na íntegra.

Esta restrição na disponibilização dos documentos é realizada no início da catalogação da produção científica. Ao fazer o registro do documento no Ainfo e antes de fazer o upload do arquivo para os repositórios digitais, o bibliotecário verifica se o documento pode ou não ser disponibilizado e seleciona uma das opções de publicação: publicar totalmente ou publicar parcialmente.

Em “publicar totalmente” são considerados todos os trabalhos que estão publicados em periódicos ou publicações que adotam o modelo Acesso Aberto, ou a maioria daqueles que são editados pela própria Embrapa.

Em “publicar parcialmente” se enquadram todos os trabalhos que não são publicados em periódicos ou publicações de Acesso Aberto, ou não possuem os termos de cessão. Esses não são disponibilizados os textos na íntegra, no entanto são incluídos nos repositórios para fins de memória institucional e têm a visualização restrita ao operador do sistema Ainfo.

3.6 Preservação Digital

Segundo Arellano (2004), a aplicação de estratégias de preservação para documentos digitais é uma prioridade, pois sem elas não existiria nenhuma garantia de acesso, confiabilidade e integridade dos documentos a longo prazo, portanto o que caracteriza a preservação do documento digital é a capacidade de o documento ser utilizado posteriormente.

Uma das estratégias de preservação de documentos digitais é a adoção de formatos abertos na criação destes documentos, pois permite o estudo e a conversão para novos formatos no futuro viabilizando seu acesso mesmo se os equipamentos e os programas de informática em que foram criados estejam obsoletos.

É nesse sentido que a Embrapa adota alguns critérios para preservar os documentos digitais depositados nos repositórios tais como o armazenamento dos documentos digitais no formato PDF e o uso de link persistente para acesso aos objetos. Além destes critérios, outros fatores como a infraestrutura de armazenamento de dados adequada, e a realização de cópia de segurança interna e

em locais diferentes também contribuem para a preservação digital dos documentos nos repositórios da Embrapa.

4 CONSIDERAÇÕES FINAIS

Os repositórios institucionais Infoteca-e e Alice foram desenvolvidos com o objetivo de dar maior visibilidade à produção técnico-científica da Embrapa e contribuir para aumentar o impacto dos resultados das pesquisas por meio da maximização do acesso e do uso da informação produzida pela instituição, além de contribuir também para a gestão da informação técnico-científica da empresa. São notórias as vantagens da criação de repositórios pois além da organização e acesso aos trabalhos científicos e acadêmicos podem ser considerados ferramentas para o progresso da ciência.

As estatísticas de consultas e downloads comprovam o alcance dos repositórios e a quantidade efetiva de acesso em suas coleções.

No entanto, o desenvolvimento e a publicação de coleções digitais nos repositórios institucionais faz surgir a preocupação quanto a preservação destas coleções e, portanto, requer a necessidade de investimentos na preservação das coleções, para garantir que continuem acessíveis ao longo do tempo. Os critérios para a preservação dos documentos digitais adotados na atualidade deverão ser constantemente reavaliados e atualizados para garantir que no futuro as ferramentas tecnológicas estejam atualizadas e assegurem o acesso a estes documentos e assim permitam que os repositórios continuem contribuindo para o desenvolvimento científico do país.

5 REFERÊNCIAS

ABREU, J. de S. Repositório institucional ou rede social de aprendizagem? **DataGramZero - Revista de Informação**, v.13 n. 2, abr. 2012. Disponível em: <http://www.dgz.org.br/abr12/Art_06.htm>. Acesso em: 12 jul. 2012

ARELLANO, M. A. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004.

BRODY, T.; HARNAD, S. **The research impact cycle**. Disponível em: <<http://opcit.eprints.org/feb19oa/harnad-cycle.ppt>>. Acesso em: 16 jul. 2012.

EMBRAPA. Secretaria de Gestão e Estratégia. **V Plano Diretor da Embrapa: 2008-2011-2023**. Brasília, DF, 2008. 43 p.

FACHIN, G. R. B.; STUMM, J.; COMARELLA, R. L.; FIALHO, F. A. P.; SANTOS, Ni. Gestão do conhecimento e a visão cognitiva dos repositórios institucionais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 14, n. 2, p. 220-236, maio/ago. 2009.

HEERY, R.; ANDERSON, S. **Digital repositories review**. Disponível em: <http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf>. Acesso em: 10 jul. 2012.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 91-102, maio/ago. 2006.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: Ibict, 2009. Disponível em: <http://repositorio.bce.unb.br/bitstream/10482/4841/1/LEITE_ComoAmpliarGerenciar.pdf>. Acesso em: 12 jul. 2012.

LEITE, F. C. L.; COSTA, S. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 206-219, maio/ago. 2006.

LOPES, I. L. Novos paradigmas para avaliação da qualidade da informação em saúde recuperada na Web. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 81-90, jan./abr. 2004.

MORENO, F. P.; LEITE, F. C. L.; ARELLANO, M. A. M. Acesso livre a publicações e repositórios digitais em ciência da informação no Brasil. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 1, p. 82-94, jan./abr. 2006.

RIBEIRO, O. B.; VIDOTTI, S. A. B. G. Otimização do acesso à informação científica: discussão sobre a aplicação de elementos da arquitetura da informação em repositórios digitais. **Biblos**, Rio Grande, v. 23, n. 2, p. 105-116, 2009.

ROSA, F. G. M. G. **A disseminação da produção científica da Universidade Federal da Bahia através da implantação do seu repositório institucional** : uma política de acesso aberto. 2011. 242 f. Tese (Doutorado) - Universidade Federal da Bahia, Salvador, 2011.

SANTOS JÚNIOR, E. R. dos. **Repositórios institucionais de acesso livre no Brasil**: estudo delfos. 2010. 177 p. Dissertação (Mestrado em Ciência da

Informação) - Universidade de Brasília, Brasília, DF, 2010.

SOUZA, J. de; VIEIRA, A. F. G. A tecnologia open archives initiative, object reuse and exchange: histórico e perspectivas. **DataGramZero - Revista de Informação** - v. 13, n. 2, abr. 2012. Disponível em: < http://www.dgz.org.br/abr12/Art_04.htm>. Acesso em: 6 jul. 2012.

SAYÃO, L. F. Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos. **Ponto de Acesso**, Salvador, v. 4, n. 3, p. 68-94, dez. 2010.