

Acesso Aberto na Embrapa: breve histórico, avanços recentes e desafios

Open Access at Embrapa: brief history, recent advances and challenges

Acceso abierto en Embrapa: breve historia, avances recientes y desafíos

Marcos Cezar Visoli | marcos.visoli@embrapa.br

Embrapa Informática Agropecuária. Campinas, Brasil.

Patrícia Rocha Bello Bertin | patricia.bertin@embrapa.br

Embrapa. Secretaria de Gestão e Desenvolvimento Institucional. Brasília, DF, Brasil.

Alessandra Rodrigues Silva | alessandra.silva@embrapa.br

Embrapa Informação Tecnológica. Brasília, DF, Brasil.

Cláudia Regina De Laia Machado | claudia.delaia@embrapa.br

Embrapa Solos. Rio de Janeiro, Brasil.

Luís Eduardo Gonzales | luis.gonzales@embrapa.br

Embrapa Informática Agropecuária. Campinas, Brasil.

Isaque Vacari | isaque.vacari@embrapa.br

Embrapa Informática Agropecuária. Campinas, Brasil.

Resumo

O trabalho objetiva relatar as principais iniciativas desenvolvidas na Empresa Brasileira de Pesquisa Agropecuária (Embrapa) sobre o acesso aberto à informação científica e tecnológica, bem como iniciativas de organização e disponibilização relacionadas aos dados de pesquisa. Para tanto, é realizado um histórico, a partir do ano de 2008, época em que a Empresa por meio de projeto interno consolidou várias ações no âmbito da gestão da informação, com destaque para a publicação de seus repositórios de acesso aberto (Alice e Infoteca-e). Em seguida, são descritas iniciativas mais recentes, como panorama sobre o acesso aberto elaborado no âmbito de projeto institucional estratégico sobre governança de dados e da informação, iniciado em 2015 e, finalizado em 2017, que apresentou uma série de diretrizes a serem observadas sobre o tema pelos gestores e pela comunidade da Embrapa, com o propósito de alinhá-la às tendências internacionais.

Palavras-chave: Acesso Aberto; Dados Abertos; Plano de Dados Abertos; Agricultura.

Abstract

The paper aims to report on the main initiatives developed by the Brazilian Agricultural Research Corporation (Embrapa) on open access to scientific and technological information, as well as initiating the organization and availability of research data. For this purpose, a history is made, starting in 2008, when the Company, through an internal project, consolidated several actions in the scope of information management, especially the publication of its open access repositories (Alice and Infoteca-e). Next, more recent initiatives are described, such as the Open Access scenario developed within the framework of a strategic institutional project on data, information and knowledge governance, initiated in 2015 and finalized in 2017, which presented a series of guidelines to be observed on the theme by the managers and the community of Embrapa with the purpose of aligning it with the international trends.

Keywords: Open Access; Open Data; Open Data Plan; Agriculture.

Resumen

El trabajo tiene como objetivo relatar las principales iniciativas desarrolladas por la Empresa Brasileña de Investigación Agropecuaria (Embrapa) sobre el acceso abierto a la información científica y tecnológica, así como iniciativas de organización y apertura de datos de investigación. El trabajo presenta un breve relato, a partir del año 2008, época en que la Empresa, a través de un prouyecto interno, consolidó varias acciones en el ámbito de la gestión de la información, con destaque para la publicación de sus repositorios de acceso abierto (Alice e Infoteca-e). A continuación se describen iniciativas más recientes, como panorama sobre el acceso abierto elaborado en el marco del proyecto institucional estratégico sobre gobernanza de datos, información y conocimiento, iniciado en 2015 y, finalizado en 2017, que presentó una serie de directrices a ser observadas sobre el tema por los gestores y por la comunidad de Embrapa con el propósito de alinearla con las tendencias internacionales.

Palabras clave: Acceso Abierto; Datos Abiertos; Plan de Datos Abiertos; Agricultura.

Introdução

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), é uma empresa pública de inovação, focada na geração de conhecimentos e tecnologias. Criada no ano de 1973, a Embrapa se destaca pelos conteúdos científicos e tecnológicos que produz no âmbito da informação agrícola.

Ao longo dos anos, esses conteúdos têm sido disponibilizados à sociedade por meio de sistemas de recuperação da informação diversificados, mediados pelo trabalho de organização e tratamento da informação realizado por profissionais de informação e de tecnologia da informação, entre outros. Sob essa perspectiva, esse trabalho objetiva relatar as principais iniciativas desenvolvidas na Embrapa sobre o acesso aberto à informação científica e tecnológica, bem como iniciativas de organização e disponibilização relacionadas aos dados de pesquisa.

Para tanto, o texto está organizado em 4 seções, incluída esta Introdução. Inicialmente é realizado um histórico, a partir do ano de 2008, época em que a Empresa por meio de projeto interno consolidou várias ações no âmbito da gestão da informação, com destaque para a publicação de seus repositórios de acesso aberto (Alice e Infoteca-e).

Em seguida, são descritas iniciativas mais recentes, como panorama sobre o acesso aberto elaborado no âmbito de projeto institucional estratégico sobre governança de dados e da informação, iniciado em 2015 e, finalizado em 2017, que apresentou uma série de diretrizes a serem observadas sobre o tema pelos gestores e pela comunidade da Embrapa com o propósito de alinhá-la às tendências internacionais. Por fim, apresentam-se algumas considerações voltadas para a reflexão do tema.

Breve histórico

As discussões institucionais sobre o acesso aberto na Embrapa tomaram fôlego entre os anos de 2008 e 2011, por meio da execução de projeto de pesquisa intitulado “Acesso Aberto na Embrapa: maximizando o impacto da pesquisa, a visibilidade e a gestão da informação científica”. Composto por equipe multidisciplinar e relacionada a diversas instâncias da Empresa, o projeto objetivou propor e implementar modelo metodológico para a gestão da informação técnico-científica da Embrapa, fundamentado nos pressupostos do Acesso Aberto e, tendo como insumo as atividades de pesquisa e desenvolvimento e a ampla disseminação da informação produzidaⁱ.

Entre os principais resultados do projeto citado estão a implantação dos dois repositórios institucionais de acesso aberto: o Acesso Livre à Informação Científica da Embrapaⁱ (Alice) e o serviço de Informação Tecnológica em Agricultura (Infoteca-e)^{ii 2,3}, que passaram a oferecer, a partir de interfaces amigáveis, o acesso à produção científica e tecnológica da Empresa. O povoamento inicial destes repositórios se deu por meio da migração dos registros existentes em sistema de gestão da documentação bibliográfica institucional, denominado Ainfo, alimentado desde a década de 1990, por bibliotecários das mais de 40 unidades da Embrapa.

Nesse sentido, o projeto foi um marco na instituição, uma vez que os repositórios dele decorrentes foram construídos no *software* DSpace e sob a arquitetura do protocolo OAI-PMH, o que possibilitou interoperabilidade com metabuscadores internacionais. Ademais, o projeto serviu de instância para o fomento de discussões sobre outros aspectos necessários para ampliação da disseminação de informações e publicações para a comunidade científica e a sociedade em geral. São exemplos: as discussões sobre a implementação de licenças públicas de uso nas publicações da Embrapa e a necessidade de uma política de acesso aberto corporativa.

Avanços recentes e desafios

Em continuidade a essas ações e, com propósitos estratégicos, durante os anos de 2015 e 2017, a Diretoria da Empresa apoiou e acompanhou o projeto “Governança de Dados e da Informação para o Conhecimento na Embrapa: Desenvolvimento de Modelo e Plano de Implantação (GovIE)”, desenvolvido por pesquisadores e especialistas com atuação relacionada à gestão da informação. O foco do projeto consistia em promover uma melhor articulação, integração e harmonização das ações relacionadas à gestão de dados e da informação na Empresa, a partir de uma perspectiva sistêmica.

Por meio de levantamentos, mapeamentos e diagnósticos exaustivos, os membros do projeto perceberam questões salienteⁱⁱⁱs caracterizadas como críticas e prioritárias – levando em conta o ambiente organizacional e o universo informacional da Empresa como um todo –, em torno das quais ações estruturantes foram delineadas e propostas, como parte de um plano estratégico para o aprimoramento da gestão de dados e da informação na organização.

Cada uma das questões assim caracterizadas foi estudada detalhadamente por especialistas de suas respectivas áreas, resultando na formulação de um modelo de governança de dados e da informação para a Embrapa, que, em síntese, consistiu na primeira fase do projeto. Uma das questões identificadas foi o

i Disponível em <http://www.embrapa.br/alice>.

ii Disponível em <http://www.embrapa.br/infoteca>.

iii Terminologias, vocabulários controlados e semântica; gestão da informação no portal e intranet corporativos; gestão, monitoramento e avaliação da produção científica; acesso aberto ao conhecimento científico; gestão de informações na escala de projetos; gestão de informações sobre profissionais da Embrapa; gestão da informação científica e tecnológica; gestão da informação arquivística; gestão de dados de pesquisa; e, gestão da informação estratégica.

‘Acesso Aberto ao Conhecimento Científico’. Posteriormente, como segunda fase, foi elaborado um plano com metas e ações objetivando operacionalizar o modelo de governança proposto.

As metas e ações propostas variam de perspectiva, algumas em nível estratégico, outras tático e, ainda, há aquelas de nível operacional. A seguir, são relacionadas as principais recomendações constantes no plano de aprimoramento do ‘Acesso Aberto ao Conhecimento Científico’ na Embrapa:

- **Inserir a perspectiva do ‘acesso aberto’ nos documentos estratégicos da Embrapa:** a temática do acesso aberto deve ser incluída nos documentos orientadores da Empresa, nas proposições de editais internos de projetos, nas orientações das estruturas de programação da pesquisa, nos documentos e nas ações relacionadas a negócios e à transferência de tecnologia.

- **Revisar a política editorial da Empresa,** com a inclusão do acesso aberto em seu contexto.

- **Revisar o modelo de avaliação de desempenho e recompensa vigente,** de modo a valorizar a publicação em fontes de informação de acesso aberto.

- **Adotar e operacionalizar o uso de licenças públicas** (a exemplo da Creative Commons), em consonância com as recomendações da Infraestrutura Nacional de Dados Abertos.

- **Definir e implantar workflow** administrativo e de publicação de documentos nos repositórios institucionais.

- **Instituir processo** que permita a elaboração de uma publicação em acesso aberto desde a sua concepção.

- **Definir estruturas organizacionais** competentes e manter um grupo permanente de acesso aberto ao conhecimento.

- **Desenvolver e implementar** política organizacional de apoio ao acesso aberto.

- **Aderir a iniciativas nacionais e internacionais** de acesso aberto, como Godan (Global Open Data for Agriculture and Nutrition) e RDA (Research Data Alliance), entre outras.

- **Implantar estratégia de comunicação interna e externa** e oferecer oportunidades de capacitação para disseminar e internalizar a cultura de acesso aberto.

- **Elaborar Plano de Dados Abertos da Embrapa** e estabelecer processo corporativo para a publicação de dados abertos, em atendimento à nova legislação no tema (Decreto 8.777, de 11 de maio de 2016).

- **Construir infraestrutura de dados abertos da Embrapa,** alinhada ao Portal Brasileiro de Dados Abertos^{iv}.

- **Disseminar boas práticas para a abertura de dados,** tais como o uso de formatos abertos, padrões abertos e uso de licenças abertas.

- **Evoluir tecnologicamente os repositórios Alice e Infoteca-e,** incluindo a revisão das interfaces, estatísticas, ferramenta de busca, assim como integração com repositórios de outras instituições.

- **Integrar e interoperar** os repositórios com sistemas de informação corporativos relevantes.

- **Revisar o conteúdo dos repositórios institucionais,** a fim de verificar não-conformidades.

- **Institucionalizar função de curadoria dos dados** disponibilizados nos repositórios.

- **Avaliar a pertinência de implantação do auto-arquivamento** da produção científica nos repositórios institucionais.

- **Estabelecer equipe mínima,** papéis e responsabilidades sobre os repositórios institucionais.

- **Instituir recursos de interatividade,** como comentários por parte de usuários, com serviços que gerem valor agregado para os autores e usuários cadastrados.

Os resultados do projeto GovIE, parcialmente relatados aqui, estão sendo analisados detalhadamente e um grupo foi constituído para discutir e priorizar as alterações necessárias nas várias estruturas

iv Disponível em: <<http://dados.gov.br>>.

organizacionais, documentos orientadores, processos e sistemas. Os desafios ainda são enormes, mas algumas ações que envolvem a publicação de dados abertos estão em andamento. A Embrapa está construindo uma Infraestrutura de dados espaciais, e desta infraestrutura um conjunto de dados estará automaticamente publicado na Infraestrutura Nacional de Dados Espaciais (INDE)^v.

Por outro lado, algumas das ações propostas já estão em desenvolvimento. Um novo projeto estratégico patrocinado pela Diretoria da Empresa contempla a elaboração do Plano de Dados Abertos da Embrapa, assim como a construção de uma plataforma piloto baseada no *software* aberto Comprehensive Knowledge Archive Network (CKAN)^{vi} para disponibilizar dados abertos, incluindo dados de pesquisa e a integração da plataforma com o Portal de Dados abertos do Governo Federal.

Os repositórios institucionais de acesso aberto (Alice e Infoteca-e) estão em processo de atualização para versões mais recentes do *software* DSpace^{vii}. Com isto, novas interfaces estarão disponíveis para os usuários. As interfaces estão sendo customizadas para refletir a identidade da Empresa, com a preocupação das customizações serem mínimas para facilitar a implementação de novas atualizações da ferramenta.

O módulo de estatísticas, baseado na contribuição dada pela Universidade do Minho, por meio do Repositorium^{viii}, já foi atualizado e, os números das estatísticas ajustados. Devido ao processo de melhoria da qualidade dos dados registrados pela equipe do Sistema Embrapa de Bibliotecas, os metadados foram aprimorados, palavras-chaves, resumo e títulos ajustados, além da inclusão de publicações mais antigas que foram digitalizadas. Contudo, a qualidade desses elementos requer monitoramento constante e minucioso, inclusive otimizado por funcionalidades vinculadas às tecnologias de informação.

Em relação às licenças de uso, os repositórios receberão a informação de que seus conteúdos estarão disponíveis sob a licença Creative Commons, atribuição não comercial, sem derivações, versão 4.0. Concomitantemente está sendo discutida a inclusão das licenças nos próprios documentos publicados, na revisão do processo de publicação técnico-científico da Empresa.

As Figuras 1 e 2, a seguir, apresentam as interfaces em desenvolvimento dos repositórios Alice e Infoteca-e:



Figura 1. Repositório Alice

v Disponível em: <<http://www.inde.gov.br>>.

vi Disponível em: <<https://ckan.org/>>.

vii Disponível em: <www.dspace.org>.

viii Repositório de acesso aberto da Universidade do Minho, disponível em: <<https://repositorium.sdum.uminho.pt/>>.



Figura 2. Repositório Infoteca-e

Considerações

Como demonstrado, ao longo dos anos as ações de acesso aberto têm se fortalecido na Embrapa. Se no lançamento da primeira versão dos repositórios de acesso aberto à informação científica e tecnológica – Alice e Infoteca-e, em 2011, o horizonte dos dados abertos ainda se mostrava distante, no momento, constitui-se necessidade premente, conforme a própria legislação federal orienta.

Nesse sentido, acredita-se que as ações em andamento, bem como aquelas identificadas por meio do panorama descrito tendem a contribuir para a consolidação de uma cultura orientada ao acesso aberto na Embrapa, o que requer a integração e o diálogo das diversas esferas institucionais.

Referências

1. Bertin PRB, Vacari I, Simão V P M, Visoli M C, Leite, F C L. An open access approach to scientific information management at the Brazilian Agricultural Research Corporation. *Scholarly and Research Communication*. 2010;1(1):1-12.
2. Vacari I, Visoli M C, Leite F C L, Pontes S D C L D, Okawachi M F, Simao V P M, Gonzales L E, Praxedes M G G. *Software* livre para implementação de repositórios digitais e provedores de serviços: experiência da Embrapa Informática Agropecuária. In: 39. Jornadas de *Software* Libre; 2010, Buenos Aires, Argentina. Buenos Aires: Sociedad Argentina de Informática, 2010. p. 2345-2360. JAIIO - JSL 2010.
3. Arruda R G, Castro R. L. Acceso abierto a la información científica y técnica en Empresa Brasileña de Investigación Agropecuária (Embrapa): relato de experiencia. In: 13. Congreso Internacional de Información; 2014, La Habana, Cuba. La Habana: Instituto de Información Científica y Tecnológica, 2014.