

# Preservação digital, gestão de dados de pesquisa e biodiversidade

Miguel Ángel Márdero Allerano

Instituto Brasileiro de Informação em Ciência e Tecnologia

[miguel@ibict.br](mailto:miguel@ibict.br)

Maria de Fátima Duarte Tavares

Instituto Brasileiro de Informação em Ciência e Tecnologia

[fatimatavares@ibict.br](mailto:fatimatavares@ibict.br)

## Resumo

Os repositórios institucionais em estágio de produção enfrentam o desafio de abrigar, preservar e dar acesso a conjuntos de dados científicos. Os repositórios de dados de pesquisa estão sendo implementados no mundo como uma nova opção para o acesso livre. Os benefícios ainda estão em experimentação e os bibliotecários, pesquisadores, instituições de financiamento e de ensino começam a conhecer seus benefícios. Entre esses benefícios está a visibilidade dos projetos e o fato de que os trabalhos de pesquisa adquiram uma nova possibilidade de gestão se articulados a sistemas de informação. Também, a transparência nos investimentos nacionais para a Ciência permite novas formas de governabilidade, relacionada com o registro de processos de gestão e disseminação. Espera-se que os indicadores resultantes dessas atividades permitam uma gestão mais equânime, atendendo às dimensões do país, mas também às condições estratégicas em que a pesquisa torna-se fundamental à preservação de seus recursos naturais e à elaboração de políticas públicas.

**Palavras-chave:** dados de pesquisa; preservação digital; políticas institucionais;

## Digital preservation, research data Management and biodiversity

### Abstract

Institutional repositories at the production stage face the challenge of sheltering, preserving, and giving access to scientific data sets. Research data repositories are being implemented worldwide as a new option for free access. The benefits are still in the pipeline, and librarians, researchers, funding and teaching institutions are beginning to see its benefits. Among these benefits is the visibility of the projects and the fact that the research works acquire a new possibility of management if articulated to information systems. Also, transparency in national investments for science allows new forms of governance, related to the registration of management and dissemination

processes. It is hoped that the indicators resulting from these activities will allow a more equitable management, taking into account the country's dimensions, but also the strategic conditions in which research becomes fundamental to the preservation of its natural resources and to the elaboration of public policies.

**Key-words:** search data; digital preservation; institutional policies

## Introdução

Os repositórios institucionais em estágio de produção enfrentam o desafio de abrigar, preservar e dar acesso a conjuntos de dados científicos. As dimensões, institucional e disciplinar, precisam ser conjugadas para a utilização dos repositórios como um componente fundamental na infraestrutura global de curadoria de dados científicos (RODRIGUES; SARAIVA, 2010). Os repositórios de dados de pesquisa estão sendo implementados no mundo como uma nova opção para o acesso livre. Os benefícios ainda estão em experimentação e os bibliotecários, pesquisadores, instituições de financiamento e de ensino começam a conhecer seus benefícios. Entre esses benefícios está a visibilidade dos projetos e o fato de que os trabalhos de pesquisa adquiram uma nova possibilidade de gestão se articulados a sistemas de informação. Também, a transparência nos investimentos nacionais para a Ciência permite novas formas de governabilidade, relacionada com o registro de processos de gestão e disseminação. Espera-se que os indicadores resultantes dessas atividades permitam uma gestão mais equânime atendendo às dimensões do país, mas também às condições estratégicas em que a pesquisa torna-se fundamental à preservação de seus recursos naturais e à elaboração de políticas públicas.

## A preservação digital na política de gestão de dados de pesquisa

A preservação digital é um processo que deve ser atualizado com o objetivo de estar sempre adaptado às mudanças do contexto, integrado às outras políticas institucionais que envolvem seu financiamento, recursos humanos e planejamento. As ações de preservação em longo prazo, dentro da política de gestão de dados, têm o propósito de garantir a permanência e acesso aos dados, como ações que, implantadas no curto prazo, antecipam, preveem, retardam ou detêm o deterioro dos dados digitais.

As políticas de gestão de dados começam a ser discutidas no âmbito dos repositórios em função da demanda de espaços para o registro dos dados científicos dos pesquisadores das instituições. A aprovação interna de uma política de gestão de dados garante o povoamento dos repositórios e o reconhecimento da sua importância dentro de um novo modelo de gestão a informação científica institucional.

No cenário informacional recente, observa-se uma visão que caracteriza os dados de pesquisas não mais somente como alicerces de pesquisas realizadas, mas como subsídios essenciais para outros pesquisadores, instituições acadêmicas e as agências de fomento, entre outros. Esses dados eram armazenados de forma inapropriada ou até descartados sem nenhuma gestão no processo de armazenamento. Na maioria das vezes ficavam na posse dos próprios geradores da informação e se perdiam no tempo e no espaço.

A ausência de um arcabouço (*framework*) para organizar, preservar e disponibilizar dados científicos por longos períodos pode resultar em uma perda ou descarte de valiosos conjuntos de dados de pesquisas no país (WITT, 2008). Na Europa as propostas de metodologias para a proteção dos dados iniciaram-se nas bibliotecas universitárias. O

objetivo delas têm sido criar um marco que permitisse a aplicação de políticas de gestão e de preservação de dados. Como mencionaram Martínez-Uribe e Macdonalds (2008) a custódia dos dados de pesquisa produzidos pelas instituições é de responsabilidade das bibliotecas. Ainda que as competências necessárias para essa função não estejam inseridas nas organizações.

O papel dos bibliotecários e arquivistas na gestão desses repositórios está começando a ser explorado (AKMON; ZIMMERMAN; DANIELS, HEDSTROM, 2011). Particularmente, ante o perigo de que os pesquisadores apliquem técnicas de gestão de informação inapropriadamente no arquivamento dos seus dados de pesquisa (JOINT, 2007).

Uma das primeiras propostas foi a das universidades inglesas que criaram a iniciativa *Data audit Framework* (JONES, 2008), nela o centro não eram os repositórios, mas sim a identificação previa dos locais e das formas de armazenamento usados pelos pesquisadores. Outras metodologias de auditoria apontaram também para essa ausência de uma política que subsidiasse os pesquisadores na preservação e arquivamento de seus dados, identificando problemas e riscos de perda e mau uso das mídias.

Por outro lado, o desenvolvimento dos repositórios de dados apresenta dificuldades técnicas e de gestão que devem ser abordados pelas equipes responsáveis, desde o momento da formatação do projeto do fluxo de vida dos dados de pesquisa (STEINHART; DIETRICH; GREEN, 2009).

Para uma realidade como a do Brasil, as iniciativas institucionais devem definir sua política de gestão como requisito para implementação dos seus repositórios de dados. Atualmente, dos 99 repositórios brasileiros registrados no *OpenDoar* apenas quatro oferecem o depósito e acesso a datasets, e nenhum deles menciona ter publicado uma política de preservação digital.

Como os projetos de pesquisa regularmente possuem um período de vida curto, uma tendência de transferência de responsabilidade está sendo observada. A curadoria de dados de longo prazo desloca-se das mãos do pesquisador para os repositórios de dados confiáveis, os quais podem ser auditados e certificados segundo uma lista de critérios. Os sistemas de auditoria e certificação como o Data Seal of Approval (DSA) contem esse catálogo de indicadores a serem cumpridos para garantir a confiabilidade dos repositórios, bibliotecas e arquivos de dados digitais (KLUMP, 2007).

Algumas propriedades arquivísticas dos dados devem ser preservadas, a autenticidade, proveniência e integridade devem ser mantidas para que futuros pesquisadores possam fundamentar seus resultados na confiança de que eles estejam armazenados em locais estáveis. Os dados precisam ser constantemente monitorados e suas propriedades verificadas, esse papel é atribuído ao curador ou especialista da área,

A responsabilidade pelo ciclo de vida dos dados de pesquisa, em particular a garantia e fornecimento de dados de pesquisa para arquivamento de longo prazo, é principalmente da responsabilidade do gestor do projeto. Assim, é inerente a cada projeto de pesquisa a proposição de um plano de gerenciamento de dados, que abordará explicitamente como a precisão, integridade, autenticidade, confidencialidade, publicação e acesso aberto de dados serão tratados.

O processo de preservação de dados está presente em todas as fases do ciclo de vida. Na etapa anterior ao depósito, a responsabilidade está nas mãos do criador do conteúdo, determinando quais dados devem ser arquivados, suas qualidades e sob qual licença seu reuso estará garantido. Já na etapa de depósito, a responsabilidade passa para o gestor do repositório, que facilita o arquivamento e sua conversão caso isso seja necessário, sem afetar a qualidade dos dados. Ele é também responsável pelas garantias de segurança e requisitos para seu reuso. Por último, a integração com recursos de informação externos, especializados nesse reuso, é de responsabilidade do provedor de conteúdo da instituição.

Na administração dos dados científicos, a atenção à preservação digital acrescenta valor ao material depositado pela descrição do contexto de produção e na vinculação da sua origem, facilitando sua reutilização, integração e promoção (RUSBRIDGE et al., 2005). A política de gestão de dados de pesquisa institucional deverá estar em conformidade com uma política de informação, além de conter indicações sobre a gestão das garantias de acesso e para a salvaguarda dos arquivos e da sua integridade no longo prazo. Portanto, a política de gestão de dados científicos deverá conter especificações de uma política de preservação dos dados de pesquisa.

A falta de políticas de preservação na maioria dos projetos de repositórios digitais sugere a carência de conhecimentos técnicos sobre a importância das estratégias de preservação de materiais digitais existentes. A literatura da área mostra que a maioria dos projetos está nos seus primeiros estágios e apenas metade deles relata seguir alguma forma de modelo de negócios e processos de auditoria planejados.

Análise do registro atual desse tipo de políticas no *OpenDoar*, evidencia o geralmente baixo interesse pelo tema da preservação digital por parte dos gestores dos repositórios institucionais. Uma explicação possível é que o considerável volume de tempo necessário para seu funcionamento faça o problema difícil de ser atendido (RYDÉN, 2018).

A implementação de políticas institucionais para a preservação dos dados de pesquisa é considerada uma das boas práticas de gestão desses materiais. As especificações técnicas da gestão dos tipos de dados, das cópias e seu acesso em longo prazo são questões a serem exemplificadas em manuais, guias e tutoriais, que podem ser incluídos em programas de letramento informacional dirigidos a estudantes e pesquisadores.

Padrões, normas e práticas de preservação são adotados, quando existe consciência e interesse por parte dos responsáveis pelos acervos em desenvolver recursos que sejam permanentemente acessados. Em nível nacional, a definição das políticas, obrigações e metodologias mais apropriadas para a preservação dos documentos eletrônicos deve levar em consideração a implementação de pacotes de software livres a fim de verificar se atendem às necessidades das instituições de ensino e pesquisa e se estão em concordância com os já testados padrões internacionais, que promovem o arquivamento digital da produção científica de longo prazo.

A fase do planejamento é um dos processos mais importante do ciclo de vida dos dados de pesquisas, pois nessa fase é formalmente realizada a descrição dos dados que serão coletados, a forma que será utilizada, as características intrínsecas dos dados, seus limites e os cuidados que eles demandam.

Tanto os pesquisadores como os responsáveis pelos repositórios de dados são os agentes determinantes acerca da preservação dos materiais por longos períodos, quando o acesso ilimitado e irrestrito é uma condição necessária. Essa distinção do acesso para usuários a curto prazo e a manutenção do acesso a longo prazo precisa ser explicitada no interesse da salvaguarda da informação, além de definida na política de gestão de dados da instituição. A manutenção correta e independente de dados de pesquisa requer mecanismos para verificar a integridade da informação depositada nos repositórios e a garantia de que os usuários possuem as permissões apropriadas para acessar os materiais.

Os dados de pesquisa requerem um esforço maior na sua descrição do que o auto arquivamento de publicações (BORGAN, 2007). A falta de metodologias adequadas para garantir a integridade dos dados chega a ser um impedimento no uso dos repositórios de dados. Os indicadores nos quais os pesquisadores podem se basear para fazer o reuso e o acesso de dados de pesquisa são aqueles mencionados nas políticas de gestão de dados, os quais direcionam para a reputação da instituição e do pesquisador depositante, à qualidade do trabalho publicado que menciona os dados, descrições, e especificações reunidas na documentação gerada na pesquisa. Segundo Karasti, Baker e Halhola (2006), os pesquisadores precisam ter opções e incentivos para produzir metadados e documentação nos seus planos de gestão de dados.

Políticas de preservação de dados como as do DataShare da Universidade de Edimburgo (<http://datashare.is.ed.ac.uk/>) e a do Dataverse de Universidade de Harvard (<http://best-practices.dataverse.org/harvard-policies/harvard-preservation>) assinalam a necessidade de que elas estejam integradas à política de gestão de dados dos repositórios.

Essas políticas de preservação digital de dados mencionam um conjunto de parâmetros sobre as funções de preservação de documentos digitais dos repositórios, quanto à garantia de legibilidade e acessibilidade:

- a) continuidade do acesso em relação ao uso corrente e posteridade;
- b) migração de itens para novos formatos;
- c) emulações de software para acessar novos formatos;
- d) atualização de novas versões;
- e) impossibilidade ou não de garantir a legibilidade de alguns formatos de arquivo incomuns.

Tratam da preservação dos bits nos aspectos relacionados com:

- a) periodicidade de Backup em servidores;
- b) verificação dos riscos para os materiais;
- c) realização de checksums periódicos para identificar perdas.

Mencionam a retenção dos dados no que se refere a:

- a) retenção de URL;
- b) retenção de bitstream (fluxo de bits);
- c) verificação de integridade ou fixity check (verificação de fixidez);
- d) adoção de ferramentas checksums (soma de controle), message digests (sumários de mensagem) e digital signatures;

- e) fluxo do bit original (se retido ou não);
- f) existência de parceiros para trabalhar com metadados de preservação.

E ressaltam a remoção de itens nas seguintes circunstâncias:

- a) violação de direitos autorais ou plágio comprovada;
- b) requisitos legais e violações comprovadas; segurança nacional; pesquisa falsificada;
- c) identificadores/URLS dos itens removidos, serão mantidos indefinidamente.

Em resumo, a colaboração de especialistas e a inclusão da gestão de dados de pesquisa de forma institucionalizada em todos os serviços dos repositórios devem criar recursos internos para que os pesquisadores usem os padrões e a documentação necessária que leve a uma preservação segura dos dados de pesquisa. Para a promoção dos princípios que devem ser de conhecimento dos pesquisadores na temática do arquivamento de dados tem-se o documento Reference Model for a Open Archival Information System (OAIS). Mesmo que esse documento apresente um consenso sobre a prática originada na comunidade de Ciências Espaciais para tratamento e arquivamento de dados, as orientações também têm sido amplamente adotadas nas ciências sociais como diretrizes.

Enfrentar os desafios inerentes à pesquisa envolve a boa gestão de dados de pesquisa; o planejamento, o cuidado na documentação e na preservação, facilitam sua produtividade e transparência, propiciando uma maior colaboração entre pesquisadores e um retorno maior do investimento das instituições de fomento.

### **Políticas de reuso de dados de pesquisa e a biodiversidade**

Um exemplo em processo de implantação no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que usa o sistema Dataverse como instrumento de institucionalização da gestão de dados de pesquisa, tem como pressuposto a reutilização de dados dos serviços nacionais dedicados à biodiversidade, mas também incorpora outros conjuntos documentais de arquivos locais e da Biblioteca Nacional, tendo em vista dar suporte à análise de relações entre práticas e saberes nem sempre institucionalizados. Outras interpretações dessa documentação histórica serão possíveis, como novas práticas de reuso, e seus fundamentos científicos não necessariamente serão os mesmos.

A memória científica digital no Brasil, estruturada em coleções de instituições públicas que têm sido objeto das principais políticas de informatização e de acesso facilitado a dados de pesquisa, está relativamente circunscrita a áreas de conhecimento em que prevalece a temática da biodiversidade.

O Programa Re flora, coordenado pelo Instituto de Pesquisa do Jardim Botânico do Rio de Janeiro, com apoio do Sistema de Informação da Biodiversidade do Brasil – SiBBR, do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), com dados de 400 coleções e 10.429.331 registros, gera a disponibilização de dados de coleções científicas brasileiras e a atualização sistemática da Lista de Espécies da Flora do Brasil, no programa Flora do Brasil 2020, que visa atender à meta brasileira da Convenção da Diversidade Biológica

– CDB, com 46.523 espécies já descritas (SiBBR, 2018; Flora do Brasil 2020, 2018). Em que pese a consolidação histórica da atividade botânica sobre outras no Brasil, note-se que esses serviços possibilitam o reuso de dados por outras áreas de conhecimento que extrapolam o domínio e as finalidades originais dos registros.

Os registros de observação e documentos de trabalhos de campo das ciências biológicas podem integrar outras análises, associados a outros materiais documentais, objetos digitais na origem ou não. No quadro recente da história ambiental do Brasil, a supressão acelerada da flora e fauna do bioma Cerrado constitui uma condicionante espacial criada e gestada no nosso tempo, e, como tal, configura processos que levam a outros métodos de investigação, em que a relação sociedade e natureza também deverá estar em foco.

O caso aplicado ao serviço do Dataverse do IBICT atenta para a pesquisa histórica do processo de ocupação territorial após a tomada de decisão de implantação da nova capital do Brasil, Brasília, em 1955. Porém, o recorte dominante de análise se dá sobre a identificação das atividades de expedições de reconhecimento botânico no DF e entorno, seus principais agentes de investigação, círculos de interlocução acadêmica e instituições de suporte. A pesquisa visa a traçar também a contribuição desses pesquisadores para o entendimento contemporâneo dos impactos gerados à paisagem multifacetada do Cerrado, tanto com a urbanização brasiliense, quanto com os efeitos de longo prazo derivados do incremento agrícola sobre a região Centro-Oeste, dirigido pelo Estado nacional. O sentido de urgência que se apreende dos registros de campo e de sua identificação botânica permite levantar hipóteses sobre a circunscrição ou não das atividades desses pesquisadores à visão do projeto de modernização desenvolvimentista do período de Juscelino Kubitschek e seus desdobramentos subsequentes nas diretrizes de ênfase tecnológica e econômica (BERTRAN, 1988).

Nos termos deste trabalho, interessa observar que essa reconstituição de relações entre saberes e práticas científicas só é possível de ser realizada pela disponibilidade informatizada das coleções botânicas brasileiras, pertencentes a instituições nacionais, em geral universitárias, e dos demais registros de flora brasileira integrados a outros acervos internacionais. Assim, a investigação explora conjuntos de dados recortados de bases múltiplas, consolidadas em serviços integradores, como o do Instituto Nacional de Ciência e Tecnologia (INCT) – Flora e Fungos, com a ferramenta specieslink, e os inter-relaciona a outras fontes documentais.

A relação entre os registros assim disponibilizados e a memória científica associada à ocupação territorial do DF pôde ser processada com ajuda de um Sistema de Informação Geográfica, mas é a aplicação de procedimentos de identificação e de documentação de materiais coletados e depositados em herbários, segundo garantidos padrões de descrição, que torna esse processamento viável.

Note-se que o conjunto de dados recortados das bases acima referidas, em procedimentos de reuso que obedecem aos critérios desse projeto de investigação, somados aos demais itens documentais relacionados a outras fontes, constituem os objetos digitais que serão abrigados no Dataverse, tendo em conta as especificidades e diferenças de acesso prescritas pelos detentores dos acervos originários.

Este estudo de caso demonstra a relevância das políticas de estruturação e disponibilização em acesso aberto de dados de pesquisa, que poderiam ser ampliadas para outras áreas de conhecimento, tendo em vista o reuso potencial de um contingente incomensurável de dados sobre o Cerrado e sobre sua ocupação, para ficar nos limites de nossa temática.

A conjuntura atual de ocupação do Cerrado, bioma que ocupa mais de 23 % do território nacional e que domina o Centro-Oeste do país, amplia a relevância do Programa Re flora para a memória científica do país, considerando-se a repatriação digitalizada de conjunto significativo de coleções de viajantes naturalistas do século XIX e outros pesquisadores do século XX, que percorreram a região central, cujos conteúdos estão acessíveis no SiBBR e no INCT – Flora e Fungos.

Para circunscrever o campo de produção de conhecimento acumulado associado ao Cerrado, foram utilizados os conjuntos de registros do SiBBR, e também do INCT– Flora e Fungos, vinculado à Universidade Federal de Pernambuco, em associação à produção acadêmica disponível em repositórios institucionais.

Na temática centrada no Cerrado, prevalece, na atualidade, o papel da Universidade de Brasília como principal agente, conforme dados quantitativos de teses e dissertações da Base Digital de Teses e Dissertações – BDTD/IBICT (BDTD, 2018). Essa prevalência local apresenta implicações de outra ordem, que dizem respeito às políticas institucionais de reconhecimento da situação regional e suas dinâmicas territoriais, em que certas espécies e seus ambientes já observados poderão ser extintos, mas seus registros em coleções científicas e séries de dados de pesquisa podem permanecer por longo prazo, enquanto outros elementos de fauna ou flora não serão nem identificados. Porém, a disponibilidade de conteúdos institucionalmente produzidos está vinculada principalmente aos serviços de repositórios de teses e dissertações, além dos dados facilmente acessíveis do seu Herbário UB, no INCT– Flora e Fungos, dependente, portanto, nesse caso, de políticas de escala nacional.

A implantação de políticas e instrumentos de infraestrutura de serviços que favoreçam o acesso aberto, as interações com outras áreas de conhecimento e a preservação de dados de pesquisa, em estratégias de longo prazo, complementam os resultados finais já normalmente objeto de difusão científica e não se contrapõem às proposições de sistemas informacionais nacionais existentes.

Já no caso em estudo, o reuso de dados de coleções botânicas está investido de diversas temporalidades e de retomadas possíveis de interlocução sobre o contexto de produção dos saberes e sobre o meio ambiente de referência. Com o suporte do Dataverse, a conjugação do papel institucional e do propósito disciplinar garante, não só a instauração da gestão de dados de pesquisa, conforme as especificidades do projeto que teve apoio de geoprocessamento, mas permitirá o tratamento no longo prazo dos dados e materiais pesquisados. Estes, em seu conjunto, são dessa forma novamente colocados em circulação no fluxo das elaborações da memória científica sobre o Cerrado e sua dinâmica de paisagens em transformação.

## Considerações Finais

Assim, uma política de gestão de dados em repositórios deve oferecer à comunidade científica e à sociedade em geral diretrizes para a gestão de repositórios de pesquisa que assegurem o acesso gratuito, online, ao conteúdo digital da produção técnico – científica, de diversas áreas do conhecimento, garantindo sua integridade e preservação assim como a observância de direitos autorais.

Algumas das vantagens da implementação de políticas de gestão de dados de pesquisa já foram assinaladas por iniciativas internacionais, entre elas estão:

- a reutilização de dados entre pesquisadores da instituição;
- o crescimento do registro da produção institucional;
- o aumento no acesso à informação científica local;
- a adoção de novos modelos de financiamento para projetos de pesquisa.

Para o cumprimento de uma política de gestão de dados a instituição deverá estabelecer estratégias que estimulem as ações de registro e depósito dos arquivos em repositórios institucionais confiáveis, contando com a colaboração dos especialistas de informação, assim como mecanismos específicos de funcionamento de plataformas de repositórios digitais. Dessa forma, realizar uma gestão eficiente a longo prazo dos dados de pesquisa possibilita o acesso universal e equitativo às informações e dados científicos de qualidade, em toda a gama de disciplinas das ciências naturais às ciências sociais e às humanas.

## Referências bibliográficas

AKMON, Dharma; ZIMMERMAN, Ann; DANIELS, Morgan; HEDSTROM, Margaret (2011) – The Application of Archival Concepts to a Data-Intensive Environment: Working with Scientists to Understand Data Management and Preservation Needs. *Archival Science* [Em linha]. 11, no. 3/4, p. 329–348. [Consult. 5 abr. 2018]. Disponível na Internet: <<http://hdl.handle.net/2027.42/86738.pdf>>

BERTRAN, Paulo (1988) – *Uma introdução à história econômica do Centro-Oeste do Brasil*. Brasília: Codeplan; Goiás: UCG.

BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES – BDTD. Instituto Brasileiro de Informação em Ciência e Tecnologia. *Busca cerrado*. [Consult. 17 abr. 2018]. Disponível na Internet: <<http://bdtb.ibict.br/vufind/>>

BORGAN, C.L. (2007) – *Scholarship in the Digital Age: Information, Infrastructure and the Internet*. Cambridge, MA: MIT Press.

FLORA DO BRASIL 2020 EM CONSTRUÇÃO. Jardim Botânico do Rio de Janeiro. *Página inicial* [Consult. 22 Abr. 2018] Disponível em: <<http://floradobrasil.jbrj.gov.br/>>

INCT – Herbário Virtual Flora e Fungos. Herbário virtual. [Consult. 04 set. 2017] Disponível na Internet: <<http://inct.florabrasil.net/>>

JOINT, Nicholas (2007) – Data Preservation, the New Science and the Practitioner Librarian. *Library Review* [Em linha]. 56, no. 6. p. 451–455. [Consult. 8 abr. 2018]. Disponível na Internet: <<http://strathprints.strath.ac.uk/7182.pdf>>

KARASTI, H.; BAKER, K.; HALKOLA, E. (2006) – Enriching the notion of data curation in e-Science: Data managing and information infrastructuring in the Long Term Ecological Research (LTER) Network. *Journal of Computer Supported Cooperative Work* [Em linha]. 15 (4), p. 321–358.

KLUMP, Jean (2011) – Criteria for the Trustworthiness of Data Centres. *D-Lib Magazine* [Em linha]. Vol. 17, no. 1/2. [Consult. 14 de apr. 2018] Disponível na Internet: <<http://www.dlib.org/dlib/january11/klump/01klump.html>>

MARTINEZ-URIBE, Luis, MACDONALD, Stuart (2018) – Um nuevo cometido para los bibliotecários académicos: data curation. *El Profesional de la Informacion*, v. 17, n. 3.

RYDÉN, Reine (2018) – Extreme long-term preservation of information – who cares? Swedish opinions about nuclear waste information. *Information Research* [Em linha]. Vol. 23 no. 1, [Consult. 28 mar. 2018]. Disponível na Internet: <<http://www.informationr.net/ir/23-1/paper782.html>>

RODRIGUES, Eloy; SARAIVA, Ricardo (2010) – *Os Repositórios de Dados Científicos: Estado da arte*. Universidade do Minho, Projeto RCAPP. [Consult. 19 abr. 2018]. Disponível na Internet: <[http://projeto.rcaap.pt/index.php?option=com\\_remository&Itemid=2&func=startdown&id=271&lang=pt\\_pdf](http://projeto.rcaap.pt/index.php?option=com_remository&Itemid=2&func=startdown&id=271&lang=pt_pdf)>

RUSBRIDGE, Chris; BURNHILL, Peter; ROSS, Seamus et al (2005) – The digital curation centre: a vision for digital curation. In: *Proceedings from local to global: data interoperability – challenges and technologies*. Forte Village Resort, Sardinia, Italy, pp. 1–11. [Consult. 19 abr. 2018]. Disponível na Internet : <<http://eprints.erpanet.org/82>>

SISTEMA DE INFORMAÇÃO SOBRE BIODIVERSIDADE BRASILEIRA. *O que é o SiBBR* [Em linha] [Consult. 22 abr. 2018]. Disponível na Internet: <<http://www.sibbr.gov.br>>

STEINHART, Gail; DIETRICH, Dianne; GREEN, Ann (2009) – Establishing Trust in a Chain of Preservation: The TRAC Checklist Applied to a Data Staging Repository (DataStaR). *D-Lib Magazine* [Em linha]. Vol 15, no. 9/10. [Consult. 14 abr. 2018]. Disponível na Internet: <<http://www.dlib.org/dlib/september09/steinhart/09steinhart.html>>

WITT, Michael (2008) – Institutional Repositories and Research Data Curation in a Distributed Environment. *Library Trends* [Em linha]. 57, no. 2, p. 191–201. [Consult. 13 abr. 2018]. Disponível na Internet : <<http://hdl.handle.net/2142/10680>>