

OS REPOSITÓRIOS DSPACE E ARCHIVEMATICA PARA DOCUMENTOS ARQUIVÍSTICOS DIGITAIS

THE DIGITAL REPOSITORIES DSPACE AND ARCHIVEMATICA FOR DIGITAL RECORDS

SÉRGIO RENATO LAMPERT | Arquivista do Centro Universitário Franciscano e Mestre em Patrimônio Cultural pela Universidade Federal de Santa Maria.

RESUMO

A crescente produção de documentos digitais suscitou nos profissionais da informação incertezas sobre como manter o acesso por longo prazo a estes documentos. Diante de diversas iniciativas existentes, o estudo apresenta uma análise com aspectos dos *softwares DSpace* e *Archivematica*, voltados para definição da ferramenta mais adequada de repositório digital para preservação e acesso de documentos arquivísticos digitais.

Palavras-chave: repositório digital; DSpace; Archivematica; documento arquivístico digital.

ABSTRACT

The growing production of digital documents has raised among information professionals uncertainties on how to maintain access for the long term to these documents. Faced of several existing initiatives, this study presents an analysis with aspects of the *DSpace* and *Archivematica* softwares aimed at defining the most appropriate tool of digital repository for preservation and access of the digital records.

Keywords: digital repository; DSpace; Archivematica; digital records.

RESUMEN

La creciente producción de documentos digitales ha provocado entre los profesionales de la información incertidumbres sobre la forma de mantener el acceso a largo plazo a estos documentos. Frente a varias iniciativas existentes, el estudio presenta un análisis de los aspectos de *DSpace* y *Archivematica* software destinados a definir la herramienta de repositório digital más adecuada para la preservación y el acceso de los documentos de archivo digitales.

Palabras clave: repositorio digital; DSpace; Archivematica; documento de archivo digital.

INTRODUÇÃO

Os documentos de arquivo visam registrar as ações ou atividades desenvolvidas por instituições ou pessoas, sendo preservados e armazenados para que produzam efeitos legais, de comprovação administrativa, jurídica ou fiscal. Também são mantidos para fins históricos e de pesquisa, ao mostrar o que se fazia no passado para as futuras gerações. Logo, considerando o aumento gradual da produção e registro dos documentos de arquivo em meio digital, este recurso informacional relevante pode ser perdido, já que os documentos digitais possuem fragilidades no que tange a sua preservação e armazenamento a médio e longo prazo, tais como: obsolescência tecnológica, de *hardware* e *software*; fragilidade dos suportes; facilidade de serem corrompidos, impossibilidade de acesso e vulnerabilidade à intervenção humana.

Neste sentido, as iniciativas de preservação digital despontam como um recurso para amenizar perdas e manter o acesso aos documentos digitais por grandes períodos. Para Rodrigues (2003, p. 56), o “principal objetivo da preservação de longo prazo é assegurar a longevidade e a manutenção da integridade intelectual dos recursos digitais”. Sob este prisma, além destas, o repositório digital destaca-se como solução para depósito de objetos digitais, à medida que possibilita armazenar, preservar e dar acesso aos mesmos. Corroborando com este entendimento, a Commission on Preservation and Access (CPA) e a Research Libraries Group (RLG) salientam que os sistemas de repositórios digitais se tornarão rapidamente os responsáveis pelo acesso de longo prazo à herança social, econômica, cultural e intelectual mundial em formato digital (CPA/RLG, 1996 apud Sayão, 2011).

Diante do desafio em manter os documentos digitais acessíveis por longo prazo, foram desenvolvidas diversos instrumentos de repositório digital com este fim, das quais, algumas são voltadas para preservação da informação digital científica, tecnológica, cultural e outras, mais recentemente, direcionadas para a preservação de documentos de arquivo. Dessa forma, este estudo apresenta duas ferramentas identificadas como sistemas de repositório digital – *DSpace* e *Archivematica*. A primeira caracteriza-se como uma solução usualmente voltada para a preservação da produção científica e cultural-institucional, e a segunda para a preservação de documentos digitais.

Logo, busca-se investigar os *softwares* apontados acima, apresentando uma análise comparativa das características e funcionalidades, a fim de dar subsídios para escolha da solução mais adequada para preservação e acesso a documentos arquivísticos digitais. Para tanto, o estudo envolve princípios e conceitos da arquivística e preservação digital, bem como apresenta os conceitos, normas e padrões de referência utilizados no desenvolvimento dos *softwares*.

A PRESERVAÇÃO DIGITAL

Os documentos digitais, ao contrário das informações registradas em suporte de papel, são instáveis e vulneráveis. Desse modo, ao considerar o crescimento da produção de documentos digitais, a preservação digital surge como um valioso recurso para garantir o acesso

a dados produzidos em meio digital. Para Ferreira (2006, p. 20), esta consiste no “conjunto de atividades ou processos responsáveis por garantir o acesso continuado a longo prazo à informação e restante patrimônio cultural existente em formatos digitais”.

Rocha e Rondinelli (2006, p. 33), no que diz respeito à preservação, destacam que o “documento digital traz uma contradição: para ser preservado, precisa ser reproduzido”. As autoras complementam que o binômio obsolescência tecnológica/fragilidade do suporte exige intervenções, como migrações, emulações, refrescamento, de maneira a garantir o acesso à informação gerada em meio digital.

Em contrapartida, a *Carta para Preservação do Patrimônio Arquivístico Digital* (Conarq, 2004) afirma que o desafio da preservação dos documentos arquivísticos digitais está na garantia do acesso contínuo a seus conteúdos e funcionalidades, por meio de recursos tecnológicos disponíveis na época em que ocorrer a sua utilização. Nessa direção, compreende-se que a preservação digital não diz respeito apenas à obsolescência tecnológica, corresponde também ao desenvolvimento de um ambiente que assegure a preservação da documentação de caráter permanente (repositório digital) e a definição das estratégias de preservação utilizadas.

OS REPOSITÓRIOS DIGITAIS: TEMÁTICO, INSTITUCIONAL E ARQUIVÍSTICO

Caracterizados como um local para armazenar documentos de maneira segura, para o presente e, principalmente para o futuro, os repositórios visam mantê-los acessíveis por longo prazo. Sob esta perspectiva, Viana et al (2005, p. 3) entendem repositório digital como “uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado, pelo uso de padrões”. Lynch (apud Café et al, 2003), por sua vez, define repositório digital como um “conjunto de serviços oferecidos por uma instituição aos membros de sua comunidade para a gestão e disseminação da sua produção técnico-científica em meio digital”.

Conforme indicado anteriormente, existe uma gama de soluções para repositório digital, que diferem-se uma da outra de acordo com as suas características. Neste sentido, Márdero Arellano (2010) destaca que o tipo de repositório digital é determinado pela aplicação e os objetivos ao qual se destina, além da ferramenta tecnológica que será adotada. Costa e Leite (2006) salientam que estes podem ser classificados como institucionais ou temáticos.

Um repositório temático, para Café et al (2003, p. 3), se constitui “em um conjunto de trabalhos de pesquisa de uma determinada área do conhecimento, disponibilizados na Internet”. Estes utilizam tecnologias abertas e seguem a filosofia da Iniciativa dos Arquivos Abertos (OAI), promovendo a maior acessibilidade à produção dos pesquisadores e à discussão entre seus pares. Márdero Arellano (2010) entende que os repositórios temáticos estão voltados para comunidades científicas específicas, ou seja, tratam, portanto, da produção intelectual de áreas do conhecimento em particular.

Já o repositório institucional é a reunião de todos os repositórios temáticos hospedados em uma organização. Para Lynch (apud Café et al, 2003, p. 5), este é caracterizado como “um

conjunto de serviços que uma universidade oferece aos membros da sua comunidade, para a gestão e disseminação de materiais digitais, criados pela instituição e pelos seus membros”.

Entende-se, a partir disso, que os repositórios digitais eram voltados, em um primeiro momento, para disseminação e preservação da produção científica e intelectual (artigos, teses, dissertações etc.), sendo amplamente utilizados para as bibliotecas digitais. Contudo, estes passaram a ser usados também para fins de arquivo, principalmente para os documentos arquivísticos digitais de valor secundário e de guarda permanente.

Sob este viés, o termo repositório digital incorporou outras funcionalidades, sendo identificado como Repositório Arquivístico Digital. Este, por sua vez, é compreendido como um repositório digital que “armazena e gerencia documentos arquivísticos, seja nas fases corrente e intermediária, seja na fase permanente” (Conarq, 2014, p. 31). Para Rocha (2015), o repositório deve estar em conformidade com as práticas e as normas de arquivo, no que tange à implementação de metadados e a gestão documental.

O DOCUMENTO ARQUIVÍSTICO DIGITAL

Tido como um “documento digital reconhecido e tratado como um documento arquivístico” (Conarq, 2014, p. 18), o documento arquivístico digital remete aos conceitos de documento arquivístico e documento digital. Logo, tem-se que documento arquivístico é o “documento elaborado ou recebido no curso de uma atividade prática como instrumento ou resultado de tal atividade, e retido para ação ou referência” (Projeto InterPARES 3, 2015).

Em contrapartida, o documento digital é compreendido como “a informação registrada, codificada em dígitos binários, acessível e interpretável por meio de sistema computacional” (Conarq, 2014, p. 19).

Sobre o documento arquivístico digital, salienta-se que este é formado por qualidades e características que visam garantir autenticidade, integridade e confiabilidade dos dados registrados por meio dos documentos. Sob este prisma, destacam-se como qualidades do documento arquivístico digital a organicidade, unicidade, confiabilidade, autenticidade e acessibilidade (Conarq, 2011). Já no que diz respeito às características do documento arquivístico digital, estas, conforme a visão da diplomática, disciplina que estuda a autenticidade e confiabilidade dos documentos, são: “forma fixa, conteúdo estável, relação orgânica, contexto identificável, ação e o envolvimento de cinco pessoas (autor, redator, destinatário, originador e produtor)” (Rondinelli, 2011, p. 227).

DEFINIÇÃO DOS SOFTWARES DE REPOSITÓRIO DIGITAL

Diante da diversidade de soluções existentes, buscou-se realizar consultas e pesquisas na *Web*, a fim de identificar *softwares* de referência utilizados como repositório digital, enquanto ambiente que possibilita garantir acessibilidade ao longo do tempo, bem como preservar os documentos. Como resultado, obteve-se cinco soluções, a saber:

- *Fedora: software open source* que fornece um repositório de livre acesso por meio de ampla infraestrutura para o armazenamento, gestão e disseminação de objetos digitais, incluindo o relacionamento entre eles (Sayão; Marcondes, 2009);
- *EPrints: software open source*, compreendido como uma das mais usadas plataformas para repositórios institucionais. É classificado como o modo mais fácil e rápido de criar repositórios de acesso livre para base de dados científica (Sayão; Marcondes, 2009).
- *DSpace: software open source* que fornece um repositório com funções de captura, distribuição e preservação da produção intelectual e científica, dando visibilidade e garantindo acessibilidade no decorrer do tempo (Sayão; Marcondes, 2009).
- *Archivematica: software open source* que fornece um sistema de preservação digital para processamento e armazenamento de objetos digitais a longo prazo, fundamentado em estratégias de preservação digital e baseado no modelo OAIS;
- *RODA: software open source* que disponibiliza um sistema de repositório digital criado para recolher, armazenar, preservar e dar acesso continuado ao patrimônio arquivístico digital a longo prazo. Foi concebido com base no modelo de referência OAIS.

Diante das características elencadas e dos materiais consultados, bem como da percepção do pesquisador no decorrer do levantamento dos *softwares* para repositório digital, pode-se identificar que as ferramentas *Fedora*, *EPrints* e *DSpace* são utilizadas predominantemente para disseminação do conteúdo digital, produzido para pesquisa (produção intelectual e acadêmica). (Lewis, 2011; tradução nossa) destaca que a maioria destas instalações poderia ser classificada como repositório institucional, pois geralmente é fornecida por um instituto de pesquisa, universidade ou departamento para uso de seus pesquisadores. No entanto, também podem ser consideradas repositórios temáticos, como por exemplo, na implementação de repositórios para banco de teses e dissertações.

Essas três soluções foram construídas dentro da filosofia da Iniciativa de Arquivos Abertos (OAI) e do Movimento de Acesso Aberto à Informação Científica (*Open Access*), os quais incentivam a publicação na *Web*, totalmente gerenciada pelo pesquisador (autoarquivamento); utilizam tecnologia aberta; e podem ser acessados por diversos provedores de serviços, disponíveis em nível nacional e internacional (Café et al, 2003). Em outras palavras, conforme Costa e Leite (2006), os repositórios permitem reunir, preservar, dar acesso e disseminar boa parte do conhecimento da instituição, aumentando a visibilidade da sua produção científica.

Em contrapartida, as ferramentas *Archivematica* e *RODA* destacam-se por serem soluções voltadas, principalmente, para preservação de documentos arquivísticos. Neste sentido, estes *softwares* podem ser classificados como soluções voltadas para construção de repositórios arquivísticos digitais. Cabe destacar que este tipo de repositório digital consiste em um ambiente que armazena documentos arquivísticos, seja nas fases corrente e intermediária ou permanente.

Diante do número de soluções encontradas, optou-se por selecionar um instrumento classificado como repositório institucional/temático – *DSpace*, e outra classificada como repositó-

rio arquivístico digital – *Archivematica*. Para definição do *DSpace*, levou-se em conta o fato deste ter sido customizado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), bem como por ser amplamente utilizado por repositórios institucionais de referência. Já a escolha do *Archivematica*, deu-se em virtude deste ser um repositório digital voltado para a preservação de documentos arquivísticos e por ser um sistema indicado pelo Conselho Internacional de Arquivos (CIA), apoiado e desenvolvido pelo Programa Memória do Mundo da Unesco.

DSPACE

Para validação da ferramenta no tocante ao armazenamento, acesso e preservação de documentos arquivísticos digitais, os testes no *software* foram realizados na versão 1.8.1, via interface *Web eXtend Mark Language User Interface (XMLUI)*, no servidor de aplicações *Tomcat* e banco de dados *PostgreSQL*.

Shintaku e Meirelles (2010) destacam que a estrutura informacional do *DSpace*, pelo qual o acervo é disponibilizado, é hierárquica, composta por comunidades, coleções e itens. E o depósito de um documento, conforme os autores, respeita um fluxo de submissão pré-estabelecido, que consiste nas etapas de catalogação, avaliação e revisão dos metadados.

Na primeira etapa do fluxo, identificada como catalogação, são preenchidos os campos do formulário de entrada, caracterizados como metadados descritores. Ainda, na primeira etapa, é inserido o objeto digital a ser armazenado e preservado, bem como deve ser aceita a licença de distribuição do conteúdo. Ao final desta etapa, é apresentado um resumo com os dados preenchidos ao longo dos passos que a compõem. Na avaliação, fase seguinte, o objeto digital inserido pode ser aceito ou rejeitado, conforme a sua pertinência em relação ao conjunto documental armazenado no repositório digital. Para tanto, pode-se levar em conta os pressupostos de um plano de preservação digital, assim como na temporalidade dos documentos. Ao final do fluxo de submissão, a última etapa consiste na revisão dos metadados informados durante o processo de catalogação.

Após submissão do objeto digital no *DSpace*, que, no âmbito desta análise, está delimitado ao exame do documento arquivístico digital, o objetivo é analisar as funcionalidades disponíveis na ferramenta para pesquisa. Logo, tem-se que a consulta é realizada através dos metadados, de acordo com o padrão *Dublin Core*, definidos no fluxo de submissão. Destaca-se que o padrão, composto por 15 elementos, é identificado como um “conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos” (Souza et al, 2000, p. 93).

Além da opção de pesquisa ou consulta através dos metadados, o *DSpace* apresenta um recurso para recuperação de conteúdo. Com isso, todas as palavras do documento podem ser pontos de recuperação para o objeto digital depositado. Entretanto, cabe destacar que esta funcionalidade não realiza a indexação das palavras, sendo apenas um recurso complementar, voltado para ampliar a quantidade de documentos recuperados, corroborando com o ideal de oferecer acesso. Portanto, mesmo que esta característica do *software* proporcione uma recuperação da informação para além dos metadados, ela é frágil, se comparada à pesquisa por termos ou palavras indexadas.

Sobre o armazenamento, o *DSpace* segue a estrutura hierárquica de comunidades, coleções e itens, sendo que os objetos digitais, ao serem submetidos, são estruturados desta forma. O acesso aos documentos digitais, por meio da pesquisa, segue a mesma estrutura. Já sobre a preservação dos documentos digitais, esta segue aspectos do modelo de referência OAIS, que, conforme Rodrigues (2003), é uma infraestrutura conceitual que descreve o ambiente, as interfaces externas, os componentes funcionais e os objetos de informação, associados a um sistema responsável pela preservação de longo prazo de materiais digitais.

De acordo com Márdero Arellano (2008), a preservação digital acontece nos níveis de preservação dos *bits*, onde o arquivo permanece igual com o passar do tempo, assim como na preservação funcional, na qual o arquivo muda conforme os meses e/ou anos passam, mas ainda continua a ser usado da mesma forma em que era originalmente. O autor salienta que a “influência do OAIS no *design* da arquitetura do *DSpace* faz dele um exemplo de repositório com funcionalidades de arquivamento digital” (Márdero Arellano, 2008, p. 140).

Após compreender os recursos da ferramenta para armazenamento, acesso e preservação, conclui-se que a estrutura do *software* é similar ao processo de editoração de um periódico científico. A respeito de suas características, Shintaku e Meirelles (2010) evidenciam que o *DSpace* gerencia e preserva objetos digitais, fornecendo facilidades de recuperação. Além disso, possui propriedades relativas à preservação e divulgação de objetos digitais, podendo adequar-se a várias finalidades. Entretanto, considerando o exposto e as prerrogativas do estudo, entende-se que o *DSpace* não contempla, em sua totalidade, o propósito de um repositório digital voltado para documentos arquivísticos. Apesar deste entendimento, ressalta-se que o *software* pode ser customizado, a fim de atender aos preceitos arquivísticos.

ARCHIVEMATICA

Para análise da ferramenta, esta foi instalada por meio do *software* de emulação/ virtualização *VirtualBox*. A versão 0.9 beta do *Archivematica* foi implantada como máquina virtual no Sistema Operacional *Linux*, distribuição *XUbuntu*.

O *Archivematica* tem como objetivo armazenar a documentação em formato digital, de acordo com os padrões exigidos em relação à preservação digital, sob a ótica arquivística, visando torná-la acessível a longo prazo (Van Garderen et al, 2012, tradução nossa). Destaca-se que foi desenvolvido sob perspectiva da preservação digital, ao implementar as estratégias de emulação, que reproduz uma estrutura de *hardware* e/ou *software* numa outra máquina; migração, que visa a transferência de objetos digitais de uma geração de tecnologia para outra subsequente; e, ainda, a normalização. Sobre esta última, compreendida como a estratégia de preservação que o norteia, Thibodeau (apud Ferreira, 2006, p. 38) destaca que objetiva “simplificar o processo de preservação através da redução do número de formatos distintos que se encontram num repositório de objetos digitais”.

O *software* também é baseado no modelo de referência OAIS/SAAI e nos padrões de metadados *Metadata Encoding Transmission Protocol (METS)*, *Preservation Metadata: Imple-*

mentation Strategies (PREMIS) e *Dublin Core*. Dos modelos e padrão apresentados, ressalta-se que o METS, para Sayão (2010), é um esquema *eXtensible Markup Language* (XML) projetado como uma infraestrutura para codificar todos os tipos de metadados associados a um objeto digital e o PREMIS é “um conjunto de elementos de meta-informação capazes de dar suporte às atividades relacionadas com a preservação digital” (Ferreira, 2009, p. 40). Já sobre o OAIS, este é organizado em um modelo funcional, que implementa as atividades de admissão, armazenamento, gestão de dados, planejamento de preservação, administração e acesso (Sayão, 2010). O OAIS também conta com um modelo funcional, que permite a geração de pacotes de informação: Pacote de Informação de Submissão – *Submission Information Package* (SIP); Pacote de Informação de Arquivamento – *Archive Information Package* (AIP) e Pacote de Informação de Disseminação – *Dissemination Information Package* (DIP).

Para a inserção de um objeto digital no *Archivematica*, a ferramenta tem como primeira etapa um processo de transferência, que pode ser do tipo: padrão, zipado, *DSpace* e *mail-dir*. Cabe destacar que o tipo de transferência *DSpace* permite receber exportação de um repositório oriundo do *DSpace*. Compreende-se, a partir disto, que os desenvolvedores do *Archivematica* consideram o *DSpace* como uma solução consolidada e amplamente utilizada para dar acesso. Na prática, esta característica do *Archivematica* permite que os *softwares* trabalhem em conjunto, sendo este utilizado como um recurso para preservação de objetos digitais, deixando para o *DSpace* a função de dar acesso aos documentos digitais.

Ainda no que diz respeito ao processo de transmissão, é nesta etapa que podem ser atribuídos os metadados do objeto digital, baseados no padrão *Dublin Core*: título, produtor, assunto, descrição, editor, colaborador, data, tipo, formato, identificador, fonte, relação, idioma, cobertura e direitos. Depois de aprovar a transferência do objeto digital, o *Archivematica* implementa uma série de microsserviços em seus processos, tais como: atribuir um identificador único para o objeto digital; gerar um arquivo METS para registro e pesquisa nos diretórios; examinar os arquivos procurando por vírus e outros *malwares*; preservar os nomes dos arquivos nos metadados do PREMIS.

Após o processamento dos microsserviços, inicia-se um *workflow* de aprovação para admissão de um pacote SIP para transferência. Com isso, o fluxo passa para a etapa de admissão, onde também são executados microsserviços, dos quais ressaltam-se os que executam: a normalização para preservação e/ou acesso dos objetos digitais; a geração de um DIP contendo cópias de acesso aos objetos; a definição de *upload* do DIP para os *softwares* ICA-Atom, destinado para descrição arquivística ou CONTENTdm, voltado para coleções digitais; a criação, indexação e armazenamento de um AIP.

Concluído o processo de transferência, bem como a geração dos pacotes de informação, parte-se para a etapa de análise de como ocorrem o armazenamento, acesso e preservação de documentos arquivísticos digitais no programa. A preservação, no *Archivematica*, é implementada no decorrer dos microsserviços, com a geração de metadados de preservação e o tipo de normalização a ser realizada. Sobre a normalização, o *software* disponibiliza quatro opções, as quais permitem criar cópias de preservação dos objetos e/ou cópias de acesso para geração do DIP, bem como a opção de não normalizar. Destaca-se que a padronização,

no *Archivematica*, visa converter os objetos digitais de acordo com um planejamento de preservação segundo a sua finalidade, se para acesso ou preservação.

Sobre o armazenamento dos objetos digitais, este ocorre via geração e indexação do AIP no diretório "Armazenamento de Arquivos". Uma vez indexado, é neste ambiente que pode-se pesquisar os documentos digitais, de acordo com o nome do arquivo ou ainda, conforme os metadados do pacote de informação. Salienta-se que a consulta dos metadados apresenta como resultado apenas os elementos descritores do pacote de informação, e não dos objetos digitais que o compõem, o que compromete o acesso via *software*. Entretanto, a solução apresenta um ambiente voltado para pesquisa e acesso através do processo de geração do pacote DIP, assim como posterior exportação e indexação nos *softwares* ICA-AtoM ou CONTENTdm.

Com base no exposto, entende-se que o *Archivematica* não tem essencialmente como foco o acesso, mesmo que propicie a geração do DIP. Compreende-se que o *software* fundamentalmente é voltado para preservação digital, ao implementar estratégias para tal, especialmente a normalização por meio do planejamento de preservação, bem como pelo fato de ter sido desenvolvido conforme as prerrogativas do modelo OAIS, e por utilizar os esquemas METS e PREMIS.

ANÁLISE COMPARATIVA DAS FERRAMENTAS

Percebeu-se ao longo da investigação que os *softwares* *DSpace* e *Archivematica* possuem semelhanças e diferenças no que tange ao armazenamento, ao acesso e à preservação de documentos digitais, as quais são destacadas a seguir.

Sobre os aspectos em comum, destaca-se que as ferramentas são *softwares open source* que utilizam plataforma *Web* voltados para preservação, operam o padrão *Dublin Core* para os metadados descritivos, podem ser customizados a fim de atender as especificidades de uma instituição, bem como aplicam o modelo OAIS no seu desenvolvimento e o XML para comunicação, o que permite interoperabilidade entre sistemas distintos.

Sobre o *DSpace*, conclui-se que é uma solução voltada para acesso, disseminação e comunicação científica e acadêmica, possuindo uma estrutura de metadados e de submissão de objetos digitais similar ao processo de editoração de um periódico científico. Neste sentido, compreende-se que o *DSpace* não contempla a preservação digital na sua totalidade, considerando todos os princípios propostos no OAIS, assim como o armazenamento de documentos arquivísticos digitais na sua essência, sendo melhor empregado para implementação de bibliotecas digitais e repositórios institucionais, com destaque para as funcionalidades de pesquisa.

Em contrapartida, conclui-se que o *Archivematica* foi desenvolvido para preservar, armazenar e dar acesso a objetos digitais, em conformidade com o esquema OAIS/SAAI. No entanto, seu destaque é a preservação digital, uma vez que ao comparar as ferramentas, percebe-se que o armazenamento de documentos digitais é equivalente nos *softwares* e o acesso via *DSpace* é melhor estruturado. Sobre o acesso, é relevante apontar que o *Archi-*

vemática o disponibiliza por meio da geração do DIP e exportação para outros *softwares*. Compreende-se, com isso, que o *Archivemática* não está disponível para ser acessado por usuários externos via *Web*. Acerca da preservação, este aplica uma série de micros serviços, como varredura de vírus, validação de formatos, captura de metadados compatíveis com PREMIS e METS, normalização de formatos para preservação, admissão de SIP, geração de DIP e armazenamento de AIP em total conformidade com o OAIS.

CONSIDERAÇÕES FINAIS

O objetivo da investigação foi analisar as características e funcionalidades dos *softwares* *DSpace* e *Archivemática*, a fim de levantar elementos que corroborem a escolha da solução mais adequada para preservação e acesso aos documentos arquivísticos digitais. Para tanto, o estudo possibilitou, em um primeiro momento, identificar cinco ferramentas de repositório digital, classificando-as em dois tipos: repositório institucional/temático e repositório arquivístico digital. Dos aplicativos identificados, buscou-se implementar os elencados acima, a fim de averiguar o seu desempenho.

Dos *softwares* de repositório utilizados, conclui-se que o *DSpace* é uma solução voltada predominantemente para dar acesso, em especial, a objetos digitais provenientes de produção e comunicação científica e acadêmica. Entende-se que a ferramenta não contempla a preservação digital e o armazenamento de documentos arquivísticos digitais, em sua totalidade. Já o *Archivemática* tem como principal característica a preservação digital, com base nas estratégias de emulação, migração e normalização e se destaca pela geração de pacotes de informação para admissão, acesso e arquivamento segundo o modelo OAIS.

Sob a perspectiva do estudo, entende-se que o *Archivemática* é mais completo que o *DSpace*, pois atende mais a aspectos da preservação digital e ainda permite o acesso, por transferência de arquivos do *DSpace* ou geração de um DIP para o ICA-AtoM ou CONTENTdm. Em contrapartida, o *DSpace* tem como foco principal o acesso, uma vez que foi idealizado segundo a filosofia da Iniciativa de Arquivos Abertos (OAI) e do Movimento de Acesso Aberto à Informação Científica (*Open Access*). Portanto, diante do exposto, compreende-se que o *Archivemática* destaca-se para preservar e que o *DSpace* evidencia-se por dar acesso.

Por fim, após análise das ferramentas, compreende-se que os *softwares* podem ser identificados como sistemas voltados para construção de repositórios digitais-referências para o armazenamento de documentos arquivísticos digitais. No entanto, o *DSpace*, se for implementado para preservação digital, além do acesso, necessitará de um mecanismo complementar para este fim. Do mesmo modo, o *Archivemática*, se for utilizado essencialmente para acesso, deverá ter um *software* em conjunto para tanto. Ainda salienta-se que as considerações, assim como os recursos e funcionalidades dos *softwares*, são pautados sob o ponto de vista do autor, considerando a problemática do estudo. Logo, tem-se que os resultados aqui apontados não visam reputar um ou outro, mas sim apresentar subsídios, qualidades e fragilidades de cada solução, a fim de auxiliar as instituições na tomada de decisão a respeito de qual *software* utilizar para preservação e acesso de documentos arquivísticos digitais.

Referências bibliográficas

CAFÉ, Lígia et al. Repositório Institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO (INTERCOM), 26., 2003, Belo Horizonte. *Anais eletrônicos...* Belo Horizonte: PUC Minas, 2003. Disponível em: <http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_ENDOCOM_TRABALHO_cafe.pdf>. Acesso em: 07 ago. 2015

Conselho Nacional de Arquivos - CONARQ (Brasil). *Carta para a preservação do patrimônio arquivístico digital*. Rio de Janeiro: Arquivo Nacional, 2004.

_____. Câmara Técnica de Documentos Eletrônicos. *Glossário*. Rio de Janeiro: Arquivo Nacional, 2014. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/imagens/ctde/Glossario/2014ctdeglossario_v6_public.pdf>. Acesso em: 03 ago. 2015.

COSTA, Sely Maria de Souza.; LEITE, Fernando César Lima. Repositórios institucionais: potencial para maximizar o acesso e o impacto da pesquisa em universidades. In: CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA, 1., 2006, Brasília. *Anais eletrônicos...* Brasília: Universidade de Brasília, 2006. Disponível em: <<http://eprints.rclis.org/10021/>>. Acesso em: 13 mar. 2015.

FERREIRA, Miguel. *Introdução à preservação digital – conceitos estratégias e actuais consensos*. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006.

_____. *Preservação de longa duração de informação digital no contexto de um arquivo histórico*. 2009. Tese (Doutorado em Tecnologias e Sistemas de Informação), Universidade do Minho/Portugal.

LEWIS, Stuart. Open repositories for scholarly communication: enhancing access to research In: *Itens Técnicos developerWorks*. [S.l.]: IBM, 2011. Disponível em: <<http://www.ibm.com/developerworks/opensource/library/os-ind-educstand3/os-ind-educstand3-pdf.pdf>>. Acesso em: 19 out. 2015.

MÁRDERO ARELLANO, Miguel Angel. *Critérios para preservação digital da informação científica*. 2008. Tese (Doutorado em Ciência da Informação), Universidade de Brasília, Brasília.

_____. Repositórios, Acesso Livre, Preservação Digital. *Encontros Bibli*, Florianópolis, v. 15, n. 29, 2010. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/13706>>. Acesso em: 25 jan. 2015.

PROJETO INTERPARES 3. Base de dados de terminologia do InterPARES 3. *Documento arquivístico*. Team Brasil. 2015. Disponível em: <http://www.interpares.org/ip3/ip3_terminology_db.cfm?letter=d&term=101>. Acesso em: 10 set. 2015.

ROCHA, Cláudia Lacombe. Repositórios para a preservação de documentos arquivísticos digitais. *Acervo*, Rio de Janeiro, v. 28, n. 2, p. 180-191, 2015. Disponível em: <<http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/608/669>>. Acesso em: 07 nov. 2015.

_____.; RONDINELLI, Rosely Curi. Gestão arquivística e preservação de documentos digitais no Brasil: estado da arte. *Arquivo & Administração*, Rio de Janeiro, v. 5, n. 1, p. 33-40, 2006.

RODRIGUES, Maria de Lurdes Tainha Saramago. *Preservação digital de longo prazo: estado da arte e boas práticas em repositórios digitais*. 2003. Dissertação (Mestrado em Estudos de Informação e Bibliotecas Digitais), Instituto Superior de Ciências do Trabalho e da Empresa, Lisboa/Portugal.

RONDINELLI, Rosely Curi. *O conceito de documento arquivístico digital frente à realidade digital: uma revisão necessária*. 2011. Tese (Doutorado em Ciência da Informação), Universidade Federal Fluminense, Niterói.

SAYÃO, Luís Fernando. Uma outra face dos metadados: informações para a gestão da preservação digital. *Encontros Bibli*, Florianópolis, v. 15, n. 30, p. 1-31, 2010. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/12528/19527>>. Acesso em: 07 jan. 2015.

_____. Repositórios Digitais Confiáveis: Conceitos, Tecnologias e Padrões. In: FÓRUM DE CIÊNCIA E TECNOLOGIA: REPOSITÓRIOS CONFIÁVEIS DE DOCUMENTOS ARQUIVÍSTICOS DIGITAIS, 2011. *Tópico temático...* Campinas: Unicamp. Disponível em: <<http://www.foruns.unicamp.br/Arquivos%20Biblioteca%20Virtual/Palestras/11-08/Prof.%20Say%C3%A3o%20-%20Repositórios-confi%C3%A1veis-agosto-2011-unicamp.pdf>>. Acesso em: 07 fev. 2015.

_____; MARCONDES, Carlos Henrique. Software livres para repositórios institucionais: alguns subsídios para a seleção. In: Sayão et al (Org.). *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. Salvador: EDUFBA, 2009.

SHINTAKU, Milton; MEIRELLES, Rodrigo. *Manual do DSPACE: administração de repositórios*. Salvador: EDUFBA, 2010.

SOUZA, Marcia Izabel Fugisawa; VENDRUSCULO, Laurimar Gonçalves; MELO, Geane Cristina. Metadados para descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. *Ciência da Informação*, Brasília, v. 29, n. 1, p. 93-102, 2000. Disponível em: <<http://revista.ibict.br/ciinf/article/view/903/940>>. Acesso em: 10 set. 2015.

VAN GARDEREN, Peter et al. The Archivematica Project: meeting digital continuity's technical challenges. In: THE MEMORY OF THE WORLD IN THE DIGITAL AGE: digitization and preservation, 2012, Vancouver/Canadá. *Anais eletrônicos...* Vancouver: University of British Columbia (Canadá), 2012. Disponível em: <http://www.ciscra.org/docs/UNESCO_MOW2012_Proceedings_FINAL_ENG_Compressed.pdf>. Acesso em: 29 jul. 2014

VIANA, Cassandra Lúcia de Maya et al. Repositórios institucionais em ciência e tecnologia: uma experiência de customização do DSpace. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 3., 2005, São Paulo. *Anais eletrônicos...* São Paulo: Universidade de São Paulo, 2005. Disponível em: <<http://eprints.rclis.org/7168/1/viana358.pdf>>. Acesso em: 14 jun. 2015.

Recebido em 31/5/2016
Aprovado em 6/9/2016