

Preservação de documentos arquivísticos digitais a longo prazo em repositórios digitais confiáveis

Elisa Maria Lopes Chaves

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Programa de Pós-Graduação em Ciência da Informação, Marília, SP, Brasil
elisachaves8@gmail.com

DOI: <https://doi.org/10.26512/rici.v16.n1.2023.44023>

Recebido/Recibido/Received: 2022-07-08

Aceitado/Aceptado/Accepted: 2022-12-30

Resumo

O processo de preservação dos documentos arquivísticos digitais devem considerar ferramentas consolidadas que consideram o ciclo de vida do documento de arquivo, além de parâmetros que considerem o armazenamento e a preservação a longo prazo. Para os documentos arquivísticos existem normas e padrões nacionais e internacionais que preconizam o uso de Repositórios Arquivísticos Digitais Confiáveis. O objeto deste trabalho introduz as premissas de um repositório confiável e seus requisitos e destaca os principais repositórios disponibilizados na área. Através de uma pesquisa exploratória e qualitativa serão analisados três repositórios, são eles: Archimática, Repositório de Objetos Digitais Autênticos e DSpace que são apresentados e avaliados no que concerne aos requisitos de um Repositório Arquivístico Digital Confiável. Na pesquisa comparativa apresentada observa-se que o Archimática e o Repositório de Objetos Digitais Autênticos são considerados Repositórios Arquivísticos Digitais Confiáveis e o DSpace é mais utilizado como um repositório de armazenamento. O Archimática destaca-se por ser mais utilizado, além da bibliografia mais extensa. Conclui-se que além da utilização de um Repositório Arquivístico Digital Confiável o mesmo deve ser compreendido dentro de uma Política de Preservação Digital bem delineada e difundida e aliado a ferramenta de acesso também compatível com padrões e normas arquivísticas.

Palavras-Chave: Archimática. Repositório Arquivístico Digital Confiável. Preservação digital.

Long-term preservation of digital archival documents in trustworthy digital repositories

Abstract

The process of preserving digital records must consider consolidated tools that consider the life cycle of the record, in addition to parameters that consider storage and long-term preservation. For archival documents, there are national and international norms and standards that advocate the use of Reliable Digital Archival Repositories. The object of this work introduces the premises of a trusted repository and its requirements and highlights the main repositories available in the area. Through an exploratory and qualitative research, three repositories will be analyzed, they are: Archimatica, Repository of Authentic Digital Objects and DSpace that are presented and evaluated regarding the requirements of a Reliable Digital Archival Repository. In the comparative research presented, it is observed that Archimatica and the Repository of Authentic Digital Objects are considered Reliable Digital Archival Repositories and DSpace is more used as a storage repository. Archimática stands out for being the most used, in addition to the more extensive bibliography. It is concluded that in addition to the use of a Reliable Digital Archival Repository, it must be understood within a well-designed and widespread Digital Preservation Policy and combined with an access tool that is also compatible with archival standards and norms.

Keywords: Archimatica. Reliable Digital Archival Repositories. Digital preservation.

Preservación a largo plazo de documentos de archivo digital en repositorios digitales confiables

Resumen

El proceso de preservación de documentos digitales debe considerar herramientas consolidadas que consideren el ciclo de vida del documento, además de parámetros que consideren almacenamiento y preservación a largo plazo. Para los documentos de archivo, existen normas y estándares nacionales e internacionales que abogan por el uso de Repositorios de Archivos Digitales Confiables. El objeto de este trabajo introduce las premisas de un repositorio de confianza y sus requisitos y destaca los principales repositorios disponibles en el área. A través de una investigación exploratoria y cualitativa, se analizarán tres repositorios, ellos son: Archivemática, Repositorio de Objetos Digitales Auténticos y DSpace que se presentan y evalúan en cuanto a los requisitos de un Repositorio Archivístico Digital Confiable. En la investigación comparativa presentada, se observa que Archivemática y el Repositorio de Objetos Digitales Auténticos son considerados Repositorios de Archivos Digitales Confiables y DSpace es más utilizado como repositorio de almacenamiento. Archivemática destaca por ser la más utilizada, además de la bibliografía más extensa. Se concluye que además del uso de un Repositorio Archivístico Digital Confiable, debe entenderse dentro de una Política de Preservación Digital bien diseñada y difundida y combinada con una herramienta de acceso que también sea compatible con los estándares y normas archivísticas.

Palabras clave: Archivemática. Repositorio Archivístico Digital Confiable. Preservación digital.

1 Introdução

Ao pesquisar sobre a gestão e a preservação de documentos arquivísticos digitais em sua fase permanente, constata-se que os documentos não devem ser armazenados em base de dados para acesso, pois estas não são ferramentas para salvaguarda de acervo. Considerando também que os sistemas computacionais não são infalíveis, as plataformas de acesso acabam sendo mais frágeis a invasões, pois muitas vezes são a porta de entrada do público externo à instituição. Deste modo, por questões de segurança, a ferramenta de preservação e armazenamento deve ser diferente da ferramenta de acesso.

Nesta perspectiva o objetivo deste artigo é discutir a relevância da implementação de um Repositório Arquivístico Digital Confiável – RDC-Arq, pois além de ser um espaço de armazenamento confiável é uma ferramenta que garante a autenticidade e integridade dos documentos arquivísticos permanentes e sua preservação a longo prazo. No entanto é importante o estabelecimento de um repositório que siga padrões estabelecidos nas normas e padrões arquivísticos nacionais e internacionais como veremos ao longo deste artigo.

Outro ponto importante da adoção de um repositório confiável a longo prazo é a garantia de migrações dos formatos dos arquivos para formatos de preservação e acesso, questão sensível principalmente na produção de documentos iconográficos e audiovisuais, onde a quantidade de formatos proprietários é mais comum. A garantia de migrações dos formatos a longo prazo tenta minimizar a obsolescência muito comum nos documentos digitais.

Para tanto, algumas questões importantes sobre a preservação digital serão apontadas, mas o objetivo principal é a apresentação e caracterização de um RDC-Arq, pois estudos na área têm direcionado para o uso dessa ferramenta para a preservação dos documentos arquivísticos

digitais. A metodologia adotada pautou-se no levantamento bibliográfico e análise qualitativa sobre os principais repositórios utilizados na área, a saber, o Archivemática, o Repositório de Objetos Digitais Autênticos - RODA e o DSpace cada um deles com suas particularidades. Ao final reuniu-se os principais pontos para a escolha de um RDC-Arq, assim como a apresentação de um quadro explicativo e comparativo entre as ferramentas presentes neste artigo.

2 Preservação Digital

A preservação digital consiste na capacidade de garantir que a informação digital permaneça acessível e com qualidade de autenticidade suficiente para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação (FERREIRA, 2006, p. 20).

Deste modo, "preservação digital" é o conjunto de atividades ou processos responsáveis por garantir o acesso continuado a longo-prazo à informação e ao patrimônio cultural existente em formatos digitais.

As instituições preocupadas com a preservação de seus documentos digitais devem implementar uma Política de Preservação Digital - PPD. A PPD tem como objetivo registrar as diretrizes e os objetivos a fim de nortear uma instituição sobre o tema e garantir o acesso, a integridade e a autenticidade dos documentos de seu acervo. A elaboração de uma PPD é fundamental para a implantação de um programa de preservação dos documentos e dos objetos digitais (GRÁCIO, 2020, p. 568-569).

Quando não existe uma PPD amplamente difundida e implementada, ou em vias de ser implementada, em uma instituição, o acervo digital institucional pode ser acometido por diversos problemas, tais como: obsolescência, insegurança, fragilidade, adulteração, corrupção da autenticidade, entre outros. Infelizmente no âmbito digital alguns desses prejuízos não podem ser revertidos, o que pode ocasionar perda de informação, de documento ou até mesmo de partes do acervo digital institucional.

Como elemento do processo de preservação digital dos documentos arquivísticos conforme as Diretrizes do Conselho Nacional de Arquivos - CONARQ (2015, p. 15-16) um RDC-Arq deve fazer parte do planejamento da preservação dos documentos sob sua custódia, a fim de enfrentar os problemas trazidos pela obsolescência tecnológica e fragilidade do suporte. Esse planejamento deve ser feito a partir de uma política de preservação digital e ser bem documentado.

Dito isso é importante, como veremos com mais detalhes, implementar Repositórios Digitais Confiáveis – RDC-Arq como um elemento essencial na política de preservação digital para assegurar o armazenamento e a preservação digital de forma segura a longo prazo.

3 Repositório Arquivístico Digital Confiável – RDC-Arq

A importância do uso dos Repositórios Arquivísticos Digitais Confiáveis – RDC-Arq se dá desde a produção documental. De acordo com CONARQ (2015, p. 4) “a preservação dos documentos arquivísticos digitais, nas fases corrente, intermediária e permanente, deve estar associada a um Repositório Digital Confiável”, ou seja, o repositório digital confiável deve ser incluído no processo de gestão, preservação e de acesso aos documentos digitais.

Dito isto antes de conceituação do termo delimita-se que os repositórios não podem ser confundidos com base de dados, *storages*, backups e outras possíveis soluções informatizadas para armazenamento que fazem parte de um dos componentes de um repositório, mas não são considerados repositórios digitais (CONARQ, 2015, p. 9).

Um RDC-Arq para documentos arquivísticos deve ser inserido dentro dos princípios básicos da Arquivologia. Nesse sentido respeitar a organização hierárquica partindo de um plano de classificação¹ e da descrição multinível² pautadas em normas de descrição como a Norma Geral Internacional de Descrição Arquivística – ISAD (G) (CONARQ, 2015, p.10). Destaque para as principais normas, requisitos e padrões arquivísticos nacionais e internacionais que um RDC-Arq deve seguir:

- *Norma Brasileira de Descrição Arquivística – Nobrade;*
- *Norma Geral de Descrição Arquivística – ISAD(G);*
- Resolução N° 43 do Conselho Nacional de Arquivos – CONARQ que estabelece as Diretrizes para a implementação de Repositórios Digitais Confiáveis – RDC-Arq;
- Modelo ISO/OAIS 14721/2012 → NBR 15.472:2007 – Modelo de referência para um Sistema Aberto de Arquivamento de Informação (SAAI);
- ISO 16363:2012–*Norma de Auditoria e Certificação de Repositórios Digitais Confiáveis;*
- Carta para a Preservação do Patrimônio Arquivístico Digital UNESCO e CONARQ;
- E-Arq Brasil Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos do CONARQ.

¹ Organização de documentos em classes elaborado a partir do estudo das estruturas e funções de uma instituição e da análise do arquivo por ela produzido (BRASIL, 2005, p. 132).

² Descrição que permite a recuperação das informações dos documentos de um arquivo em diferentes níveis, do mais abrangente ao mais específico (BRASIL, 2005, p. 67).

Repositório Digital (assim como suas especificações) de acordo com o Glossário de Documentos Arquivísticos Digitais, da Câmara Técnica de Documentos Eletrônicos – CTDE do CONARQ (2020, p.42 e 43) consiste:

Repositório digital um complexo que apoia o gerenciamento dos materiais digitais, pelo tempo que for necessário, e é formado por elementos de *hardware*, *software* e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos.

Repositório Digital Confiável repositório digital que é capaz de manter autênticos, preservar e prover acesso a materiais digitais pelo tempo necessário.

Repositório Arquivístico Digital repositório digital que armazena e gerencia documentos arquivísticos, seja nas fases corrente e intermediária, seja na fase permanente.

Assim, entende-se que um Repositório Arquivístico Digital Confiável – RDC-Arq é uma ferramenta complexa que garante a preservação de documentos arquivísticos digitais a longo prazo, dentro da cadeia de custódia (fase corrente, intermediária até a permanente) para um documento arquivístico.

O gerenciamento do documento digital de um RDC-Arq envolve segundo as Diretrizes do CONARQ (2015, p. 13) “*estar de acordo com o modelo de referência OAI³, que estabelece a formação de pacotes de informação envolvendo os documentos digitais (informação de conteúdo) e seus metadados (informação de representação)*”. De modo que para constituir um Ambiente de Preservação no Arquivo Permanente Digital das instituições, o RDC-Arq segundo IBICT (2016, p. 11) deve:

- gerenciar os documentos e metadados de acordo com os princípios e as práticas da Arquivologia, especificamente relacionados à gestão documental, descrição arquivística multinível e preservação;
- proteger as características do documento arquivístico, em especial a autenticidade (identidade e integridade) e a relação orgânica;
- preservar e dar acesso, pelo tempo necessário, a documentos arquivísticos digitais autênticos;
- estar em conformidade com a ISO 16363:2012, que lista os critérios a que um repositório digital confiável deve atender.

Um RDC-Arq prevê a independência dos repositórios no que diz respeito ao sistema de gestão responsável pela entrada do documento no RDC-Arq, interoperabilidade com padrões de metadados consistentes na área e gerenciamento de informação entendido como gestão da informação descritivas (metadados) conforme orientações do CONARQ (2015, p.11 e 16).

³*Open Archival Information System* –[Sistema de Informações de Arquivo Aberto].

São recomendados os padrões de metadados como OAI-PMH⁴, METS⁵ e o EAD⁶ que são padrões internacionalmente aceitos e que atendem a um determinado propósito de representações. A importância de metadados compatíveis e consistente na área se dá, pois de acordo com Alves (2017, p. 99) os metadados e padrões de metadados foram desenvolvidos a partir de necessidade de representar e identificar os recursos informacionais disponibilizados na *web*, possibilitando melhorar a recuperação da informação pelos mecanismos de busca, mas também seu acesso e sua preservação em longo prazo, permitindo a interoperabilidade dos dados entre ambientes de diferentes áreas.

Além da adoção de metadados nos documentos dentro de um repositório, para o mesmo ser considerado um RDC-Arq prevê-se também a possibilidade de controle e auditoria do RDC-Arq baseados na norma ISO 16363:2012⁷. De acordo com os autores Gava e Flores (2020, p.77) a ISO 16363 é a norma que permite a certificação de confiança, em nível internacional, para repositórios digitais de organizações públicas ou privadas, e amplamente usada na Ciência da Informação. Desta forma o RDC-Arq também difere de bases de dados comuns, pois podem ser auditados.

4. Cadeia de custódia

Um RDC-Arq deve ser envolvido no trâmite documental, obedecendo uma cadeia de custódia. Cadeia de custódia é “linha contínua de custodiadores de documentos arquivísticos (desde o seu produtor até o sucessor) pela qual se assegura que esses documentos são os mesmos desde o início, não sofreram nenhum processo de alteração e, portanto, são autênticos” (CONARQ, 2012, p.1).

Como dito anteriormente, para os documentos serem inseridos em uma cadeia de custódia dentro de um RDC-Arq existem padrões de modelos de referência a serem utilizados como o modelo OAIS. O modelo de referência OAIS é regulamentado através da ISO 14721:2012⁸. Trata-se de um esquema conceitual que disciplina e orienta um sistema de

⁴*Open Archives Initiative Protocol for Metadata Harvesting* – [Protocolo da Iniciativa de Arquivos Abertos para Coleta de Metadados] para coleta de registros de metadados em repositórios digitais.

⁵*Metadata Encoding and Transmission Standard* – [Padrão de codificação e transmissão de metadados] para a codificação de metadados descritivos, administrativos e estruturais.

⁶*Encoded Archival Description* – [Descrição de Arquivo Codificada] para a codificação de metadados descritivos de documentos arquivísticos.

⁷*Space data and information transfer systems – Audit and Certification of Trustworthy Digital Repositories* – [Auditoria e Certificação de Repositórios Digitais Confiáveis].

⁸*Space data and information transfer systems – Open archival information system (OAIS) – Reference model*. Criado em 2002, o modelo de referência passou a ser uma norma internacional, regulada na ISO 14721:2003 atualizada pela ISO 14721:2012. Traduzida para o português como “Sistema Aberto de Arquivamento de Informação” – SAAI através da recomendação da ABNT/NBR 15472:2007.

arquivo, contemplando a preservação e a manutenção do acesso a informações digitais por longo prazo.

O modelo OAIS toma como base quatro entidades, os produtores, a administração, os consumidores e o arquivo propriamente dito:

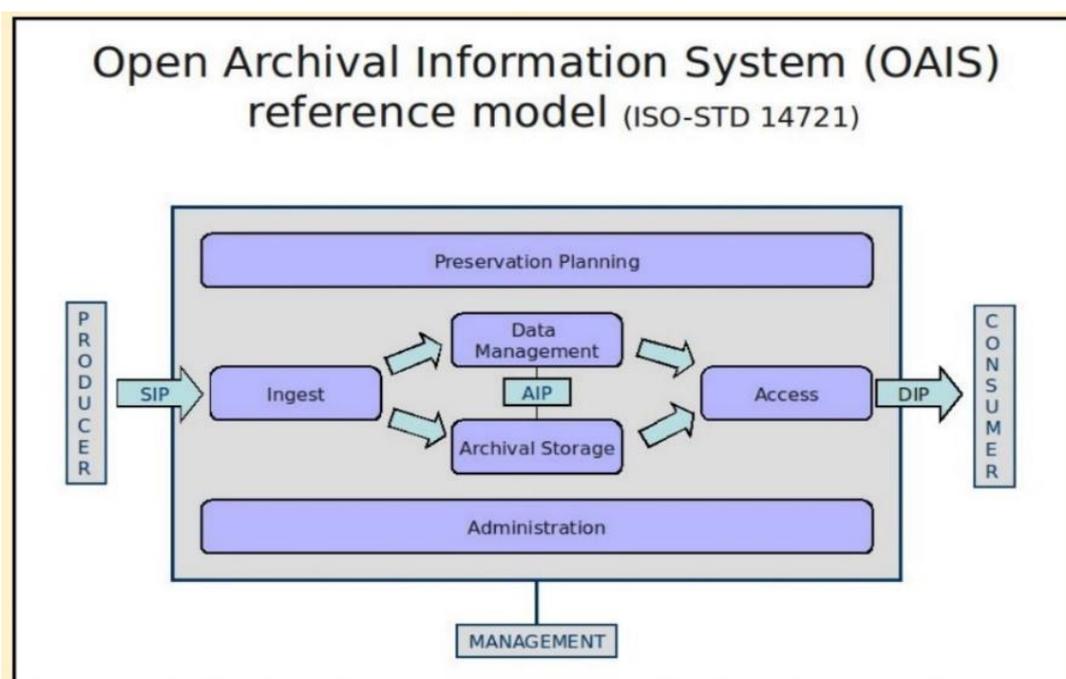
o produtor: papel desempenhado por pessoas ou sistemas que fornecem a informação a ser preservada, na fase da Submissão (SIP);

o administrador: papel desempenhado por aqueles que estabelecem as políticas gerais que governam o repositório, na fase do Arquivamento (AIP) e, por fim;

o consumidor: papel desempenhado por pessoas ou sistemas que interagem com os serviços OAIS para acessar a informação preservada na fase de Acesso e Difusão (DIP).

A figura 1 exemplifica a relação do modelo OAIS do fluxo documental dentro da cadeia de custódia, ou seja, da produção ao acesso.

Figura 1 – Fluxograma dos pacotes SIP/AIP/DIP para o modelo OAIS



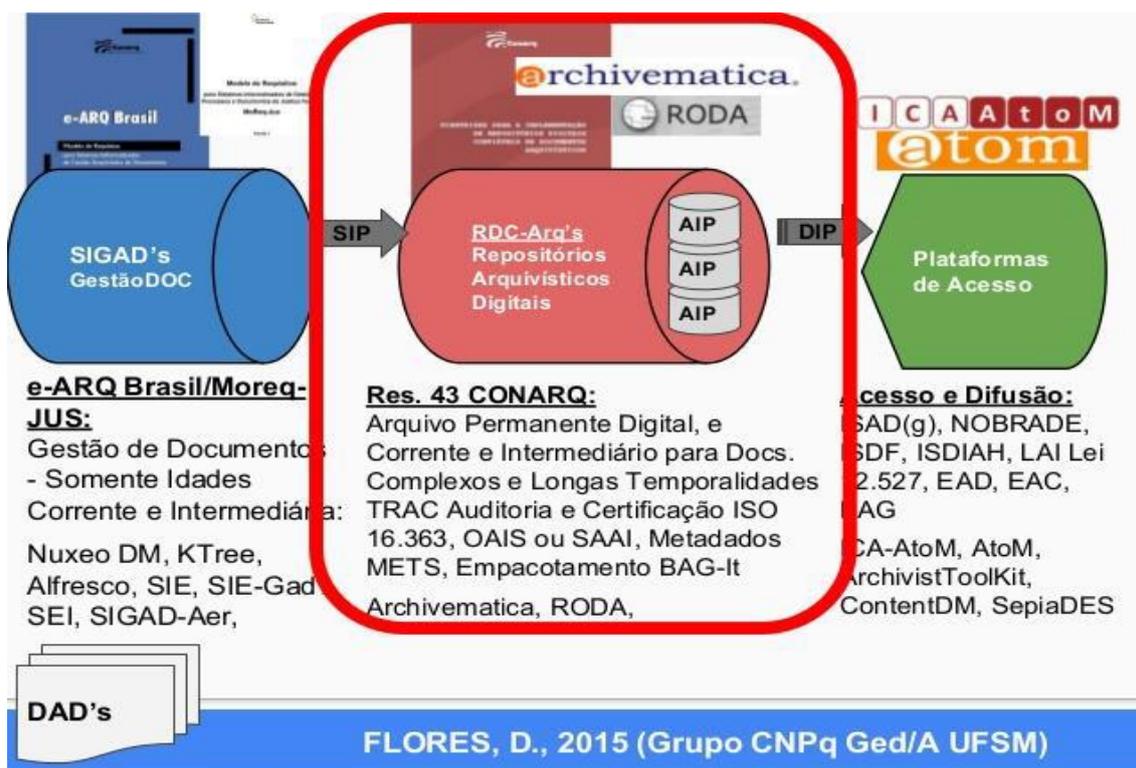
Fonte: FLORES, 2016.

No Brasil, conforme estudos de Luz e Flores (2017, p.171) a produção do documento digital se faz nos Sistemas Informatizados para a Gestão Arquivística de Documentos – SIGAD's. Os SIGADs contemplam as primeiras fases dos documentos (arquivo corrente e intermediário) o Repositório Digital Arquivístico - RDC-Arq contempla a fase de preservação, como Ambiente de Preservação Digital representado pelo Archivemática, RODA, ou qualquer outra ferramenta

que venha a contemplar os requisitos arquivísticos e ferramentas como o AtoM (do Conselho Internacional de Arquivos – CIA) atendem a descrição, acesso e difusão.

Na figura 2 a representação da cadeia de custódia com exemplos de ferramentas adequadas para cada fase do documento arquivístico. Como ferramenta de produção o uso dos SIGADs, o arquivamento através dos RDC-Arq e plataforma de acesso com o AtoM.

Figura 2 – Cadeia de custódia para os documentos digitais SIGAD, RDC-Arq e o AtoM



Fonte: Flores, 2016.

Este processo, representado no fluxograma acima, visa garantir a autenticidade do documento ao ser inserido no ciclo envolvendo uma cadeia de custódia ininterrupta.

Neste artigo são apresentados três repositórios: o Archivemática, o RODA e o DSpace. Cada repositório possui suas particularidades que foram exploradas para que fosse possível um entendimento sobre suas principais funcionalidades. Será analisado cada um deles e suas interfaces com outras ferramentas e por fim, eles serão comparados como Repositórios Arquivísticos Digitais Confiáveis – RDC-Arq.

5 Archivemática

O Archivemática é um RDC-Arq criado como um *software* livre, desenvolvido em código aberto, com acesso ao código fonte pela empresa *Artefactual System*. A Artefactual é uma

empresa canadense que desenvolveu o Archimática junto a Organização das Nações Unidas para a Educação, a Ciência e a Cultura - UNESCO.

De acordo com o IBICT (2016, p.14) o Archimática é um sistema de preservação digital projetado para manter os dados baseados em padrões de preservação digital e o acesso em longo prazo para coleções de objetos digitais. É considerado um RDC-Arq, pois foi criado com a finalidade de armazenar documentos em formato digital, dentro de uma cadeia de custódia.

O Repositório conta com uma comunidade ativa na *web* através de listas de discussão de acordo com os serviços fornecidos pela Artefactual, analistas de sistemas da empresa e usuários contribuem ativamente na discussão. As novas versões e atualizações são divulgadas na página sobre o *software*. Todo o código do Archimática é liberado, assim como todo o sistema de documentação e desenvolvimento de infraestrutura está disponível de forma gratuita e liberado sob AGPL⁹ com as licenças *Creative Commons* (IBICT, 2016, p.9).

O Archimática segue protocolos como o padrão ISO-OAIS e utiliza padrões de metadados, tais como METS, PREMIS¹⁰ e Dublin Core.

A ferramenta utiliza a especificação *BagIt* da Library of Congress¹¹ que consiste em método para empacotar diretórios de arquivos, hierarquicamente, através de transferência criada por outros programas/sistemas para armazenamento a longo prazo ou para transferência entre ambientes de armazenamento (FLORES, 2016). Desta forma a ferramenta oferece muitos fluxos de trabalho de ingestão: importação de documentação de submissão e metadados, compactação e extração de pacotes *zipados*, processamento forense digital de imagens, organização de SIP, normalização manual e gerenciamento de conjunto de dados.

O acesso ao Archimática se dá através de navegador *web*. Possui compatibilidade com diversos formatos através do Registro de Política de Formatos – FPR¹² que é integrado ao PRONOM. O PRONOM¹³ é uma base de dados com registro de formatos mantida pelo Arquivo Nacional do Reino Unido (CONARQ, 2015, p. 15). O FPR também oferece uma estrutura editável e flexível para identificação de formatos, extração de pacotes, transcrição e normatização para preservação e acesso¹⁴.

⁹ Licença A-GPL 3.0: GNU *Affero General Public License* – [Licença Pública Geral].

¹⁰ *Preservation Metadata: Implementation Strategies* – [Metadados de preservação: implementação de estratégias].

¹¹ Biblioteca do Congresso dos Estados Unidos.

¹² *Format Policy Registry* – [Registro de política de formatos].

¹³ Disponível em: <http://www.nationalarchives.gov.uk/PRONOM/>

¹⁴ Alguns objetos digitais proprietários, tais como desenhos CAD ou documentos do Microsoft Word que não possuem código aberto são mantidos em seu arquivo (formato) original, não sendo convertidos para formatos de preservação.

Ao instalar o Archivemática, os formatos, as regras e os comandos padrões são instalados através do FPR. Quando o Archivemática é atualizado, novos formatos, regras e comandos podem ser adicionados. Mas também é possível manter regras locais para adicionar novos formatos ou personalizar os formatos existentes.

Na figura 3 é exemplificado parte da base de dados de formatos acessíveis disponibilizados pelo PRONOM.

Figura 3: Política de Formatos conforme PRONOM

Tipo de media	Formato	Formato(s) de preservação	Formato(s) de acesso	Ferramenta de normalização
Audio	AC3, AIFF, MP3, WAV, WMA	WAVE (LPCM)	MP3	FFmpeg
Office Open XML	DOCX, PPTX, XLSX	Formato Original	Formato Original	Tool search in progress
Texto simples (plain text)	TXT	Formato Original	Formato Original	None
Portable Document Format	PDF	PDF/A	Formato Original	Ghostscript
Arquivos de apresentação	PPT	Formato Original	PDF	Tool search in progress
Imagens rasterizadas	BMP, GIF, JPG, JP2*, PCT, PNG*, PSD, TIFF, TGA	Uncompressed TIFF	JPEG	ImageMagick
Formato de imagem raw/negativos**	CR2, CRW, DCR, DNG, ERF, KDC, MRW, ORF, PEF, RAF, RAW	Formato Original	JPEG	ImageMagick/UF Raw
Planilhas	XLS	Formato Original	Formato Original	None
Video	AVI, FLV, MOV, MPEG-1, MPEG-2, MPEG-4, SWF, WMV	FFV1/LPCM in MKV	MP4	FFmpeg

Fonte: Canelha (2019).

Considerando as informações pesquisadas e aqui apresentadas, o Archivemática é uma ferramenta segura para armazenamento e preservação do documento arquivístico digital a longo prazo, sendo considerado um RDC-Arq.

O Archivemática também possui integração com diversos sistemas, tais como o DSpace e o AtoM exemplificados a seguir.

5.1 O Archivemática e o AtoM

O Archivemática possui como ferramenta de acesso o Atom. Nessa atuação um exemplo do ciclo vital de documentos arquivísticos digitais, ou seja, uma cadeia de custódia ininterrupta assegurada se faz através de ambientes autênticos, por exemplo, de um SIGAD, RDC-Arq e AtoM.

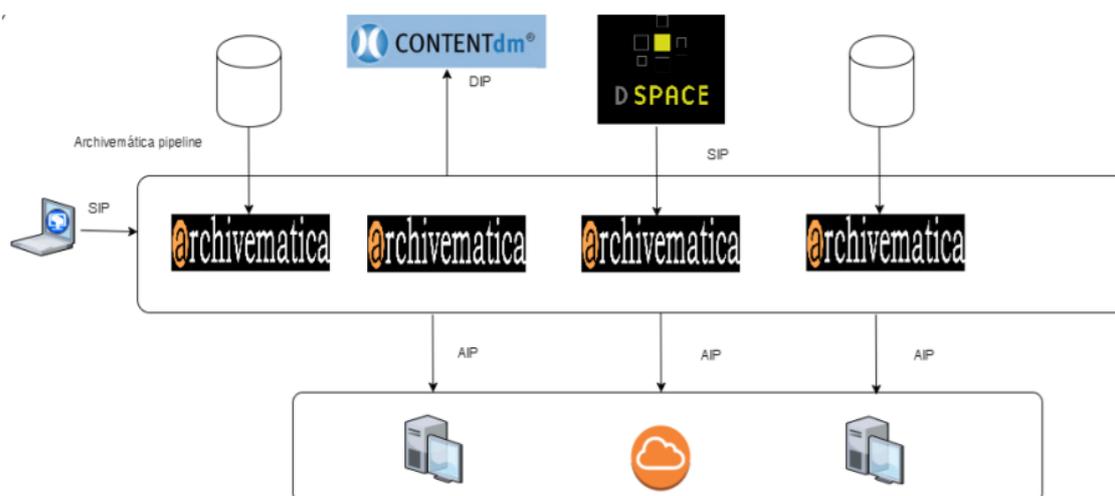
O AtoM é um sistema baseado na *web* para acesso aos seus objetos digitais, reconhecido na área de Arquivologia como uma plataforma de descrição, difusão e acesso (IBICT, 2016, p.15).

5.2 O Archimática e o DSpace

O Archimática também possui integração com diversos sistemas terceiros¹⁵, tais como: DSpace, LOCKSS, DuraCloud, OpenStack, Arkivum e das ferramentas de acesso: CONTENTdm, AtoM, Archivists' Toolkit, ArchivesSpace e Islandora.

Na figura 4 temos um exemplo de utilização de integração entre as plataformas.

Figura 4: modelo de implementação Archimática com sistemas terceiros



Fonte: Sprout y Romkey, 2016, *apud* DE GIUSTI, 2018, p. 279.

No caso da utilização do Archimática com o DSpace cabe destacar que apesar de o mesmo ser considerado um repositório. O tipo de transferência que o DSpace permite é receber exportação de um repositório oriundo do DSpace. Para a funcionalidade do DSpace junto ao Archimática, este é utilizado como um recurso para preservação de objetos digitais, deixando para o DSpace a função de acesso (LAMPERT, 2016, p.150).

5.3 BINDER

Na ausência de repositórios confiáveis na área da Museologia, o Museum of Modern Art - MoMA de Nova York utilizou o Archimática como base para construção de um repositório e ferramenta de acesso a seu acervo. Segundo o professor Arellano (2017, p. 154) junto à iniciativa

¹⁵ Para mais informações verificar site: <https://www.archivematica.org/pt-br/>

com a empresa *Artefactual Systems*, o MoMA desenvolveu uma aplicação *web* em *software* livre, chamada Binder¹⁶.

Desta forma no MoMA a estratégia de preservação de objetos digitais é facilitada pelo empacotamento de arquivos para fins de armazenamento e indexação. Essa primeira etapa de preservação de objetos digitais é fundamental para a identificação e a recuperação futura dos acervos (ARELLANO, 2017, p. 153 – 154).

De acordo com Arellano (2017, p. 154) em 2015 “*com as soluções tecnológicas do Binder, do Archimática e do Arkivum, o MoMA desenhou a arquitetura do seu Repositório Digital para Coleções de Museus – DRCM projetado para gerenciar 6,2 petabytes em 2025*”.

O Binder não é recomendado para documentos arquivísticos, mas é um exemplo do uso do Archimática em outras áreas aliando a outras funcionalidades e especificidades. De modo a destacar a importância do aproveitamento do uso de ferramentas consolidadas na área e do aproveitamento da expertise de empresas que trabalham com esta temática.

8 RODA

O Repositório de Objetos Digitais Autênticos – RODA foi desenvolvido pela Direção-Geral do Livro, dos Arquivos e das Bibliotecas – DGLAB, em Portugal é baseado em tecnologias *open source* e incorpora as normas OAIS, EAD, METS e PREMIS. Desta forma configura-se também como um repositório certificado segundo a norma ISO 16363:2012.

O projeto para a construção do repositório contou com a colaboração da Direção de Arquivos de Portugal – DGARQ e teve como suporte técnico (informática) a Universidade do Minho, desenvolvido pela empresa Keep Solutions¹⁷.

O RODA é um aplicativo que incorpora a funcionalidade exigida pelo OAIS que é um modelo de referência ao esquema conceitual, que disciplina e orienta sistemas de arquivos dedicados à preservação e difusão da informação digital a longo prazo (FLORES, 2016, p.35).

Como vantagens do uso do aplicativo temos funcionalidade compatível com os pressupostos de um RDC-Arq, tais como: autenticidade; integração com outros sistemas; suporte para múltiplos formatos. Além de ser auditável, os utilizadores do RODA têm obrigatoriedade de se registrar antes de poder utilizar o sistema.

Possui tecnologia evolutiva: através do mecanismo de *plugins* pode ser estendido sem recompilar ou reinstalar o sistema. A ferramenta dispõe de ações de preservação incorporadas:

¹⁶ Disponível em: <https://binder.readthedocs.io/en/latest/index.html#home>

¹⁷ Disponível em: <https://www.keep.pt/produtos/roda-repositorio-para-preservacao-de-informacao-digital/>

conversões de formatos, verificações de integridade, diagnóstico de riscos, ações de mitigação de riscos, produção de relatórios, verificação de vírus etc.

Isto posto com diversas vantagens e aplicações podemos concluir que o RODA é considerado um RDC-Arq usado para documentos arquivísticos.

9. DSpace

O DSpace¹⁸ Sistema para Construção de Repositórios Institucionais Digitais¹⁹ fez parte de um projeto colaborativo da *MIT Libraries*²⁰ e da empresa Hewlett-Packard Company- HP. Este é um dentre vários projetos, atualmente em operação, orientados à criação de repositórios institucionais e à preservação digital.

O DSpace é um *software* de código-fonte aberto que fornece facilidades para o gerenciamento de acervo digital, utilizado para implementação de repositórios institucionais.

O *software* foi desenvolvido para possibilitar a criação de repositórios digitais com funções de armazenamento, gerenciamento, preservação e visibilidade da produção intelectual, permitindo sua adoção por outras instituições em forma consorciada federada.

O sistema foi criado de forma a ser facilmente adaptado. Os repositórios DSpace permitem o gerenciamento da produção científica em qualquer tipo de material digital, dando-lhe maior visibilidade e garantindo a sua acessibilidade ao longo do tempo.

Possui natureza de preservar objetos digitais e é resultado da implementação de padrões internacionalmente aceitos, tais como OAI-PMH, OAIS e Dublin Core.

No entanto é importante salientar segundo Arellano (2008, p. 109) que mesmo *softwares* como Dspace que promovem longevidade dos documentos e acompanham o desenvolvimento na área da preservação digital, a motivação primeira não é a preservação de longo prazo.

Além disso, o processo descritivo diferente do Archimática e RODA não segue padrões arquivísticos. Através de um estudo comparativo entre as ferramentas DSpace e Archimática de acordo com Lampert (2016, p.151):

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.

¹⁸ Para mais informações verificar site: <http://www.ibict.br/tecnologias-para-informacao/dspace#apresentacao>

¹⁹ *Institutional Digital Repository System* – [Sistema de Repositório Digital Institucional].

²⁰ The libraries of the Massachusetts Institute of Technology.

Apesar de ter sido criado como repositório de armazenamento, ele não contempla necessariamente a fase de preservação para um documento arquivístico. O armazenamento de documentos arquivísticos não deve ser realizado neste repositório. Ele pode ser usado neste caso, como ferramenta de acesso, mas não de armazenamento.

De modo que se observou que o DSpace tem sido utilizado como ferramenta de armazenamento e/ou acesso principalmente para documentos bibliográficos, como repositório institucional e não como ferramenta de preservação, de modo que não pode ser caracterizado como um RDC-Arq.

10. Análise dos resultados

Não existe hoje um consenso sobre como se deve implementar uma estrutura adequada que permita assegurar a preservação digital. No entanto, existem boas práticas, sistemas informatizados e fluxos de trabalho que tem se mostrado bastante apropriados para boa parte das tarefas de preservação e que também resultam centrais para a evolução e auditoria de um repositório institucional (DE GIUSTI; GONZALO, 2018, p. 275, tradução nossa).

Assim como qualquer sistema, mesmo de código aberto, a implantação de um RDC-Arq requer apoio institucional, técnico e operacional de profissionais da área de tecnologia. A ferramenta, também, requer um período de teste de aplicabilidade para a operação e definição de uso, ou não, pela instituição e deve estar inserido em uma Política de Preservação Digital. Além da PPD, o uso de RDC-Arq se mostra necessário, de acordo com Luz e Flores (2017, p. 171):

O ganho de aplicar requisitos arquivísticos num repositório é a possibilidade de proporcionar a autenticidade documental daquele fundo custodiado. Como um repositório pode ter a presunção de garantir a autenticidade dos documentos? Por meio da confiança. A cadeia de custódia é importante para manter a confiança desde a criação e uso dos documentos. Em relação ao arquivo permanente digital, que futuramente será o arquivo histórico digital, este repositório digital confiável é uma plataforma tecnológica que é capaz de manter autênticos os materiais digitais, de preservá-los e prover acesso a eles pelo tempo necessário.

Isto posto no quadro 1 é apresentado a sistematização das informações reunidas neste artigo com a comparação entre os repositórios Archivemática, RODA e Dspace. Para a sistematização das informações e elaboração do quadro considerou-se a finalidade de uso para armazenamento, mas principalmente de preservação e normas e padrões considerando o modelo OAIS para documentos arquivísticos.

Quadro 1: Análise entre os repositórios Archivemática, RODA e Dspace como RDC-Arq

SOFTWARE	USO	CARACTERÍSTICAS	NORMAS / PADRÕES (OAIS)	CÓDIGO ABERTO / SOFTWARE LIVRE	COMUNIDADE DE USUÁRIOS
Archivemática	Armazenamento de documento arquivístico a longo prazo (Preservação)	Repositório Arquivístico Digital Confiável	Compatível com o modelo OAIS	Código aberto; Software livre	Possui
RODA	Armazenamento de documento arquivístico a longo prazo (Preservação)	Repositório Arquivístico Digital Confiável	Compatível com o modelo OAIS	Código aberto; Software Livre	Não encontrado
Dspace	Armazenamento para disseminação de conteúdo digital (produção intelectual e acadêmica)	Repositório institucional	Segue aspectos do modelo OAIS	Software livre	Não encontrado

Fonte: elaborado pela autora (2022).

Através do quadro é possível verificar as principais características de cada ferramenta. Considera-se o Archivemática e o RODA como RDC-Arq para documentos arquivísticos. O DSpace não possui elementos suficientes para ser considerado um RDC-Arq, mas é uma boa ferramenta para ser usada como repositório institucional e/ou para armazenamento de documentos bibliográficos. Também pode ser usado como ferramenta de acesso junto a outro RDC-Arq.

Considerações finais

Ao longo da pesquisa observou-se que existe um maior número de Sistemas Informatizados para a Gestão Arquivística de Documento - SIGADs e ferramentas de acesso, do que de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq devido às suas especificidades e complexidades. No entanto, o RDC-Arq deve estar contemplado dentro de uma cadeia de custódia e deve conversar com qualquer SIGAD que obedeça aos preceitos arquivísticos, assim como as ferramentas de acesso.

Ao longo deste artigo foi apresentado a importância da preservação digital a longo prazo por meio da utilização de repositórios, dentro de uma cadeia de custódia e para isso indicou-se opções disponíveis e documentadas de repositórios encontrados hoje na área para armazenamento e preservação. Objetivou-se demonstrar se estes repositórios contemplam os requisitos preconizados por um Repositório Arquivístico Digital Confiável - RDC-Arq

apresentando repositórios que usam padrões internacionalmente aceitos, tais como OAI-PMH, OAIS e Dublin Core.

Assim tentou-se elencar, dentre os exemplos de repositórios estudados, a sua viabilidade em termos de armazenamento a longo prazo, em detrimento do acesso, assim como à preservação de documentos arquivísticos, em detrimento à documentos bibliográficos.

Dos repositórios apresentados alguns possuíam mais informações e estudos de casos disponíveis no Brasil, do que outros, como no caso do Archivemática e Dspace, em detrimento do RODA.

O Dspace é um repositório que de alguma maneira segue o modelo OAIS, e outros padrões arquivísticos como OAI-PMH, mas não pode ser considerado um RDC-Arq, pois não contempla a preservação digital a longo prazo.

O RODA e o Archivemática podem ser considerados RDC-Arq, pois contemplam os requisitos preconizados objetivando a preservação digital a longo prazo.

O Archivemática se mostrou mais pertinente a realidade das instituições arquivísticas brasileiras, até o momento, considerando que instituições públicas e privadas que têm utilizado a ferramenta, também possui guia de instalação e configuração pelo IBICT, além de interface com a ferramenta AtoM importante base de acesso utilizada difundida e no país por diversas instituições²¹.

Este é um estudo em processo, considerando as circunstâncias, limitações e conhecimentos obtidos até o presente momento. De modo que não pretende oferecer conclusões definitivas sobre a melhor implementação de um repositório. O estudo que resulta este artigo teve como objetivo mostrar caminhos, relatar experiências e trazer as discussões ocorridas atualmente para o diálogo com as instâncias decisórias. De modo que é importante que institucionalmente o horizonte de implementação de um RDC-Arq seja factível, junto a implementação de uma Política de Preservação Digital e que pesquisas sobre o encaminhamento de implantação de repositórios seja amplamente estudada e divulgada.

Referências

ARCHIVEMÁTICA. Apresenta informações sobre o software. Disponível em: <<https://www.archivematica.org/pt-br/>>. Acesso em: 8 jul. 2022.

ALVES, R. C. V. Metadados para representação e recuperação da informação em ambiente web. In: MARINGELLI, Isabel Cristina Ayres da Silva. (Org.). In: SEMINÁRIO SERVIÇOS DE INFORMAÇÃO

²¹ Um *site* do AtoM no Brasil foi criado com todas as informações sobre a ferramenta, incluindo instituições que tem usado no mundo: <https://observatoriodocume.wixsite.com/atomnobrasil>

EM MUSEUS: INFORMAÇÃO DIGITAL COMO PATRIMÔNIO CULTURAL, 4. São Paulo: Pinacoteca de São Paulo, 2017, p. 95-106.

ARQUIVO NACIONAL (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005.

CANELHAS, T. *et al.*. **Arquivemática: instalação, integração e workflow**. Palestra. Material elaborado para a Palestra no Curso Práticas de preservação digital. Brasília – DF. 2019.

CONSELHO NACIONAL DE ARQUIVOS – CONARQ. **Diretrizes para a presunção de autenticidade de documentos arquivísticos digitais**. Rio de Janeiro, 2012. 10 p. Disponível em: <https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_presuncao_autenticidade_completa.pdf> Acesso em: 08 jul. 2022.

CONSELHO NACIONAL DE ARQUIVOS – CONARQ. **Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis: RDC-Arq**. Rio de Janeiro: CONARQ-CTDE, 2015. 31p. Disponível em: <https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_diretrizes_rdc_arq_resolucao_43.pdf>. Acesso em: 06 jul. 2022.

CONSELHO NACIONAL DE ARQUIVOS – CONARQ. **Glossário: Documentos Arquivísticos Digitais**. Versão 8.0. Rio de Janeiro: Arquivo Nacional. 2020. 62 p. Disponível em: <https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf>. Acesso em: 10 jul. 2022.

DE GIUSTI, M. R.; LUJÁN VILLARREAL, G. Revisión de distintas implementaciones para preservación digital: hacia una propuesta metodológica para la preservación y la auditoría de confiabilidad de repositorios institucionales. In: **Revista Digital Biblioteconomia e Ciência da Informação**, Campinas. v. 16, p. 273 – 292, mai./ago. 2018. Disponível em: <<https://brapci.inf.br/index.php/res/download/114216>>. Acesso em: 30 jun. 2022.

FERREIRA, M. **Introdução à preservação digital – Conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 88p.

FLORES, D. **Cadeia de custódia dos documentos arquivísticos digitais: do SIGAD ao RDC-Arq**. Palestra. Material elaborado para a Palestra no Curso de Arquivologia. Vitória, ES: Universidade Federal do Espírito Santo, 2016. Disponível em: [http://www.arquivologia.ufes.br/sites/arquivologia.ufes.br/files/field/anexo/cadeia_de_custodia_dos_documentos_arquivisticos_digitais- do_sigad_ao_rdc-arq.pdf](http://www.arquivologia.ufes.br/sites/arquivologia.ufes.br/files/field/anexo/cadeia_de_custodia_dos_documentos_arquivisticos_digitais-_do_sigad_ao_rdc-arq.pdf). Acesso em: 16 jun. 2022.

GAVA, T. B. S.; FLORES, D. Repositórios Arquivísticos Digitais Confiáveis (RDC-ARQ) como plataforma de preservação digital em um ambiente de gestão arquivística. **Informação & Informação**, Londrina, v. 25, n. 2, p. 74-99, abr./jun. 2020.

GRÁCIO, J. C. A., *et al.* Modelo para elaboração de políticas de preservação digital de documentos de arquivo por instituições de ensino superior: o caso da Unesp. **Reciis – Revista Eletrônica Comunicação Informação & Inovação em Saúde**, Rio de Janeiro, v. 14, n. 3, p. 563-579, jul./set. 2020. Disponível em: <<https://www.reciis.ict.fiocruz.br/index.php/reciis/article/view/2111>>. Acesso em: 21 jun. 2022.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT. Milene Costa *et al.* **Guia do usuário Archivemática**. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2016.

LAMPERT, S. R. Os repositórios DSpace e Archivemática para documentos arquivísticos digitais. In: **Revista Acervo**: Rio de Janeiro, v. 29, n. 2, p. 143-154, jul./dez. 2016.

LUZ, C.; FLORES, D. Cadeia de Custódia e de preservação: autenticidade nas plataformas de gestão e preservação de documentos arquivísticos. In: MARINGELLI, I. C. A. S. (Org.). In: **SEMINÁRIO SERVIÇOS DE INFORMAÇÃO EM MUSEUS: INFORMAÇÃO DIGITAL COMO PATRIMÔNIO CULTURAL**, 4. São Paulo: Pinacoteca de São Paulo, 2017, p. 171-185.

MÁRDERO ARELLANO, M. Á. A preservação digital da documentação museológica. In: MARINGELLI, Isabel Cristina Ayres da Silva. (Org.). In: **SEMINÁRIO SERVIÇOS DE INFORMAÇÃO EM MUSEUS: INFORMAÇÃO DIGITAL COMO PATRIMÔNIO CULTURAL**, 4. São Paulo: Pinacoteca de São Paulo, 2017, p. 149-157.

MÁRDERO ARELLANO, M. Á. **Critérios para a preservação digital da informação científica**. 2008. 354f. Tese (Doutorado) – Universidade de Brasília, Departamento de Ciência da Informação, 2008. Disponível em: <https://repositorio.unb.br/bitstream/10482/1518/1/2008_MiguelAngelMarderoArellano.pdf>. Acesso em: 9 jul. 2022.

RODA. Apresenta informações sobre o software. Disponível em: <<https://roda.arquivos.pt/#welcome>>. Acesso em: 8 jul. 2022.