



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO-PPGCI
MESTRADO EM CIÊNCIA DA INFORMAÇÃO**

WELLINGTON DA SILVA GOMES

**CARACTERÍSTICAS, FUNÇÕES E PRESERVAÇÃO DIGITAL DO
ARCHIVEMATICA À LUZ DA RESOLUÇÃO Nº 43 DO CONSELHO NACIONAL DE
ARQUIVOS**

**João Pessoa - PB
2019**

WELLINGTON DA SILVA GOMES

**CARACTERÍSTICAS, FUNÇÕES E PRESERVAÇÃO DIGITAL DO
ARCHIVEMATICA À LUZ DA RESOLUÇÃO Nº 43 DO CONSELHO NACIONAL DE
ARQUIVOS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, na linha de Pesquisa Ética, Gestão e Políticas de Informação, como requisito à obtenção do título de Mestre em Ciência da Informação.

Orientadora: Professora Doutora Marynice de Medeiros Matos Autran

**João Pessoa - PB
2019**

Catálogo na publicação
Seção de Catalogação e Classificação

G633c Gomes, Wellington da Silva.
Características, funções e preservação digital do
Archivematica à luz da Resolução nº 43 do Conselho
Nacional de Arquivos / Wellington da Silva Gomes. -
João Pessoa, 2019.
146 f. : il.

Orientação: Marynice de Medeiros Matos Autran.
Dissertação (Mestrado) - UFPB/CCSA.

1. Archivematica. 2. Repositório Digital Arquivístico
Confiável. 3. Software livre. I. Autran, Marynice de
Medeiros Matos. II. Título.

UFPB/BC

WELLINGTON DA SILVA GOMES

CARACTERÍSTICAS, FUNÇÕES E PRESERVAÇÃO DIGITAL DO
ARCHIVEMÁTICA À LUZ DA RESOLUÇÃO Nº 43 DO CONSELHO NACIONAL
DE ARQUIVOS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, na linha de Pesquisa Ética, Gestão e Políticas de Informação, como requisito à obtenção do título de Mestre em Ciência da Informação, defendido em 2019 e apresentado à banca examinadora constituída dos seguintes professores:



Professora Doutora Marynice de Medeiros Matos Autran - UFPB
Orientadora/Presidente



Professor Doutor Marckson Roberto Ferreira de Sousa – UFPB
Membro Interno Titular



Professora Doutora Manuela Eugênio Maia – UEPB
Membro Externo



Professora Doutora Rosa Zuleide de Brito – UFPB
Membro Externo Suplente

À minha mãe e meus verdadeiros amigos, pela
força, companheirismo e perseverança nesta
trajetória acadêmica, DEDICO.

AGRADECIMENTOS

Primeiramente a Deus, por me conceder o dom da vida e acompanhamento nessa trajetória.

À professora Marynice de Medeiros de Matos Autran pelas leituras sugeridas ao longo dessa orientação com dedicação. E também, por toda a paciência durante o percurso até aqui.

À minha mãe pelo companheirismo, força e motivação de querer alçar, sempre, maiores voos.

A todos do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, que contribuíram ao longo deste mestrado.

À Professora Manuela Eugênio Maia, por continuar me auxiliando com discernimento na academia.

Ao amigo Marcelo Dias Perrayon, pela permanência e força no acompanhamento das longas noites de escrita da dissertação.

À amiga Rebeca de Oliveira Patrício, por estar junto nas disciplinas, discussões e vida neste roteiro acadêmico.

Ao amigo Evllon Évrig Sá Sousa, por sua amizade e companheirismo em todos os momentos.

Ao amigo Jefferson Higinio da Silva, por sua parceria acadêmica.

A cultura fornece um enorme equipamento cognitivo aos indivíduos. A cada etapa de nossa trajetória social, a coletividade nos fornece línguas, sistemas de classificação, conceitos, analogias, metáforas, imagens, evitando que tenhamos que inventá-las por conta própria. As regras jurídicas ou administrativas, a divisão do trabalho, a estrutura hierárquica das grandes organizações e suas normas de ação são tipos de memória, de raciocínio e de tomada de decisão automáticas, incorporadas à máquina social e que economizam certa quantidade e atividade intelectual dos indivíduos. Uma vez que são convencionais e historicamente datadas, é claro que as tecnologias intelectuais (a escrita, a informática...) são instituições. (LÉVY, 2008, p. 87).

RESUMO

A emergência da Tecnologia da Informação e Comunicação (TIC) fez surgir um novo paradigma na sociedade Pós-industrial: o da evolução tecnológica sobre a informação. Isto acarretou a dinamização dos meios de comunicação instrumentalizados pela informática e Internet, especialmente, na difusão do que se conhece como o *ciberespaço*. Tais alterações também tiveram impacto nos campos científicos que têm como práxis a relação entre a análise e objeto: a informação, passando a se adaptarem às transformações causadas pelo progresso tecnológico, produzindo instrumentos para racionalizar e controlar o aumento dos seus registros. Destarte, a transição dos documentos analógicos para os digitais, prova incontestemente que as TICs vêm possibilitando diferentes formas de interação entre os indivíduos, tem alterado a forma de pensar o gerenciamento da informação no ambiente virtual, sobretudo, pela intermediação do hardware e software. Dessa forma, percebe-se, que os softwares livres possuem papel essencial no armazenamento e preservação da informação, dando ênfase ao *Archivematica*. Para adentrar os aspectos de funcionalidades e características do referido sistema que possuem oito guias responsáveis pelo processo de pacotes de informações (SIP, AIP, DIP), tomou-se como ponto de análise à Resolução nº 43 do CONARQ, a qual estabelece diretrizes para implementação do chamado Repositório Digital Arquivístico Confiável (RDC-Arq). Assim, esta pesquisa teve como objetivo avaliar o *Archivematica* a partir dos elementos da Resolução nº 43. Como método, apresentou caráter descritivo, documental e qualitativo, no viés comparativo, confrontando as incumbências do software em avaliação aos requisitos de confiabilidade postados da referida normativa. Os resultados demonstram o total cumprimento dos requisitos preconizados pela Resolução nº 43, o que respalda o *Archivematica* como um repositório digital arquivístico confiável. Conclui-se, então, que todos os elementos inseridos na citada diretiva foram englobados a tal software, o caracterizando como um RDC-Arq.

Palavras-chave: *Archivematica*. Repositório Digital Arquivístico Confiável. Software livre.

ABSTRACT

The emergence of Information and Communication Technology (TIC) has given rise to a new paradigm in post-industrial society: that of technological evolution over information. This has led to the dynamization of the media used by computers and the Internet, especially in the diffusion of what is known as cyberspace. These changes also had an impact in the scientific fields that have as praxis the relation between the analysis and object: the information, adapting to the transformations caused by the technological progress, producing instruments to rationalize and control the increase of its registers. Thus, the transition from analog to digital documents proves unquestionably that TICs have enabled different forms of interaction between individuals, has altered the way of thinking information management in the virtual environment, mainly through the intermediation of hardware and software. In this way, free software plays an essential role in the storage and preservation of information, emphasizing Archivemática. In order to penetrate the aspects of functionalities and characteristics of said system that have eight guides responsible for the information packet process (SIP, AIP, DIP), it was taken as a point of analysis to Resolution n° 43 of CONARQ, which establishes guidelines for implementation of the so-called Trusted Archival Digital Repository (RDC-Arq). Thus, this research had the objective of evaluating Archivemática from the elements of Resolution n° 43. As a method, it presented a descriptive, documentary and qualitative character in comparative bias, confronting the tasks of the software in evaluation to the reliability requirements posted of the mentioned regulation. The results demonstrate the full compliance with the requirements of Resolution 43, which supports Archivemática as a Trusted Archival Digital Repository. It is concluded, then, that all the elements inserted in the mentioned directive were encompassed to such software, characterizing it as an RDC-Arq.

Keywords: Archivemática. Trusted Archival Digital Repository. Free software.

LISTA DE QUADROS

Quadro 1 - Cumprimento dos requisitos do <i>Archivematica</i> à luz da Resolução nº 43	23
Quadro 2 - Estrutura das Revoluções Científicas de Kuhn	25
Quadro 3 - Processo histórico relativo aos arquivos	34
Quadro 4 - Disciplinas e áreas interdisciplinares da CI	37
Quadro 5 - Tipos de documentos	38
Quadro 6 - Estratégias de preservação digital	44
Quadro 7 - Síntese da tipologia dos repositórios digitais	53
Quadro 8 - Classes dos metadados	56
Quadro 9 - Tipos de repositórios de acordo com Resolução de nº 43	65
Quadro 10 - Tipos de <i>software</i>	70
Quadro 11 - Principais características do <i>Archivematica</i>	75
Quadro 12 - Cumprimento dos requisitos do <i>Archivematica</i> à luz da Resolução nº 43 do CONARQ (1ª seção)	80
Quadro 13 - Cumprimento dos requisitos do <i>Archivematica</i> à luz da Resolução nº 43 do CONARQ (2ª seção)	84
Quadro 14 - Cumprimento dos requisitos do <i>Archivematica</i> à luz da Resolução nº 43 do CONARQ (3ª seção)	91
Quadro 15 - Definições dos elementos do Plano de Preservação <i>Archivematica</i>	140

LISTA DE FIGURAS

Figura 1 - Estrutura dos objetos digitais	43
Figura 2 - Os dez mandamentos da preservação digital	43
Figura 3 - Modelo do ambiente OAIS/SAAI	48
Figura 4 - Entidades funcionais do OAIS/SAAI	49
Figura 5 - Integrantes do SINAR/CONARQ	61
Figura 6 - Intersecção dos repositórios digitais	65
Figura 7 - Referências para desenvolvimento de um RDC-Arq	66
Figura 8 – Página inicial do <i>Archivematica</i>	73
Figura 9 – Página <i>wiki</i> do <i>Archivematica</i>	73
Figura 10 - Ciclo de vida dos documentos em cadeia de custódia ininterrupta	77
Figura 11 - Ambiente de preservação e acesso de uma cadeia de custódia	78
Figura 12 - Fluxograma dos pacotes de informação no <i>Archivematica</i>	79
Figura 13 - Painel baseado na <i>web</i> do <i>Archivematica</i>	116
Figura 14 - Campos de preenchimento da guia Transferência	118
Figura 15 - Processo de criação da Transferência	120
Figura 16 - Exemplo de lista de transferência no <i>Backlog</i>	123
Figura 17 - Exclusão de transferências no <i>Backlog</i>	124
Figura 18 - Transferências do <i>Backlog</i> na guia de Avaliação	125
Figura 19 - Adição de <i>Tags</i>	126
Figura 20 - Painel de análise de Objetos (<i>Objects</i>)	127
Figura 21 - Guia <i>Tags</i>	127
Figura 22 - Guia Examinar conteúdo (<i>Examine contents</i>)	128
Figura 23 - Guia Lista de arquivo (<i>File list</i>)	129

Figura 24 - Painel Archivespace	130
Figura 25 - Guia Ingerir (<i>Ingest</i>)	131
Figura 26 - Organizando um SIP na guia Ingerir	132
Figura 27 - Etapa de normalização	133
Figura 28 - Execução dos microsserviços na guia Ingerir	134
Figura 29 - Guia Armazenamento Arquivístico	135
Figura 30 - Exclusão de um AIP	136
Figura 31 - Guia Acesso	137
Figura 32 - Página do <i>upload</i> DIP/AtoM	138
Figura 33 - Guia Plano de Preservação	141
Figura 34 - Regras de políticas de formato	142
Figura 35 - Guia Administração	144

LISTA DE SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AIP	Pacote de Informação de Arquivo
ANCIB	Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
ARIST	<i>Annual Review for Information Science and Technology</i>
BOAI	<i>Budapest Open Access Initiative</i>
CCSDS	Consultative Committee for Space Data Systems
CI	Ciência da Informação
CONARQ	Conselho Nacional de Arquivos
COPPUL	Council of Prairie and Pacific University Libraries Acesso to Memory
CPU	<i>Central Process Unit</i>
CSV	Valores Separados por Vírgula
CTDE	Comissão Técnica de Documentos Eletrônicos
DIAS	<i>Digital Information Archive System</i>
DIP	Pacotes de Informação e Disseminação
EAD	<i>Encoding Archival Description</i>
EDC	<i>Environmental Data Catalogue</i>
FDGC	Federal Geographic Data committee
FMI	Fundo Monetário Internacional
FPR	<i>Format Policy Registry</i>
FSF	Free Software Foundation
GILS	<i>Government Information Locator Service</i>
GNU	<i>General Public License</i>
GNU 3.0	<i>Affero General Public License</i>
GT	Grupos de Trabalho
IAFA	<i>Internet Anonymous Ftp Archive</i>
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia

ICA	Conselho Internacional de Arquivos
ISAD(G)	Norma Geral Internacional de Descrição Arquivística
ISO	Organization for Standardization
LOCKSS	<i>Lots Of Copies Keep Stuff Safe</i>
MARC	<i>Machine Readable Catalogue</i>
METS	<i>Metadata Encoding Transmission Protocol</i>
MIT	Massachusetts Institute of Technology
NARA	National Archives and Record Administration
NARS	National Archives and Records Service
NOBRADE	Norma Brasileira de Descrição Arquivística
OA	<i>Open Access</i>
OAI	<i>Open Archive Initiative</i>
OAI-PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
OAIS	<i>Reference Model for an Open Archival</i>
OCLC	Online Computer Library Center
PHM	<i>Protocol for Metadata Harvesting</i>
PII	<i>Personally Identifiable Information</i>
PREMIS	<i>Preservation Metadata: Implementation Strategies</i>
RDC-Arq	Repositório Arquivístico Digital Confiável
RLG	Research Library Group
RLG/OCLC	Research Library Group/Online Computer Library Center
RODA	Repositório de Objetos Digitais Autênticos
SAAI	Sistema Aberto de Arquivamento de Informação
SDUM	Serviços de Documentação da Universidade do Minho
SIGAD	Sistema Informatizado de Gestão Arquivística de Documentos
SINAR	Sistema Nacional de Arquivos
SIP	Pacote de Submissão de Informação
SOIF	<i>Summary Object Interchange Format</i>
SPARC	Scholarly Publishing and Academic Resources Coalition

TEI	<i>Text Encoding Initiative</i>
TI	Tecnologia da Informação
TIC	Tecnologia da Informação e Comunicação
TRAC	<i>Trustworthy Repository Audit & Certification: Criteria and Checklist</i>
UFSM	Universidade Federal de Santa Maria
UK	United Kingdom
UMIC	Unidade de Missão Inovação e Conhecimento
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura
UUID	<i>Universally Unique Identifier</i>
XML	<i>eXtensible Markup Language</i>

SUMÁRIO

1	INTRODUÇÃO	16
2	PERCURSO METODOLÓGICO	20
2.1	OBJETO DE ESTUDO	21
2.2	MÉTODO UTILIZADO	22
3	O CAMPO CIÊNCIA DA INFORMAÇÃO: OBJETO E DISCIPLINA	25
4	TRAJETÓRIA DE UMA ÁREA DOCUMENTAL: EVOLUÇÃO DOS ARQUIVOS E A FORMAÇÃO DA ARQUIVOLOGIA	29
4.1	RELAÇÕES INTERDISCIPLINARES: ARQUIVOLOGIA NO PERÍODO PÓS-MODERNO E A CIÊNCIA DA INFORMAÇÃO	33
4.2	ARQUIVÍSTICA E A TECNOLOGIA DIGITAL	38
5	PRESERVAÇÃO DIGITAL: CONCEITO E ESTRUTURA	41
5.1	MODELO DE REFERÊNCIA OAIS/SAAI: NORMA ISO 14721:2003/15427:2007	46
5.2	REPOSITÓRIOS DIGITAIS: ANTESCEDENTES E CARACTERIZAÇÃO	50
5.2.1	Repositórios Digitais Arquivísticos	56
6	DIRETRIZES PARA IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS (RDC-Arq): RESOLUÇÃO N° 43	60
7	SOFTWARES LIVRES	68
8	SOFTWARE <i>ARCHIVEMATICA</i>	71
8.1	ESPECIFICIDADES E CARACTERÍSTICAS DO <i>ARCHIVEMATICA</i>	76
9	RESULTADOS E DISCUSSÃO	80
10	CONSIDERAÇÕES FINAIS	95
	REFERÊNCIAS	98
	ANEXO A – RESOLUÇÃO N° 43, DE 4 DE SETEMBRO DE 2015	106
	APÊNDICE A – REQUISITOS PARA UM REPOSITÓRIO DIGITAL CONFIÁVEL	109
	APÊNDICE B – FUNCIONALIDADES DO <i>ARCHIVEMATICA</i>	115

1 INTRODUÇÃO

Os campos científicos que têm como práxis a relação estudo e objeto: informação, passaram a se adaptar às transformações causadas pelo progresso tecnológico, produzindo instrumentos para racionalizar e controlar o aumento dos seus registros. A ciência em sua contemporaneidade tornou-se a principal forma de conhecimento, fundando a base para que as tecnologias digitais transformassem os meios de produção, e por isso deve ser vista como uma estrutura que atende às necessidades de organização da sociedade, seja em seu sistema de comunicação, econômico ou tecnológico (FREIRE; FREIRE, 2009).

Na atualidade, vivencia-se a transição dos documentos analógicos para os digitais, prova incontestável de que a Tecnologia da Informação e Comunicação (TIC) vem possibilitando diferentes formas de interação entre os indivíduos, graças à intermediação de hardware e software, representados pela máquina/computador e pelos programas lógicos, responsáveis por gerenciar informação no ambiente virtual.

As formas de pensar e conviver estão sendo reformuladas nessa era das tecnologias da inteligência, transformando a relação homem/trabalho, pois a própria compreensão do viver no mundo atual depende da metamorfose dos dispositivos de informação em suas diferentes tipologias: escrita, leitura, criação e aprendizagem que são capturados por uma modernização cada vez mais complexa (LÈVY, 2008).

Observa-se que se encontra em processo de consolidação a utilização do documento digital em grande parte das atividades humanas, tendo em vista as vantagens que sua empregabilidade proporciona na produção, organização, acesso, armazenamento e uso da informação.

Essas mudanças alteram, também, os métodos aplicados aos arquivos, bibliotecas e museus, demandando dos profissionais a ampliação de suas competências e habilidades na utilização dos dispositivos eletrônicos para a realização de suas atividades laborais. Faz-se imperativo que os especialistas nas áreas mencionadas, dominem as funcionalidades e os sistemas de gerenciamento da informação digital frente às suas demandas.

A inserção das TICs possibilitou o surgimento de novas técnicas aplicadas aos diversos domínios científicos, dentre esses a Arquivologia, Biblioteconomia e Ciência da Informação (CI). Nesse caso, grandes avanços ocorreram, principalmente no que se refere aos processos de organização, representação, gestão da informação e gestão documental, especialmente no que tange aos sistemas informatizados. A utilização desses recursos encontra eco nos softwares livres, entendidos como programas de computador que podem ser executados,

copiados, customizados ou redistribuídos.

Para a área da Arquivologia, especificamente, existe um número considerável desse tipo ferramenta disponível no mercado, tais como aqueles desenvolvidos para administração de documentos, responsáveis por abarcar todo seu ciclo, desde sua primeira fase (corrente e intermediário) a segunda (permanente). Em relação à primeira citam-se, por exemplo, o Agorum Core, ArchivistaBox, Alfresco e Nuxeo DM, tendo como base os preceitos do e-Arq Brasil, pois são ferramentas identificadas como sistemas de gestão documental). No que tange a segunda, identifica-se o Access to Memory (AtoM), que tem como função a descrição arquivística, baseada na web, para acesso a objetos digitais em uma plataforma de difusão da informação

Atenta-se para os softwares de repositório digital: DSpace, Fedora, Repositório de Objetos Digitais Autênticos (RODA) e *Archivematica*. Optou-se por este último, uma vez que possui características de Repositório Arquivístico Digital Confiável (RDC-Arq) para preservação da informação na web, no qual instituições podem preservar a autenticidade e integridade de seus registros digitais em longo prazo. Sua composição é ligada a diretrizes e normas, a saber: o *Reference Model for an Open Archival* (OAIS), *Trustworthy Repository Audit & Certification: Criteria and Checklist* (TRAC) e a Organization for Standardization (ISSO) 16363:2012, existindo tutorial traduzido e adaptado à realidade brasileira por meio da Resolução nº 43 do Conselho Nacional de Arquivos (CONARQ). (Ver Anexo A).

Nesta pesquisa, aprofundou-se os estudos sobre o referido software a respeito a seus aspectos de funcionalidades (ver Apêndice B), especificidades e características, e sobre a Resolução nº 43 do CONARQ, que estabelece diretrizes para implementação de repositórios digitais confiáveis. (Ver Apêndice A).

Ao manifestar a preocupação sob a perspectiva arquivística, para conhecer pormenorizadamente os meandros de um repositório digital confiável e o atendimento aos requisitos da supracitada Resolução, colocou-se a seguinte questão: *as características e funções do software livre Archivematica para repositórios arquivísticos digitais confiáveis atende aos postulados preconizados pela Resolução nº 43 do CONARQ?*

De acordo com essa pergunta, elegeram-se os seguintes objetivos, os quais foram analisados à luz dos preceitos estabelecidos pela citada resolução: avaliar o *Archivematica* a partir das perspectivas arquivísticas dos requisitos da Resolução nº 43 do CONARQ. E para responder a essa finalidade, determinou-se: analisar as funcionalidades do *Archivematica* de acordo com o *Guia do usuário* e sua página eletrônica; caracterizar o referido sistema; verificar se suas funções preenchem os requisitos preconizados pela Resolução nº 43.

Ressalta-se que o objeto e fenômeno informação é ponto basilar de investigação entre CI e Arquivologia, uma vez que tratam desta propriedade como elemento constituinte para suas atuações. Silva e Ribeiro (2008) afirmam que as duas áreas, desde a sua constituição, tiveram por essência a preocupação com a análise do conhecimento gerado pela sociedade. Logo, a CI é uma disciplina que investiga as propriedades do comportamento da informação e seu fluxo para obter acesso (BORKO, 1968). É uma ciência interdisciplinar relacionada a diversos campos, como Computação, Biblioteconomia, Tecnologia, e, especialmente, a Arquivologia, disciplina cujos princípios, conceitos e técnicas perpassam a produção, organização, guarda, preservação e uso da informação registrada orgânica (RONDINELLI, 2013).

Destaca-se que a escolha da Resolução nº 43 não ocorreu aleatoriamente. Identificou-se 27 publicações técnicas no site do CONARQ acerca de orientações, dentro dos preceitos arquivísticos, bem como recomendações para criação de arquivos e sugestões de softwares. Dessa forma, a opção por esta normativa se deu por envolver as necessidades referentes à confiabilidade da informação de repositórios digitais.

Esta pesquisa, no âmbito da Arquivologia, justifica-se por envolver uma ferramenta normativa e de dimensão tecnológica sob a ótica dos sistemas de informação digital, os quais possibilitam a eficiência, o acesso, difusão e confiança da informação, no intuito de dirimir os aspectos que envolvem sua insegurança.

Na CI, torna-se importante no que diz respeito aos aspectos fundamentais da estruturação, organização e representação da informação em sistemas de informação, tendo como respaldo seu escopo teórico acerca de esquemas, características e gerenciamento informacional no meio digital, assim como nos preceitos arquivísticos, destacando sua natureza interdisciplinar.

Este trabalho está estruturado em dez seções, além das referências, anexo e apêndices. Assim está disposto:

Na primeira seção, a Introdução, explana a amplitude do assunto, ou seja, os aspectos que levaram a escolha da temática; A segunda, o Percorso metodológico, delineando o objeto de estudo, utilizando o método comparativo para verificar as similitudes existentes; A terceira, o campo da CI, exibindo conceitos de seu objeto e definição de sua área enquanto disciplina científica. A quarta, a evolução e a formação da Arquivologia, com suas relações interdisciplinares frente à CI, na era digital e no contexto Pós-moderno. A quinta, explana a importância da Preservação digital e suas diretrizes, que dão vazão à confiabilidade da

informação no repositório, dispostas pelo modelo de referência OAIS/SAAI. Assim também, as características relacionadas aos repositórios digitais e repositórios digitais arquivísticos.

Na sexta seção, mostra então, a Resolução nº 43 do CONARQ (objeto de nosso estudo), a qual estabelece diretrizes para implementação de repositórios arquivísticos digitais confiáveis, através de elementos que dão segurança e proteção à informação neste tipo de sistema. A seção sétima manifesta o conceito e importância dos softwares livres, elencando seus tipos. A oitava, também do objeto deste estudo (o *Archivematica*), descreve sua criação, especificidades e características. A nona seção, se encarrega dos Resultados e discussões, cuja atribuição justapõe a completude do objeto aqui estudado, então, o *Archivematica* à luz da Resolução nº 43, empregado por quadros na análise de suas verossimilhanças. Na décima estão as considerações finais, assim, imbuído dos dizeres e sugestões finais do trabalho como todo.

Por fim, foram postos o Anexo e Apêndices, no qual o primeiro relata toda escrita da Resolução nº 43 (ver Anexo A). No Apêndice A, o quadro com todos elementos de análise ao *Archivematica* retirado da alusiva diretiva. E no Apêndice B, as funcionalidades e características do *Archivematica*, usado também para o exame e comparação dos resultados.

2 PERCURSO METODOLÓGICO

A preocupação com o conhecimento sobre a realidade sempre foi uma questão de inquietação nas ciências. Explicar e entender os fenômenos naturais e sociais tem sido um desafio constante para os pesquisadores, que buscam gradativamente a solução para os problemas globais e locais. Esses problemas, quando resolvidos, geram outras lacunas, outros questionamentos, e assim ocorre o desenvolvimento científico, baseado na quebra de paradigmas (KUHN, 2017).

O ato de pesquisar constitui a busca pelo conhecimento, traçando um caminho de ação, de pensamento e, conseqüentemente, do que se compreende sobre uma determinada realidade. Para se chegar a um fim, a metodologia é o caminho que se deve percorrer, ao mesmo tempo que se faz imprescindível uma teoria de abordagem (o método) dos instrumentos de operacionalização (técnicas) e da criatividade do pesquisador (sua experiência, capacidade pessoal e sensibilidade) (DESLANDES, 2009; GOMES, 2004).

Em contrapartida, essa organização ou sistematização de procedimentos significa muito mais que técnicas, pois inclui concepções, articulações e pensamentos teóricos sobre a realidade em que se vive, ancorando-se em uma das atividades básicas do fazer científico: a pesquisa. Por isso, pesquisar tanto é o procedimento chave para o conhecimento, quanto de aprendizagem, sendo assim parte integrante de todo o processo de análise de um determinado objeto, na busca ou descoberta de novos fatos ou dados, vinculando pensamento e ação (DEMO, 2000; MINAYO, 2011).

O objeto que concentra as Ciências Sociais é histórico, isso significa que cada sociedade existe e se constrói em um determinado espaço, se organizando de maneira particular, uma diferente da outra (DESLANDES, 2009; GOMES, 2004). Portanto, é a provisoriedade, o dinamismo e a especificidade de qualquer questão social que torna seu traço fenomenológico, ou seja, um fato consideravelmente observável e passível de ser estudado, pois detém elementos de crises, reflexos e complexidade no seu desenvolvimento.

2.1 O OBJETO DE ESTUDO

A questão central que norteia a condução deste estudo está diretamente relacionada com o software livre *Archivematica* (suas características e funcionalidades) avaliado à luz dos requisitos preconizados pela Resolução de nº 43 do CONARQ.

Tal Resolução está dividida em três (3) partes:

1. Apresentação: objetivo; escopo e definições;
2. Repositório Arquivístico Digital Confiável (RDC-Arq) - principais requisitos: 1 Considerações sobre um repositório digital de documentos arquivísticos 2. Requisitos para um repositório digital confiável;
3. Padrões e normas de referência.

A pesquisa se debruçou sobre a segunda parte, a qual se refere aos principais requisitos para um repositório digital.

O CONARQ disponibiliza digitalmente esta normativa para download na seção de “publicações técnicas”, assim como outros manuais de recomendações arquivísticas. A versão analógica encontra-se disponível para comercialização.

A junção das características e funções do sistema *Archivematica* e os requisitos da Resolução nº 43 compõem a completude do objeto de estudo deste trabalho, quando se investigou se o referido software cumpre com as seguintes exigências:

1. Infraestrutura organizacional;
2. Gerenciamento do documento digital e
3. Tecnologia, infraestrutura técnica e segurança.

O primeiro grupo subdivide-se em: a) Governança e viabilidade organizacional b) Estrutura organizacional e de pessoal; c) Transparência de procedimentos e arcabouço político; d) Sustentabilidade financeira; e) Contratos, licenças e passivos.

O segundo em: a) Admissão (captura de documentos digitais); b) Admissão (criação do pacote de arquivamento); c) Planejamento da preservação; d) Armazenamento e preservação/manutenção do AIP Gerenciamento e informação e) Gerenciamento de acesso.

E o terceiro contempla: a) Infraestrutura de sistema; b) Tecnologias apropriadas: c) Segurança.

Cada elemento possui uma peculiaridade, como armazenagem da informação digital, preservação, interoperabilidade, recomendações de normas e padrões de metadados. Tais requisitos estão definidos em nível conceitual e devem ser cumpridos no desenvolvimento de um repositório digital considerado confiável. Reitera-se que esses elementos estão baseados

na norma ISO 16363:2012 e abrangem todos os tipos de materiais digitais, principalmente os documentos arquivísticos.

A segmentação das funções do *Archivematica* está elencada em oito guias ou painéis projetados(as) por uma interface na web direcionada à perspectiva do usuário final: Transferência (*Transfer*), Ingerir (*Ingest*), Armazenamento Arquivístico (*ArchivalStorage*), Planejamento de Preservação (*Preservation Planning*), Acesso (*Access*), Administração (*Administration*). Lista de pendências (*Backlog*) e Avaliação (*Appraisal*). Os seis primeiros estão dispostos no *Guia do Usuário Archivematica* na versão 1.6 e os dois últimos em seu site do oficial (versão 1.8) (ARTEFACTUAL, [201-a]); COSTA *et al*, 2016).

A pesquisa teve como fonte de análise o *Guia do Usuário* e o site¹ do sistema. Destaca-se que esses registros estão legitimados por instituições de influência nacional e internacional na área da Ciência, Tecnologia e Educação, como por exemplo, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), refletido em publicações na área da TI-CI-Arquivologia.

2.2 MÉTODO UTILIZADO

O objetivo de conhecer e explicar fenômenos é a função significativa da pesquisa científica, pois percorre, desde a formulação do problema até a apresentação dos resultados (PRODANOV; FREITAS, 2013).

Desse modo, buscou-se em Marconi e Lakatos (2003) uma conceituação de método, o qual é assim definido pelas autoras: procedimento regular, explícito e passível de ser repetido para conseguir-se alguma coisa, seja em nível material ou conceitual. Nessa mesma linha de raciocínio, pode-se elencar também que o método é constituído de processos empregados na investigação e demonstração de uma realidade.

Na escolha por um método que conduzisse esta pesquisa, optou-se pelo comparativo, acostados em autores como Marconi e Lakatos (2003), as quais situam tal estratégia para fazer comparações de grupos ou variáveis, podendo ser utilizado nas diferentes etapas da pesquisa, ou seja: nos estudos descritivos pode analisar a analogia ou, ainda, os elementos de uma estrutura e apontar a ausência ou presença desses elementos. Justifica-se, portanto, a opção deste método para analisar o *Archivematica* à luz dos requisitos da Resolução nº 43 do CONARQ.

¹ Disponível em: <http://livroaberto.ibict.br/bitstream/123456789/1063/4/Manual-Archivematica.pdf>.

Do ponto de vista lógico, caracteriza-se como indutivo, no qual partiu-se de dados suficientemente constatados. O objetivo deste, é levar a conclusões cujo conteúdo é muito mais amplo do que o das premissas nas quais se basearam, e que sustentam ou atribuem certa verossimilhança à sua conclusão (MARCONI; LAKATOS, 2003).

Para as autoras, existem três elementos fundamentais para indução, isto é: a) Observação dos fenômenos (finalidade de descobrir as causas de sua manifestação); b) Descoberta de relação entre eles (entende-se por intermédio da comparação, aproximar fatos e ou fenômenos, com a finalidade de descobrir a relação existente entre eles); c) Generalização da relação (generalização da relação encontrada na precedente, entre fenômenos e fatos semelhantes ainda não observados). Isso se insere na abordagem de nosso objeto: a avaliação e relação entre as variáveis deste estudo.

Segundo os objetivos, trata-se de uma pesquisa descritiva, cuja finalidade é pormenorizar situações, acontecimentos e feitos, retratando como se dá a manifestação dos fenômenos (DESLANDES, 2009; GOMES, 2004). Possui a finalidade de definir o problema, proporcionar intuições de soluções, descrever comportamentos, fatos e variáveis, o que se encontra em consonância com a análise do *Archivematica* em relação à sua composição, estruturação e funcionalidades, mediante os requisitos da Resolução nº 43 do CONARQ.

Quanto ao delineamento, caracterizou-se como documental, pois teve fundamento em materiais registrados, no caso em evidência, a Resolução nº 43. Segundo Prodanov e Freitas (2013) a pesquisa documental se destaca no momento em que se organiza informações que se encontram dispersas, conferindo-lhe uma nova importância como fonte de consulta.

De acordo com a natureza, empregou-se a abordagem qualitativa, analisando a tríade observação-descrição-sistematização, dando significado aos requisitos postulados pela Resolução nº 43, na verificação da compatibilidade dessas premissas com as características do *Archivematica*.

Dessa forma, sistematizaram-se os dados coletados em um quadro para análise dos resultados, conforme mostra o Quadro 1.

Quadro 1 - Cumprimento dos requisitos do *Archivematica* à luz da Resolução nº 43

Seção	Item	Tópico	Descrição do tópico	Cumprimento dos requisitos por parte do <i>Archivematica</i>
Infraestrutura organizacional	1,2, 3	a) Governança e viabilidade organizacional, b)... c)...	Definição de cada tópico.	Sim () Não ()

Fonte: Elaborado pelo autor (2019).

Na aba “Seção”, estão inseridos os elementos dos três requisitos descritos na segunda parte do Resolução nº 43; no “Item”, foi representado um formato numérico atribuído a cada seção; no “Tópico”, dispuseram-se os requisitos a serem cumpridos pelo *Archivematica*; na “Descrição do tópico”, colocaram-se os requisitos de cada tópico, e por fim, no “Cumprimento dos requisitos por parte do *Archivematica*”, verificou-se se os mesmos estão em consonância, através da opção “sim” ou “não”.

3 O CAMPO DA CIÊNCIA DA INFORMAÇÃO: OBJETO E DISCIPLINA

O progresso dos registros informacionais em nosso estado burocrático possibilitou a intensificação e propagação das ciências, sendo assim, consequência para sua evolução nas diversas camadas sociais (GOODY, 1988). Nessa linha, Chibeni (2006) acorda que, a crença generalizada do conhecimento proporcionado pela ciência se distingue dos demais tipos, como: o religioso, tácito, entre outros; por exemplo, teorias, métodos, técnicas e produtos contam com a aprovação de sua comunidade para ser legitimado como tal. As várias atividades que se autoqualificavam “científicas”, buscavam se afirmar apenas nos seus próprios campos de atuação, correspondendo, no entanto, às ciências duras (CHIBENI, 2006).

Com o Pós-modernismo, a partir dos anos de 1950, esta afirmação tomou maior amplitude, principalmente com a interdisciplinaridade cuja atribuição expandiu sua abrangência. Dessa forma, as ciências passaram a se comunicar com maior intensidade (SANTOS, 2008). Para uma área ser considerada científica, faz-se necessário um conjunto de objetos consistentes, ou seja, a reunião de fatos, teorias e métodos, o que por muito tempo foi status privilegiado da Física, Biologia, Química e Matemática (KUHN, 2017; SANTOS, 2008).

No livro *A Estrutura das Revoluções Científicas*, Kuhn (2017) se refere ao desenvolvimento da ciência através de um viés estrutural, dando-se de forma cíclica, pela “ciência normal” (ciência), o paradigma, o estado de anomalia, a crise e as revoluções científicas. Para melhor explanar essas ideias, sintetizam-se os princípios citados acima de acordo com os postulados do referido autor (Quadro 2).

Quadro 2 - Estrutura das Revoluções Científicas de Kuhn

Ciência normal, direcionada a um campo científico, então possuindo relação com o estado de “anormalidade”, mas reconhecimento e aceitação pela maioria da comunidade científica.
O paradigma, peça-chave e comum à “ciência normal”, é o modelo coerente e específico de teoria a se seguir.
A anomalia, o início de uma crise, ou seja, a irregularidade, contrariando leis e expectativas da pesquisa científica.
A crise se desloca no estado de mudança de paradigma para um novo, pois os “antigos” não proliferam mais.
As revoluções, episódios de desenvolvimento, no qual o paradigma antigo é total ou parcialmente substituído por um novo.

Fonte: Kuhn (2017).

Esses postulados demonstram os estágios de composição de uma ciência. O autor pondera essas características em uma concepção de existência, amadurecimento, estabilidade e renovação de teorias.

Desse modo, com o avanço das TICs, observa-se a exigência e adaptação que a Era da Informação lançou nas diversas áreas, sobretudo com a emergência dos computadores, diminuindo a ideia de espaço e tempo, a exemplo, da Ciência da Computação e CI. Para esta última, atribuir um conceito integral a seu objeto de estudo: a informação, é uma tarefa complexa, visto que se caracteriza como elemento comum a diversos campos.

Capurro e Hjørland (2007) consideram que tal objeto não deve ser considerado como componente estático, nem ao menos conceituado de maneira isolada, mas fundamentado a partir do que já existe. Ampliando, afere a informação relacionada a um conhecimento comunicado, podendo desempenhar papel central nas relações sociais, e que por meio dela torna-se possível a evolução e desenvolvimento da sociedade.

Em seu artigo intitulado *O conceito de Informação*, os autores levantam a questão de revisar o status do próprio conceito da informação na CI sob o prisma da interdisciplinaridade, abrangendo um horizonte temporal desde a antiguidade. Sua significação é o que a faz heterogênea, pois para entendê-la é necessário revisitar a história da escrita, dos suportes e da comunicação. Capurro e Hjørland (2007) afirmam que:

Para uma ciência como a CI, é, sem dúvida, importante a forma como seus termos fundamentais são definidos e, assim como em outros campos, na CI a questão sobre como definir informação é frequentemente levantada. [...]. No discurso científico, conceitos teóricos não são elementos verdadeiros ou falsos ou reflexos de algum outro elemento da realidade; em vez disso, são construções planejadas para desempenhar um papel, da melhor maneira possível. (CAPURRO; HJORLAND, 2007, p. 149).

Para tanto, caracterizar “informação” é compreender que sua designação está agrupada a diversos significados e interpretações. Le Coadic (1996) afirma que a informação é um conhecimento inscrito como um elemento de sentido e significado em um suporte espaço-temporal, através de um sistema de signos de linguagem que associa um significante a um significado.

Silva e Ribeiro (2008) apontam que essa matéria-prima é um fenômeno imbuído de fato ou notícias na geração do conhecimento, portanto, atribuída de propriedades e atribuições em um contexto social.

Buckland (1991) apresenta três significados para informação: informação-como-processo, informação-como-conhecimento e informação-como-coisa. O significado dessa tríplice envolve processos que relacionam o conhecimento à comunicação. Segundo o autor, informação-como-processo se dá pela emissão de informação geradora de impacto em um receptor, e esse resultado é atribuído a conhecimento. Informação-como-conhecimento

também é usado para aquilo que é percebido (informação-como-processo), ou seja, é um conhecimento comunicado referente a algum fato particular, assunto ou evento. Já informação-como-coisa é àquela atribuída a objetos, assim como para documentos.

Vê-se que a caracterização da informação detém em seu escopo uma multiplicidade de atribuições acerca de sentidos e conceitos, podendo possuir interferência dos contextos que estão inseridas. Destaca-se que, desde sua gênese, a CI recebeu forte influência dos pilares de algumas disciplinas, como é o caso da Biblioteconomia e Documentação, e sobretudo, da própria tecnologia, a exemplo de Vannevar Bush e Moeers (1951) com o Memex, uma máquina de otimização na estruturação de dados.

A CI nasceu e se consolidou no contexto europeu, soviético e anglo-saxão entre 1940 e 1960. Nesse sentido, as propostas iniciais para a fundação dessa disciplina foram empreendidas, em 1961 e 1962, com o encontro denominado: Conferences on training Science information specialists no Georgia Institute of Technology. Após este evento, passou a incorporar outras teorias e problemáticas informacionais, ou seja, tendências na União Soviética, Inglaterra e Estados Unidos (ARAÚJO, 2018). Entretanto, foi a mudança de nomenclatura do American Documentation Institute para American Society for Information Science and Technology, que levou Borko (1968, p. 3) a publicar o artigo *Information science: what is it?*, considerado um dos pioneiros na consolidação da CI.

A ciência da informação é uma disciplina que investiga as propriedades e o comportamento da informação, as forças que governam o fluxo de informação e os meios de processar a informação para obter acesso e facilidade de uso. Preocupa-se com esse corpo de conhecimento relacionado à origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação. Isso inclui a investigação de representação da informação em sistemas naturais e artificiais, o uso de códigos para transmissão eficiente de mensagens e o estudo de dispositivos e técnicas de processamento de informações, como computadores e seus sistemas de programação. É uma ciência interdisciplinar derivada e relacionada a campos como Matemática, Lógica, Linguística, Psicologia, Tecnologia da Computação, pesquisa operacional, Artes gráficas, Comunicação, Biblioteconomia, Administração e outras áreas similares.

Isso demonstra que a informação tem como base dessas ações, não mais a necessidade de se ter a posse de documentos, mas a circulação, fluxo e atendimento às necessidades dos cientistas e usuários de forma geral (ARAÚJO, 2018).

Portanto, trata-se de uma disciplina que detém suas propriedades ligadas ao comportamento e fluxos informacionais, visando o acesso e o uso. Na qualidade de ciência se

relaciona com diversas áreas, tendo um componente tanto de ciência pura, baseando-se na aquisição de novos conhecimentos que pode gerar novas teorias, quanto aplicada na resolução de problemas práticos cotidianos (BORKO, 1968).

Emergida no contexto do Pós-guerra, os estudos da CI posicionavam-se no viés da informação científica e tecnológica, tendo como uma de suas causas a explosão informacional, marcada pela gama de conhecimento e registros produzidos.

Na perspectiva russa, ditada pelo termo *Informatika* de Mikhailov, Chernyi e Gilyarevski é caracterizada como uma disciplina social que estuda a estrutura e as propriedades gerais da informação científica, bem como seus processos de comunicação. Essa estrutura se conecta com os aspectos semânticos e formais, ou seja, atribuída a conceitos, significados e teorias. Suas propriedades atrelam-se a essas mesmas estruturas com atributo de sentido e descrição da informação científica (MIKHAILOV; CHERNYI; GILYAREVSKYI, 1980).

Nesse cenário, observa-se que a origem da CI sofreu influência da Documentação, Biblioteconomia e *Informatika*. Destacando a íntima relação dessa primeira, sobretudo, quando relacionado ao conceito de documento, o qual abrange uma multiplicidade de materiais de informação, divergindo da ideia estática de exclusividade do papel. No processo de criação dos instrumentos de organização, armazenamento e recuperação Otlet e La Fontaine focalizaram o “conteúdo” dos documentos, ou seja, a informação em si, procedimento ainda sem observação à época (FREIRE; FREIRE, 2009).

Ademais, como campo interdisciplinar possui estreita vinculação com outras áreas, a exemplo da Arquivologia, disciplina ancorada no gerenciamento da informação orgânica, independentemente do suporte, possuindo elementos que as justapõem, assim, às instituições que atuam no fluxo do conjunto informacional e no lugar de memória (arquivo).

4 TRAJETÓRIA DE UMA ÁREA DOCUMENTAL: EVOLUÇÃO DOS ARQUIVOS E A FORMAÇÃO DA ARQUIVOLOGIA

O surgimento da escrita remonta também à gênese dos arquivos e da Arquivologia. Servindo como baliza para o desenvolvimento da sociedade, permitiu que a humanidade por toda sua trajetória registrasse suas ações e evoluísse, sendo o suporte artefato especial para esse feito.

Há uma pluralidade de significados no entorno da palavra “arquivo”. Porém, existem controvérsias em relação a seu nascimento, pois sua aceção foi concebida em diversos lugares. Rosseau e Couture (1994) acordam que, inicialmente, foi cunhada pelos gregos nos séculos III e II a. C., com o termo *archeion*, possuindo designação de: *general administrator*, *record office* e *government place*, por exemplo. Tais proposições, de imediato, referiam-se às atividades administrativas, visando estabelecer as relações entre governo, organizações e pessoas (ROSSEAU; COUTURE, 1994).

Nesse sentido, os arquivos constituem documentos, os quais desempenham papel de prova na sociedade e podem ser entendidos como informação registrada em um suporte: físico ou digital (RONDINELLI, 2005; SCHELLENBERG, 2006).

Para melhor compreensão, Rousseau e Couture (1994, p 34) esclarecem que:

Os documentos, qualquer que seja o seu caráter, pessoal, administrativo, financeiro, são portadores de uma informação particular diferente da obra literária, da escrita científica ou da reportagem fatural. [...], trata-se de uma informação pessoal [cada pessoa é dona dos seus arquivos], a sua utilidade é própria a cada organismo, o seu tratamento é igualmente específico. Os arquivos constituem uma fonte de informação única sobre as pessoas e as organizações e, por esse motivo, constituem materiais indispensáveis à história ou a qualquer outra disciplina cujo objeto seja o passado.

Verifica-se, então, as distintas facetas que os documentos possuem. Vale destacar que as práticas arquivísticas datam épocas passadas; como já exposto, os arquivos remontam ao início da sociedade. Os chamados guardiões da informação eram cargos determinados por reis a pessoas altamente letradas durante o curso da História antiga e medieval (POSNER, 1972). Destarte, ocupar este tipo de cargo era privilégio de poucos, a exemplo dos intelectuais da época.

O surgimento da Arquivologia foi concebido em meados do século XVII, considerando os seus primeiros elementos, enquanto área do conhecimento, no ano de 1681 com a Diplomática (DURANTI, 1993). Isto posto, Rosseau e Couture (1994, p. 48, grifo nosso) tecem algumas considerações sobre o desenvolvimento da área arquivística:

Ela é constituída por um *savoir-faire* que se foi acumulado ao longo dos anos. Os métodos de trabalho mudaram, mas encontramos geralmente as mesmas preocupações funcionais. A história permite definir quatro grandes setores principais que foram objetos dos trabalhos dos especialistas de arquivos, ou seja, o tratamento, a conservação, a criação e a difusão.

No entanto, pode-se considerar a publicação do *Manual dos Arquivistas Holandeses*, no século XIX (1898), como o marco inaugural de um campo autônomo dos arquivos (SCHELLEBERG, 2006). É a partir do estabelecimento desse manual que surgem alguns tópicos importantes para a análise das características da área, isto é, suas bases teóricas e conceituais, que para Fonseca (2005) se deram na:

- a) importância da Arquivologia na *episteme* (conjunto de relações que podem unir práticas discursivas que dão lugar à ciência) da sociedade, principalmente na esfera política;
- b) importância das instituições arquivísticas para lidar com problemas de eficiências junto à Administração Pública;
- c) subordinação da disciplina em relação a seu objeto, ou seja, se a concepção de arquivo estiver clara, também estará a de Arquivologia;
- d) tradução dos manuais da área, suas limitações e as tentativas de generalizar o particular.

O século XIX foi o período de maior desenvolvimento, progresso e evolução dos preceitos arquivísticos, uma vez que se assistia a uma concentração diferente dos arquivos nos estados, em grandes depósitos, no período da Revolução Francesa. A legislação vigente à época teve alguns aspectos inovadores, como destacam Silva *et al* (2009): a criação de um órgão nacional e independente para superintendência dos arquivos, disposta pelo art. 3º da Lei do 7 Messidor e, o segundo, a intenção de liberar o acesso dos arquivos à população. Fonseca (2005, p. 36) evidencia o início dos arquivos enquanto instituições de guarda de documentos importantes à Administração Pública:

As instituições arquivísticas, como hoje a concebemos, remontam à criação, em 1789, do Arquivo Nacional da França, primeiramente como Arquivo da Assembleia Nacional e depois transformando, em 24 de junho de 1794, no estabelecimento central dos arquivos do Estado, ao qual foram subordinados os depósitos existentes nas províncias. Neste depósito deveriam ser recolhidos os documentos produzidos pelos diferentes níveis da administração pública na França.

A política da França, na fase sob a ascensão e estratégia expansionista de Napoleão Bonaparte para a consolidação do seu império, ofereceu aos arquivos importante reflexo,

devido à promulgação de leis que tornavam obrigatória a transferência a Paris, dos arquivos dos países dominados e territórios ocupados por ele. Depois da Era napoleônica houve a devolução desses acervos aos países de origem, o que acarretou na destruição de muitos documentos.

A concentração desse conjunto documental (por Napoleão) foi tratado como único. Assim, o Arquivo Nacional da França passou a classificar, como conjunto unitário, sua massa documental. Os documentos eram reunidos em cinco seções metódica e cronologicamente: legislativa; administrativa, dominial, judiciária e histórica (JARDIM; FONSECA, 1992).

Estes marcos formalizaram também conceitos e princípios importantes até hoje, como o de respeito aos fundos por Natalis de Wailly, em 1841. Sua concepção é a reunião de documentos por fundos (de famílias e/ou instituições), dispondo em ordem os fundos sem se misturar, o que destaca também a premissa da ordem original (DUCHEIN, 1986).

A partir desses postulados, estabeleceram-se ferramentas que permanecem no campo da Arquivologia, como a classificação e organização de arquivos, que a *posteriori* servirão para arquitetura de informação nos repositórios digitais disponíveis na web. No transcurso de sua evolução, sobretudo nessa época de consolidação de seu campo disciplinar como auxiliar da História, destacam-se os autores Hilary Jenkinson, em 1922 e Eugênio Casanova, em 1928, em uma perspectiva tecnicista de gestão administrativa no fazer arquivístico (ARAÚJO, 2011).

Em meados do século XX, começam as primeiras preocupações com a eliminação de documentos, situação que, mesmo antes da Segunda Guerra, já se mostrava peculiar. Isso era o prenúncio da gestão documental e da teoria das três idades. Alguns estudos se faziam presentes, no intuito de identificar modelos de responsabilidade administrativa para eliminação de documentos. O britânico Jenkinson (1965), defendia que a eliminação de documentos não deveria fazer parte das atribuições dos arquivistas, diferentemente do modelo alemão, que considerava o arquivista como peça central nessa atividade.

O modelo de Schellenberg (2006) na questão da avaliação documental atribui valor aos documentos: primário, quando ligado aos preceitos da criação dos documentos em face à entidade produtora e, secundário, na perspectiva do seu caráter informativo à investigação científica. As ideias deste autor estão inseridas no ápice da explosão documental, também acontecida na CI, sobretudo, pelo avanço das TICs.

Por sua vez, a consolidação sobre a avaliação documental se dá na prática da gestão documental nos Estados Unidos. Na vivência do pós-guerra emerge o conceito de *record group* - adaptação americana ao princípio de proveniência. Foi estabelecido em 1941, no

Nationale Archives de Washington, para possibilitar uma flexibilidade na organização dos documentos produzidos por diferentes entidades (SILVA *et al*, 2009). Este conceito foi difundido nos EUA e Canadá e propiciou o *record management*, traduzido na França, Espanha, Brasil e outros países por gestão documental.

Segundo o Dicionário Brasileiro de Terminologia Arquivística, gestão documental é o conjunto de procedimentos e operações técnicas que perpassa as fases do seu ciclo (corrente, intermediário, permanente), produção, tramitação, uso, avaliação e arquivamento, visando sua eliminação ou recolhimento (BRASIL, 2011). Entende-se por arquivo corrente aquele referente às atividades imediatas da entidade produtora; intermediário, os que aguardam seu uso ou sua locomoção para o permanente, possuidor de valor histórico e de prova (SCHELLENBERG, 2006).

De acordo com Fonseca (2005), a efetivação da gestão documental foi resultado dos relatórios da Commission on Organization of the Executive Branch of the Government, mais conhecida pela Comissão Hoover. Uma das recomendações desses relatórios foi a criação National Archives and Records Service (NARS) em 1950, o que se transformou no National Archives and Record Administration (NARA) em 1985.

Inicialmente, a gestão documental foi estabelecida como uma área apartada da Arquivologia, tanto é que, na tentativa de superar esta cisão, os arquivistas canadenses intitularam o conceito de “Arquivologia integrada” (justapondo as duas), extrapolando assim, já naquela época, as suas fronteiras disciplinares (FONSECA, 2005).

Em paralelo a esses acontecimentos cria-se, no âmbito da UNESCO, o Conselho Internacional de Arquivos (ICA), preocupando-se em sua gênese com problemas relativos à terminologia, mas a *posteriori* amplia-se, no sentido de fazer chegar às entidades/organizações/instituições, informações sobre as práticas arquivísticas, como gestão e difusão informacional em nível mundial.

Nesse período, a Arquivologia vivia sua fase moderna, desde o surgimento do Manual dos Holandeses, considerado o pilar de seu campo disciplinar (COOK, 1998). Porém, na afirmação desse autor, atualmente a Arquivologia se encontra no período pós-moderno, modelo de teoria que vem sendo disseminado em consequência das rupturas dos paradigmas atuais da área. Dessa forma, tem-se uma disciplina integrativa, na ótica canadense de Terry Cook, disseminando a informação arquivística como seu objeto de estudo e não, necessariamente, o arquivo em si (ARAÚJO, 2013), voltada para a interação do usuário da informação e seus contextos, bem como na utilização das tecnologias digitais.

4.1 RELAÇÕES INTERDISCIPLINARES: ARQUIVOLOGIA NO PERÍODO PÓS-MODERNO E A CIÊNCIA DA INFORMAÇÃO

Pode-se destacar que a Revolução industrial foi um período de progresso nos diversos âmbitos da sociedade, acarretando uma significativa evolução das tecnologias no campo científico. Tal evento, teve influência que perpassaram as grandes guerras, enfatizando: a maior produção de armas bélicas, de construções civis, e, conseqüentemente, do conhecimento. O desenvolvimento das características das ciências conhecidas como moderna, nesse contexto, avançaram consideravelmente, assim explica Santos (2008, p. 20-21):

O modelo de racionalidade que preside a ciência moderna constituiu-se a partir da Revolução científica do século XVI e foi desenvolvido nos séculos seguintes basicamente no domínio das ciências naturais. Ainda que com alguns prenúncios no século XVIII, é só no século XIX que este modelo de racionalidade se estende às ciências sociais emergentes. A partir de então pode falar-se de um modelo global de racionalidade científica que admite variedade interna, mas que se distingue e defende, por via de fronteiras ostensivas e ostensivamente policiadas, de duas formas de conhecimento não científico (e, portanto, irracional) potencialmente perturbadoras e intrusas: o senso comum e as chamadas humanidades ou estudos humanísticos (em que se incluíram, entre outros, os estudos históricos, filológicos, jurídicos, literários, filosóficos e teológicos).

Uma das causas advindas desse período, por conseguinte, foi a comunicação, entre as diversas áreas, colaborando para eloquência da era Pós-moderna, marcada pelas ideias de questionamento crítico, político e social.

Esta passagem começou desde pelo menos o final dos anos 50, marcando para a Europa o fim de sua reconstrução. Foi mais ou menos rápida conforme os países e, nos países, conforme os setores de atividade: donde uma diacronia geral, que não torna fácil o quadro de conjunto. Uma parte das descrições não pode deixar de ser conjectural. E sabe-se que é imprudente conceder um crédito excessivo à futurologia. (LYOTARD, 2009, p. 3).

Haja vista a contradição na atribuição do período Pós-moderno por parte de alguns autores, pode-se caracterizá-la atribuindo sua implicação na sociedade Pós-industrial, como formas de reflexão e fortalecimento da ideia de liberdade e igualdade.

A Arquivologia insere-se nessa acepção, uma vez que traz consigo as rupturas de uma disciplina estática custodial, isto quer dizer, seu objeto se transformou e ampliou sua conotação para um tipo de informação voltada ao contexto da Era tecnológica digital e ao próprio conteúdo do documento.

Assim, Silva *et al* (2009), a caracteriza em três momentos: a fase sincrética e custodial, técnica e custodial e científica e pós-custodial, conforme se apresenta no Quadro 3.

Quadro 3 - Processo histórico relativo aos arquivos

Fase sincrética/custodial (séc. XVIII - 1898) - marcada pelos arquivistas paleógrafos, como auxiliar da História	<ul style="list-style-type: none"> • Incorporação maciça da documentação de organismos extintos • Arquivo Nacional • Noção de fundo de Natalis de Wailly em 1841 • Revolução industrial • Revolução política e social • Relacionalismo, Positivismo, Historicismo • Estado Nação
Fase técnica/custodial (1898 - 1980) - marcada pela especialização profissional do arquivista	<ul style="list-style-type: none"> • Arquivos históricos • Arquivos administrativos • Gestão de documentos • Normalização arquivística (terminologia, classificação) • Reformulação crítica da noção de fundo • 2ª e 3ª vagas de industrialização • Evolução tecnológica e científica • Democracia e Totalitarismos
Fase científica/pós-custodial (1980...) - marcada pelo arquivista como profissional da informação/cientista da informação	<ul style="list-style-type: none"> • Arquivos como sistemas de informação • Conhecimento arquivístico • Normalização do acesso aos arquivos e à informação • Consequência da pós-industrialização • Sociedade da informação • Globalização da economia

Fonte: Silva *et al* (2009).

Salienta-se as ponderações de Kuhn (2017) nesse sentido, quando se mostrou a existência e rupturas das fases relativas ao processo histórico da Arquivologia no quadro acima, destacando o que Kuhn chamou de Revolução científica, nesse caso, as quebras de paradigmas.

Thomassen (2001) também transporta tal ideia à arquivística nos tempos de hoje, evidenciando que o desenvolvimento das TICs está conduzindo com eficiência às mudanças de paradigmas nas suas fases. Pois se tomarmos como experiência suas primeiras práxis, teremos as tradições europeias com a Paleografia e Diplomática, voltadas a um conjunto de decifração de documentos, e a segunda advinda da América do Norte pela gestão documental, incumbida do entendimento do fluxo do documento, insumo este, ligado ao contexto de produção de instituições e pessoas (THOMASSEN, 2001).

Nessa acepção, o objeto da Arquivologia também sofreu transformação, pois o “arquivo” propriamente dito, aquele que era caracterizado como produto final da instituição, passa a incorporar uma nova concepção, a conotação de “informação arquivística ou orgânica”, gerada naturalmente pelos processos das atividades de uma pessoa ou entidades em função de suas necessidades (FONSECA, 2005). Nessa perspectiva, Thomassen (2001) refere que esse objeto possui duas características: a informação arquivística (conteúdo semântico dos

documentos) e o seu contexto de geração/processo de criação dos documentos (conteúdo diplomático do documento), premissas estas advindas também da “era digital”.

Tais mudanças, ligada as ideias pós-modernas, são atribuídas aos preceitos de COOK (1998), baseado no racionalismo e no método científico, pois nada é neutro ou imparcial, tudo é conformado, apresentado e representado, sendo assinado por aquele que fala, fotografa ou escreve.

[...] como decorrência tanto do impacto das reflexões da Arquivologia integrada, como das tecnologias da informação, [...], o arquivo passou a ser cada vez mais caracterizado como um sistema de informação, e o objeto de estudo da Arquivologia, como a informação arquivística. (ARAÚJO, 2013, p. 77-78).

Esta perspectiva transborda a noção de arquivo para além dos limites institucionais. As instituições arquivísticas deixam de ser lugares onde usuários vão obter informações para serem provedores de acesso remoto a milhares de sistemas arquivísticos interligados (KETELLAR, 2002).

A inserção da fase pós-custodial na Arquivologia caracteriza-se em uma questão de nomenclatura e fatores que integram os estudos sobre arquivos. Silva *et al* (2009), em complemento a essa ideia, corroboram que temas ainda considerados recentes na área, como a elaboração de normas descritivas, têm servido para dar estabilidade à sua comunidade frente a disciplinas já consolidadas, como a Biblioteconomia, fortalecendo seu domínio no que se refere à sensibilidade dos arquivistas em relação a problemas técnico-científicos.

Essas ponderações expõem uma característica pontual na conjuntura pós-custodial, a visibilidade dos arquivistas em suas demandas. Diferentemente das fases históricas e tecnicistas, esse período, considerado ainda recente, tem promovido cada vez mais a inserção da CI no campo dos arquivos, principalmente nos sistemas de informação.

Os métodos tradicionais de preservação de documentos, também baseados em padrões, tornam-se obsoletos mais rapidamente, na medida em que passam a migrar seus conteúdos antes de se deteriorarem, resultado decorrente da tecnologia digital.

Como consequência desses acontecimentos, os pressupostos tradicionais de proveniência, originalidade e funcionalidade dos documentos também precisam ser reformulados. A preservação não será voltada à restauração e conservação do arquivo físico, ao contrário, seu principal objetivo ocorre na migração e emulação constantes para novos softwares (COOK, 1998), aproximando cada vez mais a Arquivologia da CI pela relação com a informação no ambiente digital.

Tanus e Araújo (2012) abordam a aproximação da CI com a Arquivologia, destacando pontos específicos, tais como: as instituições (que, por essência, utilizam a informação/documento nas suas atividades), a memória (elemento comum caracterizado como lugar de guarda informacional), o documento propriamente dito (informação independentemente do suporte), e a própria interdisciplinaridade (a qual comunica-se uma com a outra).

A interdisciplinaridade é um adjetivo que qualifica o que é comum a duas ou mais disciplinas no processo de ligação entre elas. A partir dos anos de 1970, em especial, ouvia-se muito desse discurso, mas de início, servia apenas para estudos de congressos (HEEMANN, 2004).

A CI, ainda em fase de maturação, permeia diversas áreas do conhecimento na busca da construção de fundamentos teóricos para sua consolidação. Observa-se, a expressiva relação que seu campo possui com os aspectos biblioteconômicos e tecnológicos. Autran (2014, p. 107), expressa que as contribuições sobre a interdisciplinaridade na CI também se destacam nos contextos

[...] do ensino e suas implicações com os conteúdos curriculares dos cursos de graduação e pós-graduação *stricto sensu* [...] além da pesquisa, manifestada na Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB), através dos Grupos de Trabalho (GT) a exemplo dos seguintes GTs: Estudos Históricos e Epistemológicos da Ciência da Informação, Organização e Representação do Conhecimento e Mediação, Circulação e Apropriação da Informação.

A autora, em sua tese, refere à preocupação que existe com a interdisciplinaridade em nível nacional e internacional. Relata que no Brasil, a interdisciplinaridade da CI manifestou-se a partir de trabalhos pioneiros desenvolvidos por Zaher e Gomes em 1972, Braga em 1995 e Pinheiro e Loureiro em 1995. No plano internacional, destacam-se Borko em 1968, Foskett em 1980, Wersig em 1993 e Saracevic em 1996.

Pinheiro e Loureiro (1995) também empreenderam uma pesquisa sobre interdisciplinaridade, que culminou com um esquema de 12 subáreas com as quais a CI se relaciona. Em trabalhos posteriores, Pinheiro (1997) continuou a pesquisa tendo como fonte de consulta o *Annual Review for Information Science and Technology (ARIST) 1966-1995*, mapeando, além das anteriores, mais cinco subáreas e em complemento atualizou os dados utilizando a mesma fonte de informação, com a baliza temporal 1996-2004. De acordo com a autora, as 13 primeiras são áreas já consolidadas, as demais são emergentes (Quadro 4).

Quadro 4 - Disciplinas e áreas interdisciplinares da CI

Disciplinas	Áreas interdisciplinares
Sistemas de informação	Administração; Ciência da Computação
Tecnologia da informação	Ciência da Computação
Sistemas de recuperação da informação	Biblioteconomia; Ciência da Computação; Linguística
Políticas de informação	Administração; Ciência Política; Direito
Necessidades e usos de informação	Arquivologia; Biblioteconomia; Museologia; Psicologia
Representação da informação	Arquivologia; Biblioteconomia; Filosofia; Linguística; Museologia
Teoria da Ciência da Informação	Epistemologia; Filosofia; Filosofia da Ciência; Matemática
Formação e aspectos profissionais	Educação; Ética; Direito
Gestão da informação	Administração; Economia; Estatística
Bases de dados	Ciência da Computação
Processamento automático da linguagem	Biblioteconomia; Ciência da Computação; Linguística
Economia da informação	Administração; Economia
Bibliometria	Estatística; História da Ciência; Matemática; Sociologia da Ciência
Inteligência competitiva e Gestão do conhecimento	Administração; Economia
Mineração de dados	Ciência da computação
Comunicação científica eletrônica	Ciência da Computação; Comunicação; História da Ciência; Filosofia da Ciência
Bibliotecas digitais/virtuais	Biblioteconomia; Ciência da computação; Comunicação

Fonte: Adaptado de Pinheiro (2006).

Saracevic (1996, p. 42), em suas contribuições, compreende três características fundamentais da natureza da CI: 1. A interdisciplinaridade; 2. Ligação inexorável com as tecnologias informacionais; 3. Participação ativa e deliberada na sociedade da informação. O autor tece as seguintes considerações acerca das características fundamentais da CI:

O imperativo tecnológico determina a CI, como ocorre também em outros campos. Em sentido amplo, o imperativo tecnológico está impondo a transformação da sociedade moderna em sociedade da informação, era da informação ou sociedade pós-industrial. [...], a CI é, juntamente com muitas outras disciplinas, uma participante ativa e deliberada na evolução da sociedade da informação. A CI teve e tem um importante papel a desempenhar por sua forte dimensão social e humana, que ultrapassa a tecnologia. Essas três características ou razões constituem o modelo para compreensão do passado, presente e futuro da CI e dos problemas e questões que ela enfrenta.

O autor reforça a ideia da relação com outros campos disciplinares científicos, que na Arquivologia pode-se atrelar ao domínio tecnológico na análise e compreensão da informação em meio digital.

Salienta-se que a emergência do ambiente digital não intervém na essência do conceito de documento de arquivo, isto é, sua estrutura como documento digital permanece inalterada as características arquivísticas. Rondinelli (2005) afirma que o documento arquivístico digital é uma unidade indivisível de informação, constituída por mensagem fixada em um suporte digital codificado em dígito binário e acessível por um sistema computacional.

Existe diferença de significados no que se refere aos documentos: digitais, eletrônicos e digitalizados. Expõe-se essas diferenças no Quadro 5.

Quadro 5 - Tipos de documentos

Documento digital: caracteriza-se pela informação registrada, codificada em dígitos binários, sendo acessível por um sistema computacional.
Documento eletrônico: informação registrada, codificada em forma analógica ou em dígitos binários acessível por um equipamento eletrônico.
Documento digitalizado: informação registrada e/ou convertida para o formato digital, por um dispositivo apropriado.

Fonte: Brasil (2011).

Segundo Araújo (2013), a terceira fase da Arquivologia, (tecnologia e documento em formato digital) ou fase pós-custodial surge após uma primeira ligação com a História e uma segunda com a Administração. Assim, como consequência tanto da integração da Arquivologia quanto das TICs, além dos tipos especiais de arquivo, este “[...] passou a ser cada vez mais caracterizado como um sistema de informação, e o objeto de estudo da Arquivologia, como informação arquivística” (ARAÚJO, 2013, p. 77-78), ampliando seu escopo, caracterizando o documento como elemento de tomada de decisão, gerado de maneira natural pela sociedade referente a uma função.

4.2 ARQUIVÍSTICA E A TECNOLOGIA DIGITAL

Atualmente, vive-se em plena era da informação e do avanço tecnológico em todos os setores da sociedade. Pode-se até afirmar que essa modernização tem ocasionado a dependência do ser humano em relação às ferramentas disponíveis, principalmente os dispositivos móveis.

O significado das TICs corresponde a qualquer tecnologia, cuja interferência se direciona aos processos informacionais e comunicativos. De acordo com Farias (2013), esse tipo de tecnologia pode ser considerado como um conjunto de recursos, que permite maior facilidade ao acesso e disseminação das informações, presentes no cotidiano da sociedade.

Nesse cenário, Lopez (2004) aponta que existe a necessidade de ampliar as discussões sobre o impacto das tecnologias digitais no universo arquivístico. O autor destaca, entre

outros elementos, a gestão eletrônica de documentos e a importância dos metadados para essa nova era da Arquivologia. Os metadados são mais uma das inovações impostas pela Informática e têm a finalidade de garantir que o documento, em meio digital, mantenha sua fidedignidade, assegurando seu valor probatório legal (LOPEZ, 2004). Como forma de representação e recuperação da informação, os metadados são indispensáveis, pois, além de garantir a autenticidade e fidedignidade do registro, independentemente do suporte, passa a ser um elemento que define a organicidade dos mesmos, conforme afirma Rondinelli (2013, p. 231): “no mundo digital tudo é codificado em linguagem binária e, para se tornar acessível aos olhos humanos precisa da intermediação de programas computacionais, igualmente codificados em bits, numa sofisticação tecnológica que passa despercebida pela maioria dos usuários.”.

Saracevic (1996) pondera que a máquina “computador” causou grande impacto e transformações no uso, descrição e transferências das informações pela interação com o homem através do gerenciamento com software.

É válido ressaltar que ambas as áreas, Arquivologia e CI têm mantido uma interlocução profícua na comunicação de seus aportes teórico-metodológicos. Ao analisar a contribuição da CI para a Arquivologia, Araújo (2011, p. 119) ressalta:

De um lado, a CI ofereceu à Arquivologia possibilidade de construção de conhecimentos propriamente científicos, indo além da dimensão de produção de manuais de “como fazer”. Ao mesmo tempo, abriu portas para que a Arquivologia problematizasse questões que vão além da instituição arquivo: as políticas de informação, os arquivos pessoais, as realidades documentais não tratadas do ponto de vista arquivístico, entre outras.

A Arquivologia trouxe suas contribuições para CI, indo além de uma extensa acumulação de técnicas, conceitos e visões sobre os documentos arquivísticos, mas transportando a noção sobre organicidade, ciclo vital dos documentos, patrimônio e memória.

Destaca-se que a Gestão Eletrônica de Documentos (GED) teve como ponto de partida apenas a digitalização, não abarcando os procedimentos necessários à gestão e à legislação arquivística (LOPEZ, 2004). Porém, com a evolução da tecnologia da informação e o aumento na produção dos documentos chamados nato-digitais, emergiu a necessidade de se realizar uma gestão voltada para esse novo tipo de suporte digital, como destaca Flores e Lampert (2013, p. 49) “[...] com o passar dos anos, a crescente produção de documentos suscitou nos profissionais da informação uma preocupação em como estes estavam sendo geridos enquanto documentos arquivísticos”.

Ao falar sobre esses tipos de procedimentos relacionados ao gerenciamento informacional, tal propositura ocorre em sistemas de informação, como repositórios digitais, pois sua importância se insere na alçada da preservação digital (FLORES; HEDLUND, 2014). O resultado da constante e rápida evolução tecnológica faz brotar uma série de preocupações em relação à longevidade dos documentos digitais, bem como sua segurança, autenticidade e integridade, elementos estes caracterizados como propriedades indispensáveis para a preservação digital.

5 PRESERVAÇÃO DIGITAL: CONCEITO E ESTRUTURA

Desde o surgimento das civilizações, a escrita possibilitou a produção de vestígios sobre sua história. Para a perpetuação desse conhecimento durante o curso das épocas, a prática de preservar foi um dos mecanismos centrais. Dessa maneira, guardar e manter artefatos é uma ação remota, pois se assim não fosse, pouco tínhamos conhecido do passado (POSNER, 1972).

Pode-se observar que tais afirmações na antiguidade não foram necessariamente intencionais, pois a concepção de preservação que se tem na atualidade é considerada recente, haja vista que museus, bibliotecas e arquivos assumem papel determinante no viés dessa atividade, principalmente posterior à revolução francesa (ROSSEAU; COUTURE, 1994).

Com as TICs, e a conseqüente eclosão do ambiente digital, as premissas de preservação passaram a ganhar conotação diferente das ideias tradicionais, pois a preocupação de conservar e/ou proteger a informação em um objeto não palpável tornou-se um desafio imediato para os que trabalhavam com o gerenciamento informacional (FLORES; HEDLUND, 2014). O acesso à informação, o advento da Internet e, sobretudo, a preservação digital, fundamentam-se no desenvolvimento e aplicação de técnicas capazes de viabilizar a longevidade dos documentos (FLORES; PRADEBON; CÉ, 2017).

Para Ferreira (2006), o material digital carrega consigo um problema de estrutura, colocando sua perenidade em risco através de um fator elementar: a obsolescência. O documento digital pode ser copiado inúmeras vezes sem perder a qualidade, mas para isso é imprescindível a presença de um contexto tecnológico nos parâmetros da preservação, bem como de diferentes estratégias. Porém, a dependência tecnológica sem manutenção torna o material digital vulnerável à rápida obsolescência, fator intrínseco à tecnologia (FERREIRA, 2006).

Durante o curso da história, presencia-se a mudança dos suportes de informação, como: fita *VHS*, passando pelo disquete, *CD-ROM*, *DVD-ROM*, *pen-drive* e armazenamento na nuvem. O transcurso dessa transição não se dá apenas nos materiais que as compõem, mas na composição de seus formatos, permitindo que as aplicações de softwares sejam capazes de interpretá-los, pois na medida que a tecnologia avança, os suportes se modificam, tanto no formato quanto na capacidade de armazenamento. Estes exemplos estão relacionados com a evolução do significado de preservação, uma vez que tal termo objetiva garantir o acesso a longo prazo à informação digital.

Santos e Flores (2015, p. 203-204) tecem considerações sobre os procedimentos estruturais e operacionais de preservação:

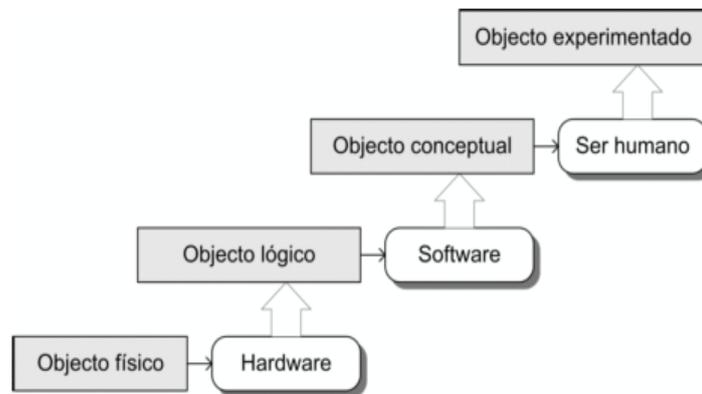
A preservação digital é composta por procedimentos de manutenção, que podem ser denominados estruturais e por procedimentos de preservação dos componentes, os quais podem ser denominados como operacionais. Os procedimentos estruturais se referem aos estágios iniciais, definindo questões como a infraestrutura, normas, escolha de padrões, financiamentos e definição de metadados. Já os procedimentos operacionais se referem a atividades que objetivam a preservação física, lógica e intelectual dos documentos digitais através de ações como a migração, a emulação e o encapsulamento.

Os objetos digitais podem ser entendidos como a informação representada por uma sequência de dígitos binários (linguagem computacional), isto inclui informação nascida no contexto tecnológico digital (objetos natos-digitais), como também a partir de suportes analógicos (objetos digitalizados). Documentos de texto, fotografias digitais, bases de dados, sequências de vídeo e áudio, modelos de realidade virtual, páginas web e aplicações de software são apenas alguns exemplos desses objetos (FERREIRA, 2006).

Segundo o mesmo autor existe uma anatomia ou estrutura que compõe tal objeto, que é promovida pelo tipo físico, lógico, conceitual e experimentado:

- **Objeto físico:** constitui aquilo que o hardware é capaz de interpretar, transformando símbolos inscritos no suporte físico no conjunto de dados que o software será capaz de manipular;
- **Objeto lógico:** estruturação de dados designado por formato do objeto digital. O software é responsável por preparar o objeto digital para que seja apresentado a um receptor humano através de um periférico de saída;
- **Objeto conceitual:** assume a forma ou concepção familiar aos seres humanos, como livros, filmes ou fotografias (documento(s)), constituindo aquilo que deve ser preservado;
- **Objeto experimentado:** experiência, interpretação ou entendimento feita pelo emissor/receptor do objeto.

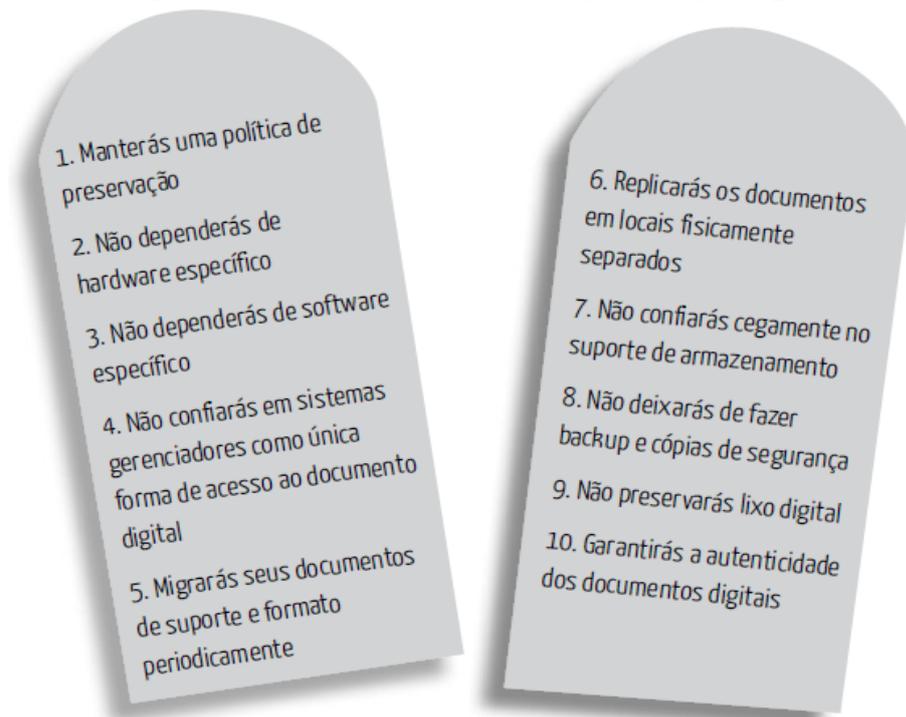
Essa estrutura mostra os diferentes tipos de objetos digitais que, conseqüentemente, em se tratando de sistemas de informação são inseridos em um repositório, visando seu armazenamento, fluxo e acesso. Para melhor visualizar esta configuração, veja-se a Figura 1:

Figura 1 - Estrutura dos objetos digitais

Fonte: Ferreira (2006).

A partir dessa figura, pode-se compreender como ocorre uma das etapas da atividade de preservação, pois ela é responsável por garantir a comunicação entre o sistema-usuário. Isto quer dizer que a preservação digital está estritamente relacionada com o armazenamento de informação na adoção de políticas de longevidade de seu material.

Nesse sentido, Innarelli (2012) apresenta uma série de recomendações que visam manter a utilização e acesso à informação digital, as quais são chamadas por ele de os dez mandamentos (Figura 2).

Figura 2 - Os dez mandamentos da preservação digital

Fonte: Innarelli (2012, p. 320).

As recomendações dessa figura estão voltadas para a orientação acerca da não dependência de hardware e software; da migração de suporte (físico) e do formato (lógico) do objeto digital periodicamente; manter a desconfiança na mídia de armazenamento; adoção de rotinas de backup; manutenção dos documentos digitais (INNARELLI, 2012).

Essas são apenas algumas estratégias que fazem parte do discurso da preservação digital. Ferreira (2006) aponta outras recomendações, agrupando suas ideias em onze classes (Quadro 6).

Quadro 6 - Estratégias de preservação digital

Preservação de Tecnologia	Consiste na conservação e manutenção de todo o hardware e software necessários à correta apresentação dos objetos digitais. Trata-se da criação de museus de tecnologia, onde a preservação não se concentra no objeto conceitual, mas sim no objeto digital em sua forma original.
Refrescamento	O refrescamento de suporte é a transferência de informação de um suporte físico de armazenamento para outro mais atual antes que o primeiro se deteriore ou se torne obsoleto. Essa estratégia não constitui uma preservação por si só. Deverá ser entendido como um pré-requisito para o sucesso de qualquer estratégia de preservação.
Emulação	Baseia-se na utilização de um software, designado emulador, capaz de reproduzir o comportamento de uma plataforma de hardware e software, em outra que a partida seria incompatível. Existem dois tipos de emuladores: sistemas operativos e de hardware. Os primeiros focam-se na reprodução de um sistema operativo por completo, permitindo a execução de diversas aplicações no contexto de um único emulador. Os segundos direcionam-se ao comportamento de uma plataforma de hardware, possibilitando que vários sistemas operativos e correspondentes aplicações possam ser executados no contexto de um único emulador.
Migração/conversão	Transferência periódica de material digital de uma dada configuração de hardware/software para uma outra, ou de uma geração de tecnologia para outra subsequente. A migração tem como objetivo manter os objetos digitais compatíveis com tecnologias atuais de modo que um utilizador comum seja capaz de os interpretar sem necessidade de recorrer a artefatos menos convencionais, por exemplo, emuladores. A migração ainda continua sendo a estratégia de preservação mais aplicada. Existem diversos tipos de migração: para suportes analógicos, atualização de versões, conversão para formatos concorrentes, distribuída. Normalização, migração a-pedido e migração distribuída.

Migração para suportes analógicos	Consiste na conversão de objetos para suportes não digitais com o intuito de aumentar a sua longevidade. Se dá na reprodução de um objeto digital em papel, microfilme ou qualquer outro suporte analógico de longa duração e concentra os esforços de preservação em torno do novo suporte. Esta estratégia, apenas pode ser aplicada a objetos digitais que possuam uma representação aproximada em suportes analógicos, como por exemplo, documentos de texto ou imagens. Objetos interativos são excluídos deste tipo de estratégias.
Atualização de versões	Estratégia de preservação mais vulgarmente utilizada. Resume-se em atualizar os materiais digitais produzidos por um determinado software recorrendo-o a uma versão mais atual
Conversão para formatos concorrentes	Processo e atualização de versões controlados pela organização que desenvolveu a respectiva aplicação de software. A qualidade da migração depende da capacidade dos importadores fornecidos pelo fabricante e do grau de retrocompatibilidade oferecido pelo novo formato. Existem formatos que não são dependentes de qualquer aplicação de software. Tal, acontece com grande parte dos formatos de imagem (e.g. JPEG, TIFF, PNG).
Normalização	Objetiva simplificar o processo de preservação através da redução do número de formatos distintos que se encontram no repositório de objetos digitais. Havendo um número controlado de formatos, uma mesma estratégia de preservação poderá ser aplicada transversalmente a um maior número de objetos digitais, o que poderá conduzir a uma redução generalizada dos custos de preservação. Por exemplo, existe um leque variado de opções no que diz respeito a formatos para representação de imagens bidimensionais (BMP, GIF, JPEG, PNG, TARGA). Se durante o processo de ingestão, todas as imagens digitais forem convertidas para um único formato, futuras intervenções ao nível da sua preservação poderão ser realizadas de forma mais simples e, conseqüentemente, mais econômica.
Migração a-pedido	Neste tipo de migração, ao invés de as conversões serem aplicadas ao objeto mais atual, estas são sempre aplicadas ao objeto original. Deste modo, se de uma dada conversão resultar um objeto substancialmente diferente do original, numa futura conversão, o problema poderá ser resolvido recorrendo a um conversor de melhor qualidade ou a um formato de destino mais adequado.
Migração distribuída	Neste tipo de migração, existe um conjunto de serviços de conversão que se encontram acessíveis através da Internet e que poderão ser invocados remotamente recorrendo a uma pequena aplicação-cliente. Um repositório de informação digital pode facilmente conter milhares de itens, atingindo níveis de armazenamento na ordem dos múltiplos Terabytes. Transferir através da Internet um volume de informação desta natureza acarreta custos que poderão ser impeditivos para muitas organizações. Para além disso, requisitos em termos de largura de banda, segurança dos dados e tempo de transferência poderão ser fatores determinantes no sucesso deste tipo de estratégias.
Encapsulamento	A estratégia de encapsulamento consiste em preservar, juntamente com o objeto digital, toda a informação necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores. Esta informação poderá consistir, por exemplo, numa descrição formal detalhada do formato do objeto preservado.

A Pedra de Rosetta digital	A descoberta da Pedra de Rosetta conduziu à descodificação de inúmeros outros textos egípcios encontrados nos mais variados locais e suportes (monumentos, rochas, papiros). Nesta estratégia, em vez de se preservar as regras que permitem descodificar o objeto digital, são reunidas amostras de objetos que sejam representativas do formato que se pretende recuperar. Estas amostras deverão existir num formato que possa ser diretamente interpretado pelo ser humano. Trata-se do conjunto de referência, versão grega do decreto inscrito na Pedra de Rosetta. Com esta informação seria possível inferir as regras necessárias para traduzir/converter o objeto original para um qualquer formato contemporâneo. Tal estratégia deverá ser considerada apenas em situações em que todos os esforços de preservação falharam. Trata-se de uma ferramenta de arqueologia digital e não propriamente de uma estratégia de base para preservação de objetos digitais.
-----------------------------------	---

Fonte: Ferreira (2006, p. 31-45).

As estratégias descritas por Ferreira (2006) formam um panorama genérico acerca dos tipos existentes de preservação. Outras estratégias surgiram desde esses postulados, porém as principais utilizadas e disseminadas ainda hoje são: emulação, migração e encapsulamento.

5.1 MODELO DE REFERÊNCIA OAIS/SAAI: NORMA ISO 14721:2003/15427:2007

Na década 1990, o Consultative Committee for Space Data Systems (CCSDS) em conjunto com a International Organization for Standardization (ISO) apresentaram recomendações voltadas para o armazenamento das informações digitais produzidas no âmbito das missões espaciais. Como consequência, foi divulgado em 1999, o primeiro trabalho intitulado: *Reference Model for an Open Archival* (OAIS). Em 2002, o Comitê Consultivo procedeu a uma revisão, avançando os estudos do OAIS para o Padrão ISO 14721:2003. O intuito dessa norma tinha como fim o uso pelas agências espaciais, mas tornou-se modelo de referência universal (CONSULTATIVE..., 2012).

O CCSDS foi criado em 1982 com a finalidade de discutir e desenvolver padrões para o sistema de dados espaciais. A ISO, por sua vez, surgiu 1947 em Genebra, Suíça, sendo a entidade internacional responsável pela aprovação de normas internacionais, tais como: normas técnicas, normas de procedimentos e processos e classificações de países. No Brasil, tem como representante a Associação Brasileira de Normas Técnicas (ABNT).

O OAIS é um modelo conceitual que tem por finalidade identificar componentes funcionais para um sistema de informação de preservação da informação digital a longo prazo. O significado de “longo prazo” está relacionado à preocupação com os impactos de mudanças tecnológicas. Descreve ainda configurações internas e externas do sistema e objetos de informação, bem como padronização de terminologias, conceitos e técnicas de preservação. O objetivo desse modelo é facilitar a compreensão de conceitos para o

arquivamento de objetos digitais, independentemente da tipologia documental, mais ainda, tem a possibilidade de ser adotado por instituições que pretendem padronizar os formatos de arquivo, visando acesso a longo prazo (FONTANA *et al*, 2014).

Traduzida e adaptada no Brasil, a norma ISO 14721:2003 é estandardizada pela ABNT NBR 15472:2007, sendo o Sistema Aberto de Arquivamento de Informação (SAAI) (ABNT NBR 15472, 2007). Este dispositivo pode ser compreendido como mecanismo de organização de informação de pessoas e sistemas, com a responsabilidade de preservar informação e torná-la disponível à comunidade. As informações mantidas no SAAI necessitam de preservação em longo prazo. E este prazo longo é o suficiente para se preocupar com os impactos de mudanças tecnológicas, incluindo suporte a novas mídias, formato de dados, ou mudanças na comunidade usuária (ABNT NBR 15472, 2007). Esta norma é um documento que aborda diversas funções de armazenamento, gerenciamento, acesso e disseminação de pacotes de informações para arquivo. Em sua introdução explana alguns esquemas importantes em seus modelos de referências, tais como:

- a) Esquema para amplo entendimento e compreensão de conceitos arquivísticos necessários para preservação por longo prazo e acesso à informação digital;
- b) Conceitos necessários para que organizações não arquivísticas participem efetivamente da função de preservação;
- c) Esquema, incluindo terminologia e conceitos, para descrever e comparar arquiteturas e operações de arquivos existentes e futuros;
- d) Esquema para descrever estratégias e técnicas de preservação por longo prazo;
- e) Base para comparar modelos de dados de informação digital preservada por arquivos e para discutir como esses modelos e a informação relacionada podem mudar ao longo do tempo;
- f) Fundamento que podem ser expandidos por outros esforços dirigidos à preservação por longo prazo de informações que não se encontram em formato digital (por exemplo, meios físicos e amostras físicas);
- g) Consenso sobre elementos e os processos para preservação por longo prazo e acesso à informação, fomentando um mercado mais amplo para fornecedores;
- h) Identificação e a produção de padrões relacionados ao SAAI.

Observa-se que esses postulados estão em consonância com as funções arquivísticas para preservação da informação, incluindo admissão, arquivamento, gerenciamento de dados, acesso e disseminação. Além disso, aborda ainda a estratégia de migração para novos formatos e mídias, modelo de dados para representação da informação, software para

preservação da informação e o intercâmbio de informação digital entre arquivos. Basicamente define um ambiente de preservação digital por meio de um conjunto mínimo de responsabilidades, para um arquivo ser chamado de OAIS. No Brasil, um arquivo OAIS é interpretado e representado pelo RDC-Arq da Resolução nº 43 do CONARQ (GONÇALEZ, 2017).

Nesse sentido, faz-se necessário entender como estão organizadas as funcionalidades desse modelo e suas interações em um repositório. Primeiro, é importante definir os papéis das entidades OAIS/SAAI que são assinaladas pelo produtor, administração e consumidor (Figura 3).

Figura 3 - Modelo do ambiente OAIS/SAAI



Fonte: ABNT NBR 15472 (2007, p. 9).

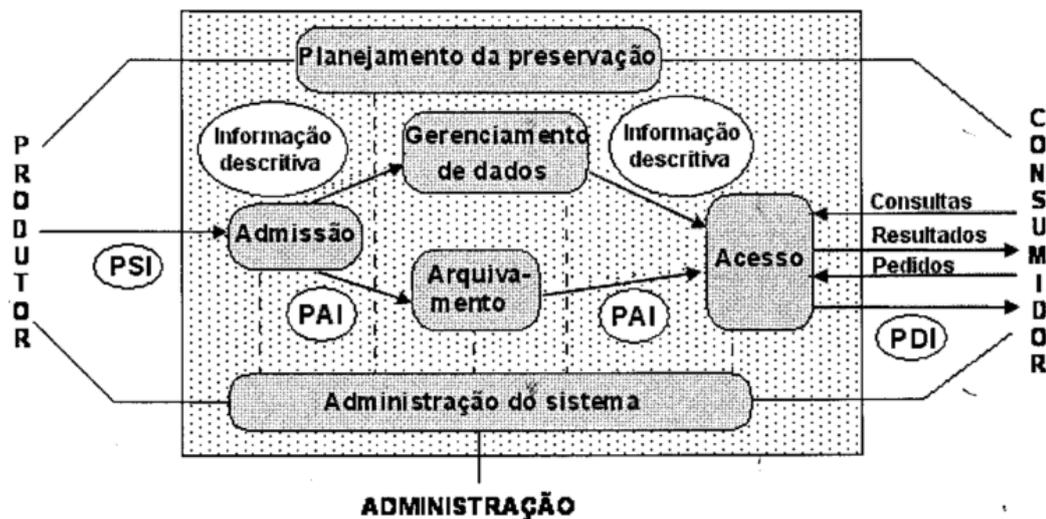
A figura ilustra as entidades do modelo OAIS/SAAI em relação ao usuário e sistemas. Essas entidades representam os pontos de comunicações pelos quais é possível fluir informações, neste caso o produtor, o administrador e o consumidor, terminologias usadas pelo próprio modelo (ABNT NBR 15427, 2007):

- a) O produtor desempenha o papel direcionado por pessoas ou sistemas externos que fornecem informação a ser preservada. Esses mesmos produtores tanto podem ser sistemas de produção de documentos digitais, quanto indivíduos ligados à sua produção;
- b) A administração desempenha o papel que estabelece políticas gerais. Em outras palavras, esta política de gestão de preservação digital deve ser desenvolvida pelo arquivo como apoio à administração a fim de ser a principal fonte de financiamento de um OAIS. Porém, a administração não está envolvida nas operações rotineiras do arquivo;

- c) O consumidor desempenha o papel de pessoas ou sistema que interagem, procurando encontrar informações de seu interesse, através de pedidos de informações ou plataformas que viabilizam o acesso aos documentos.

A Figura 4 ilustra a composição dos principais fluxos de informação do modelo OAIS/SAAI, o qual conecta as entidades, identificando os caminhos de comunicação em ambas as direções em um repositório.

Figura 4 - Entidades funcionais do OAIS/SAAI



Fonte: ABNT NBR 15472 (2007, p. 19).

O produtor é classificado como uma entidade externa ao repositório que se responsabiliza pela submissão do material. Esse material de submissão é o arquivo, entendido aqui como Pacote de Submissão de Informação (SIP).

No processo de submissão ou incorporção (admissão/ingestão), o repositório torna-se responsável por garantir à integridade da informação recebida. Nesta mesma fase se produz toda a informação descritiva que suporte a descoberta e localização do material depositado. Essa informação, também denominada metainformação, é armazenada e gerida pelo componente de gerenciamento de dados. Dessa forma, o material a preservar é denominado Pacote de Informação de Arquivo (AIP)² e será conservado no arquivamento ou repositório de

² Os pacotes de informações para conceber a confiabilidade dos repositórios devem estar de acordo com o modelo de referência OAIS, e se tipificam em três, os quais têm nomenclatura usual em inglês: *Submission Information Package* (SIP), *Archival Information Package* (AIP), *Dissemination Information Package* (DIP).

dados. O componente de admissão constitui a interface entre o arquivo OAIS/SAAI e os produtores de informação.

No que tange ao planejamento de preservação, este se encarrega das políticas de preservação, que é responsável por monitorar o ambiente externo ao repositório, as tendências comportamentais da sua comunidade de interesse ou identificar formatos que se encontram obsoletos.

O componente de acesso, por sua vez, estabelece a comunicação entre o arquivamento ou repositório e sua comunidade de interesse, ou seja, o conjunto de usuários. Este componente facilita a descoberta e localização dos objetos digitais, bem como o seu preparo para a entrega ao usuário.

Os pacotes levados ao consumidor são denominados de Pacotes de Informação e Disseminação (DIP). O último componente, administração do sistema, se responsabiliza pelas operações cotidianas de manutenção, principalmente na monitorização dos processos no interior do arquivamento. Esse interage com todo o restante, de modo a assegurar o correto funcionamento.

Nesse sentido, cada entidade funcional é responsável pelos procedimentos citados. As principais são: admissão, administração de sistema, área de armazenamento, gerenciamento de dados, acesso e planejamento de preservação. Restam, ainda, outros serviços considerados como entidade funcional, porém sua análise não demanda maior aprofundamento dos fluxos de dados OAIS/SAAI.

5.2 REPOSITÓRIOS DIGITAIS: ANTESCEDENTES E CARACTERIZAÇÃO

O surgimento dos repositórios está diretamente relacionado com a denominada “crise dos periódicos” em virtude do aumento do preço das assinaturas, desestabilizando o sistema de comunicação científica em nível mundial. Este acontecimento teve início no final década de 1970, se fortificou nos idos de 1980, e provocou turbulência no sistema formal de comunicação da ciência. Mueller (2006, p. 31) tece considerações sobre o fenômeno e afirma que:

O gatilho da crise foi a impossibilidade de as bibliotecas universitárias e de pesquisa americanas continuarem a manter suas coleções de periódicos e a corresponder a uma crescente demanda de seus usuários, impossibilidade decorrente da falta de financiamento para a conta apresentada pelas editoras, cada ano mais alta, até mesmo que a inflação e outros índices que medem a economia. Isso já vinha acontecendo nos países em desenvolvimento, inclusive no Brasil, cujas bibliotecas já não conseguiam manter suas

coleções atualizadas, mas a crise só detonou quando atingiu as universidades norte-americanas.

Autran e Borges (2014) apontam que os editores, que têm o domínio do mercado editorial, decidiram fundir e/ou adquirir empresas, passando a ter o monopólio, entre elas citam-se: a Reed Elsevier (hoje também conhecida por Relx Group), Harcourt, incluindo a Academic Press; porém o destaque vai para Elsevier, líder no mercado nas áreas das Ciências, Medicina, Tecnologia e Negócios. Em 2003, a editora Betermans Springer foi vendida à Holding, proprietária da Kluwer Academic Publishers, que, ao se unir com a primeira, emergiu com o nome Springer. Estes são verdadeiros conglomerados internacionais que monopolizam, principalmente, a comercialização dos periódicos científicos.

A explosão dessa crise obrigou a busca de alternativas para a comunicação da ciência. Nesse sentido, Stevan Harnard escreveu o artigo *Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry*, que preconizava uma visão futurista de “escrevendo nos céus” – (*skywriting*), prevendo que a disseminação da palavra escrita seria como escrever nos céus, possibilitando que todos vissem e adicionassem comentários como se fosse grafite, mas em uma escala galáctica, ou seja, de uma abrangência que possibilitasse englobar vários autores (MUELLER, 2006).

Essa crise foi o estopim para a emergência do *Open Access* (OA) (acesso livre ou aberto). Por iniciativa de Paul Ginsparg, foi criado o arquivo ou repositório de *preprints* em Los Alamos, no ano de 1991, no Novo México, Estados Unidos. Essa iniciativa conclamava os pesquisadores da área da Física a enviarem seus trabalhos para um repositório central, de onde poderiam ser recuperados por outros interessados. De início, o intuito do projeto era incluir apenas a literatura da área de Física, mas devido a sua aceitação, ampliou-se para outros domínios do conhecimento, como, por exemplo: Ciência da Computação e Matemática. Em 1998, o projeto passou a ser denominado de ArXiv. Neste sistema, as comunicações *preprints* e *posprints* são depositadas por seus autores, os quais também são responsáveis pelo preenchimento dos metadados. Concebido para receber cerca de 100 submissões anuais, seu rápido crescimento em escopo e número de usuários fez com que recebesse, nos seus 18 primeiros meses, 400 submissões - número considerável na ocasião.

Dada a sua aceitação pela comunidade acadêmica, o Arxiv, além das áreas anteriormente referidas, inclui atualmente comunicações sobre Astrofísica, Biologia, Estatística e Finanças. Evidentemente, a introdução desses novos campos científicos contribuiu para a expansão substancial do número de submissões. Após mais de 22 anos de existência deste repositório, os dados de 30 de janeiro de 2014 se referem ao depósito de 911.700 documentos, uma média de 40.250 anuais. (AUTRAN; BORGES, 2014, p. 127).

Atualmente, quem gerencia e financia o ArXiv é a Biblioteca da Universidade de Cornell, e conta com o patrocínio da Fundação Simons.

Em meio a esses acontecimentos emergem novos suportes de informação sob o paradigma do *AO*. Björk (2004) elenca quatro desses tipos de canais:

- a) Periódicos científicos eletrônicos com avaliação prévia pelos pares;
- b) Servidores de *e-prints* para áreas específicas - repositórios para assuntos específicos;
- c) Repositórios institucionais de universidades específicas;
- d) Autoarquivamento em páginas pessoais dos autores.

No paradigma do *OA*, destacam-se três ações pioneiras, que ficaram conhecidas como 3B: Budapest Open Access Initiative (BOAI), Bethesda Statement on Open Access Publishing e Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.

Uma das recomendações da BOAI preconizava a criação de alternativas para publicação: revistas de acesso aberto e repositórios institucionais ou temáticos. Dentro dessas estratégias destacam-se: as vias verde e dourada (*greenroad* e *goldenroad*). Os repositórios são considerados via verde e os periódicos de acesso aberto, via dourada, conforme refere Alves (2008, p. 133, grifo nosso):

A primeira estratégia é a de autoarquivamento – via verde (*greenroad*), que trata do arquivamento que poderá ser realizado pelos próprios autores de artigos científicos já publicados ou aceitos para publicação, obtendo autorização (sinal verde) dos editores que os aceitaram para que possam disponibilizar em um servidor de arquivo aberto. A segunda estratégia trata de via dourada (*goldenroad*), que abrange os periódicos científicos eletrônicos, cujo acesso aberto a seus conteúdos é garantido pelos próprios editores. Sendo assim, a publicação em ambiente de acesso aberto está assegurada no próprio periódico. São essas duas estratégias norteadoras das discussões sobre arquivos abertos.

O OpenDOAR (2006) classifica os repositórios em distintas categorias: temáticos, institucionais, governamentais (repositórios que contêm dados governamentais) e agregadores (arquivo que agrega dados de repositórios subsidiários). Há ainda outros, como o *e-Learning*, da Universidade do Minho, em Portugal.

De acordo com Autran e Borges (2014), os repositórios institucionais foram descritos pela primeira vez por Crow (2002), em documento sobre o posicionamento do Scholarly Publishing and Academic Resources Coalition (SPARC). Os definindo como coleções digitais que capturam e preservam à produção intelectual de uma ou de várias universidades (CROW, 2002).

Outro conceito de repositório encontra-se no relatório da Research Library Group/Online Computer Library Center (RLG/OCLC), sendo aquele composto de informação digital que são coletivamente responsáveis pela acessibilidade em longo prazo do patrimônio social, econômico, cultural e intelectual da nação instanciado em forma digital (RESEARCH..., 2002).

Na concepção de Lynch (2003, p. 89, tradução nossa), os repositórios constituem

[...] um conjunto de serviços que uma universidade oferece aos membros de sua comunidade para a gestão e disseminação de materiais digitais criados pela instituição e os membros da comunidade. É essencialmente um compromisso organizacional para a administração destes materiais digitais, incluindo a preservação a longo prazo, quando apropriado, bem como organização e acesso ou distribuição.

O *e-Learning* constitui-se de conteúdos educativos da TecMinho e da Universidade do Minho, e tem como finalidade armazenar, divulgar e dar acesso aos recursos didáticos e pedagógicos como: apresentações, exercícios, questionários, *packages*, fichas pedagógicas e resumos. Sua implementação iniciou em abril de 2006 no âmbito do projeto “Concepção de e-Conteúdos para *e-Learning*”.

O relatório do *e-Learning* destaca que esse repositório reúne documentos de quase todas as áreas científicas existentes na Universidade do Minho e das áreas de formação da TecMinho, estando em crescimento constante, com novos documentos depositados regularmente e previsivelmente com novas comunidades aderentes (BAPTISTA *et al*, 2006). Apresenta-se no Quadro 7 uma síntese da tipologia dos repositórios.

Quadro 7 - Síntese da tipologia dos repositórios digitais

Repositório temático: voltado à produção de uma instituição, especialmente universidades e institutos de pesquisa. Os quais servem para disseminar de forma ampla o conhecimento científico.
Repositório institucional: voltado às comunidades científicas específicas. Tratam, portanto, da produção intelectual de áreas do conhecimento em particular, podendo armazenar informação de cunho arquivístico, ou seja, da produção acumulada orgânica de uma instituição, como por exemplo, universidade.
Repositório dissertações e teses: lida exclusivamente com o armazenamento e difusão de dissertações e teses.
Repositório e-Learning: objetiva contribuir para preservação e divulgação de produção didática pedagógica, tal como preservar a memória intelectual da Universidade do Minho.
Repositório governamental: refere-se a repositórios que contêm em si dados governamentais.
Repositório agregador: agrega dados de outros repositórios.

Fonte: Adaptado de Sousa Filho *et al* (2012); OpenDOAR (2006).

A rápida obsolescência da tecnologia coloca em risco o acesso e a integridade dos documentos. Com isso, a falta de políticas que garantam a preservação digital no escopo dos arquivos, bem como sua implementação, tanto em nível internacional quanto nacional, tem ameaçado a capacidade humana de continuar utilizando esse mecanismo como fonte de informação confiável, sobretudo, quando se fala em software de gerenciamento informacional.

Os repositórios digitais possuem papel relevante na construção de um espaço arquivístico confiável para manter seus objetos e possuem várias características que dão sustentação à preservação da informação. Dentre as principais, segundo Corujo (2014), citam-se: autenticidade, fidedignidade, integridade e inteligibilidade:

- a) Autenticidade: está direcionada à segurança da informação, avaliando se o cumprimento de seus requisitos legais são coerentes;
- b) Fidedignidade: volta-se para o aspecto da avaliação, se a informação é fidedigna, de acordo com o contexto em que foi criada;
- c) Integridade: relaciona-se àquela informação que não sofreu nenhuma alteração em seu fluxo, isto é, da produção até a chegada ao repositório digital;
- d) Inteligibilidade: permite manter a capacidade de apresentar os elementos essenciais de objetos digitais autênticos.

Essas características se entrelaçam com aquelas atribuídas ao arquivo que, segundo o referido autor, pode levar a uma linha de ambiguidade com o que vem a ser um repositório. Em contrapartida, salienta-se que essas ponderações estão relacionadas ao tratamento do repositório como sinônimo de arquivo (local de armazenamento de informação), pois a subsistência entre os dois são dicotômicas, visto que o arquivo, de acordo com os princípios da Arquivologia, possui elementos de natureza própria, diferindo dos demais: proveniência, organicidade e unicidade, ou seja, um documento digital pode ser ou não arquivístico, dependendo de sua essência. O repositório, por sua vez, pode ser um arquivo ou sistema (local) com caráter de guarda, visando o acesso à informação.

Corujo (2014) afirma que, para dirimir essa ambiguidade, os especialistas da área de tecnologias digitais utilizaram o termo “programa de informação” para substituir os termos arquivos e repositórios digitais, no sentido de conjunto articulado de operações com o objetivo de preservar matérias digitais, incluindo todas as responsabilidades de preservação, desde a política e a estratégia, até a execução. Como exemplo, tem-se o próprio modelo OAIS, enfatizando as principais funções dos repositórios, ou seja, a garantia e continuidade

do seu objeto a longo prazo, fornecendo acesso à informação arquivada de maneira consistente.

Em meio a todo esse escopo de conceitos e significados de repositório, existe um item indispensável, que garante a autenticidade e fidedignidade de suas informações: os metadados. Este foi cunhado para dar significado à informação necessária aos dados armazenados em sistema computacional. De maneira usual esta mesma atribuição passou a designar um dado sobre outro, analogamente, índices, catálogo e inventários (RONDINELLI, 2005).

Barreto (1999) classifica os metadados em três tipos básicos:

- a) metadados para catalogação bibliográfica, cujo objetivo é identificar e descrever os materiais bibliográficos, sendo os mais importantes: o formato *Machine Readable Catalogue* (MARC), o formato eletrônico *Text Encoding Initiative* (TEI) e o formato *Encoding Archival Description* (EAD);
- b) metadados para descoberta de recursos na web, que dão suporte aos motores de busca na indexação dos recursos informacionais existentes na Internet, dentre os quais estão o padrão *Internet Anonymous Ftp Archive* (IAFA), o padrão *Summary Object Interchange Format* (SOIF) e o padrão *Dublin Core* baseado no MARC;
- c) metadados para infraestrutura global de informação, que colecionam recursos informacionais e suas localizações e formas de acesso a estes recursos, destacando-se, nessa categoria de metadado, o *Government Information Locator Service* (GILS) e os padrões para sistemas de informação ambiental como o *Federal Geographic Data Committee* (FGDC) e o *Environmental Data Catalogue* (EDC).

Estes três padrões de metadados são considerados os mais utilizados para descrição da produção científica, na web e no acesso à informação. Ademais, existe outra divisão, relativa ao emprego de documentos de características arquivísticas segmentada em três classes: administrativo, que se subdivide em: descritivo, de preservação, de uso e técnico (estruturais) (Quadro 8).

Quadro 8 - Classes dos metadados

TIPO	DEFINIÇÃO	EXEMPLOS
Administrativo	Metadado utilizado na administração de recursos de informação	<ul style="list-style-type: none"> ▪ Aquisição de informação ▪ Direitos de reprodução ▪ Critérios de seleção para digitalização etc.
Descritivo	Metadado para descrição de recursos de informação	<ul style="list-style-type: none"> ▪ Catalogação de registros ▪ Índices especializados etc.
De Preservação	Metadado utilizado para preservação de recursos de informação	<ul style="list-style-type: none"> ▪ Documentação das condições físicas dos recursos etc
Técnico	Metadado utilizado para conhecer as funções de um sistema ou o comportamento dos metadados	<ul style="list-style-type: none"> ▪ Hardware e software ▪ Dados de segurança ▪ Documentação etc.
De uso	Metadado relativo ao nível e tipo de uso de um recurso de informação	<ul style="list-style-type: none"> ▪ Registros de exibição ▪ Sumário de re-uso e de versões etc.

Fonte: Lourenço (2007, p. 68).

De acordo com Lourenço (2007, p. 68, grifo nosso):

[...] os padrões de metadados incluem metadados com estas três funções distintas: os descritivos, que descrevem o recurso informacional (autor, título, assunto etc); os estruturais utilizados para estruturação visual da biblioteca digital (basicamente *tags* de marcação de HTML); e os administrativos, que controlam datas de acesso ao recurso informacional, de direitos autorais etc. (senhas, data de digitalização, legislação sobre direitos de reprodução etc).

Esses elementos garantem a autenticidade dos documentos, isto é, os aspectos de sua confiabilidade. Nessa perspectiva, os profissionais da informação precisam estar atentos e atualizados à dinâmica das tecnologias, pois, no meio digital, tudo tende a se tornar obsoleto rapidamente.

5.2.1 Repositórios Arquivísticos Digitais

O advento da Internet proporcionou o surgimento de ferramentas que permitem a comunicação e compartilhamento da informação em tempo real.

Nesse sentido, a iniciativa do acesso aberto ao conhecimento científico tem alterado as formas de produção, acesso, disseminação e uso da informação. Tal protagonismo serviu para criação de suportes voltados ao ciclo produtivo e difusão da comunicação: os repositórios temáticos e institucionais (WEITZEL, 2006). A autora ainda reverbera, discorrendo que a configuração dessas mudanças possui, às vezes, problema de caráter estrutural no âmbito dos fluxos informacionais nas instituições científicas, demanda que pode ser sanada por

profissionais especializados para a manutenção de tais sistemas, cujas implicações necessitam de maiores reflexões no escopo da CI.

Destaca-se que os softwares para repositórios digitais devem possuir a recomendação da iniciativa do *Open Archive Initiative* (OAI), adotando o protocolo OAI-PMH (*Protocol for Metadata Harvesting*), no compartilhamento de metadados, tornando seus conteúdos interoperáveis, conforme refere Weitzel (2006, p. 60):

Para isso foi desenvolvido um conjunto de padrões técnicos e tecnológicos (incluindo a OAI-PMH, padrão de metadados, etc.) que sustentou uma infraestrutura para publicação na web a partir de modelos de próprios, conforme o de teses, revistas, etc. que, como vistos anteriormente, são implementados com softwares específicos e no seu agrupamento em repositórios digitais. Segundo a OAI, esses repositórios digitais funcionam como provedores de dados, possibilitando a coleta integrada dos metadados pelas ferramentas de busca comuns, tais como o Google e o Google Scholar, e também por provedores de serviços próprios segundo os moldes da OAI, tais como a OAIster.

A OAI foi criada em 1999 por um grupo de pesquisadores europeus e norte-americanos, sendo responsável pela implementação de um novo modelo de acesso à publicação científica, a partir da integração de soluções tecnológicas e também pela busca de mecanismos para garantir sua legitimação (RODRIGUES, 2015).

Segundo Arellano e Oliveira (2016), as ferramentas para repositórios institucionais não foram originalmente destinadas a projetos de preservação digital, o mote era a garantia da longevidade dos conteúdos digitais. Porém, como a demanda de seu uso serve para representar à produção intelectual de uma instituição, caracterizando-se como sistemas para armazenar, preservar e difundir o acesso à informação, ampliando o acesso e facilitando a preservação da memória institucional (COSTA; LEITE, 2010). Pode-se afirmar que seus objetivos são responsáveis pela visibilidade, tanto institucional como das pesquisas realizadas, as quais impactam por serem disponibilizadas em um canal de comunicação aberto (COCCO; RODRIGUES, 2014).

As organizações que utilizam os repositórios digitais, ditos confiáveis, só recebem esse status quando demonstram concordância com padrões e normas, por exemplo: o modelo OAIS/SAAI, Resolução nº 43 do CONARQ, ISO 16363:2012 e TRAC. Para que isso se torne possível, faz-se necessário a existência da certificação³, assim como de mecanismos de colaboração que facilitem o intercâmbio entre a comunidade envolvida no tratamento de

³ Dispositivos de confiança em um repositório digital, como critérios específicos a serem cumpridos. Se faz aplicável por meio de normas e/ou diretrizes: modelo OAIS; ISO 16363:2012; Resolução nº 43 do CONARQ e a TRAC.

materiais digitais (ARELLANO; OLIVEIRA, 2016). A necessidade do estabelecimento de programas de certificação e critérios para serem usados em auditorias têm sido identificadas no escopo da CI (LYNCH, 2003).

Nesse sentido, vários softwares que envolvem os repositórios digitais são direcionados às instituições, primeiro por sua aquisição nas diversas organizações como protótipo de guarda e racionalização da informação e, também, por serem lugares adequados para testar e formular as metodologias e políticas a serem adotadas. Este axioma se baseia na ideia de que a preservação de objetos digitais deve estar localizada em instituições confiáveis e capazes de armazenar, migrar e dar acesso às coleções como: teses, relatórios, documentos governamentais, para além de materiais suplementares, conjuntos de dados, imagens, visualizações e simulações, comunicação informal como: e-mails, blogs, *podcasts*, websites, *wikis* e apresentações (RESEARCH..., 2002). Nesse sentido, ponderam Arellano e Oliveira (2016, p. 468-469, grifo nosso):

Com o desenvolvimento de pacotes de software, pode-se afirmar que a área de preservação digital está chegando à maturidade. Algumas dessas soluções saíram da área da ciência da informação, como ferramentas e serviços para bibliotecas e arquivos. Entre os primeiros sistemas citados estavam o LOCKSS (*Lots Of Copies Keep Stuff Safe*), o PANDAS, o OCLC Digital Archive, e o DIAS (*Digital Information Archive System*) da IBM. Todos eles procediam de organizações de diferentes perfis: provedores de serviço para biblioteca, grupos de pesquisa de bibliotecas especializadas, equipes de repositórios universitários e bibliotecas nacionais trabalhando em parceria com empresas terceirizadas, mostrando a necessidade de integração entre todos os interessados no assunto.

Ressalta-se que os preceitos arquivísticos estão inseridos no objetivo dos modelos de certificação e normas dos repositórios (CONSELHO..., 2015), especialmente, por envolver a utilização de documentos produzidos e recebidos de instituição, destacando seus princípios, como a organicidade, unicidade, por exemplo. Como afirma Ferreira (2006), o ambiente que trabalha com informação registrada é, por essência, campo de investigação dos arquivos, sobretudo em relação ao elemento de preservação.

Destaca-se que a imersão no ambiente digital não intervém na essência do documento de arquivo, isto é, sua estrutura como documento digital não se altera. Rondinelli (2013) afirma que o documento arquivístico digital é uma unidade indivisível de informação constituída por mensagem fixada em um suporte (registro), o qual é produzido por uma pessoa física ou jurídica, em suas atividades codificadas em dígito binário e interpretável por um sistema computacional.

Os documentos digitais possuem o obstáculo advindo da rápida obsolescência tecnológica do hardware, software, formatos, entre outros problemas, como a manutenção desses, o que pode levar a comprometer sua autenticidade. Dessa forma, a Lei de Arquivos, 8159/1991, informa no seu Art. 25 (p. 1):

Ficará sujeito à responsabilidade penal, civil e administrativa, na forma da legislação em vigor, aquele que desfigurar ou destruir documentos de valor permanente ou considerado como de interesse público e social.

Fazendo-se necessário atentar para importância da implementação de estratégias de confiabilidade para integridade dos documentos, analogamente, os RDC-Arqs. O repositório digital arquivístico deve organizar e recuperar a documentação, mantendo a relação orgânica entre eles, agrupando-os de forma hierárquica como disposto pelo plano de classificação e, para dar acesso, os instrumentos de pesquisa precisam estar pautados na Norma Brasileira de Descrição Arquivísticas (NOBRADE), principalmente no contexto brasileiro (DISTRETTI, [201-], p. 1).

A preservação é um processo contínuo e deve começar na concepção de um documento. Outro procedimento importante é reconhecer que o registro, em forma de metadados, das intervenções da preservação auxiliam na presunção da autenticidade do documento arquivístico digital. [...]. Um documento arquivístico deve estar sempre associado a um repositório arquivístico digital para ser preservado e continuar acessível com o passar do tempo. [...]. É importante salientar que isso não exclui a necessidade de um Sistema Informatizado de Gestão Arquivística de Documentos (SIGAD), pois é através dele que é feita a gestão arquivística de documentos. Um RDC-Arq deve ser interoperável com outros repositórios e com outros sistemas informatizados que tratem de documentos arquivísticos. Para que isso ocorra, o repositório deve estar de acordo com as normas e padrões estabelecidos no Brasil e no mundo. No caso brasileiro, essas normas são estabelecidas pelo CONARQ e nos Padrões de Interoperabilidade do Governo Eletrônico – e-PING.

Nesse contexto, os repositórios vêm se destacando na construção de um espaço arquivístico digital responsável pela custódia de documentos. Entretanto, é fundamental a implementação de medidas para incorporá-los a certificação, no intuito de gerar maior confiabilidade, visto a inconstância do ambiente virtual.

6 DIRETRIZES PARA IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS (RDC-Arq): RESOLUÇÃO Nº 43

O CONARQ é um órgão, vinculado ao Arquivo Nacional do Ministério da Justiça, responsável pela política nacional de arquivos nas esferas pública e privada. Atua como órgão central do Sistema Nacional de Arquivos (SINAR), o qual exerce orientação normativa sobre a gestão documental e a proteção a documentos de arquivo (BRASIL, 2014).

No site⁴ do CONARQ, encontra-se toda a documentação acerca de leis, decretos e resoluções referentes às políticas de arquivo, bem como coletânea de manuais, no intuito de estabelecer diretrizes para a área da Arquivologia, conforme Brasil (2014).

Desde sua criação, o CONARQ tem concentrado seus esforços no sentido de suprir o Brasil de um importante corpus de atos normativos que regulam matérias arquivísticas sobre diversos temas relativos à gestão, à preservação e ao acesso aos documentos públicos. O CONARQ é responsável pela edição de decretos regulamentadores da Lei n. 8.159, e de resoluções que tratam de temas diversos relativos à gestão de documentos convencionais e digitais, microfilmagem, digitalização, transferência e recolhimento de documentos de qualquer suporte, classificação, temporalidade e destinação de documentos, acesso aos documentos públicos, capacitação de recursos humanos, terceirização de serviços arquivísticos públicos, dentre outros. (BRASIL, 2014, p. 1).

Além disso, promove cursos e desenvolve ações técnico-científicas por intermédio de suas câmaras técnicas e setoriais. Seu desempenho é considerado referência para a prática de arquivos tanto em instituições nacionais, quanto na América Latina.

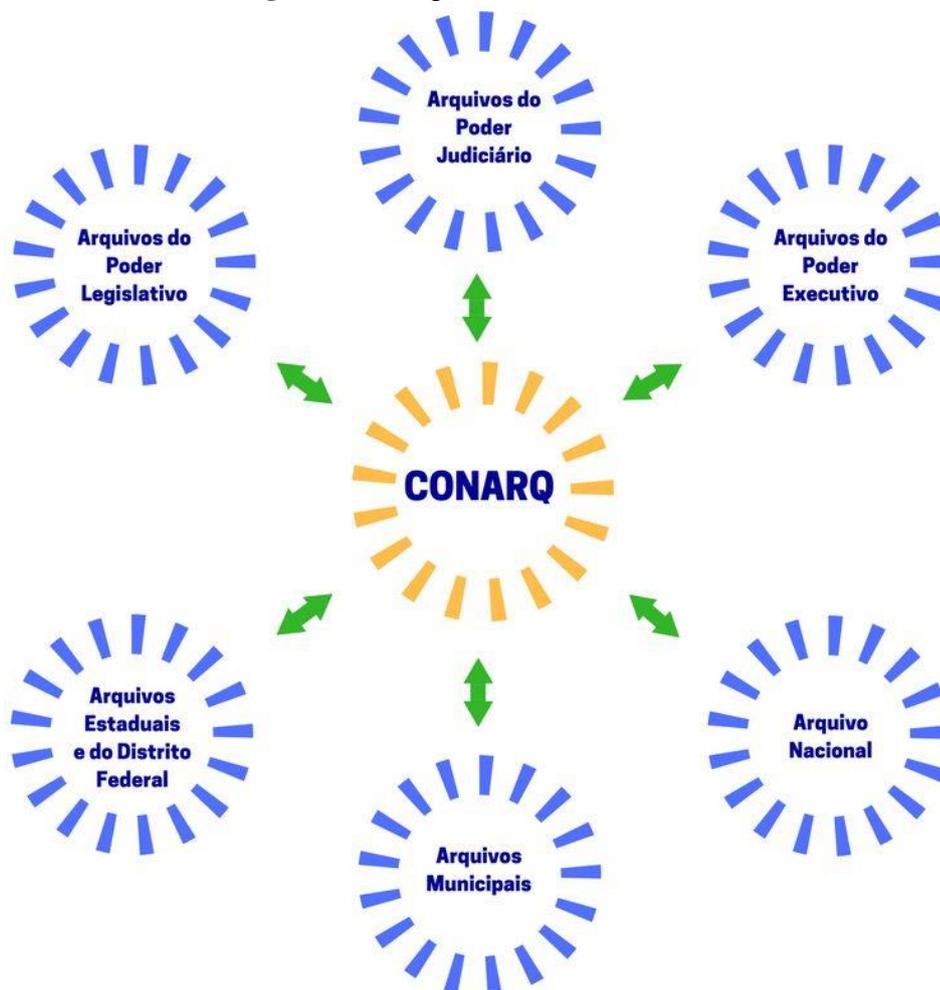
Seu conselho compõe-se por 17 membros, que também formam o plenário: o diretor-geral do Arquivo Nacional (presidente), representantes dos poderes Executivo, Judiciário e Legislativo, em nível federal, dos arquivos públicos estaduais, municipais e do Distrito Federal, das entidades mantenedoras de curso superior de Arquivologia, das associações de arquivistas e das instituições que congreguem profissionais que atuem nas áreas de ensino e pesquisa.

Tal estrutura espelha a convergência de interesse do Estado e da sociedade, compatibilizando questões de responsabilidade funcional e social do poder público perante a gestão e preservação do patrimônio arquivístico. Para a operação eficiente do CONARQ e do SINAR, ficou prevista a criação de câmaras técnicas e comissões especiais para elaborar estudos e normas necessárias à implementação da política nacional de arquivos (BRASIL, 2014).

⁴ Disponível em: <http://conarq.arquivonacional.gov.br/>.

Nesse sentido, o SINAR, criado em 25 de Setembro de 1978, apresenta compromisso com o funcionamento do CONARQ, uma vez que é responsável pela elaboração das normas necessárias à administração nacional dos arquivos. Assim, cabe ao Arquivo Nacional dar suporte administrativo e técnico ao Conselho. Possui como órgão central: o Arquivo Nacional; os arquivos dos poderes Executivo, Legislativo e Judiciário, em âmbito federal, estadual e do Distrito Federal. E quanto ao âmbito municipal, apenas os arquivos dos poderes Executivo e Legislativo (BRASIL, [201-a]). Os integrantes do SINAR seguem as diretrizes e normativas emanadas do CONARQ, como destacado na Figura 5.

Figura 5 - Integrantes do SINAR/CONARQ



Fonte: Brasil ([201-a]).

Os arquivos acima referidos, exceto o Arquivo Nacional, quando organizados sistematicamente, passam a integrar o SINAR por intermédio de seus órgãos centrais. As pessoas físicas e jurídicas de direito privado, detentoras de arquivos, podem integrar o SINAR mediante acordo ou ajuste com o órgão central (BRASIL, [201-a]).

A CTDE, criada em 1995, é um grupo de trabalho que conta com especialistas de vários setores do governo e da sociedade civil, tendo por objetivo definir e apresentar ao CONARQ normas, diretrizes, procedimentos técnicos e instrumentos legais sobre gestão arquivística e preservação de documentos digitais, em conformidade com os padrões nacionais e internacionais (BRASIL, [201-b]). Esses estudos envolvem diferentes áreas do conhecimento, cuja formação tem caráter multidisciplinar, a exemplo dos profissionais de Arquivologia, CI, Biblioteconomia, Administração, Direito e TI, inseridos nas abordagens arquivísticas de documentos digitais. Em suma, a CTDE incumbe-se de prerrogativas voltadas à orbe dos arquivos em âmbito científico e como parâmetro para todas as outras instituições emanadas dos poderes públicos e privados, assim:

Atualmente, a Câmara está tratando dos documentos arquivísticos digitais, isto é, produzidos em formato digital, processados e armazenados por computador. Para isso, desenvolve estudos e análises sobre as iniciativas internacionais e a literatura especializada a fim de subsidiar o desenvolvimento dos seus trabalhos. (BRASIL, [201-b], p.1).

Para a presente pesquisa, utilizou-se a Resolução nº 43 de 04 de Setembro de 2015, que altera a Resolução do CONARQ de nº 39, de 29 de Abril de 2014, a qual estabelece diretrizes para a implementação de repositórios digitais confiáveis, na transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas dos órgãos e entidades integrantes do SINAR.

As mudanças nessa Resolução ocorreram na nomenclatura, pois tais modificações foram representadas pelos termos Repositórios Arquivísticos Digitais Confiáveis, acrescida da sigla “RDC-Arq”, ficando assim: Repositórios Arquivísticos Digitais Confiáveis – RDC-Arq, acrônimo inexistente na Resolução nº 39 (GONÇALEZ, 2017). A autora afirma, ainda, que os repositórios arquivísticos se diferenciam dos demais pela especificidade, ou seja: o cumprimento de requisitos que, obrigatoriamente, assegurem o armazenamento e acesso em longo prazo.

Dessa forma, a Resolução nº 43, está dividida em três partes:

- I. Apresentação: 1. Objetivo 2. Escopo 3. Definições;
- II. Repositório Arquivístico Digital Confiável (RDC-Arq) – principais requisitos: 1. Considerações sobre um repositório digital de documentos arquivísticos 2. Requisitos para um repositório digital confiável;
- III. Padrões e normas de referência: 1. Modelo de referência OAIS; 2. Relatório da Research Library Group (RLG) e da Online Computer Library Center (OCLC) –

Repositórios digitais confiáveis: atributos e responsabilidades; 3. Certificação e auditoria de repositórios confiáveis: critérios e *checklist* – TRAC; 4. Requisitos técnicos para entidades de auditoria e certificação de organizações candidatas a serem repositórios digitais confiáveis; 5. Metadados de preservação – PREMIS; 6. Norma Geral Internacional de Descrição Arquivística – ISAD(G); 7. Norma Brasileira de Descrição Arquivística – NOBRADE; 8. Metadados do e-ARQ Brasil; 9. Protocolo para coleta de metadados – OAI-PMH; 10 Padrão de codificação e transmissão de metadados – METS; 11. Descrição arquivística codificada – EAD (CONSELHO..., 2015).

Na primeira divisão, trata de conceitos referente ao documento de arquivo digital, SIGAD e RDC-Arqs. O SIGAD tem como finalidade gerenciar os documentos digitais em fase corrente e intermediária, garantindo o controle de seu ciclo para o cumprimento da destinação e manutenção da autenticidade e relação orgânica (BRASIL, 2011).

Para a guarda permanente, ocorre alteração na cadeia de custódia, passando a responsabilidade pela preservação dos documentos dos produtores para a instância de guarda. Em fase permanente, os documentos são dependentes de um eficiente sistema informatizado que apoie o tratamento técnico adequado, incluindo arranjo, descrição e acesso, de forma a assegurar a manutenção da autenticidade e da relação orgânica desses documentos (CONSELHO..., 2015).

Nas fases corrente, intermediária e permanente, a preservação desses documentos digitais deve estar associada a um repositório digital confiável nos parâmetros arquivísticos, segundo à Resolução supracitada.

Algumas iniciativas em nível internacional indicam a importância do desenvolvimento dos repositórios digitais confiáveis. Dentre elas, aponta-se o grupo de trabalho liderado pelo RLG/OCLC.

Para que um repositório desempenhe suas funções de maneira coerente faz-se necessário um significativo investimento em infraestrutura tecnológica para o arquivamento e para a preservação digital. Para isso, o CONARQ apresentou diretrizes na Resolução nº 43 de RDC-Arq, assim como a iniciativa para criação do OAIS:

Desde a década de 1990, a comunidade internacional tem se envolvido com atividades no sentido de orientar a implementação de repositórios digitais e de apontar requisitos coerentes à sua composição. Tanto que, essas premissas são fundamentais para assegurar a preservação, acesso e autenticidade de longo prazo dos materiais digitais. A preocupação com a confiabilidade dos repositórios digitais foi evidenciada no relatório da Task

Force on Archiving of Digital Information, uma ação cooperativa do RLG e da Commission on Preservation and Access, publicado em 1996, no qual se declarou que ‘um componente crítico da infraestrutura de arquivamento digital é a existência de um número suficiente de instituições confiáveis, que sejam capazes de armazenar, migrar e prover acesso a acervos digitais’ (CONSELHO..., 2015, p. 8, grifo do autor).

Esse relatório estimulou a colaboração do RLG/OCLC, em março de 2000, no intuito de definir sua formalidade como bases conceituais e os atributos de um repositório digital confiável. Como resultado desse trabalho obteve-se outro relatório sob o título TRAC, publicado em 2002. Mais adiante, o RLG, dando continuidade a esse trabalho celebrou parceria com o NARA, no objetivo de definir critérios para certificação de repositórios confiáveis (CONSELHO..., 2015).

Em 2007, o relatório TRAC foi publicado, apresentando um conjunto de critérios e um *checklist* para a certificação de repositórios digitais confiáveis. Esse mesmo documento serviu de base para a elaboração da norma ISO 16363:2012, a qual lista os critérios que um repositório digital confiável deve atender. O CONARQ pautou-se nessa norma ISO para a criação dos requisitos para um repositório digital considerado confiável (GONÇALEZ, 2017).

A ISO 16363:2012 viabiliza auditoria, avaliação e certificação dos repositórios digitais. Ou seja, aponta a documentação necessária para o cumprimento de auditoria: os requisitos dos auditores para o processo de certificação de confiança relacionado ao repositório. A norma possui três categorias que abrangem 105 critérios: Infraestrutura organizacional; Gerenciamento de objetos digitais e Infraestrutura; e Gestão de riscos de segurança (CONSULTATIVE..., 2012).

A finalidade dessa norma é avaliar a responsabilidade da gestão dos objetos digitais de um repositório. Gonzalez (2017) reitera que as interfaces dos repositórios possibilitam não somente o acesso rápido e confiável, mas conectam seus serviços e coleções às comunidades usuárias.

A Resolução nº 43 especifica os diferentes tipos de repositórios, conforme ilustra o Quadro 9.

Quadro 9 - Tipos de repositórios de acordo com Resolução de nº 43

Repositório digital: ambiente de armazenamento e gerenciamento de materiais digitais, constituindo uma solução informatizada em que os objetos são capturados, armazenados, preservados e acessados. É formado por hardware, software e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos. Pode ser empregado em: arquivo corrente e intermediário (em associação com um SIGAD), arquivo permanente, biblioteca digital, acervo de obras de arte digitais, depósito legal de material digital e curadoria de dados digitais de pesquisa.

Repositório digital arquivístico: armazena e gerencia documentos nas fases corrente, intermediária e permanente. Deve: gerenciar os documentos e metadados de acordo com as práticas e normas da Arquivologia, especificamente relacionadas à gestão documental, descrição arquivística multinível e preservação e proteger as características do documento arquivístico, em especial a autenticidade (identidade e integridade) e a relação orgânica entre os documentos.

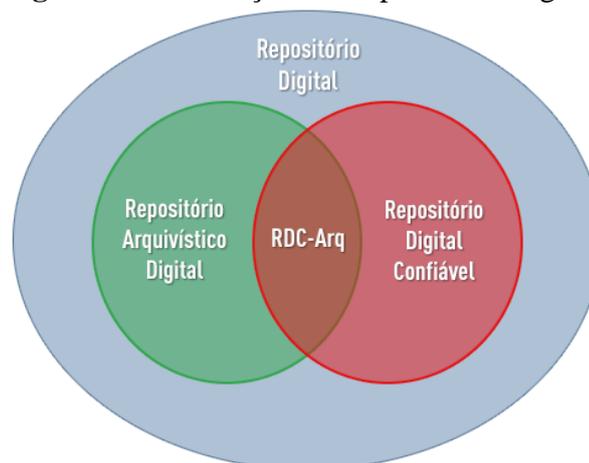
Repositório digital confiável: capaz de manter autênticos os materiais digitais, de preservá-los e prover acesso a eles pelo tempo necessário. Respalda-se segundo o relatório TRAC.

Repositório arquivístico digital confiável: deve ser capaz de atender aos procedimentos arquivísticos em suas diferentes fases e aos requisitos de um repositório digital confiável.

Fonte: CONSELHO... (2015, grifo nosso)

O Quadro 9 mostra as peculiaridades dos tipos de repositórios digitais, de confiança e arquivístico. Note-se que os elementos descritos nos RDC-Arqs se sobrepõem aos demais, englobando todas as suas características. Desse modo, a Figura 6 exibe à orbe referente a convergência dos repositórios digitais:

Figura 6 - Intersecção dos repositórios digitais

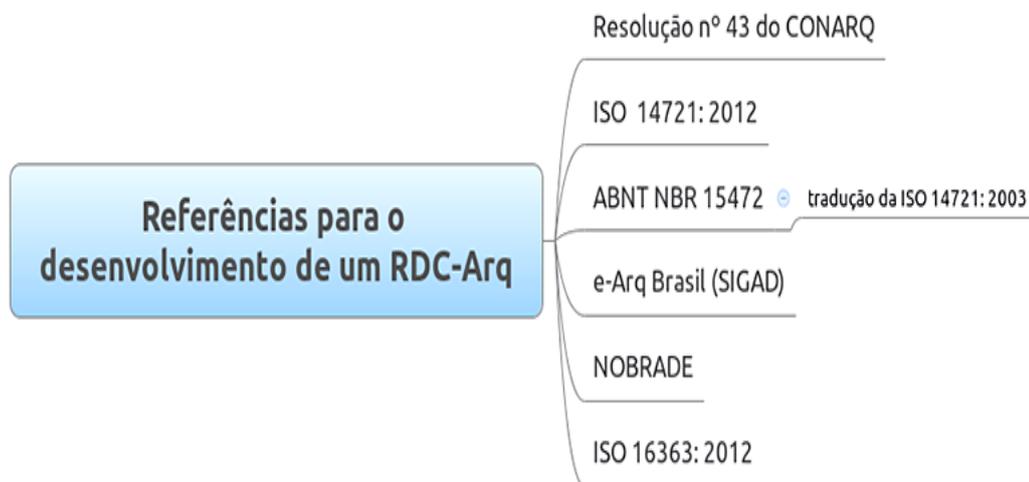


Fonte: Distretti ([201-], p. 1).

Observou-se, na Figura acima, a intersecção do RDC-Arq. O repositório digital funciona como ambiente de armazenamento e gerenciamento de materiais digitais, sendo comum aos sistemas de informação. O repositório digital arquivístico armazena e gerencia documentos nas fases corrente, intermediária e permanente, mas não necessariamente compõe

os aspectos de confiabilidades, isto é, as normas de preservação. O repositório digital confiável, é capaz de manter autênticos os materiais digitais, respaldando-se em diretrizes, cujas características arquivísticas são inexistentes. Porém, os RDC-Arqs abrangem todos os elementos citados. As referências acerca do RDC-Arq apontam os critérios que precisam ser colocados em prática para criação e desenvolvimento deste tipo repositório (Figura 7).

Figura 7 - Referências para desenvolvimento de um RDC-Arq



Fonte: Distretti ([201-], p. 1).

Os requisitos apresentados na segunda parte da Resolução nº 43, definem-se em nível conceitual e devem ser cumpridos no desenvolvimento de um repositório, os quais estão baseados na norma ISO 16363:2012, envolvendo todos os tipos de materiais digitais, inclusive os documentos arquivísticos. Esses requisitos dividem-se em três conjuntos: infraestrutura organizacional; gerenciamento do documento digital; e tecnologia, infraestrutura técnica e segurança.

O primeiro grupo subdivide-se em: a) Governança e viabilidade organizacional b) Estrutura organizacional e de pessoal; c) Transparência de procedimentos e arcabouço político; d) Sustentabilidade financeira; e) Contratos, licenças e passivos. Esse grupo tem como objetivo listar critérios, dispondo das funcionalidades dos repositórios referentes às políticas e estruturação extrínseca.

O segundo subdivide-se em: a) Admissão (captura de documentos digitais); b) Admissão (criação do pacote de arquivamento); c) Planejamento da preservação; d) Armazenamento e preservação/manutenção do AIP; e) Gerenciamento e informação f) Gerenciamento de acesso. Seu objetivo é indicar que o gerenciamento dos documentos deve estar de acordo com o modelo de referência OAIS, que estabelece a formação de pacote de

informação dos documentos digitais (informação de conteúdo) e seus metadados (informação de representação) (CONSELHO..., 2015).

O terceiro e último grupo diz respeito: a) Infraestrutura do sistema; b) Tecnologias apropriadas; c) Segurança. Tais requisitos não determinam hardware e software específicos para garantir a preservação a longo prazo dos AIPs, contudo, descrevem as melhores práticas das áreas de gestão de dados e segurança, que devem ser atendidas por um repositório digital confiável (BRASIL, 2015).

A terceira parte da norma apresenta documentos de referência para construções de RDC-Arq, os quais já foram mencionados: o modelo OAIS/SAAI, a ISO 16363:2012 e o TRAC, que na redoma dos repositórios listam seus aspectos para certificação de confiabilidade. Assim também, como a utilização de metadados para interoperar com outros sistemas.

7 SOFTWARES LIVRES

A aplicação e manuseio de um repositório digital se dá por meio de sistema computacional, como banco de dados e software livre, por exemplo. A qualificação deste último insere-se na gratuidade e repasse do código-fonte, o que difere de um sistema considerado proprietário, privativo, ou não livre, no qual há uma licença com direitos exclusivos ao produtor. A liberdade de uso, cópia, modificações e redistribuição, se torna efetiva através da distribuição do código-fonte, disponível para utilização pelos interessados (HEXSEL, 2002).

Desde 1997, os softwares livres se tornaram conhecidos em virtude da iniciativa dos pesquisadores para o livre acesso ao conhecimento e também devido a reportagens frequentes em diversos veículos de comunicação. Apesar de se apresentar como um fenômeno recente, a disponibilização do código-fonte remonta à década de 60, e a partir de então, seu desenvolvimento tem sido contínuo, com períodos de maior produtividade em meados de 1980 até hoje (HEXSEL, 2002).

No início dos anos 1980, Richard Matthew Stallman, cientista do Laboratório de Inteligência Artificial do Massachusetts Institute of Technology (MIT) (Instituto de Tecnologia de Massachusetts), vivenciou o fim das comunidades que tentavam sobreviver a uma sociedade da informação coletiva, onde havia compartilhamento de programas de computador e difusão do conhecimento tecnológico. Esse pesquisador, caracterizado como fundador do software livre assistiu à difusão, por grandes empresas privadas de softwares, de contratos de não-compartilhamento, que passaram a coibir a colaboração entre desenvolvedores de programas de computador (STALLMAN, 2009).

Nesse cenário, Stallman (2009) estimulou um trabalho em prol da cooperação e colaboração de programação por todo mundo, a fim de que contribuíssem com conhecimento e compartilhamento de experiências. Sua intenção era manter a cooperação entre comunidades, onde todos teriam acesso ao código-fonte dos programas desenvolvidos, surgindo assim o software. Stallman (2009), fundou a Free Software Foundation (FSF) e o projeto de sistema operacional livre, chamado *General Public Licence* (GNU).

Para que um programa seja considerado aberto, Stallman (2009) lista alguns requisitos específicos, denominados de quatro liberdades: liberdade de usar o programa para qualquer propósito (liberdade 0), de estudar como o programa funciona e modificá-lo de forma a que ele venha a atender às suas necessidades específicas (liberdade 1), de redistribuir cópias, com a finalidade de ajudar sua comunidade (liberdade 2), e a liberdade de distribuir cópias da

versão modificada (liberdade 3). Para que se possa alcançar a liberdade 1 e 3, precisa-se da disponibilização do código-fonte. A liberdade não inviabiliza sua comercialização, um programa desenvolvido em modelo livre, pode ser pago (LESSIG, 2004).

A liberdade para usar, copiar, modificar e redistribuir software livre lhe confere uma série de vantagens sobre o software proprietário, uma vez que a disponibilidade do código-fonte evita que usuários se tornem reféns de tecnologias proprietárias.

A comunidade de desenvolvimento desses programas está espalhada por todo mundo. Estimava-se, no ano de 2002-2003, a participação de mais de 100 mil programadores e projetistas, a maioria trabalhando voluntariamente, para além de mais de 10 milhões de usuários regulares de sistemas operacionais e aplicativos distribuídos, o que conseqüentemente está compondo um novo e fundamental ramo da economia (OLIVEIRA *et al*, 2004). Nessa perspectiva Hexsel (2002, p. 3, grifo nosso) afirma:

A importância estratégica do software livre é evidenciada pela publicação de dois estudos sobre o assunto, um encomendado pela Comunidade Europeia, e outro encomendado pela Presidência dos Estados Unidos. Os dois relatórios contêm recomendações ao incentivo da ampliação do uso de software livre, e ao financiamento pelo governo, direta ou indiretamente, do desenvolvimento de sistemas baseados em software livre. O software proprietário (não-livre) geralmente é produzido com a finalidade de obtenção de lucro e, portanto, está sujeito a três tipos principais de pressões de mercado: [1] inclusão de funcionalidades imprescindíveis e frequentemente inúteis, [2] obsolescência programada para possibilitar a venda de novas versões, e [3] prazos de desenvolvimento e testes muito curtos para atender às pressões já mencionadas.

Note-se que Hexsel (2002) chama a atenção do papel que o governo deve desempenhar no sentido de incentivar não apenas a adoção, mas também, estimular o financiamento para o desenvolvimento do software livre. Alerta o autor que:

O poder de compra dos três níveis de governo, e em especial o de esfera federal, deve ser empregado para criar a demanda pelo desenvolvimento de aplicativos que atendam às suas necessidades nas áreas de gestão, tributação, educação e governo-eletrônico, entre outras (HEXSEL, 2002, p. 1).

Saliente-se, ainda, que existem alguns tipos de software que podem ser utilizados pelas instituições, em nível público ou não, e sua utilização depende da necessidade (Quadro 10).

Quadro 10 - Tipos de softwares

Software de domínio público	Software sem <i>copyright</i> . Alguns tipos de cópia, ou versões modificadas, podem não ser livres porque o autor permite que restrições adicionais sejam impostas na redistribuição do original ou de trabalhos derivados.
Softwares semi-livre	Software que é não-livre, mas é concedida a permissão para que indivíduos o usem, copiem, distribuam e modifiquem, incluindo a distribuição de versões modificadas, desde que o façam sem o propósito de auferir lucros. Exemplos de softwares semi-livre são as primeiras versões do <i>internet explorer</i> da Microsoft, algumas versões dos <i>browsers</i> da Netscape, e o StarOffice.
Freeware	O termo <i>Freeware</i> não possui uma definição amplamente aceita, usado com programas que permitem a redistribuição, mas não a modificação, e seu código-fonte não é disponibilizado. Estes programas não são software livre.
Shareware	Software disponibilizado com a permissão para que seja redistribuído, mas à sua utilização implica no pagamento pela sua licença. Geralmente, o código fonte não é disponibilizado e, portanto, modificações são impossíveis.
Software proprietário	Software proprietário é aquele cuja cópia, redistribuição ou modificação são em alguma medida proibidos pelo seu proprietário. Para usar, copiar ou redistribuir, deve-se solicitar permissão ao proprietário, ou pagar para poder fazê-lo.
Software Comercial	Software desenvolvido por uma empresa com o objetivo de lucrar com sua utilização. ‘comercial’ e ‘proprietário’ não são o mesmo. A maioria do software comercial é proprietário, mas existe software livre que é comercial, e existe software não-livre não-comercial.

Fonte: Adaptado de Hexsel (2002).

Observa-se, assim, que é possível obter qualquer um desses programas para utilização de repositório digitais, gratuito ou não. As instituições de domínio público utilizam dos softwares livres para dar acesso à sua produção intelectual, em especial às universidades (INSTITUTO..., 200-?). Com isso, subsiste um leque de opções no sentido dos softwares livres de acordo com os princípios arquivísticos, então: Agorum Core, Alfresco, Archivisticabox, Maarch Network, Owl Intranet, Nexeo DM, ArchOn, Roda, AtoM, SepiaDES, Dspace, Eprints, Greenstone, Fedora, Nou-Rau, *Archivematica*, incluindo a primeira e segunda fase documental.

Constata-se que os repositórios digitais são, hoje, um veículo de comunicação democrático, alcançando os diversos setores da sociedade no acesso livre, impondo o reconhecimento da importância dos atributos de preservação digital e suas políticas para o estabelecimento de sistemas de gerenciamento confiável de informação digital.

8 SOFTWARE ARCHIVEMATICA

No ano de 2007, a UNESCO publicou um relatório intitulado *Rumo ao Repositório e Preservação de Sistema Open Source*. Neste documento, recomendava-se que preservar objetos digitais era imprescindível para futuras gerações, sugerindo o uso de ferramentas que trabalhassem nesse sentido.

Emergia, assim, a necessidade de criação de um sistema de código aberto que englobasse a ideia de preservação digital como benefício para comunidades e instituições em nível mundial. Para este feito, a UNESCO ofereceu apoio financeiro para iniciar o que se tornaria o *Archivemática*. Este sistema torna possível a implementação de preservação da informação na web, servindo como um ambiente de repositório confiável, onde instituições podem salvaguardar a autenticidade de seus registros digitais ao longo do tempo (VAN GARDEREN, 2010).

Em 2008, o arquivo municipal da cidade de Vancouver aderiu às recomendações do mencionado relatório contratando a empresa canadense Artefactual Systems para projetar e desenvolver um sistema utilizando o modelo de referência ISO-OAIS.

Costa e Castro (2016, p. 113, grifo do autor) sinteticamente relatam como ocorreu tal criação:

Foi desenvolvido pela empresa canadense Artefactual Systems, tendo a colaboração de algumas instituições e projetos: da UNESCO, por meio do *Memory of the World's – Subcommittee on Technology*; do Arquivo Municipal da cidade de Vancouver – Canadá; da Biblioteca da Universidade de British Columbia - Canadá; do Arquivo Central de Rockefeller - Canadá; do Arquivo da Universidade Simon Fraser University - Canadá, e de outros colaboradores.

No ano de 2010, Peter Van Garderen, arquivista da Artefactual Systems, apresentou o *Archivemática* à comunidade internacional através de seus estudos na iPres⁵ com o trabalho: *Archivemática: using micro-services and open-source software to deliver a comprehensive digital curation solution*. A repercussão pela comunidade foi das melhores, culminando com o apoio financeiro de várias instituições (VAN GARDEREN, 2010). Em 2012, Van Garderen e Courtney Mumma, envolvidos no projeto do *Archivemática*, publicaram, também no iPres, o artigo intitulado: *the Community-driven Evolution of the Archivemática Project*, no intuito de disseminar o conhecimento acerca desse software como forma de proporcionar maior visibilidade na comunidade universal (VAN GARDEREN, COURTNEY, 2013).

⁵A iPres (Conferência Internacional sobre Preservação de Objetos Digitais) é organizadora de livros que reúne diversos trabalhos científicos em formato de artigo para disseminar estudos sobre a temática preservação digital.

Os requisitos para o desenvolvimento do *Archivematica*, desde sua gênese, foram elaborados na base de um diálogo permanente dentro de sua comunidade, analisando normas e estratégias que melhor compusessem sua estrutura e a capacidade para que os profissionais que a usassem e pudessem implementá-las (RODRIGUES, 2015).

Além do OAIS, os principais padrões de metadados inseridos nesse sistema são: *Metadata Encoding Transmission Protocol* (METS), *Preservation Metadata: Implementation Strategies* (PREMIS) e *Dublin Core*. Dos modelos apresentados, ressalta-se que o METS é um esquema *eXtensible Markup Language* (XML) projetado como uma infraestrutura para codificar todos os tipos de metadados associados a um objeto digital; o PREMIS é um conjunto de elementos de meta-informação capazes de dar suporte às atividades relacionadas à preservação digital e o *Dublin Core* é um conjunto de metadados composto por 15 elementos para descrição (LAMPERT, 2016).

Um dos destaques do uso do *Archivematica* é a sua integração com diferentes sistemas de acesso e armazenamento. Esses princípios contam com a interoperabilidade⁶ a outras ferramentas de gestão (AtoM e CONTENTdm) e diferentes arquiteturas de armazenamento (dispositivos de armazenamento de rede, LOCKSS, armazenamento em nuvem).

Vale salientar que o projeto-piloto do *Archivematica* tinha, como um de seus pressupostos, a integração de *pipeline* (técnica de hardware que permite que a *Central Process Unit* - CPU realize a busca de uma ou mais instruções, além da próxima a ser executada) de processamento com sistemas como o DSpace, CONTENTdm, Fedora, AtoM (que foi desenvolvido em conjunto e já conta com o pacote *Archivematica*) e Archivists' Toolkit arquivistas (MARTINEZ *et al*, 2017).

Como esses programas são interoperáveis, os indivíduos que interagem com os sistemas descritos têm a garantia do acesso à informação, ao mesmo tempo em que as instituições possuem a garantia da preservação do documento. Além disso, o *Archivematica* realiza, automaticamente, a migração ou conversão de formatos dos documentos digitais armazenados para preservação digital, e apresenta um diferencial: a possibilidade de construção de um planejamento para preservação (JORENTE *et al*, 2016).

A página eletrônica do *Archivematica*⁷ oferece aos usuários uma *interface* dinâmica e intuitiva, composta por abas seções/links, tornando fácil a navegabilidade (Figura 8).

⁶ Capacidade de um sistema (informatizado ou não) de se comunicar com outro sistema (semelhante ou não), utilizando de padrões abertos ou ontologias através de plataformas de hardware e software.

⁷ Disponível em: <https://www.archivematica.org/pt-br/>.

Figura 8 - Página inicial do Archivematica



Fonte: Artefactual ([entre 2008 e 2019]).

Encontram-se, ainda, informações sobre o sistema, como o tutorial de instalação, material para downloads, desenvolvimento, notícias, *wiki* e demo. Atenta-se para aba *wiki*, cujo conteúdo está disponibilizado aos usuários: fóruns de discussão e direcionamento a redes sociais (Twitter), ou seja, é o local onde utilizadores podem trocar informações e experiências sobre o programa (Figura 9).

Figura 9 - Página wiki do Archivematica



Fonte: Archivematica (201-).

Essa página também direciona a todos os links do *Archivematica*, conforme se pode observar no lado direito da imagem. Apesar de ainda não possuir tradução em língua portuguesa, a arquitetura do site facilita a navegabilidade em função do modo como as abas estão dispostas.

Destaca-se a utilização de três pacotes distintos: o SIP pelo produtor a um OAIS para construção de um ou mais AIP, o DIP, derivado de um ou mais AIP, recebido pelo consumidor em resposta a uma requisição dirigida ao OAIS.

O SIP é o produtor da informação com o objetivo de criação ou atualização dos AIPs, que consiste nas informações do conteúdo a serem preservadas. A partir dos AIPs é gerado o DIP, sendo este utilizado para o acesso. Entendendo o conceito de pacotes do modelo OAIS, o *Archivematica* é administrado através das entidades funcionais (COSTA; CASTRO, 2016).

Costa *et al* (2016, p. 14-15, grifo nosso) descrevem sucintamente o *Archivematica*:

Archivematica é um sistema de preservação digital de software livre, gratuito e de código aberto projetado para manter os dados baseados em padrões de preservação digital e o acesso em longo prazo para coleções de objetos digitais. O *Archivematica* é desenvolvido com o gerenciador de conteúdo AtoM, sistema baseado na web para acesso aos seus objetos digitais, reconhecido arquivisticamente como uma plataforma de descrição, difusão e acesso.

Compreende-se, assim, que sua função é armazenar a documentação em formato digital, de acordo com os padrões exigidos para preservação. Por possuir código aberto, funciona sob a licença do *Affero General Public License* (GNU 3.0) e licença *Creative Commons* (Atribuição-Compartilhamento). O principal objetivo destas duas é a liberdade de copiar, distribuir e modificar sistemas.

O formato de software livre do *Archivematica* configura-se como eficiente estratégia de preservação digital, pois a obsolescência tecnológica é uma preocupação atual em relação à condição dos documentos eletrônicos. Como já visto, outra vantagem é a interoperabilidade desse sistema, ou seja, é possível criar vias de comunicação com outros tipos de software desde que os códigos dos sistemas estejam disponíveis. Isto se dá entre o AtoM e o *Archivematica* de maneira muito fluida (JORENTE *et al*, 2016).

A página eletrônica do *Archivematica*⁸ oferece aos usuários uma *interface* dinâmica e intuitiva, composta por abas seções/links, tornando fácil a navegabilidade. Enquanto sistema de preservação digital confiável, o *Archivematica* possui características que atestam seus aspectos de preservação: modelo OAIS, consonância com as normas de arquivos,

⁸ Disponível em: <https://www.archivematica.org/pt-br/>.

interoperabilidade, acesso e difusão de informação, estratégias e planejamento de preservação, conforme mostra a Quadro 11.

Quadro 11 - Principais características do *Archivematica*

Característica \ Repositório	<i>Archivematica</i>
Desenvolvimento	Desenvolvido pela <i>Artefactual Systems</i> com o apoio do projeto Memória do Mundo, da UNESCO, e indicado pelo CIA.
Modelo OAIS	Desenvolvido de acordo com a norma.
Preservação	Estratégias de preservação digital.
Planejamento da Preservação	Possibilidade de customização de acordo com as necessidades da entidade produtora.
Navegação multinível	Não possui
Interoperabilidade	SIGAD, <i>ICA-AtoM</i>
Acesso e difusão	<i>Upload</i> do pacote DIP para o <i>ICA-AtoM</i>

Fonte: Rodrigues (2015, p. 102).

Atualmente, o sistema possui 11 versões, sendo a última 1.8.1. Como principais colaboradores têm-se: UNESCO, City of Vancouver Archives, Harvard Business School Baker Library, Museum of Modern Art (MoMA), University of Alberta Libraries, University of British Columbia Library, Rockefeller Archive Center, Simon Fraser University Archives and Records Management, Yale University Library, Zuse-Institute Berlin, Council of Prairie and Pacific University Libraries (COPPUL), Bentley Historical Library, University of Michigan, Duraspace, MIT Libraries, Ontario Council of University Libraries, National Library of Wales, Canadian Council of Archives (ARCHIVEMATICA, 201-).

No Brasil, a utilização do *Archivematica* pelas instituições arquivísticas ainda é recente (FONTANA *et al*, 2014; JORENTE *et al*, 2016; LAMPERT, 2016). Estes ainda supõem, que tal fator é uma das razões para a escassez de literatura em língua portuguesa.

8.1 ESPECIFICIDADES E CARACTERÍSTICAS DO ARCHIVEMATICA

A composição e arquitetura do *Archivematica* está baseada no modelo de referência OAIS e em padrões de metadados para o estabelecimento de um RDC-Arq. Reitera-se que este ambiente possibilita a preservação e o acesso, pelo tempo que se fizer necessário, nas idades corrente, intermediária e permanente dos documentos de arquivo.

Segundo Costa *et al* (2016), para o arquivo em segunda fase, esse tipo de repositório deve gerenciar documentos e metadados de acordo com os princípios e as práticas da Arquivologia, especificamente relacionados à gestão documental, descrição arquivística multinível e preservação (identidade e integridade) e a relação orgânica; proteger as características desses documentos; preservar e dar acesso por tempo indeterminado, estando em conformidade com a ISO 16363:2012.

A tradução do *Guia do usuário Archivematica* realizado pelo IBICT em 2016 teve o intuito de disponibilizar às instituições brasileiras um tutorial para facilitar a compreensão e o manuseio do software. Com essa mesma finalidade, cita-se também o *Guia de instalação e configuração* do referido, o qual tem o objetivo de auxiliar os profissionais de TI/Computação (MARTINEZ *et al*, 2017). Conforme refere Aguiar (2018), manipular este tipo de tecnologia requer know-how de agentes capacitados.

Outro fator a destacar nos ambientes de sistemas de preservação digital é o da ininterrupção da cadeia de custódia, o qual faz um documento de arquivo corrente tornar-se íntegro. Flores, Rocco e Santos (2016) adaptam o conceito de cadeia de custódia de Jenkinson (1965), para o novo cenário do ambiente digital no qual perpassa todo seu o ciclo. Esta significação é uma maneira de lumiar a confiabilidade dos documentos arquivísticos, compreendendo suas três fases, promovida por intermédio da instituição de custódia e seus mecanismos de preservação digital, ou seja, os sistemas de RDC-Arqs e as plataformas de acesso interoperáveis. A Figura 10 esquematiza o ciclo de vida em cadeia de custódia dos documentos:

Figura 10 - Ciclo de vida dos documentos em cadeia de custódia ininterrupta

Ciclo de vida dos documentos, as 3 idades, Lei 8.159 (Lei de Arquivos), a 12.527, a LAI de acesso à informação, Princípios Arquivísticos, Normas, Metodologias, sua epistemologia, etc



Cadeia de custódia ininterrupta: linha contínua de custodiadores de documentos arquivísticos (desde o seu produtor até o seu legítimo sucessor) pela qual se assegura que esses documentos são os mesmos desde o início, não sofreram nenhum processo de alteração e, portanto, são autênticos.

Fonte: Costa *et al* (2016).

Verifica-se na Figura a cadeia de custódia ininterrupta de sistemas de informação digital, por um SIGAD (considerando também um RDC-Arq) nas fases corrente e intermediária (cor verde), e na fase permanente por um RDC-Arq (cor vermelha) (entende-se aqui o *Archivematica* e AtoM). Este último, como plataforma de acesso e difusão da informação.

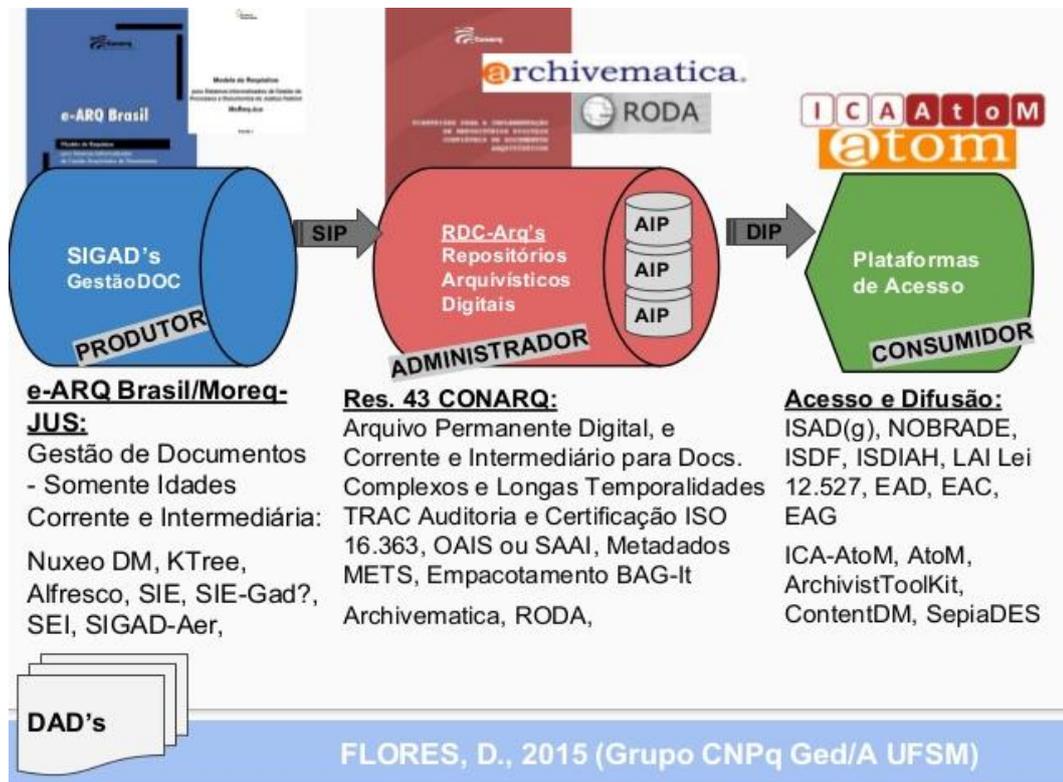
Existem cenários distintos para implementação de um RDC-Arq, podendo englobar o ciclo completo documental, além de haver um RDC-Arq para as fases corrente e intermediária e outro para a fase permanente (FLORES; ROCCO; SANTOS, 2016).

Ao identificar as plataformas digitais ou ambientes de cadeia de custódia arquivística, considera-se em primeiro plano a gestão documental (produtor), em segundo o RDC-Arq (administrador), e em terceiro a plataforma de descrição, difusão e acesso (consumidor), constituindo a totalidade do ambiente de preservação e acesso por meio de sistemas de informação.

Na Figura 11, coexiste uma representação da cadeia de custódia representada pelos manuais dispostos no site do CONARQ e a exposição de suas diretrizes em cada fase: e-Arq

Brasil para os SIGADs, Resolução nº 43 para os RDC-Arqs (*Archivematica* e RODA), e por último o AtoM para as plataformas de acesso, difusão e descrição de documentos arquivísticos.

Figura 11 - Ambiente de preservação e acesso de uma cadeia de custódia

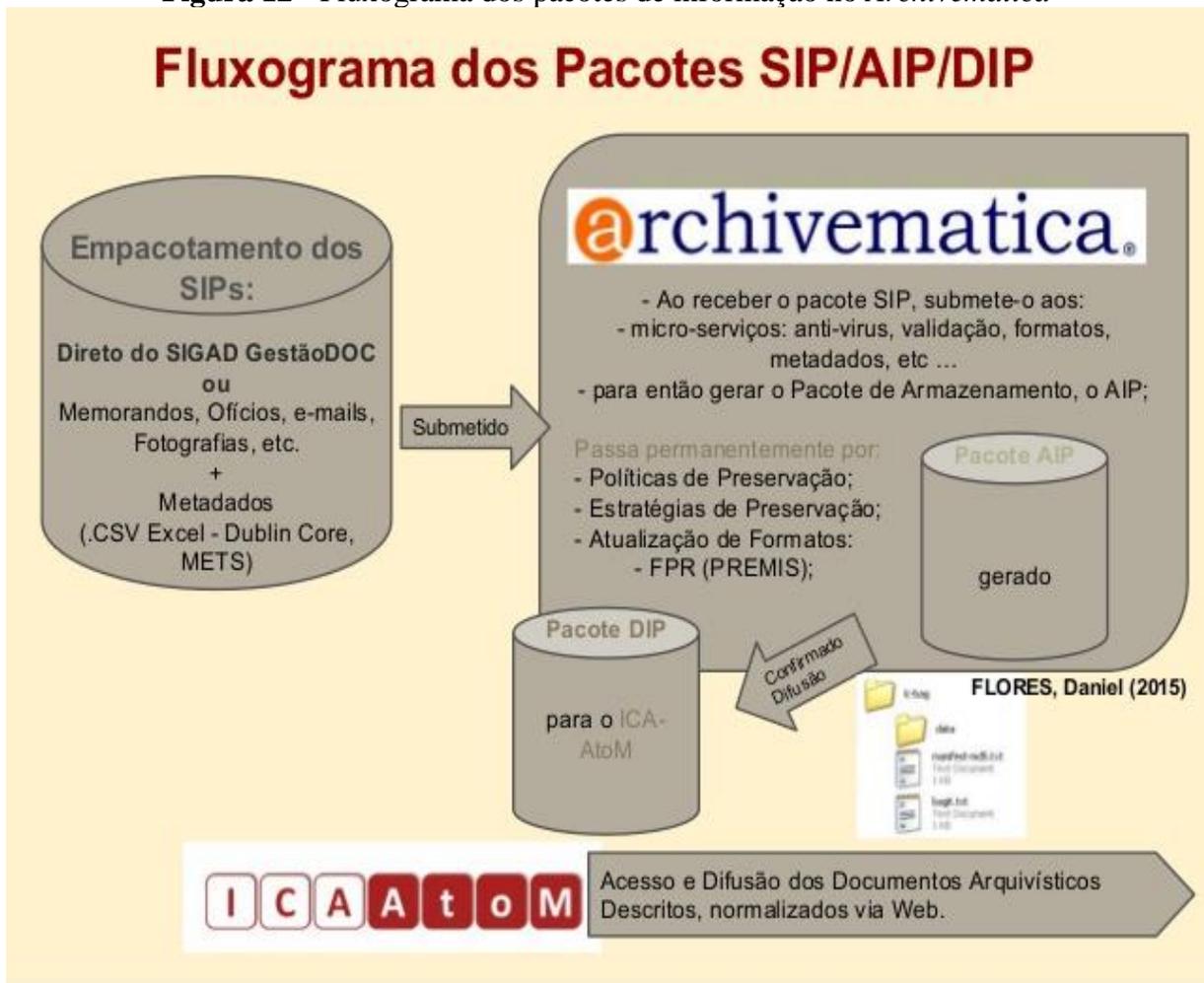


Fonte: Costa *et al* (2016).

No Brasil, ainda é incipiente a implementação do *Archivematica* como solução para local de armazenamento confiável de informação. Entretanto existem esforços neste sentido, podendo-se mencionar as investigações técnico-científicas do Grupo de Pesquisa (CNPQ-UFSM): Gestão Eletrônica de Documentos Arquivísticos – GED/A, vinculado ao curso de Arquivologia do Programa de Pós-graduação em Patrimônio Documental da Universidade Federal de Santa Maria (UFSM) (AGUIAR, 2018).

Aguiar (2018) ainda argumenta, o *Archivematica* foi projetado por “empacotamento”: SIP, AIP, DIP, ou seja, seu processamento é feito a partir de pacotes de informação de arquivo, isto é, para a entrada de documentos, armazenamento e preservação no repositório e acesso (Figura 12).

Figura 12 - Fluxograma dos pacotes de informação no *Archivematica*



Fonte: Costa *et al* (2016).

Flores, Rocco e Santos (2016) recomendam que não se deve acessar diretamente a informação dentro do ambiente do repositório digital para manter sua autenticidade, pois como o documento encontra-se em uma “zona militarizada”, acessada pelo ambiente de administração disposta pelo OAIS, não permitir acesso de agentes externos reforça seus aspectos de confiabilidade.

9 RESULTADOS E DISCUSSÃO

Nesta seção, apresentam-se os resultados e a discussão sobre a análise do *Archivematica* à luz dos requisitos da Resolução nº 43 do CONARQ.

Para esta pesquisa utilizou-se dos elementos dessa normativa elencados em sua segunda seção, que é composta por: Seção 1 – Infraestrutura organizacional; Seção 2 – Gerenciamento do documento digital; Seção 3 – Tecnologia, infraestrutura técnica e segurança. Tais elementos basearam-se na diretriz ISO 16363:2012 de preservação digital da informação para longo prazo, e estão alinhados em uma perspectiva que abrange os repositórios, adaptando-se à realidade brasileira.

Como resultado inicial, relacionamos no Quadro 12 os efeitos encontrados quanto ao cumprimento dos requisitos da 1ª seção da Resolução nº 43.

O Quadro é composto pelos itens: Seção, Item, Tópico, Descrição do Tópico e Cumprimento dos requisitos por parte do *Archivematica*, mostram os resultados da comparação da primeira seção da Resolução, com as características e funcionalidades do referido software.

Quadro 12 – Cumprimento dos requisitos do *Archivematica* à luz da Resolução nº 43 do CONARQ (1ª seção)

Seção	Item	Tópico	Descrição do tópico	Cumprimento dos requisitos por parte do <i>Archivematica</i>
Infraestrutura organizacional	1	a) Governança e viabilidade organizacional	O repositório tem como missão o compromisso com a preservação, o gerenciamento e o acesso de longo prazo dos documentos digitais. Essa missão é claramente identificada por todos os interessados no repositório e envolve: mandato legal, contexto organizacional e requisitos regulatórios. O repositório tem um plano de sucessão formal, planos de contingência e/ou acordos estabelecidos para garantir a continuidade do serviço, no caso de o repositório parar de operar ou de a instituição responsável e/ou financiadora mudar seu escopo.	Sim (X) Não ()
		b) Estrutura organizacional e de pessoal	O repositório tem uma equipe dotada de qualificação e formação necessárias, e em número suficiente, para garantir todos os serviços e funcionalidades pertinentes ao repositório. Além disso, deve manter	Sim (X) Não ()

			um programa de desenvolvimento profissional contínuo.	
		c) Transparência de procedimentos e arcabouço político	O repositório deve demonstrar explicitamente seus requisitos, decisões, desenvolvimento e ações que garantem a preservação de longo prazo e o acesso a conteúdos digitais sob seus cuidados. Dessa forma, assegura aos usuários, gestores, produtores e certificadores que está cumprindo plenamente seu papel enquanto um repositório digital confiável.	Sim (X) Não ()
		d) Sustentabilidade financeira	Um repositório digital confiável deve demonstrar sustentabilidade financeira.	Sim (X) Não ()
		e) Contratos, licenças e passivos	Os contratos, licenças e passivos firmados pelo repositório devem ser claros e mensuráveis; delinear papéis, responsabilidades, prazos e condições; e ser facilmente acessíveis ou disponíveis aos interessados. Esses contratos, licenças e passivos podem envolver tanto a relação entre o repositório e os produtores de documentos digitais, como a relação entre o repositório e fornecedores de serviços. Esses mesmos instrumentos devem especificar todos os direitos e obrigações do repositório sobre os documentos digitais a ele confiados, em especial no que diz respeito à propriedade intelectual e a restrições de uso.	Sim (X) Não ()

Fonte: Adaptado do Conselho... (2015).

No tópico “a” está a incumbência da missão e o compromisso que um repositório precisa possuir no que se refere à preservação digital em longo prazo. Sendo assim, o *Archivematica* cumpre com os preceitos arquivísticos, bem como políticas e operações por parte de sua comunidade usuária. Tem a finalidade de armazenar e preservar tais informações em instituições públicas e privadas, especialmente, por um período prolongado. Isso vem ao encontro do que afirmam Jorente *et al* (2016) e Martinez *et al* (2017), ao ratificarem que a preocupação com o período prolongado é indispensável, visto que os impactos tecnológicos de hardware, software e formato, seguem os padrões do modelo OAIS/SAAI nas fases corrente, intermediária e permanente.

O tópico “b” se refere a um repositório alinhado em estrutura organizacional e de pessoal. Isto significa que o sistema demanda uma equipe qualificada com formação e competências necessárias para garantir suas funcionalidades e serviços de maneira contínua.

Desde a sua criação, o *Archivematica* teve como meta manter um diálogo constante com a comunidade usuária, Para isso, estimula os fóruns de discussão e as redes sociais. De acordo com Van Garderen (2010), a maioria dos profissionais envolvidos nesse projeto são oriundos das áreas da Computação e da Arquivologia, o que indica que a expertise desses profissionais, cada um trazendo os aportes de sua especialidade, são fundamentais para manter a estrutura e capacidade de funcionamento do software.

O tópico “c” preconiza que os repositórios de confiança precisam abranger: requisitos, decisões, desenvolvimento e ações de garantia de preservação de longo em prazo e acesso a objetos digitais, enfatizando assim, sua transparência e arcabouço político. Para tanto, deve:

- 1) definir a comunidade alvo e sua base de conhecimento;
- 2) possuir políticas e definições, acessíveis publicamente, que demonstrem como os requisitos do serviço de preservação serão contemplados;
- 3) possuir políticas, procedimentos e mecanismos de atualização, na medida em que o repositório cresce e a tecnologia e práticas da comunidade evoluem;
- 4) documentar permissões legais – por meio de acordos de custódia, normas de procedimentos e outros – que o isentem de responsabilidade, no caso de alterações passíveis de ocorrer em estratégias de preservação digital;
- 5) fazer o registro histórico das mudanças de procedimentos, de software e hardware;
- 6) relacionar o registro histórico, acima referido, com as estratégias de preservação digital, e descrever os potenciais efeitos dessas mudanças sobre os documentos digitais;
- 7) demonstrar que está, sistematicamente, avaliando a satisfação das expectativas dos produtores e dos usuários, e buscando atendê-las;
- 8) estar comprometido com a definição, coleta, auditoria e fornecimento (sob demanda) de mecanismos de controle da integridade dos documentos digitais sob sua custódia;
- 9) estar comprometido em realizar regularmente uma autoavaliação de seu funcionamento e renovar sua certificação;
- 10) estar comprometido em notificar as entidades certificadoras sobre as mudanças operacionais que afetarão seu status de certificação (no caso de repositórios certificados) (CONSELHO..., 2015, p. 12-13, grifo nosso):

Com referência aos itens “1”, “7”, “9” e “10”, o *Archivematica* viabiliza às instituições públicas e privadas sua utilização, desde que cumpram com as políticas e padrões de preservação. As relações, colaborações, autoavaliações e as mudanças operacionais ocorrem

por intermédio de seus usuários. Isto é, as respectivas informações sobre o software, são discutidas em seus veículos de comunicação explanado no tópico “a”.

Em relação ao “2”, “3”, “5” e “6”, possuem suas políticas dispostas pelas diretrizes OAIS/SAAI, ISO 16:363 e TRAC, concebendo as estratégias de preservação digital (emulação e migração), que ditam os aspectos de confiabilidade e manifestando preocupação com a obsolescência e os impactos tecnológicos.

Quanto aos itens “4” e “8”, apresentam procedimentos que permitem a identificação de qualquer modificação/alteração, o que está de acordo com os preceitos das diretrizes mencionadas nos itens anteriores, certificando e dando respaldo à confiança da informação.

Observa-se no tópico “d”, que o repositório deve demonstrar sustentabilidade financeira. Destacando um plano de gestão nos seguintes aspectos:

- 1) demonstração da capacidade de obter recursos financeiros estáveis e contínuos para sustentá-lo, seja por meio de prestação de serviço, parcerias, doações, verba da própria instituição, dentre outros;
- 2) revisão e ajustes anuais;
- 3) transparência dos procedimentos para obtenção dos recursos e auditoria dos mesmos, de acordo com o sistema jurídico no qual o repositório se insere; e
- 4) compromisso dos ciclos de planejamento com o equilíbrio dos riscos, benefícios, investimentos e gastos (CONSELHO..., 2015, p. 13).

Como o *Archivematica* é um sistema de uso totalmente livre, sua aquisição não demanda utilização de recursos financeiros. Nesse sentido, destaca-se que o software cumpre com os quatro itens acima referenciados, visto que, para sua sustentabilidade, torna-se fundamental o serviço da entidade criadora e a atuação de profissionais especializados como tecnólogos da informação e arquivistas para manusear e manter o referido sistema pelo período que se fizer necessário, o que dirime custos, quando comparado a um software proprietário.

Requer, ainda, revisão periódica de suas funcionalidades, tais como: atualização, mudanças, extinção e inserção de guias, backups e atualização de políticas de preservação. Funções estas, acontecidas por parte da atuação de quem manuseia o referido sistema.

O tópico “e” que os contratos, licenças e passivos são aqueles que dão delineamento às responsabilidades e condições, podendo envolver a relação entre o repositório e os produtores de documentos digitais.

Acentua-se, que a descrição deste ponto, insere-se no *Archivematica*, sob seus termos de licenças gratuitas, por exemplo a GNU (3.0) e a *Creative Commons*, possuindo como

principal objetivo, a liberdade de copiar, distribuir e modificar sistemas. Stallman (2009), denomina esses princípios de quatro liberdades: de usar o programa para qualquer propósito, de estudar como o programa funciona e modificá-lo de forma a que ele venha a atender as suas necessidades específicas, de redistribuir cópias, com a finalidade de ajudar sua comunidade e, de distribuir cópias da versão modificada.

De acordo com a análise realizada, à luz da Resolução nº 43, nessa primeira seção, Infraestrutura organizacional, o cumprimento dos requisitos por parte do *Archivematica* apresenta o percentual de 100%, o que significa dizer que sua totalidade se encontra em consonância com as características e funcionalidades do sistema em apreço.

Quadro 13 - Cumprimento dos requisitos do *Archivematica* à luz da Resolução nº 43 do CONARQ (2ª seção)

Seção	Item	Tópico	Descrição do tópico	Cumprimento dos requisitos por parte do <i>Archivematica</i>
Gerenciamento documento do digital	2	a) Admissão (captura de documentos digitais)	Consiste na entrada dos documentos e seus metadados no repositório digital. Os requisitos de admissão variam dependendo do tipo de material, do contexto legal e da relação entre o produtor de documento e o repositório. Independentemente dessas variações, pode-se afirmar que a admissão se inicia com o recebimento de um SIP, que é convertido em AIP, e termina quando um AIP está seguro no repositório, incluindo a criação de cópias de segurança.	Sim (X) Não ()
		b) Admissão (criação do pacote de arquivamento)	O repositório deve completar o processo de admissão, criando um pacote de informação apropriado para arquivamento (AIP), com toda a informação recebida do produtor.	Sim (X) Não ()
		c) Planejamento da preservação	Um repositório digital deve fazer o planejamento da preservação dos documentos sob sua custódia, a fim de enfrentar os problemas trazidos pela obsolescência tecnológica e fragilidade do suporte.	Sim (X) Não ()
		d) Armazenamento e preservação/manutenção do AIP	Um repositório deve atender a um conjunto de condições para garantir o bom desempenho da preservação de longo prazo dos AIPs.	Sim (X) Não ()
		e) Gerenciamento de informação	Uma funcionalidade essencial de um repositório digital confiável é o gerenciamento da informação, aqui	Sim (X) Não ()

			entendido como a gestão das informações descritivas (metadados) dos documentos admitidos no repositório. O principal objetivo desses metadados é apoiar o acesso e a recuperação dos documentos, e vão além das informações descritivas mais usuais (autor, título, data), envolvendo outras informações descritivas úteis aos usuários, tais como tamanho do arquivo disponível para download ou informação sobre a aplicação necessária para ler o arquivo	
		f) Gerenciamento de acesso	Todo repositório deve produzir pacotes de disseminação de informação (DIP).	Sim (X) Não ()

Fonte: Adaptado do Conselho... (2015).

A seção de Gerenciamento de documentos para um repositório digital confiável deve estar estabelecida pelo modelo de referência OAIS e com a formação de pacotes de informação (SIP, AIP, DIP), envolvendo documentos digitais (informação e conteúdo) e seus metadados (informação de representação). Os seis grupos do Quadro 13 apresentam os requisitos da TRAC, baseados nas funcionalidades de: Admissão (captura de documentos digitais); Admissão (criação do pacote de arquivamento; Planejamento da preservação; Armazenamento e preservação /manutenção do AIP; Gerenciamento de informação e Gerenciamento de acesso (CONSELHO..., 2015).

O tópico “a”⁹ se caracteriza pela admissão de documentos e metadados no repositório (entrada). No caso do *Archivematica*, seu percurso ocorre com o que está disposto na Resolução nº 43. Inicia-se com o recebimento de um SIP, que pode ser convertido em AIP, incluindo a criação de cópias de segurança, e após o DIP. Dessa forma, apresentam-se os requisitos gerais a serem cumpridos pelo repositório, cuja adequação deve ser avaliada de acordo com a missão e as necessidades de cada repositório:

- 1) identificar as propriedades do documento que serão preservadas (ex.: o conteúdo, *layout*, tabela de cor, resolução da imagem, canais de som etc.);
- 2) especificar claramente a informação que deve estar associada ao documento (metadados associados) no momento da sua submissão;
- 3) ter mecanismos para autenticar a origem dos documentos que estão sendo admitidos no repositório, de forma a garantir sua proveniência;

⁹ No caso de um repositório para documentos arquivísticos, a definição dos metadados deve observar o e-ARQ Brasil (nas fases corrente e intermediária) e a NOBRADE (na fase permanente). Para a admissão de documentos no repositório, no caso de transferência ou recolhimento, devem-se observar os procedimentos indicados na Resolução nº 24, de 3 de agosto de 2006, do CONARQ (CONSELHO..., 2015, p. 14).

- 4) ter procedimentos para verificar a integridade do SIP, o que pode ser feito por meio de procedimentos automatizados e/ou checagem humana;
- 5) ter o controle físico (controle completo dos *bits*) dos documentos transmitidos com cada SIP, a fim de preservá-los;
- 6) fornecer ao produtor/depositante relatórios do andamento dos procedimentos durante todo o processo de admissão;
- 7) demonstrar em que momento a responsabilidade pela preservação do documento submetido (SIP) é formalmente aceita pelo repositório;
- 8) ter registros de todas as ações e processos administrativos que ocorrem durante o processo de admissão e são relevantes para a preservação (CONSELHO..., 2015, p. 14).

Pode-se aferir que o *Archivematica* consegue englobar todos os itens acima, pois identifica e especifica as propriedades dos documentos digitais, sobretudo, os admitidos/submetidos, expondo o tipo, formato, metadados e no ato do download, a visualização dos arquivos, (imagético, audiovisual, convencional), ressaltando seus mecanismos de autenticidade, integridade e preservação, bem como o controle de ações em andamento dos pacotes de informação através de microsserviços.

Isto posto, observam-se as funcionalidades do software em questão, a exemplo da guia de Transferência (processo de transformar qualquer conjunto de objeto de digital ou diretório em um SIP à guia de Administração) responsável pela configuração dos componentes do aplicativo e gerenciamento dos usuários, uma espécie de painel de controle (COSTA *et al*, 2016). (Ver Apêndice B).

O tópico “b” menciona a admissão, como a criação de um AIP, a fim de garantir que o pacote de informação recebido do produtor e verificado pelo repositório, seja convertido para o formato de arquivamento (AIP) e armazenado para preservação de longo prazo. Para tanto, um repositório deve atender aos seguintes requisitos:

- 1) descrever cada classe de informação (texto estruturado, imagem matricial, banco de dados, imagem em movimento e outras) a ser preservada pelo repositório, e como ela está implementada – essa descrição deve apontar os componentes-chave do AIP: o documento arquivístico, sua informação de representação (informação estrutural e semântica) e as várias categorias de informação descritiva de preservação (fixidade, proveniência e contexto), e ainda como esses componentes se relacionam;
- 2) descrever minuciosamente as diferentes classes de informação e como os AIPs são implementados, nos casos em que a especificidade daquelas classes exigir ações de preservação diferentes (por exemplo, a imagem TIFF que é processada por um sistema pode necessitar de ações de preservação diferentes das ações necessárias à imagem TIFF que é apresentada para o olho humano);
- 3) descrever como os AIPs são construídos a partir dos SIPs, ou seja, apontar todas as transformações pelas quais passarão os documentos e os metadados submetidos, e os metadados a serem adicionados no momento da formação do AIP;

- 4) ser capaz de demonstrar se os SIPs foram aceitos e transformados em um AIPs integralmente ou em parte, ou ainda se foram recusados;
- 5) atribuir aos AIPs, identificadores que sejam únicos, persistentes e visíveis aos gestores e auditores, de acordo com padrões reconhecidos (por exemplo: *Handle System*, DOI, URN, PURL);
- 6) no caso de o documento já possuir um identificador único, a ele atribuído no SIP, o repositório deverá mantê-lo no AIP, ou criar um outro identificador, que deverá ser associado, de maneira persistente, ao do SIP;
- 7) ter acesso a ferramentas amplamente reconhecidas para apoiar o monitoramento dos componentes digitais dos documentos, tais como diretórios de formatos de arquivos (ex.: PRONOM – base de dados com registro de formatos mantida pelo arquivo nacional do Reino Unido) e registros de outras informações de representação;
- 8) registrar, em um banco de dados local, a informação de representação dos documentos admitidos, quando essa informação não estiver disponível nas ferramentas mencionadas no ponto anterior;
- 9) registrar metadados de preservação associados aos documentos admitidos, de maneira a apoiar sua integridade, localização, legibilidade e proveniência, dentre outros;
- 10) ter procedimentos para testar se os documentos são compreensíveis pela comunidade-alvo e, em caso negativo, adequá-los às necessidades dessa comunidade (ex.: documentos voltados para deficientes visuais);
- 11) verificar a completude e a correção de cada AIP no momento em que é gerado, isto é, no momento em que o SIP é convertido em AIP;
- 12) ter um mecanismo independente para verificar a integridade do conjunto do seu acervo, ou seja, verificar que todos os documentos previstos foram, de fato, admitidos no repositório, justificando possíveis lacunas; e
- 13) documentar todas as ações relevantes à preservação dos documentos e que estão relacionadas à criação do AIP (CONSELHO..., 2015, p. 14-15).

Os itens elencados possuem uma estrutura voltada ao percurso do AIP. Tal composição também está compreendida com os preceitos do *Archivematica*, principalmente no que diz respeito às características das guias Ingerir, Armazenamento Arquivístico e Plano de Preservação, por onde perpassa o SIP, executando vários microsserviços no AIP. (Ver Apêndice B).

Desse modo, destaca-se que os mecanismos de representação da informação terão maior expressividade nas três primeiras guias citada, pois já passaram pela fase de submissão de informação e estão no processo de arquivamento, partindo para o acesso. As funções do SIP-AIP referentes à criação, percurso, relação e implementação no repositório, atingem o caráter de confiança nos documentos digitais arquivísticos quando passam pela conferência e aprovação dos microsserviços (COSTA *et al*, 2016).

Tais preceitos se consolidam com alguns elementos, a exemplo das especificidades de formato (TIFF), metadados, *Universally Unique Identifier* (UUID) e ferramentas de preservação (PRONOM), comprovando sua integridade enquanto documento de arquivo no *Archivematica*, que por ser *open source* há a verificação de êxitos e correções por sua comunidade. Todas essas ações são documentadas no próprio sistema, podendo ser analisadas no processo de cada guia. (Ver Apêndice B).

O tópico “c” enfatiza a importância do planejamento da preservação em um repositório digital, como enfrentamento da obsolescência tecnológica e fragilidade do suporte. Esse planejamento deve ser feito a partir de uma política de preservação digital, ser bem documentado e incluir:

- 1) estratégias de preservação bem definidas e periodicamente atualizadas, apontando e detalhando cada procedimento a ser adotado, como, por exemplo, a normalização de formatos;
- 2) mecanismos para monitoramento e notificação quando alguma informação de representação dos documentos no repositório estiver se tornando obsoleta ou inviável (ex.: um formato de arquivo que esteja entrando em desuso, um suporte que esteja no final de sua vida útil);
- 3) mecanismos de mudanças do plano de preservação como resultado do monitoramento;
- 4) fornecimento de evidências sobre a eficácia do plano de preservação (CONSELHO..., 2015, p. 15-16).

Dessa forma, reitera-se o cumprimento por parte do *Archivematica* em relação a esse tópico. No que tange ao plano político, tem-se a guia Plano de Preservação, demonstrando a importância e o papel da normalização para formatos de preservação e acesso por meio da admissão, os quais possibilitam que arquivos originais sejam mantidos na permissão de ações futuras, tais como a emulação e migração (ARTEFACTUAL, [201-b]; COSTA *et al*, 2016). (Ver Apêndice B).

Bem verdade, que os mecanismos de Plano de Preservação são monitorados por seus usuários, ademais as políticas utilizadas nesse software são consistentes, pois foram planejadas pelo *Format Policy Registry* (FPR) e PRONOM, alinhadas à TRAC e de acordo com a norma ISO 16363:2012.

O tópico “d” elenca o armazenamento, preservação e manutenção de um AIP. Significa dizer que um repositório deve atender a um conjunto de condições para garantir o bom desempenho da preservação em longo prazo dos AIPs, conforme se descreve a seguir:

- 1) utilização das estratégias previstas no planejamento da preservação, que podem ser várias e devem ser registradas nos metadados de preservação;

- 2) atender minimamente a dois aspectos da preservação digital – os cuidados com armazenamento (controle dos suportes, dos formatos e da localização de cópias) e a eventual necessidade de migração (atualização de suportes e conversão de formatos);
- 3) preservação do documento digital (informação de conteúdo do AIP) originalmente admitido no repositório e daquele resultante da última migração;
- 4) monitoramento constante da integridade dos AIPs, por meio do registro de metadados de fixidade e de *logs* de checagem dessa integridade (por exemplo, *checksum*);
- 5) registro de todas as ações de preservação realizadas nos AIPs (CONSELHO..., 2015, p. 16).

Após a análise, verifica-se que todos os tópicos estão se encontram condizentes com as funcionalidades do *Archivematica*. Note-se que, para o armazenamento dos documentos, alguns cuidados são requeridos, dentre estes: a atribuição de metadados, tipos de políticas, formatos, cópias e backups. Tal sistema atende às premissas de preservação quando diz respeito a preocupação com a obsolescência tecnológica, que se utiliza de estratégias que não só permitem a migração de suporte ou formato, mas também a emulação, encapsulamento, refrescamento, etc. Logo, corrobora com os elementos de autenticidade relacionados aos pacotes de informação, os registros, os metadados de preservação adequados e os *checksums* (código de verificação para conferir a integridade de um arquivo). Desse modo, a Resolução nº 43 tece algumas considerações:

As migrações podem provocar alterações na forma e no conteúdo do documento, entretanto, no caso de documentos arquivísticos, não se admite a alteração de conteúdo. As migrações e quaisquer alterações da forma documental daí decorrentes devem ser registradas como metadados, a fim de apoiar a presunção de autenticidade do documento. (CONSELHO..., 2015, p. 16).

O tópico “e” mostra que o gerenciamento da informação é uma funcionalidade essencial do repositório digital confiável, assim entendida como gestão das informações descritivas (metadados), os quais possuem como elemento principal a recuperação da informação. De maneira geral, as informações descritivas mais comuns são: autor, título e data. Entretanto, em se tratando de repositórios e documentos, outros itens são acrescentados, como: tamanho do arquivo ou quaisquer outras necessárias para lê-los. O gerenciamento da informação descritiva envolve os seguintes aspectos:

- 1) metadados mínimos que permitam a busca e localização dos documentos – esses metadados devem ser identificadores conhecidos pela comunidade-alvo de usuários (ex.: número de matrícula do servidor público, título de livro numa biblioteca, número de processo);

- 2) captura ou criação dos metadados mínimos pelo repositório, durante o processo de admissão, e associação desses metadados ao AIP correspondente;
- 3) integridade referencial entre os AIPs e sua informação descritiva (metadados), ou seja, todo AIP deve ter uma informação descritiva, e toda informação descritiva deve apontar para um AIP; e
- 4) permanência da integridade referencial, mesmo no caso de quebra temporária da relação entre o AIP e seus metadados descritivos – nesse caso, o repositório deve ser capaz de restaurar a relação rompida (CONSELHO..., 2015, p. 16).

As guias do *Archivematica* possuem partes específicas que fazem menção aos metadados. Na guia *Backlog*, por exemplo, o usuário pode explorar arquivos visualizando-os na forma de relatório, isto inclui a busca e localização de dados pessoais *Personally Identifiable Information* (PII) e número de cartão de crédito. (Ver Apêndice B). Em outra analogia tem-se os padrões de metadados como *Dublin core* e PREMIS, que fornecem descrição para livro, jornal e afins, mantendo sua autenticidade, funcionando como instrumento de recuperação da informação, o que torna íntegro os AIPs originais e suas cópias de segurança. Dessa forma o *Archivematica* também cumpre com o que determina a Resolução nº 43.

O tópico “f” diz respeito ao Gerenciamento de acesso, ou seja, todo repositório tem que produzir pacotes de disseminação da informação (DIPs). Atendendo aos seguintes requisitos:

- 1) divulgação, para a comunidade de usuários, das opções disponíveis de acesso aos documentos e de entrega;
- 2) implementação de uma política de registro dos acessos ocorridos que esteja de acordo com as necessidades de controle desses acessos, tanto da parte do repositório como dos produtores dos documentos nele admitidos;
- 3) concessão de acesso a cada AIP, para os usuários autorizados e da forma devida (ex.: autorização de “somente leitura”, ou acesso a um número limitado de itens por período), em conformidade com o acordo estabelecido entre o repositório e o produtor/depositante;
- 4) documentação e implementação de políticas de acesso (identificação e autenticação de usuários), em conformidade com os acordos estabelecidos entre o repositório e o produtor/depositante – essas políticas de acesso podem variar, desde a isenção da necessidade de identificação de usuário até o controle rígido da identificação e autenticação do usuário;
- 5) registro de falhas de controle de acesso (como, por exemplo, um acesso indevidamente negado) e uso desse registro para avaliar eventuais falhas no sistema de segurança;
- 6) demonstração de que o processo que gera o DIP atende completamente à requisição do usuário (ex.: se o usuário pediu um conjunto de documentos, receberá o conjunto completo; se ele pediu um documento, receberá apenas este único documento);

- 7) demonstraco de que o processo que gera o DIP est correto em relao ao pedido do usurio (ex.: se o repositrio oferece imagens nos formatos JPG e PNG, o usurio deve receber, dentre esses, o formato que solicitou);
- 8) demonstraco de que todos os pedidos de acesso resultam em uma resposta de aceitao ou rejeio; e
- 9) garantia da autenticidade dos DIPs, por meio da entrega de cpias autnticas dos originais ou da viabilidade de rastreamento auditvel da relao entre o DIP e o objeto original – para isso, um repositrio deve ser capaz de demonstrar o processo de construo do DIP a partir de um AIP (CONSELHO..., 2015, p. 17).

A guia de Acesso explica o DIP de maneira pormenorizada, pois  nela que o usurio carrega o pacote de informao e o armazena para uso futuro. (Ver Apndice B). Quando o DIP  preparado, o *Archivematica* tem a funo de mov-lo para o diretrio de *UploadDIP*, ou seja, o envio para uma plataforma de acesso, se preciso for. As seguintes plataformas de descrio suportam *uploads* de DIPs: *Atom*, *ArchivesSpace*, *CONTENTdm* e *Archivists' Toolkit* etc. (COSTA *et al*, 2016; MARTINEZ *et al*, 2017).

Os guias de instalao e de usurio do *Archivematica* deixam claro que tais sistemas de descrio precisam ser interoperveis, observando o compromisso com as polticas de registros e de preservao, neste ltimo, a normalizao. Aps esta tarefa, o SIP  executado atravs do processamento da documentao, ou seja: submisso, gerao de arquivo METS, indexao, gerao do DIP e de acondicionamento do AIP. Todas as alteraes que acontecem so documentadas e dispostas pelos aspectos da normalizao para a preservao e acesso: criao de cpias de preservao dos objetos e cpias de acesso que sero utilizadas para gerar o DIP (COSTA *et al*, 2016).

A anlise pormenorizada desta seo permite afirmar que todos os requisitos foram contemplados pelo *Archivematica*, encontrando-se de acordo com o que prescreve a Resoluo n 43.

Quadro 14 - Cumprimento dos requisitos do *Archivematica*  luz da Resoluo n 43 do CONARQ (3ª seo)

Seo	Item	Tpico	Descrio do tpico	Cumprimento dos requisitos por parte do <i>Archivematica</i>
Tecnologia, infraestrutura tcnica e segurana	3	a) Infraestrutura de sistema	um repositrio deve possuir uma infraestrutura tecnolgica robusta, de maneira a apoiar a confiabilidade dos AIPs nele mantidos.	Sim (X) No ()
		b) Tecnologias apropriadas	o repositrio deve adotar uma tecnologia de hardware e software apropriada para os servios que presta, procedimentos para o recebimento e monitoramento de notificaes e para a avaliao da necessidade de mudanas	Sim (X) No ()

			na tecnologia utilizada.	
		c) Segurança	a segurança do repositório não se limita a aspectos de tecnologia, mas abrange também instalações físicas e ações de pessoas.	Sim (X) Não ()

Fonte: Adaptado do Conselho... (2015).

O objetivo desta seção é mostrar as melhores práticas da gestão de dados e segurança, os quais devem atender a um repositório digital confiável (CONSELHO..., 2015).

O tópico “a” enfatiza a infraestrutura do sistema, de maneira que o repositório possua tecnologia robusta para apoiar a confiabilidade dos AIPS. Dessa forma observam-se os seguintes aspectos:

- 1) funcionamento do repositório com base num sistema operacional e outros softwares de infraestrutura que tenham um bom suporte do mercado e da comunidade de usuários;
- 2) adequação dos processos, do hardware e do software do sistema de *backup* às necessidades do repositório;
- 3) gerenciamento do número de cópias de todos os documentos mantidos no repositório, e a localização de cada uma delas;
- 4) mecanismos para garantir o sincronismo entre as cópias de um mesmo documento, ou seja, garantir que as mudanças intencionais feitas em uma cópia sejam propagadas para todas as outras;
- 5) mecanismos efetivos para a detecção de corrupção ou perda de *bits*;
- 6) relato dos incidentes de corrupção ou perda de dados eventualmente ocorridos e adoção de medidas para reparação ou substituição desses mesmos dados;
- 7) previsão de procedimentos de atualização de suporte (*refreshing*) e de migração decorrentes do cumprimento do prazo de vida do suporte ou da obsolescência dos componentes de hardware;
- 8) documentação da gestão de mudanças capaz de identificar alterações em processos críticos que afetem a capacidade de o repositório cumprir com suas responsabilidades obrigatórias;
- 9) previsão de procedimentos para testar o efeito de mudanças críticas no sistema;
- 10) ponderação entre os riscos e os benefícios nas decisões de atualização de software de segurança (CONSELHO..., 2015, p. 17-18).

A contemplação desses requisitos está imersa no *Archivematica*, uma vez possui infraestrutura técnica e de compartilhamento entre sua rede usuária. As questões de adequação a sistemas operacionais são descritas nas instruções de download (em seu site oficial), assim também como os requisitos para sua execução. As estratégias de políticas e preservação permitem que, para cada geração de pacotes de informação, neste caso o AIP, exista uma cópia de segurança, sobretudo por meio da normalização, preocupando-se com os impactos de

obsolescência tecnológica. Para tanto, os quesitos de risco passam por rigorosa política de segurança no intuito garantir seu caráter confiável.

O tópico “b” evidencia que o repositório precisa adotar uma tecnologia apropriada para seus serviços, procedimentos para o recebimento e monitoramento de notificações e avaliação da necessidade de mudança na tecnologia (CONSELHO..., 2015).

O *Archivematica* está inserido em tal proposta, pois permite adequar-se aos impactos tecnológicos, e conseqüentemente, à evolução da tecnologia. A característica de sistema aberto o permite enquadrá-lo nesse tipo de adaptação, ficando a cargo de sua comunidade usuária, a responsabilidade pelos procedimentos técnicos e avaliação do software.

O tópico “c” mostra que a segurança do repositório não deve se limitar a aspectos de tecnologia, mas também de instalações físicas e de pessoas. Para isso, incluem:

- 1) análise sistemática de dados, sistemas, pessoas e instalação física;
- 2) adoção de procedimentos de controle para tratar adequadamente as necessidades de segurança;
- 3) delineamento de papéis, responsabilidades e autorizações relativas à implementação de mudanças no sistema; e
- 4) plano de prevenção de desastres e de reparação, que inclua, ao menos, um *backup, off-site*, de tudo o que é mantido no repositório (documentos, metadados, trilhas de auditoria etc.), inclusive do próprio plano de reparação (CONSELHO..., 2015, p. 18).

Nesta etapa pode-se mencionar os aspectos descritos na guia Plano de Preservação para os AIPS. As funcionalidades do *Archivematica*, desde sua criação, seguem um rigoroso esquema de segurança, baseados em norma de padrão internacional (OAIS), testado ainda em fase de protótipo.

Os procedimentos de controle de segurança são declarados em suas diretrizes de instalação e manuseio, - a comunidade usuária participa desta prática com afimco compartilhando as melhores práticas para seu desempenho.

No que toca ao plano de preservação, é sempre aconselhável fazer o *backup* dos dados antes de atualizações, dando a possibilidade de introduzir tal recurso de segurança para a recuperação das informações futuras. Em caso de problema durante o processo de atualização, é possível recuperar o banco de dados do mysql a partir deste *backup* e tentar atualizar novamente, fazendo que os dados não se percam (COSTA *et al*, 2016).

As análises empreendidas nessa seção também contemplam todas as propriedades do *Archivematica* à luz dos requisitos da Resolução nº 43.

Percebe-se que os elementos elencados em cada seção se dão de forma ordenada, observando as diretrizes que um repositório deve atender, desde o escopo da missão à segurança.

Os resultados das análises apresentadas nos Quadros 12, 13 e 14 integralizam as atribuições em que o *Archivematica* se projeta para tornar-se substancial em relação aos preceitos da alusiva normativa. Compreende-se então, que esse software segue os parâmetros de preservação em cenário internacional e nacional, fixado nas diretrizes que atualmente são consideradas como primordiais aos objetos digitais em longo prazo (GONÇALEZ, 2017).

Baseados na avaliação das três seções, divididas em 14 tópicos, verifica-se total compatibilidade do *Archivematica* à Resolução nº 43 do CONARQ, o que vem a confirmar o respaldo teórico e legal no que tange à perspectiva de confiança que um repositório deve possuir.

Entende-se que a citada diretiva consegue compreender e alcançar repositórios de cunho gratuito e/ou privado, porém não se faz menção a nenhum desses dois tipos, logo, destaca que para se ter tal sistema, um recurso se faz importante: a sustentabilidade financeira. O que não necessariamente está ligado apenas ao custo monetário, mas a profissionais especializados e instituições que queiram utilizar um sistema considerado parâmetro na confiança, proteção, segurança e preservação da informação digital na instabilidade do ambiente web.

10 CONSIDERAÇÕES FINAIS

Em razão de sua característica multidisciplinar, a CI dialoga com vários domínios científicos, inclusive com a tecnologia. Nessa perspectiva, a preservação digital é um tema discutido no âmbito de diferentes áreas como a Arquivologia, Administração, Contabilidade, Direito, dentre outras, respeitando o contexto e a necessidade de cada instituição, conforme corrobora Fontana *et al* (2014).

A efetivação dessa tarefa se faz necessária, uma vez que torna acessível os objetos digitais, os quais, na órbita do ambiente computacional, se configuram de modo instável/infixo/inconstante, por seu caráter dinâmico e efêmero, para além da obsolescência da informação tecnológica.

Nessa perspectiva, conforme refere Ferreira (2006), se posta à implementação de estratégias para a preservação digital, que não se baseia apenas no armazenamento de informações, mas na manutenção dos suportes, formatos, hardware e software. A propósito, em todo processo envolvente dos elementos associados à manutenção da possibilidade de acesso. Para tanto, é imprescindível a adoção de tratamentos técnicos dos objetos digitais em seus níveis, indo além do físico, por isso, a capacidade de acessar documento ao longo do tempo vai além da conservação dos suportes para garantir sua integridade (RODRIGUES, 2015).

Na análise desses preceitos, compreende-se que os softwares podem se caracterizar como instrumentos voltados à construção de sistemas de informações dedicados a um espaço confiável através do tripé gerenciamento-armazenamento-acesso. Tem-se então os RDC-Arqs, especificamente o *Archivematica*, que recebe respaldo da Resolução nº 43 do CONARQ para certificação da autenticidade e integridade de documentos digitais.

Apresentou-se tal diretiva no intuito de comparar seu conjunto de elemento, definido por requisitos de confiabilidade que um repositório precisa acordar. Corrobora Gonzalez (2017), que os estudos dos RDC-Arqs ainda são poucos na vertente da certificação, pois as discussões, na área da CI, são mais frequentes no que se refere à curadoria digital e à utilização de software. Nesse sentido, a Resolução nº 43 serve como âncora para dar visibilidade à Arquivologia e a arquivistas como profissionais no mundo, de tendências tecnológicas cada vez mais profícuas.

O *Archivematica* faz parte do objeto deste estudo e está afinado como uma estratégia de preservação, representado por software, que visa afastar o risco de inacessibilidades dos documentos, pois há o monitoramento de sua comunidade, sugerindo ao administrador do

repositório, a constante migração e atualização dos formatos que estão se tornando obsoletos (FONTANA *et al*, 2014).

Destaca-se que o principal papel deste sistema é a preservação em longo prazo e, através de um plano de preservação, atribuir padrões para que a recuperação da informação seja feita de forma eficaz, mesmo havendo a obsolescência de sua extensão. Torna-se, contudo, fundamental (re)conhecer e observar se suas funcionalidades estão cumprindo as devidas responsabilidades, verificando se outros dados foram inseridos ao software, bem como guias e quais estão tendo êxitos, os problemas e dificuldades, mudanças de hardware e software e etc., isto quer dizer que a esta incumbência deve-se partir de sua coopertiva de usuários.

Apesar do *Archivematica* ter visibilidade internacional, existe ainda um desafio a ser superado: seu uso e visibilidade no Brasil. Esta afirmativa se baseia no fato da escassez da literatura brasileira sobre a temática até a presente data. Visto que se configura como eficiente estratégia de preservação digital, com vantagem da interoperabilidade para criar vias de comunicação com outros tipos de dispositivos, já empregados por arquivos, biblioteca, museus e instituições de ensino.

Dessa forma, pode-se sugerir o seu emprego, ou sua adoção pelos vários tipos de entidades em razão do baixo custo caracterizado pela contratação de profissionais aptos à manuseá-lo, a exemplo dos tecnólogos da informação e arquivistas; pela autenticidade e integridade da informação posta nos sistemas de informação; por possuir certificação de diretivas nacionais e internacionais de preservação de digital nos padrões arquivísticos, indicados pelo ICA; por ser um sistema livre, e garantir a confiabilidade dos documentos.

A importância dos profissionais citadosse faz presente pelo fato do *Archivematica* exigir competências e habilidades inerentes a esses, tais como: os procedimentos técnicos de instalação e manutenção, no auxílio de gerenciamento da informações do repositório, da descrição arquivística, compreensão de terminologias. Logo, é uma consonância de via mútua entre esses dois especialistas.

Desse modo, pode-se dizer que os objetivos propostos neste trabalho foram alcançados, uma vez que o *Archivematica* cumpriu em sua avaliação com os requisitos estabelecidos pela Resolução nº 43 do CONARQ. As funcionalidades e características do referido software de acordo com o *Guia do usuário* e sua página eletrônica estão de acordo com o que preconiza tal normativa, assim, respondendo ao problema proposto para esta dissertação.

Supõe-se que este estudo pode servir de inspiração para a continuidade de outras pesquisas e almeja-se, inclusive, que venha contribuir com a literatura brasileira sobre a temática.

Conclui-se que a análise das características e funcionalidades do *Archivematica* à luz da Resolução nº 43, foram cumpridos em sua totalidade, ou seja, todos os elementos desta diretriz se encontram previstos no software, o que definitivamente pode-se caracterizá-lo como um RDC-Arq, respondendo a nossa questão de investigação.

Por fim, sugere-se, vivamente que o *Archivematica* seja adotado e aplicado pelas instituições brasileiras em todas as suas esferas, sobretudo nas de ensino, incluindo, principalmente, os cursos de Arquivologia.

REFERÊNCIAS

AGUIAR, FRANCISCO LOPES. **Dspace e archivematica**: concepção e criação de um protótipo de repositório digital aplicado no domínio da SBPC- sob uma perspectiva interdisciplinar entre arquivística e organização do conhecimento. 2018. Tese (Doutorado em Ciência da Informação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2018.

ALVES, Virginia Barbara Aguiar. Open archives: via verde o via dourada? **Pontodeacesso**, Salvador, v. 2, n. 1, p. 127-137, ago./set. 2008.

ARAÚJO, Carlos Alberto Ávila. Ciência da Informação, Biblioteconomia, Arquivologia e Museologia: relações teóricas e institucionais. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 16, n. 31, p. 110-130, 2011.

_____. Correntes teóricas da Arquivologia. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 18, n. 37, p. 61–82, maio/ago., 2013.

_____. **O que é ciência da informação**. Belo Horizonte: KMA, 2018.

ARELLANO, Miguel Ángel Mardero; OLIVEIRA, Alexandre Faria de. Gestão de repositórios de preservação digital. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campina, v. 14, n. 3, p. 465–483, set./out. 2016.

ARCHIVEMATICA. Canadá, 201-. Disponível em: https://wiki.archivematica.org/Main_Page. Acesso em: 31 dez. 2018.

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

ARTEFACTUAL. (Canadá). Versão 1.8.1. **ARCHIVEMATICA**. Canadá, [entre 2008 e 2019]. Disponível em: <https://www.archivematica.org/en/>. Acesso em: 25 jan. 2019.

_____. (Canadá). Versão 1.8.1. **ARCHIVEMATICA**. Canadá, 201-a. Disponível em: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/#user-manual>. Acesso em: 24 jan. 2019.

_____. (Canadá). Versão 1.8.1. **ARCHIVEMATICA**. Canadá, 201-b. Disponível em: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/user-manual/preservation/preservation-planning/#preservation-planning>. Acesso em: 03 jan. 2019.

_____. (Canadá). Versão 1.8.1. Versão 1.8.1. **ARCHIVEMATICA**. Canadá, 201-c. Disponível em: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/user-manual/backlog/backlog/#backlog>. Acesso em: 03 jan. 2019.

_____. (Canadá). Versão 1.8.1. Versão 1.8.1. **ARCHIVEMATICA**. Canadá, 201-d. Disponível em: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/user-manual/appraisal/appraisal/#appraisal>. Acesso em: 03 jan. 2019.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 15427**: Sistemas espaciais de dados e informações: modelo de referência para um sistema aberto de arquivamento de informação (SAAI). Rio de Janeiro, 2007.

AUTRAN, Marynice de Medeiros Matos. **Comunicação da ciência, produção científica e rede de colaboração acadêmica**: análise dos Programas brasileiros de Pós-Graduação em Ciência da Informação. 2015. Tese (Doutorado em Informação e Comunicação em Plataformas Digitais) - Universidade de Aveiro Departamento de Comunicação e Arte, Porto, 2014.

_____; BORGES, Maria Manuel. Comunicação da ciência: (r)evolução ou crise? **RECIS**, p. 122-138, 08 jun. 2014.

BAPTISTA, Ana Alice *et al.* **Descrição do repositório de conteúdos educativos**: concepção de e-conteúdos e-learning. TecMinho/Gabinete de Formação Contínua da Universidade do Minho: Guimarães, 2006. 29 p. Disponível em: https://elearning.iefp.pt/pluginfile.php/49476/mod_resource/content/0/produtos/Descricao_Geral_do_Repositorio.pdf. Acesso em: 20 maio 2018.

BARRETO, Cássia Maria. **Modelo de metadados para a descrição de documentos eletrônicos na web**. 1999. Dissertação (Mestrado em Ciências em Sistemas de Computação) - Instituto Militar de Engenharia, Rio de Janeiro, 1999.

BELLOTTO, Heloísa Liberalli. **Arquivos permanente**: tratamento documental. 4. ed. Rio de Janeiro: FGV, 2006.

BJÖRK, Bo-Christer. Open access to scientific publications: an analysis of the barriers to change? **Information Research**, v. 9, n. 2, jan. 2004. Disponível em: <http://www.informationr.net/ir/9-2/paper170.html>. Acesso em: 15 maio 2018.

BORKO, Haroldo. Information Science: What is it? **American Documentation**, v. 19, n.1, p.3-5, jan. 1968.

BRASIL. Conselho Nacional de Arquivos. Câmara Técnica de Documentos Eletrônicos. **e-ARQ Brasil**: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos. Rio de Janeiro. Arquivo Nacional. 2011.

_____. **Guia livre**: referência de migração para software livre do governo federal. Brasília: Governo Federal, 2005.

_____. **Lei Federal Nº 8.159, de 8 de Janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8159.htm. Acesso em: 16 out. 2016.

_____. Conselho Nacional de Arquivos. Ministério da Justiça e Segurança Pública. O CONARQ: Rio de Janeiro, 2014. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/o-conselho.html>. Acesso em: 02 nov. 2019.

_____. Conselho Nacional de Arquivos. Ministério da Justiça e Segurança Pública. O SINAR: Rio de Janeiro: 201-a. Disponível em: <http://conarq.arquivonacional.gov.br/o-sinar.html>. Acesso em: 02 nov. 2019.

_____. Conselho Nacional de Arquivos. Ministério da Justiça e Segurança. A CTDE: Rio de Janeiro, 201-b. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/documentos-eletronicos-ctde.html>. Acesso em: 02 nov. 2018.

BUCKLAND, Michael Keeble. Information as thing. Tradução Luciane Artêncio. **Journal of the American Society for Information Science (JASIS)**, v. 45, n. 5, p. 351-360, 1991.

CAPURRO, Rafael; HJORLAND, Birger. O conceito de informação. **Perspectiva em Ciência da Informação**, v. 12, n. 1, p. 148-207, jan./abr. 2007.

CHIBENI, Silvio Seno. **O que é ciência?** 2006. Disponível em: <http://www.unicamp.br/~chibeni/textosdidaticos/ciencia.pdf>. Acesso em: 20 set. 2018.

COCCO, Ana Paula; RODRIGUES, Rosângela Schwarz. Repositórios institucionais de acesso aberto: cenário nos países ibero-americanos. **Informação & Sociedade: Estudos**, João Pessoa, v. 24, n. 2, p. 111-120, mais/ago. 2014.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). Ministério da Justiça. Resolução nº 43, de 04 de Setembro de 2015. Altera a redação da Resolução do CONARQ nº 39, de 29 de abril de 2014, que estabelece diretrizes para a implementação de repositórios digitais confiáveis para a transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas dos órgãos e entidades integrantes do Sistema Nacional de Arquivos - SINAR. **Diretriz para a implementação de repositório arquivísticos digitais confiáveis - RDC-Arq**, Rio de Janeiro, 2015.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Reference model for an Open Archival Information System (OAIS): recommended practice CCSDS 650.0-M-2**. Washington, DC: CCSDS, 2012. Disponível em: <https://public.ccsds.org/pubs/650x0m2.pdf>. Acesso em: 02 maio 2018.

COOK, Terry. Arquivos pessoais e arquivos institucionais: para um entendimento arquivístico comum da formação da memória em um mundo pós-moderno. **Revista Estudos Históricos**, v. 11, n. 21, p. 129-149, 1998.

CORUJO, Luis Miguel Nunes. **Repositórios Digitais e Confiança - um exemplo de repositório de preservação digital: o RODA**. 2014. Dissertação (Mestrado em Ciência da Documentação e Informação) - Faculdade de Letras, Universidade de Lisboa, Lisboa, 2014.

COSTA, Milene *et al.* **Guia do usuário archivematica**. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2016.

COSTA, Milene Aparecida da Silva; CASTRO, Priscilia de Paiva. Archivematica uma Ferramenta de Software livre para preservação de documentos arquivísticos digitais. **Revista Intercâmbio dos Congressos Internacionais de Humanidades**, Brasília, n. 6, p. 106-119, 2016.

COSTA, Sely Maria de Souza; LEITE, Fernando César Lima. Insumos conceituais e práticos para iniciativas de repositórios institucionais de acesso aberto à informação científica em bibliotecas de pesquisa. 2010. Disponível

em: <http://repositorio.unb.br/bitstream/10482/5470/1/Costa%20%26%20Leite%202010.pdf>.

Acesso em: 26 out. 2018.

CROW, Raym. The Case for institutional repositories : a SPARC position paper. **American Research Libraries: Bimonthly Report**, v. 223, n. 37, 2002.

DEMO, Pedro. **Metodologia do conhecimento científico**. São Paulo: Atlas, 2000.

DESLANDES, Suely Ferreira. O projeto de pesquisa como exercício científico e artesano intelectual. In: MINAYO, Maria Cecília de Souza (Org.). **Pesquisa social: teoria, método e criatividade**. São Paulo, Petrópolis: Vozes, 2009.

DISTRETTI, Henrique Alvares. **Repositório arquivístico digital confiável - RDC-Arq. SERPRO**. 201-. Disponível em: <http://intra.serpro.gov.br/tema/artigos-opinioes/repositorio-arquivistico-digital-confiavel-2013-rdc-arg>. Acesso em: 27 out. 2018.

DUCHEIN, Michel. O respeito aos fundos em arquivística: princípios teóricos e problemas práticos. **Arquivo & Administração**, Rio de Janeiro, v. 10-14, n. 2, p. 01-16, 1986.

DURANTI, Luciana. The archival body of knowledge: archival, theory, method and practice, and graduate and continuing education. **Journal of Education Library and Information Science**, v. 34, n. 1, p. 8-24, 1993.

FARIAS, Suelen Conceição. Os benefícios das Tecnologias da Informação e Comunicação (TIC) no processo de Educação a Distância (EAD). **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 11, n. 2, p. 15-29, set./dez. 2013.

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

FLORES, Daniel; HEDLUND, Dhion Carlos. A preservação do patrimônio documental através da produção de instrumentos de pesquisa arquivísticos e da implementação de repositórios arquivísticos digitais. **Série Patrimônio Cultural e Extensão Universitária**, n. 3, p. 33, fev. 2014

_____; LAMPERT, Sérgio Renato. As funções de produção, classificação e avaliação de documentos arquivísticos no software Nuxeo Document Management. **Informação Arquivística**, Rio de Janeiro, v. 2, n. 1, p. 41-64. jan/jun. 2013.

_____; PRADEBON, Daiane Segabinazzi; CÉ, Graziella. Análise do conhecimento teórico- metodológico da preservação digital sob a ótica da OAIS, SAAI, ISO 14721 e NBR 15472. **Brazilian Journal of Information Science: Research Trends**, v. 11, n. 4, p. 72-80, 2017.

_____. ROCCO, Brenda Couto de Brito; SANTOS, Henrique Machado dos. Cadeia de custódia para documentos arquivísticos digitais. **Acervo**, Rio de Janeiro, v. 29, n. 2, p. 117-132, jul./dez. 2016.

FONSECA, Maria Odila. **Arquivologia e ciência da informação**. Rio de Janeiro: FGV, 2005.

FONTANA, Fabiana Fagundes *et al.* Archivemática como ferramenta para acesso e preservação digital à longo prazo. **Ágora**, Florianópolis, v. 24, n. 48, p. 62-82, 2014.

FREIRE, Gustavo Henrique de Araújo; FREIRE, Isa Maria. **Introdução à Ciência da Informação**. João Pessoa: Editora Universitária da UFPB, 2009.

GOMES, Romeu. Análise e interpretação de dados de pesquisa qualitativa. *In*: MINAYO, Maria Cecília de Souza. (Org.). **Pesquisa social: teoria, Método e Criatividade**. São Paulo: Petrópolis: Vozes, 2004. p. 79-108. 2004.

GONÇALEZ, Paula Reginas Ventura Amorim. Recomendações para certificação ou medição de confiabilidade de repositórios arquivísticos digitais com ênfase no acesso à informação. **Informação & Informação**, v. 22, n. 1, p. 215-241, 2017.

GOODY, Jack. **The logic of writing and the organization of society**. Cambridge: Cambridge University Press, 1988.

HEEMANN, Ademar. Considerações sobre alguns obstáculos à interdisciplinaridade na pós-graduação. **Desenvolvimento e Meio Ambiente**, n. 10, p. 47-51, jul./dez. 2004.

HEXSEL, Roberto André. **Software Livre: propostas de ações de governo para incentivar o uso de software livre**. Curitiba: Universidade Federal do Paraná, 2002.

INNARELLI, Humberto Celeste. Os dez mandamentos da preservação digitais: uma brevíssima introdução. *In*: II SEMINÁRIO SERVIÇOS DE INFORMAÇÃO EM MUSEUS. 2012, São Paulo. **Anais[...]**. São Paulo: SESC, 2012. p. 318-325.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). Brasília, [200-?]. Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorio-institucional-do-ibict%28ridi%29/?searchterm=reposit%C3%B3rio>. Acesso em: 05 nov. 2018.

JARDIM, José Maria; FONSECA, Maria Odila. As relações entre a Arquivística e a Ciência da Informação. **Cadernos Bad**, Lisboa, v. 2, p. 29-45, 1992.

JENKINSON, Hilary. **A manual of archive administration**. 3. ed. Oxford: Clarendon Press, 1965.

JORENTE, Maria José Vicentini *et al.* O marco civil da internet e a ciência da informação: uma discussão sobre os softwares livres atom e archivemática. **Liinc em Revista**, Rio de Janeiro, v. 12, n. 1, p. 90-106, maio 2016.

KETELLAR, Eric. Archival temples, archival prisons: modes of power and protection. **Archival Science**, v. 2, n. 4 p. 221-238, 2002.

KUHN, Thomas Samuel. **A estrutura das revoluções científicas**. Tradução Beatriz Vianna Boeira e Nelson Boeira. 13. ed. São Paulo: Perspectiva, 2017.

LAMPERT, Sérgio Renato. Os repositórios DSpace e Archivematica para documentos arquivísticos digitais. **Acervo**, Rio de Janeiro, v. 29, n. 2, p. 143-154, jul./dez. 2016.

LE COADIC, Yves-François. **A ciência da informação**. Brasília: Briquet de Lemos, 1996.

LESSIG, Lawrence. **Free culture: How big media uses technology and the law to lock down culture and control creativity**. New York: The Penguin Press, 2004.

LÈVY, Pierre. **As tecnologias da inteligência: o futuro do pensamento na era da informática**. Tradução Carlos Irineu da Costa. 34. ed. Rio de Janeiro: Letras, 2008.

LYNCH, Clifford. Institutional repositories: essential infrastructure for scholarship in the digital age. **American Research Libraries: Bimonthly Report**, v. 226, n. 11, p. 1-7. 2003.

LOPEZ, André Porto Ancona. Princípios Arquivísticos e documento digitais. **Revista Arquivo Rio Claro**, n. 2, p. 70-85. 2004.

LOURENÇO, Cíntia Azevedo. Metadados: o grande desafio na organização da web. **Ciência da Informação**, v. 17, n. 1, p. 65-72, jan./abr. 2007.

LYOTARD, Jean François. **A condição pós-moderna**. 12. edição. Tradução Ricardo Corrêa Barbosa. Rio de Janeiro: José Olympo, 2009.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MARTIZEZ, Ninfa *et al.* **Guia de instalação e configuração archivematica/atom**. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2017.

MIKHAILOV, Alexander Ivanovich; CHERNYI, Arkadii Ivanovich; GILYAREVSKYI, Rudzero. Estrutura e principais propriedades da informação científica: a propósito do escopo da Informática. *In*: GOMES, Hagar Espanha (Org.). **Ciência da Informação ou Informática?** Rio de Janeiro: Calunga, 1980. p. 70-89.

MINAYO, Maria Cecília de Sousa (Org.). **Pesquisa social: teoria, método e criatividade**. 30. ed. Rio de Janeiro: Vozes, 2011.

MOOERS, Northrup Calvin. Zatoncoding applied to mechanical organization of knowledge. **American Documentation**, v. 2, n.1 p. 20-32. 1951.

MUELLER, Suzana Pinheiro Machado. A Comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006.

OpenDOAR. Diretório de repositórios de acesso aberto. Reino Unido, 2006. Disponível em: <http://v2.sherpa.ac.uk/pendoar/search.html>. Acesso em: 24 out. 2018.

OLIVEIRA, Vicente Matellán *et al.* **Sobre software livre**: compilación de ensaio sobre software livre. Madri: Universidad Rey Juan Carlos, 2004.

PINHEIRO, Lena Vania Ribeiro. **A Ciência da Informação entre sombra e luz** : Domínio Epistemológico e Campo Interdisciplinar. Rio de Janeiro: UFRJ, 1997.

_____. Ciência da Informação: Desdobramento Disciplinares, Interdisciplinares e Transdisciplinares. *In*: GÓMEZ, Maria Nelida Gonzáles de; ORICO, Dill, Goyannes, Evelyn (Org). **Políticas de memória e informação**. Natal: EDUFRN, 2006. p. 111-142.

_____; LOUREIRO, José Mauro Matheus. Traçados e limites da ciência da informação. **Ciência da Informação**, v. 24, n. 1, p. 42–53, 1995.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico**: métodos e técnicas da pesquisa do trabalho acadêmico. 2 ed. Novo Hamburgo: Feevale, 2013.

POSNER, Ernst. **Archives in the ancient world**. Cambridge: Harvard University Press. 1972.

RESEARCH LIBRARY GROUP (USA). **Trusted Digital Repositories**: attributes and responsibilities. California, 2002, 62 p. Disponível em: <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>. Acesso em: 25 out. 2018.

RODRIGUES, Mateus de Moura. **Repositório arquivístico digital confiável para o patrimônio documental oriundo do processo judicial eletrônico**. 2015. Dissertação (Mestrado em Patrimônio Cultural) - Universidade Federal de Santa Maria, Centro de Ciências Sociais e Humanas, Programa de Pós-Graduação Profissionalizante em Patrimônio Cultural, Rio Grande do Sul, 2015.

RONDINELLI, Rosely Curi. **Gerenciamento Arquivístico de documentos eletrônicos**: uma abordagem teórica da diplomática contemporânea. Rio de Janeiro: FGV, 2005.

_____. **O Documento arquivístico ante a realidade digital**: uma revisão conceitual necessária. 4 ed. Rio de Janeiro: FGV, 2013.

ROSSEAU, Jean-Yves; COUTURE, Carol. Os fundamentos da disciplina arquivística. Tradução Magda Bigotte de Figueiredo. Lisboa: Publicações Dom Quixote, 1994.

SANTOS, Boaventura de Sousa. **Um discurso sobre as ciências**. 5. ed. São Paulo: Cortez, 2008.

SANTOS, Henrique Machado dos; FLORES, Daniel. Políticas de preservação digital para documentos arquivísticos. **Perspectivas em Ciência da Informação**, v. 20, n. 4, p. 197–217, out./dez. 2015.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas da Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SCHELLENBERG, Theodore Roosevelt. **Arquivos modernos: princípios e técnicas**. Tradução: Nilza Teixeira Soares. 6. ed. Rio de Janeiro: FGV, 2006.

SILVA, Armando Malheiros da; RIBEIRO, Fernanda. **Das ciências documentais à ciência da informação: ensaio epistemológico para um novo modelo curricular**. 2. ed. Porto: Afrontamento, 2008.

_____. *et al.* **Arquivística: teoria e prática de uma ciência da informação**. 3. ed. Porto: Afrontamento, 2009.

SOUSA FILHO, Agenor Leandro de *et al.* Importância dos repositórios institucionais na preservação intelectual: Em foco a gestão do conhecimento. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 2, n. 2, p. 1-17, out. 2012.

TANUS, Gabrielle Francinne de Souza Carvalho; ARAÚJO, Carlos Alberto Ávila. Proximidades Conceituais Entre Arquivologia, Biblioteconomia. **Biblionline**, João Pessoa, v. 8, n. 2, p. 27-36, 2012.

THOMASSEN, Theo. A first Introduction to archival science. **Archival Science**, v. 1, n. 4, p. 373-385, dez. 2001.

VAN GARDEREN, Peter. Archivematica: using micro-services and open-source software to deliver a comprehensive digital curation solution. *In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON PRESERVATION OF DIGITAL OBJECTS*, 7, 2010. Austria. **Anais [...]**. Austria: Comitê Austríaco da Sociedade de Computação para Relações Públicas, 2010. p. 145-149.

_____; COURTNEY, Mumma. Realizing the archivematica vision: delivering a comprehensive and free OAIS implementation. *In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON PRESERVATION OF DIGITAL OBJECTS*, 10, 2013. Portugal. **Anais [...]**. Portugal: Biblioteca Nacional de Portugal, 2013. p. 84-87.

WEITZEL, Simone da Rocha. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 54-71, jan./jun. 2006.

ANEXO A –RESOLUÇÃO Nº 43, DE 4 DE SETEMBRO DE 2015

Altera a redação da Resolução do CONARQ nº 39, de 29 de abril de 2014, que estabelece diretrizes para a implementação de repositórios digitais confiáveis para a transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas dos órgãos e entidades integrantes do Sistema Nacional de Arquivos - SINAR.

O PRESIDENTE DO CONSELHO NACIONAL DE ARQUIVOS - CONARQ, no uso de suas atribuições, previstas no item IX do art. 23 de seu Regimento Interno, aprovado pela Portaria nº 2.588, do Ministério da Justiça, de 24 de novembro de 2011 e de acordo com a deliberação adotada na 80ª Reunião Plenária, realizada no dia 12 de agosto de 2015, Resolve:

Art. 1º - A ementa da Resolução do CONARQ Nº 39, de 29 de Abril de 2014, passa a vigorar com a seguinte alteração:

Estabelece diretrizes para a implementação de repositórios arquivísticos digitais confiáveis para o arquivamento e manutenção de documentos arquivísticos digitais em suas fases corrente, intermediária e permanente, dos órgãos e entidades integrantes do Sistema Nacional de Arquivos (SINAR).

Art. 2º - O art. 1º da Resolução do CONARQ Nº 39, de 29 de Abril de 2014, passa a vigorar com a seguinte alteração:

Art. 1º - Aprovar as Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq, anexas a esta Resolução, e recomendar sua adoção aos órgãos e entidades integrantes do Sistema Nacional de Arquivos - SINAR, para o arquivamento e manutenção dos documentos arquivísticos em suas fases corrente, intermediária e permanente em formato digital, e de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade e a preservação desses documentos.

Art. 3º - A redação do anexo da Resolução N.º 39, de 29 de Abril de 2014, passa a vigorar com as seguintes alterações:

Na página 1, onde se lê: "DIRETRIZES PARA A IMPLEMENTACAO DE REPOSITARIOS DIGITAIS CONFIÁVEIS DE DOCUMENTOS ARQUIVÍSTICOS".

Nova redação: "DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS - RDC-Arq".

Na página 2, onde se lê: "DIRETRIZES PARA A IMPLEMENTACAO DE REPOSITORIOS DIGITAIS CONFIÁVEIS DE DOCUMENTOS ARQUIVÍSTICOS".

Nova redação: "DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS - RDC-Arq".

Na página 3, onde se lê: "II. Repositório digital confiável de documentos arquivísticos - principais requisitos".

Nova redação: "II. Repositório Arquivístico Digital Confiável - RDC-Arq: principais requisitos".

Na página 5, onde se lê: "Assim, em face da necessidade de implantação de repositórios digitais confiáveis para documentos arquivísticos digitais, nas fases corrente, intermediária e permanente, o CONARQ apresenta estas diretrizes".

Nova redação: "Assim, em face da necessidade de implantação de repositórios digitais confiáveis para documentos arquivísticos digitais, nas fases corrente, intermediária e permanente, o CONARQ apresenta estas diretrizes de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq".

Na página 5, onde se lê: "Indicar parâmetros para repositórios confiáveis de documentos arquivísticos digitais, de forma a garantir a integridade, a autenticidade, a confidencialidade, a disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente".

Nova redação: "Indicar parâmetros para repositórios arquivísticos digitais confiáveis, de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente".

Na página 10, onde se lê: "Um repositório digital confiável de documentos arquivísticos deve ser capaz de atender aos procedimentos arquivísticos e aos requisitos de um repositório digital confiável."

Nova redação: "Um repositório arquivístico digital confiável deve ser capaz de atender aos procedimentos arquivísticos em suas diferentes fases e aos requisitos de um repositório digital confiável".

Na página 19, onde se lê: "A seguir, são apresentados documentos de referência para a construção de repositórios digitais confiáveis de documentos arquivísticos".

Nova redação: "A seguir, são apresentados documentos de referência para a construção de repositórios arquivísticos digitais confiáveis - RDC-Arq."

Art. 4º Esta Resolução entra em vigor na data de sua publicação.

JAIME ANTUNES DA SILVA

Presidente do CONARQ

APÊNDICE A – REQUISITOS PARA UM REPOSITÓRIO DIGITAL CONFIÁVEL

Os requisitos apresentados a seguir estão definidos em nível conceitual e devem ser cumpridos no desenvolvimento de um repositório digital confiável. Reitere-se que esses requisitos estão baseados na norma ISO 16363: 2012, e abrangem todos os tipos de materiais digitais, inclusive os documentos arquivísticos.

Os requisitos estão organizados em três conjuntos: infraestrutura organizacional; gerenciamento do documento digital; e tecnologia, infraestrutura técnica e segurança.

1 Infraestrutura organizacional

O ambiente em que o repositório digital vai se estabelecer tem que cumprir determinados requisitos, conforme descrito a seguir.

a. Governança e viabilidade organizacional: o repositório tem como missão o compromisso com a preservação, o gerenciamento e o acesso de longo prazo dos documentos digitais. Essa missão é claramente identificada por todos os interessados no repositório e envolve: mandato legal, contexto organizacional e requisitos regulatórios.

O repositório tem um plano de sucessão formal, planos de contingência e/ou acordos estabelecidos para garantir a continuidade do serviço, no caso de o repositório parar de operar ou de a instituição responsável e/ou financiadora mudar seu escopo.

b. Estrutura organizacional e de pessoal: o repositório tem uma equipe dotada de qualificação e formação necessárias, e em número suficiente, para garantir todos os serviços e funcionalidades pertinentes ao repositório. Além disso, deve manter um programa de desenvolvimento profissional contínuo.

c. Transparência de procedimentos e arcabouço político: o repositório deve demonstrar explicitamente seus requisitos, decisões, desenvolvimento e ações que garantem a preservação de longo prazo e o acesso a conteúdos digitais sob seus cuidados. Dessa forma, assegura aos usuários, gestores, produtores e certificadores que está cumprindo plenamente seu papel enquanto um repositório digital confiável. Para tanto, o repositório deve:

- definir a comunidade alvo e sua base de conhecimento;
- possuir políticas e definições, acessíveis publicamente, que demonstrem como os requisitos do serviço de preservação serão contemplados;
- possuir políticas, procedimentos e mecanismos de atualização, na medida em que o repositório cresce e a tecnologia e práticas da comunidade evoluem;
- documentar permissões legais – por meio de acordos de custódia, normas de procedimentos e outros – que o isentem de responsabilidade, no caso de alterações passíveis de ocorrer em estratégias de preservação digital;
- fazer o registro histórico das mudanças de procedimentos, de software e hardware;
- relacionar o registro histórico, acima referido, com as estratégias de preservação digital, e descrever os potenciais efeitos dessas mudanças sobre os documentos digitais;
- demonstrar que está sistematicamente avaliando a satisfação das expectativas dos produtores e dos usuários, e buscando atendê-las;
- estar comprometido com a definição, coleta, auditoria e fornecimento (sob demanda) de mecanismos de controle da integridade dos documentos digitais sob sua custódia;
- estar comprometido em realizar regularmente uma autoavaliação de seu funcionamento e

<p>renovar sua certificação;</p> <ul style="list-style-type: none"> • estar comprometido em notificar as entidades certificadoras sobre as mudanças operacionais que afetarão seu status de certificação (no caso de repositórios já certificados).
<p>d. Sustentabilidade financeira: um repositório digital confiável deve demonstrar sustentabilidade financeira. Para isso, deve ter um plano de gestão que observe os seguintes aspectos:</p> <ul style="list-style-type: none"> • demonstração da capacidade de obter recursos financeiros estáveis e contínuos para sustentá-lo, seja por meio de prestação de serviço, parcerias, doações, verba da própria instituição, dentre outros; • revisão e ajustes anuais; • transparência dos procedimentos para obtenção dos recursos e auditoria dos mesmos, de acordo com o sistema jurídico no qual o repositório se insere; e • compromisso dos ciclos de planejamento com o equilíbrio dos riscos, benefícios, investimentos e gastos.
<p>e. Contratos, licenças e passivos: os contratos, licenças e passivos firmados pelo repositório devem ser claros e mensuráveis; delinear papéis, responsabilidades, prazos e condições; e ser facilmente acessíveis ou disponíveis aos interessados. Esses contratos, licenças e passivos podem envolver tanto a relação entre o repositório e os produtores de documentos digitais, como a relação entre o repositório e fornecedores de serviços. Esses mesmos instrumentos devem especificar todos os direitos e obrigações do repositório sobre os documentos digitais a ele confiados, em especial no que diz respeito à propriedade intelectual e a restrições de uso.</p>
<p style="text-align: center;">2 Gerenciamento do documento digital</p> <p>O gerenciamento dos documentos de um repositório digital confiável deve estar de acordo com o modelo de referência OAIS, que estabelece a formação de pacotes de informação envolvendo os documentos digitais (informação de conteúdo) e seus metadados (informação de representação).</p> <p>São três os tipos de pacotes de informação:</p> <ul style="list-style-type: none"> • Pacote de informação para submissão <i>Submission Information Package</i> (SIP) – refere-se à admissão dos documentos digitais e seus metadados associados. • Pacote de informação para arquivamento <i>Archival Information Package</i> (AIP) – refere-se ao acondicionamento e armazenamento dos documentos digitais e seus metadados associados. • Pacote de informação para disseminação <i>Dissemination Information Package</i> (DIP) – refere-se ao acesso aos documentos digitais e seus metadados associados. <p>A TRAC apresenta os requisitos para gerenciamento do documento no repositório digital, categorizados em seis grupos, com base nas funcionalidades, conforme detalhado a seguir:</p> <p>a. Admissão (captura de documentos digitais): consiste na entrada dos documentos e seus metadados no repositório digital. Os requisitos de admissão variam dependendo do tipo de material, do contexto legal e da relação entre o produtor de documento e o repositório. Independentemente dessas variações, pode-se afirmar que a admissão se inicia com o recebimento de um SIP, que é convertido em AIP, e termina quando um AIP está seguro no repositório, incluindo a criação de cópias de segurança.</p> <p>A seguir, apresentam-se requisitos gerais a serem cumpridos pelo repositório, cuja adequação deve ser avaliada de acordo com a missão e as necessidades de cada repositório:</p> <ul style="list-style-type: none"> • identificar as propriedades do documento que serão preservadas (ex.: o conteúdo, <i>layout</i>,

tabela de cor, resolução da imagem, canais de som etc.);

- especificar claramente a informação que deve estar associada ao documento (metadados associados) no momento da sua submissão;
- ter mecanismos para autenticar a origem dos documentos que estão sendo admitidos no repositório, de forma a garantir sua proveniência;
- ter procedimentos para verificar a integridade do SIP, o que pode ser feito por meio de procedimentos automatizados e/ou checagem humana;
- ter o controle físico (controle completo dos *bits*) dos documentos transmitidos com cada SIP, a fim de preservá-los;
- fornecer ao produtor/depositante relatórios do andamento dos procedimentos durante todo o processo de admissão;
- demonstrar em que momento a responsabilidade pela preservação do documento submetido (SIP) é formalmente aceita pelo repositório; e
- ter registros de todas as ações e processos administrativos que ocorrem durante o processo de admissão e são relevantes para a preservação.

No caso de um repositório para documentos arquivísticos, a definição dos metadados deve observar o e-ARQ Brasil (nas fases corrente e intermediária) e a NOBRADE (na fase permanente). Para a admissão de documentos no repositório, no caso de transferência ou recolhimento, devem-se observar os procedimentos indicados na Resolução nº 24, de 3 de agosto de 2006, do CONARQ.

b. Admissão (criação do pacote de arquivamento): o repositório deve completar o processo de admissão, criando um pacote de informação apropriado para arquivamento (AIP), com toda a informação recebida do produtor.

A fim de garantir que o pacote de informação recebido do produtor, e verificado pelo repositório, seja convertido para o formato de arquivamento (AIP) e armazenado para preservação de longo prazo, um repositório deve atender os seguintes requisitos:

- descrever cada classe de informação (texto estruturado, imagem matricial, banco de dados, imagem em movimento e outras) a ser preservada pelo repositório, e como ela está implementada – essa descrição deve apontar os componentes-chave do AIP: o documento arquivístico, sua informação de representação (informação estrutural e semântica) e as várias categorias de informação descritiva de preservação (fixidade, proveniência e contexto), e ainda como esses componentes se relacionam;
- descrever minuciosamente as diferentes classes de informação e como os AIPs são implementados, nos casos em que a especificidade daquelas classes exigir ações de preservação diferentes (por exemplo, a imagem TIFF que é processada por um sistema pode necessitar de ações de preservação diferentes das ações necessárias à imagem TIFF que é apresentada para o olho humano);
- descrever como os AIPs são construídos a partir dos SIPs, ou seja, apontar todas as transformações pelas quais passarão os documentos e os metadados submetidos, e os metadados a serem adicionados no momento da formação do AIP;
- ser capaz de demonstrar se os SIPs foram aceitos e transformados em um AIPs integralmente ou em parte, ou ainda se foram recusados;
- atribuir aos AIPs, identificadores que sejam únicos, persistentes e visíveis aos gestores e auditores, de acordo com padrões reconhecidos (por exemplo: *Handle System*, DOI, URN, PURL);
- no caso de o documento já possuir um identificador único, a ele atribuído no SIP, o repositório deverá mantê-lo no AIP, ou criar um outro identificador, que deverá ser associado, de maneira persistente, ao do SIP;
- ter acesso a ferramentas amplamente reconhecidas para apoiar o monitoramento dos componentes digitais dos documentos, tais como diretórios de formatos de arquivos (ex.: PRONOM – base de dados com registro de formatos mantida pelo arquivo nacional do Reino Unido) e registros de outras informações de representação;

- registrar, em um banco de dados local, a informação de representação dos documentos admitidos, quando essa informação não estiver disponível nas ferramentas mencionadas no ponto anterior;
- registrar metadados de preservação associados aos documentos admitidos, de maneira a apoiar sua integridade, localização, legibilidade e proveniência, dentre outros;
- ter procedimentos para testar se os documentos são compreensíveis pela comunidade-
- alvo e, em caso negativo, adequá-los às necessidades dessa comunidade (ex.: documentos voltados para deficientes visuais);
- verificar a completude e a correção de cada AIP no momento em que é gerado, isto é, no momento em que o SIP é convertido em AIP;
- ter um mecanismo independente para verificar a integridade do conjunto do seu acervo, ou seja, verificar que todos os documentos previstos foram, de fato, admitidos no repositório, justificando possíveis lacunas; e
- documentar todas as ações relevantes à preservação dos documentos e que estão relacionadas à criação do AIP.

c. Planejamento da preservação: um repositório digital deve fazer o 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, ser bem documentado e incluir:

- estratégias de preservação bem definidas e periodicamente atualizadas, apontando e detalhando cada procedimento a ser adotado, como, por exemplo, a normalização de formatos;
- mecanismos para monitoramento e notificação quando alguma informação de representação dos documentos no repositório estiver se tornando obsoleta ou inviável (ex.: um formato de arquivo que esteja entrando em desuso, um suporte que esteja no final de sua vida útil);
- mecanismos de mudanças do plano de preservação como resultado do monitoramento;
- fornecimento de evidências sobre a eficácia do plano de preservação.

d. Armazenamento e preservação/manutenção do AIP: um repositório deve atender a um conjunto de condições para garantir o bom desempenho da preservação de longo prazo dos AIPs. Tais condições são:

- utilização das estratégias previstas no planejamento da preservação, que podem ser várias e devem ser registradas nos metadados de preservação;
- atender minimamente a dois aspectos da preservação digital – os cuidados com armazenamento (controle dos suportes, dos formatos e da localização de cópias) e a eventual necessidade de migração (atualização de suportes e conversão de formatos);
- preservação do documento digital (informação de conteúdo do AIP) originalmente admitido no repositório e daquele resultante da última migração;
- monitoramento constante da integridade dos AIPs, por meio do registro de metadados de fixidade e de logs de checagem dessa integridade (por exemplo, *checksum*);
- registro de todas as ações de preservação realizadas nos AIPs.

As migrações podem provocar alterações na forma e no conteúdo do documento, entretanto, no caso de documentos arquivísticos, não se admite a alteração de conteúdo. As migrações e quaisquer alterações da forma documental daí decorrentes devem ser registradas como metadados, a fim de apoiar a presunção de autenticidade do documento.

e. Gerenciamento de informação: uma funcionalidade essencial de um repositório digital confiável é o gerenciamento da informação, aqui entendido como a gestão das informações descritivas (metadados) dos documentos admitidos no repositório. O principal objetivo desses metadados é apoiar o acesso e a recuperação dos documentos, e vão além das informações descritivas mais usuais (autor, título, data), envolvendo outras informações descritivas úteis aos usuários, tais como tamanho do

arquivo disponível para download ou informação sobre a aplicação necessária para ler o arquivo. O gerenciamento da informação descritiva envolve os seguintes aspectos:

- metadados mínimos que permitam a busca e localização dos documentos – esses metadados devem ser identificadores conhecidos pela comunidade-alvo de usuários (ex.: número de matrícula do servidor público, título de livro numa biblioteca, número de processo);
- captura ou criação dos metadados mínimos pelo repositório, durante o processo de admissão, e associação desses metadados ao AIP correspondente;
- integridade referencial entre os AIPs e sua informação descritiva (metadados), ou seja, todo AIP deve ter uma informação descritiva, e toda informação descritiva deve apontar para um AIP; e
- permanência da integridade referencial, mesmo no caso de quebra temporária da relação entre o AIP e seus metadados descritivos – nesse caso, o repositório deve ser capaz de restaurar a relação rompida.

f. Gerenciamento de acesso: todo repositório deve produzir pacotes de disseminação de informação (DIP), atendendo aos seguintes requisitos:

- divulgação, para a comunidade de usuários, das opções disponíveis de acesso aos documentos e de entrega dos mesmos;
- implementação de uma política de registro dos acessos ocorridos que esteja de acordo com as necessidades de controle desses acessos, tanto da parte do repositório como dos produtores dos documentos nele admitidos;
- concessão de acesso a cada AIP, para os usuários autorizados e da forma devida (ex.: autorização de “somente leitura”, ou acesso a um número limitado de itens por período), em conformidade com o acordo estabelecido entre o repositório e o produtor/depositante;
- documentação e implementação de políticas de acesso (identificação e autenticação de usuários), em conformidade com os acordos estabelecidos entre o repositório e o produtor/depositante – essas políticas de acesso podem variar, desde a isenção da necessidade de identificação de usuário até o controle rígido da identificação e autenticação do usuário;
- registro de falhas de controle de acesso (como, por exemplo, um acesso indevidamente negado) e uso desse registro para avaliar eventuais falhas no sistema de segurança;
- demonstração de que o processo que gera o DIP atende completamente à requisição do usuário (ex.: se o usuário pediu um conjunto de documentos, receberá o conjunto completo; se ele pediu um documento, receberá apenas esse único documento);
- demonstração de que o processo que gera o DIP está correto em relação ao pedido do usuário (ex.: se o repositório oferece imagens nos formatos JPG e PNG, o usuário deve receber, dentre esses, o formato que solicitou);
- demonstração de que todos os pedidos de acesso resultam em uma resposta de aceitação ou rejeição; e
- garantia da autenticidade dos DIPs, por meio da entrega de cópias autênticas dos originais ou da viabilidade de rastreamento auditável da relação entre o DIP e o objeto original – para isso, um repositório deve ser capaz de demonstrar o processo de construção do DIP a partir de um AIP.

3 Tecnologia, infraestrutura técnica e segurança

Esses requisitos não prescrevem hardware e software específicos para garantir a preservação de longo prazo dos AIPs, mas apenas descrevem as melhores práticas das áreas de gestão de dados e segurança, que devem ser atendidas por um repositório digital confiável.

a. Infraestrutura de sistema: um repositório deve possuir uma infraestrutura tecnológica robusta, de maneira a apoiar a confiabilidade dos AIPs nele mantidos. Para tanto, deve observar os seguintes aspectos:

- funcionamento do repositório com base num sistema operacional e outros softwares de infraestrutura que tenham um bom suporte do mercado e da comunidade de usuários;
- adequação dos processos, do hardware e do software do sistema de backup às necessidades do repositório;
- gerenciamento do número de cópias de todos os documentos mantidos no repositório, e a localização de cada uma delas;
- mecanismos para garantir o sincronismo entre as cópias de um mesmo documento, ou seja, garantir que as mudanças intencionais feitas em uma cópia sejam propagadas para todas as outras;
- mecanismos efetivos para a detecção de corrupção ou perda de bits;
- relato dos incidentes de corrupção ou perda de dados eventualmente ocorridos e adoção de medidas para reparação ou substituição desses mesmos dados;
- previsão de procedimentos de atualização de suporte (*refreshing*) e de migração decorrentes do cumprimento do prazo de vida do suporte ou da obsolescência dos componentes de hardware;
- documentação da gestão de mudanças capaz de identificar alterações em processos críticos que afetem a capacidade de o repositório cumprir com suas responsabilidades obrigatórias;
- previsão de procedimentos para testar o efeito de mudanças críticas no sistema;
- ponderação entre os riscos e os benefícios nas decisões de atualização de software de segurança.

b. Tecnologias apropriadas: o repositório deve adotar uma tecnologia de hardware e software apropriada para os serviços que presta, procedimentos para o recebimento e monitoramento de notificações e para a avaliação da necessidade de mudanças na tecnologia utilizada.

c. Segurança: a segurança do repositório não se limita a aspectos de tecnologia, mas abrange também instalações físicas e ações de pessoas. Os aspectos de segurança incluem:

- análise sistemática de dados, sistemas, pessoas e instalação física;
- adoção de procedimentos de controle para tratar adequadamente as necessidades de segurança;
- delineamento de papéis, responsabilidades e autorizações relativas à implementação de mudanças no sistema; e
- plano de prevenção de desastres e de reparação, que inclua, ao menos, um *backup, off-site*, de tudo o que é mantido no repositório (documentos, metadados, trilhas de auditoria etc.), inclusive do próprio plano de reparação.

Fonte: Adaptado do Conselho... (2015).

APÊNDICE B – FUNCIONALIDADES DO ARCHIVEMATICA

As funcionalidades descritas neste Apêndice - B estão de acordo com o que preconiza o *Guia do usuário Archivematica* publicado pelo IBICT e também pelo sítio eletrônico do *Archivematica*. O primeiro documento elenca as incumbências do software na versão 1.6, apresentando seis guias/paineis, a saber: Transferência (*Transfer*), Ingerir¹⁰ (*Ingest*), Armazenamento Arquivístico (*ArchivalStorage*), Planejamento de Preservação (*Preservation Planning*), Acesso (*Access*) e Administração (*Administration*). Na versão 1.8 (do site oficial) são introduzidos mais dois itens: Lista de pendências (*Backlog*) e Avaliação (*Appraisal*).

As funções do *Archivematica* são projetadas por uma interface na web direcionada à perspectiva do usuário final (Figura 13).

¹⁰ Pode-se encontrar variações do termo “Ingerir”, como sinônimo de “Ingestão” ou “Admissão”. Na literatura brasileira sobre o *Archivematica*, tais terminologias podem apresentar quaisquer um desses três tipos.

Figura 13 - Painel baseado na web do Archivematica

archivematica Transferência 1 Lista de pendências Avaliação Ingerir 1 Armazenamento arquivístico Planejamento de preservação Acesso Administração demo@example.com Conectado

Padrão
Tipo de transferência

Nome da transferência

Acesso não.

ID do sistema de acesso

Squeaky toy

Iniciar transferência

Aprovar automaticamente

Transferir	UUID	Hora de início da transferência
1	77723ce2-cd07-4a6e-a4c8-fa0b1d09de5e	2018-12-18 05:41
<ul style="list-style-type: none"> Microservice: transferência falhada Microservice: limpar nomes Microservice: gerar relatório de estrutura de transferência Microservice: Varredura por vírus Microservice: Quarentena Microservice: Gerar documento METS.xml Microservice: Verificar as somas de verificação de transferência Microservice: formatar arquivos de metadados Microservice: Atribui arquivos UUIDs e checksums Microservice: Incluir Processamento de Transferência padrãoMCP.xml Microservice: Verifique a conformidade da transferência Microservice: Renomear com transferência UUID 		
jhbjhubf_2	fa67593e-8401-499e-87c8-7cfff5d97231	2018-12-18 05:45
<ul style="list-style-type: none"> Microservice: transferência falhada Microservice: Pacotes de extração Microservice: Identifique o formato do arquivo Microservice: limpar nomes Microservice: gerar relatório de estrutura de transferência Microservice: Varredura por vírus Microservice: Quarentena 		

Fonte: <https://sandbox.archivematica.org/transfer/>.

Vê-se na imagem acima que as abas compreendidas por oito guias estão organizadas em uma interface para o usuário que manuseia o software, isto quer dizer seus utilizadores podem submeter e alterar informações.

a) Guia Transferência (*Transfer*)

Ao painel de Transferência se atribui a responsabilidade pelo processo de transformar qualquer conjunto de objeto de digital ou diretório em um SIP. Costa *et al* (2016) afirmam que:

A transformação pode incluir avaliação, arranjo, descrição e a identificação restrita do doador, privado ou de conteúdos confidenciais. A guia Transferência prepara seu conteúdo para a preservação no *Archivematica*. Na guia Transferência do painel, o usuário move objetos digitais a partir de diretórios de origem acessíveis através do serviço de armazenamento no *Archivematica*. Uma vez carregada no painel de instrumentos, a transferência executa vários microsserviços: Atribuição UUID; verificação de *checksum* (se as somas de verificação estão presentes); extração de pacote (ou seja, de descompactação de arquivos compactados ou não embalados); verificação de vírus; indexação; identificação formato e validação; e extração de metadados. (COSTA *et al*, 2016. p. 22, grifo nosso).

Para criar uma Transferência faz-se necessário o login do usuário, e, além disso, o preenchimento de seus campos: selecionar o tipo de transferência (padrão, saco descompactado, saco com zíper, Dspace, imagem de disco, Dataverse); nomear esta transferência; inserir um número de acesso se for preciso; e iniciar a transferência de fato no quadro tarjado verde (Figura 14).

Figura¹¹ 14 - Campos de preenchimento da guia Transferência

archivematica Transferência 3 Lista de pendências Avaliação Ingerir 2 Armazenamento arquivístico Planejamento de preservação

Padrão ▼ Teste 01 Acesso não. 001 Squeaky toy

Tipo de transferência Nome da transferência ID do sistema de acesso

Iniciar transferência

Aprovar automaticamente

Transferir	UUID	Hora de início da transferência	
Teste_2	4d17ddf9-02c3-44de-8b62-a30330c3b79d	2018-12-26 01:28	
<ul style="list-style-type: none"> Microservice: criar SIP a partir de transferência 			
Job : Criar SIP (s) [?]		Aguardando decisão	⚙️ Ações ▼
: Opções de Carga de Trabalho para criar SIPs		Completado com sucesso	⚙️
: Diretório de transferência de verificação de trabalhos para objetos		Completado com sucesso	⚙️
<ul style="list-style-type: none"> Microservice: transferência completa Microservice: Examine o conteúdo Microservice: Validação Microservice: caracterize e extraia metadados Microservice: atualizar o documento METS.xml Microservice: Pacotes de extração Microservice: Identifique o formato do arquivo Microservice: limpar nomes Microservice: gerar relatório de estrutura de transferência Microservice: Varredura por vírus Microservice: Quarentena Microservice: Gerar documento METS.xml Microservice: Verificar as somas de verificação de transferência Microservice: formatar arquivos de metadados Microservice: Atribui arquivos UUIDs e checksums 			

Fonte: <https://sandbox.archivematica.org/transfer/>.

Após preencher os campos da Transferência, clica-se em *squeaky toy*¹² acima de “Iniciar transferência (em verde)” como mostra a figura, para localizar os diretórios, pastas e

¹¹ As figuras, nesta seção, traduzidas para a língua portuguesa em seu layout se originam de *printscreens* da versão demo do *Archivematica*.

arquivos presentes e disponíveis (se for na versão demo¹³). É preciso estruturar o diretório¹⁴ caso não seja a versão completa do software, podendo este possuir qualquer nomenclatura. Os objetos digitais serão preservados através desta pasta. Adicionar subdiretórios é permitido, mas não obrigatório (COSTA *et al*, 2016).

Ao selecionar uma pasta e pressionar em “Adicionar”, o diretório é localizado e realizada a Transferência. Para processá-la, um ícone de engrenagem aparece, o que significa a decisão por parte do usuário sobre aprovação da tarefa que quer realizar. Com isso, pode-se ver os microsserviços (ações desempenhadas pelo software). Para efetuar a Transferência é necessário aprová-la ou rejeitá-la, caso deseje interromper o processamento (Figura 15).

¹² Também conhecido como *Browser* na interface do *Archivematica* em versão inglês. Possui a função de encontrar ou procurar diretórios.

¹³ O único serviço possível dentro do *Archivematica* na versão demo é a Transferência. Tem-se, então ações limitadas, como é o caso da inserção de arquivo no software. Só é possível executar com os diretórios e arquivos já disponíveis.

¹⁴ Funciona como uma estrutura para organizar arquivos que contém outros arquivos, servindo, por exemplo, para organizar mídias de computador: disquetes, CDs, DVDs, e outras mídias.

Figura 15 - Processo de criação da Transferência

archivematica

Transferência 3 Lista de pendências Avaliação Ingerir 2 Armazenamento arquivístico Planejamento de preservação Acesso

Padrão ▼ Teste 01 Acesso não. 001 Squeaky toy

Tipo de transferência Nome da transferência ID do sistema de acesso

Iniciar transferência

Aprovar automaticamente

artefactual (3757 objetos)

- archivematica-sampleddata (3744 objetos)
 - createtransfers (7 objetos)**
 - Corpus de formato OPF (1561 objetos)
 - Transferências de amostra (150 objetos)
 - TestTransfers (1898 objetos)
 - recursos de treinamento (68 objetos)
 - issue_template.md (143 bytes)
 - Makefile (611 bytes)
 - README.md (814 bytes)
- ubuntu (4 objetos)

Adicionar

Transferir	UUID	Hora de início da transferência	
Teste_2	4d17ddf9-02c3-44de-8b62-a30330c3b79d	2018-12-26 01:28	
Microservice: criar SIP a partir de transferência			
Job : Criar SIP (s) [?]		Aguardando decisão	⚙️ Ações ▼
: Opções de Carga de Trabalho para criar SIPs		Completado com sucesso	⚙️ Ações
: Diretório de transferência de verificação de trabalhos para objetos		Completado com sucesso	⚙️ Ações

- Crie um único SIP e continue processando
- Enviar para o backlog
- Rejeitar transferência

Fonte: <https://sandbox.archivematica.org/transfer/>.

Na anuência da Transferência suas funções ocorrem através dos seguintes microsserviços: verificar se a transferência foi devidamente estruturada, com metadados e pastas de objetos; atribuir (UUID) (Identificador Único Universal); criar um arquivo METS que captura a ordem original da transferência; colocar em quarentena o pacote de transferência por um período fixo para permitir que as definições de vírus sejam atualizadas antes de uma verificação; averiguar vírus e *malware*; gerar uma árvore de diretórios da transferência original e lugares como um arquivo de texto no AIP; remover caracteres proibidos de pasta e nomes de arquivos; identificar o formato de arquivo baseado na normalização; caracterizar e extrair metadados. (COSTA *et al*, 2016).

Caso as ações tenham sido exitosas, já se pode “embalar” os objetos digitais em forma de SIP para Ingestão, ou enviá-los, indexá-los e armazená-los para recuperação posterior. Assim, é possível criar um SIP na opção "Criar único SIP e continuar o processamento" ou

"Enviar transferência para o *backlog*" (Figura 15). Nesse caso, a Transferência será armazenada em um *backlog* no mesmo local que seu AIP, de modo que se possa recuperá-la na guia Ingerir para processamento em data posterior (COSTA *et al*, 2016).

Os referido autores concordam que a importação de metadados precisa está estruturada para englobar um diretório que contenha qualquer tipo de metadados. Com a inclusão de um arquivo *metadata.csv*¹⁵ estruturado, o *Archivematica* torna-se capaz de analisar os metadados para o *dmdSec*¹⁶ (seção de metadados descritivos) do arquivo METS, que ficarão disponíveis para pesquisa na guia Armazenamento arquivístico, conforme descrevem Costa *et al* (2016):

Todos os elementos *Dublin Core* são usados para gerar um *dmdSec* para cada diretório ou arquivo com *MDTYPE* = "DC", ou seja, um tipo de descrição nos arquivos. [...]. Os objetos simples fornece arquivo CSV e METS e apresenta arquivos individuais que não são páginas em um objeto composto, como um livro ou um problema de jornal. Os objetos compostos fornece arquivo CSV e METS de várias páginas de objetos digitais, como jornais e livros. [...]. O Envio de transferências através dos microserviços na guia Transferência permite que o repositório tenha o mínimo de controle sobre seus objetos digitais. O arquivo de conteúdo e METS de transferência são indexados para facilitar a recuperação. Esse *workflow* pressupõe que o usuário enviou transferências ao *backlog* a partir do final do *workflow* de transferência no painel de instrumentos. (COSTA *et al*, 2016, p. 31-33, grifo nosso).

Assim, o pacote *BagIt* permite a preservação da Ingestão de informação no formato da Biblioteca do Congresso Americano (Library of Congress).

Os pacotes de preservação “normalmente” são usados para armazenar sistemas de arquivos. O *BagIt File Packaging Format* é uma especificação de estrutura de um pacote desenvolvida pela Biblioteca do Congresso Americano que envolve especificações combinadas com o METS e PREMIS, usada em ambientes de preservação digital como *Archivematica* (AGUIAR, 2018).

Na sintetização desses termos, o usuário consegue um SIP (criação) de uma ou mais transferências padrão e chegar ao seu objetivo final, o qual pode ser movido para Ingestão ou enviado para um *backlog* (em posterior processamento). Das relações que o próprio *Archivematica* reconhece pode-se gerar um DIP (a interoperabilidade), sobretudo, por meio da

¹⁵ Valores Separados por Vírgula (CSV) é um arquivo de texto que usa uma vírgula para separar valores. Armazena dados tabulares (números e texto) em texto simples. O uso da vírgula como um separador de campo é a origem do nome desse formato de arquivo.

¹⁶ *dmdSec*: descrição seção de metadados. Esta seção registra todos os metadados descritivos de todos os itens no objeto *METS*. Vários elementos *dmdSec* são permitidos para que metadados descritivos possam ser registrados para cada item separado dentro do objeto *METS*. Disponível em: http://metsnavigator.sourceforge.net/project/Doc/javadoc/edu/indiana/dlib/metsnav/mets/v1_5/DmdSec.html.

integração com vários sistemas de acesso e armazenamento: AtoM, CONTENTdm, Archivists' Toolkit, Dspace, LOCKSS, Duracloud e Arkivum.

b) Guia Lista de pendências (*Backlog*)

A lista de pendências exibe conteúdos que foram adicionados ao *Backlog* do *Archivematica*, permitindo que os usuários visualizem, pesquisem, solicitem, baixem e excluam as Transferências (ARTEFACTAL, [201-c]).

De maneira abrangente, *Backlog* se refere a um *log* (resumo histórico) de acumulação de trabalho em um determinado período. Para melhor compreensão, tome-se como exemplo, a representação de uma equipe para finalizar um serviço acumulado em uma “pilha de pedidos” em espera (Figura16).

Figura 16 - Exemplo de lista de transferência no *Backlog*

The screenshot shows the Archivematica interface with the 'Backlog' tab selected. At the top, there is a navigation bar with the Archivematica logo and various menu items: Transfer, Backlog, Appraisal, Ingest, Archival storage, Preservation planning, Access, Administration, and sromkey. Below the navigation bar, there is a search area with a text input field, a dropdown menu set to 'Any', a 'Keyword' dropdown, a 'Search transfer backlog' button, and a 'Show files?' checkbox. Below the search area, there is a table with the following data:

Name	Transfer UUID	File count	Ingest date	Actions
coll_1234_test_3-b765fef9-3703-4030-ac5d-00d471b6b619	b765fef9-3703-4030-ac5d-00d471b6b619	17	2016-11-28	Download
coll_9000_test-01bfi47c-8ae0-4824-a9b1-43370ba2771a	01bfi47c-8ae0-4824-a9b1-43370ba2771a	10	2016-11-28	Download
coll_5678_test-0a7c5509-5440-484d-8cdd-355c4bfbf74d	0a7c5509-5440-484d-8cdd-355c4bfbf74d	16	2016-11-28	Download
coll_9000_test_2-46a266a4-9e08-42ae-8d10-1dfbeae91368	46a266a4-9e08-42ae-8d10-1dfbeae91368	4	2016-11-28	Download
coll_1234_test-faba33fa-fcfd-48ee-8369-beca56eee448	faba33fa-fcfd-48ee-8369-beca56eee448	13	2016-11-28	Download
coll_1234_test_2-527b44d6-e029-4d45-b57b-d11584de9082	527b44d6-e029-4d45-b57b-d11584de9082	15	2016-11-28	Download

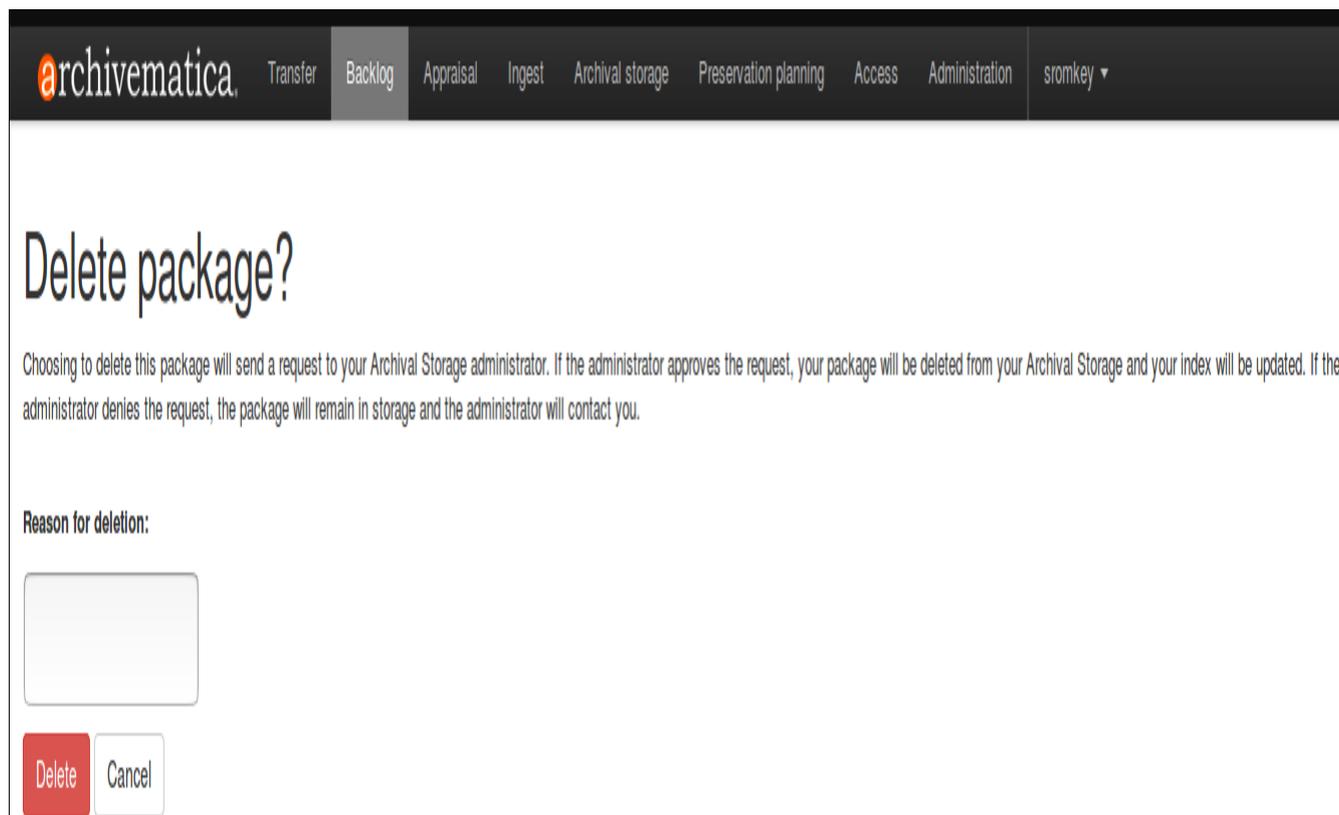
Below the table, there is a pagination bar showing 'Showing 1 to 6 of 6 entries' and buttons for 'First', 'Previous', '1', 'Next', and 'Last'.

Fonte: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/ images/backlog tab.png>.

As Transferências armazenadas usando a funcionalidade de *Backlog* só podem ser acessadas usando o mesmo *pipeline* no qual foi criada. Os arquivos podem ser baixados no “Download” (lado direito da Figura 16). Quando baixados serão empacotadas como um arquivo *tar*.

É possível excluir uma Transferência ao pressionar o ícone na cor vermelha “Delete”, sendo solicitada uma justificativa por parte do usuário. Para concluir a exclusão, o administrador deve ter efetuado o login. Até que esta atividade seja aprovada, a Transferência permanecerá visível na guia do *Backlog* e disponível para download (Figura 17).

Figura 17 - Exclusão de transferências no *Backlog*



Fonte: https://www.archivematica.org/pt-br/docs/archivematica1.8/images/backlog_delete.png.

c) Guia de Avaliação (*Appraisal*)

Esta guia analisa e organiza o material que foi colocado no *Backlog*, direcionado para recursos no ArchivesSpace, assim:

Fornece um espaço entre a Transferência e Ingestão para que arquivistas decidam quais arquivos devem se tornar SIPs e como estes podem ser organizados. Inclui recursos para examinar as características e o conteúdo dos arquivos, bem como para conectá-los a ferramentas de pesquisa criadas no ArchivesSpace. (ARTEFACTUAL, [201-d], p. 1, tradução nossa).

As transferências enviadas para o *Backlog* podem ser encontradas no seu próprio painel da guia Avaliação conforme a Figura 18, mostrando o nome do arquivo, o SIP, UUID, e logo abaixo no canto esquerdo da imagem um pasta (example1-2fbbb597...).

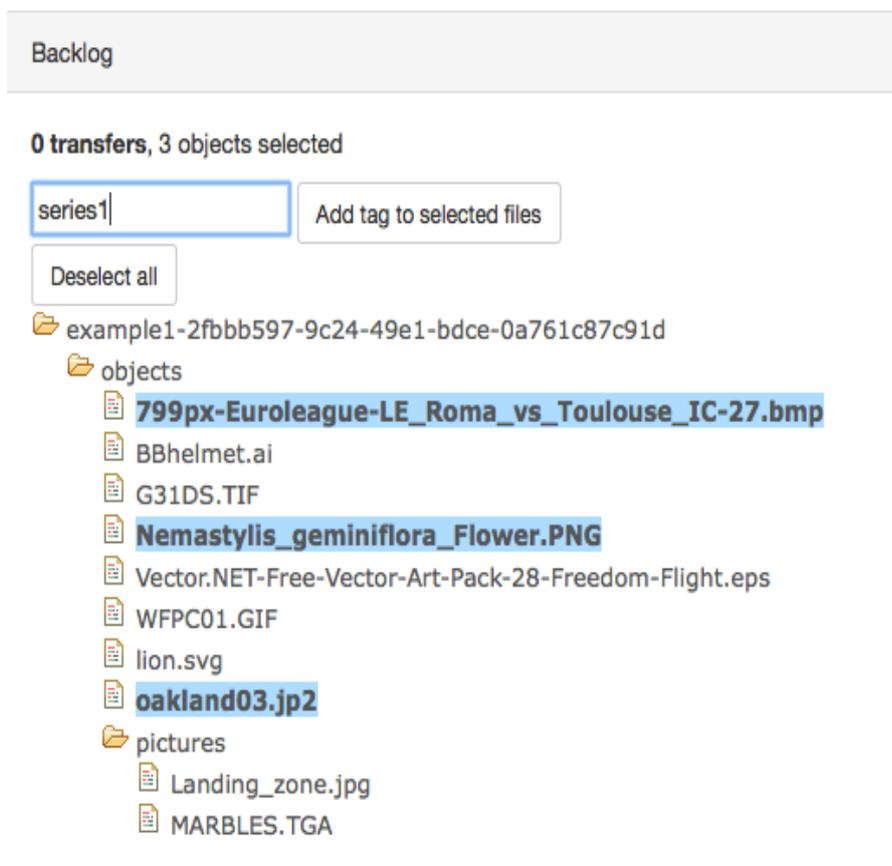
Figura 18 - Transferências do *Backlog* na guia de Avaliação

The screenshot displays the Archivematica web interface. At the top, a dark navigation bar contains the Archivematica logo and several menu items: Transfer, Appraisal (highlighted with a red notification badge), Ingest, Backlog, Archival storage, Preservation planning, Access, Administration, and artefactual. Below the navigation bar, there is a search area with a text input field containing 'e1-bdce-0a761c87c91d', a dropdown menu for 'SIP UUID', another dropdown for 'Phrase', and a blue 'Search transfer backlog' button. Below the search area, there is a link for 'Add New' and a 'Tags' dropdown menu set to 'All'. A secondary navigation bar contains buttons for 'Backlog' (active), 'Analysis', 'File list', 'ArchivesSpace', and 'Arrangement'. Below this, the 'Backlog' section shows '0 transfers, 0 objects selected'. It includes a text input field, an 'Add tag to selected files' button, a 'Deselect all' button, and a list item with a folder icon and the ID 'example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d'. To the right, a panel contains links for 'Objects', 'Tags', 'Examine contents', and 'Preview file'.

Fonte: https://www.archivematica.org/pt-br/docs/archivematica 1.8/ images/search_backlog.png.

As *Tags* podem ser adicionadas aos arquivos selecionados no *Backlog*, inserindo um nome clicando no botão "Adicionar *tag* aos arquivos selecionados" (*Add tag to selected files*) (Figura 19).

Figura 19 - Adição de *Tags*



Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/tags_backlog.png.

As *Tags* que já foram criadas aparecem no menu. A escolha de uma determinada *Tag* filtra os arquivos exibidos nas transferências no painel *Backlog*. Os filtros adicionados em outros painéis da Avaliação também aparecerão listados nesse mesmo menu.

Chama-se atenção também para o painel de análise, o qual permite ao usuário explorar arquivos em Transferências enviadas ao *Backlog*. Assim, existem quatro guias: *Objetos*, *Tags*, *Examinar conteúdo* e *Visualizar arquivo*. Em *Objetos*, as informações sobre os arquivos podem ser visualizadas na forma de um relatório e, também, através de visualizações com opções para representar o número e tamanho total de arquivos em cada formato (Figura 20).

Figura 20 - Painel de análise de Objetos (Objects)

The screenshot displays the Archivematica interface for object analysis. At the top, there are navigation tabs: Backlog, Analysis, File list, ArchivesSpace, and Arrangement. The main area is split into two panels.

Backlog Panel: Shows '0 transfers, 5 objects selected'. It includes a search box, a 'Deselect all' button, and a file tree structure:

- example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d
 - objects
 - 799px-Euroleague-LE_Roma_vs_Toulouse_IC-27.bmp series1
 - BBhelmet.ai
 - G31DS.TIF
 - Nemastylis_geminiflora_Flower.PNG series1
 - Vector.NET-Free-Vector-Art-Pack-28-Freedom-Flight.eps
 - WFPC01.GIF
 - lion.svg
 - oakland03.jp2 series1
 - pictures
 - Landing_zone.jpg
 - MARBLES.TGA

Report Panel: Shows a table with the following data:

Format	PUID	Group	Number of files	Size
1989a	fmt/4	Image (Raster)	1 object	0.108 MB
Acrobat PDF 1.5	fmt/19	Portable Document Format	1 object	1.030 MB
JPEG 1.01	fmt/43	Image (Raster)	1 object	1.298 MB
Scalable Vector Graphics 1.0	fmt/91	Image (Vector)	1 object	0.017 MB
TIFF	fmt/353	Image (Raster)	1 object	0.120 MB

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/analysis_report.png.

A guia *Tags* lista todas as *tags* que foram adicionadas e quantas delas estão presentes nos arquivos selecionados.

Figura 21 - Guia Tags

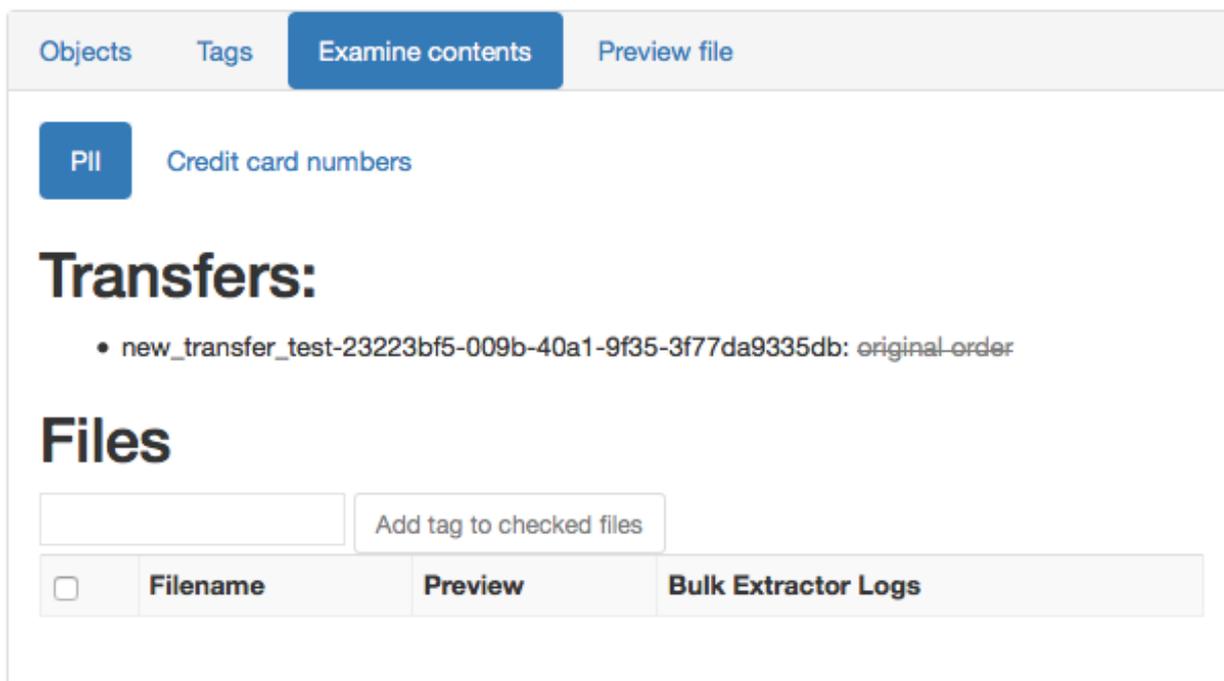
The screenshot shows the 'Tags' panel in the Archivematica interface. It features a table with the following data:

Tag	Count
dataset1	1
images	3
researchData	2
test	5

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/analysis_tags.png.

A guia Examinar conteúdo fornece os relatórios criados durante este microserviço. Inclui opções para listar arquivos que potencialmente contêm PII (Informações de Identificação Pessoal) ou número de cartão de crédito. Também integra uma opção para adicionar *Tags* a esses arquivos.

Figura 22 - Guia Examinar conteúdo (*Examine contents*)



Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/analysis_examine_contents.png.

Os arquivos selecionados no painel Lista de arquivos (Figura 23) podem ser exibidos na guia Visualizar arquivo. Fornece informações sobre os nomes de arquivos, tamanhos, e datas modificadas(os) pela última ou quaisquer marcas que foram adicionadas.

Figura 23 - Guia Lista de arquivo (*File list*)

The screenshot shows a 'File List' interface with the following elements:

- Header: File List
- Status: 7 objects in list, 1 object selected
- Date range start: 2009
- Date range end: 2015
- Dates: 2009 - 2015 ✕
- Series 3
- Add tag to selected files
- Table with columns: Filename, Size, Last modified, Tags

<input type="checkbox"/>	Filename	Size	Last modified	Tags
<input type="checkbox"/>	example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d/objects/799px-Euroleague-LE_Roma_vs_Toulouse_IC-27.bmp	1.371 MB		series1 ☰
<input type="checkbox"/>	example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d/objects/G31DS.TIF	0.120 MB		series2 ☰
<input type="checkbox"/>	example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d/objects/pictures/Landing_zone.jpg	1.298 MB		series2 ☰
<input checked="" type="checkbox"/>	example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d/objects/pictures/MARBLES.TGA	4.064 MB		
<input type="checkbox"/>	example1-2fbbb597-9c24-49e1-bdce-0a761c87c91d/objects/Nemastylis_geminiflora_Flower.PNG	1.956 MB		series1 ☰

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/analysis_file_list.png.

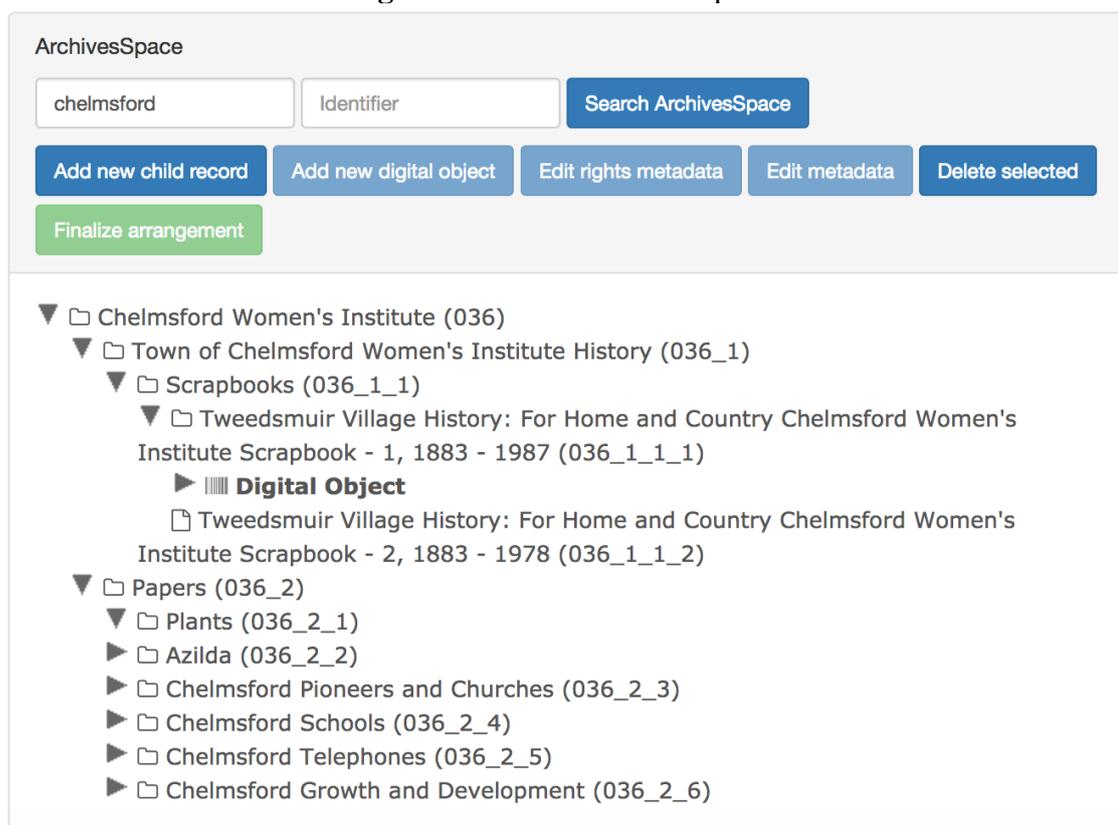
Todos os filtros que foram aplicados no painel Análise ou no menu *Tags* afetarão quais arquivos serão exibidos na Lista de arquivos.

Dessa forma, coexiste outro recurso importante a destacar na Avaliação: o painel ArchivesSpace¹⁷. As configurações para o *upload* do ArchivesSpace (Figura 24) devem ser dispostas na guia de Administração do *Dashboard* para associar este repositório. Seus recursos também são chamados de fundos, coleções, grupos de registros etc. Os níveis mais baixos de descrição são chamados de “objetos de arquivamento” (séries, arquivos, itens etc.).

Os metadados são gerenciados separadamente como “objetos digitais”, que podem ser relacionados a recursos/objetos de arquivamento em qualquer nível de descrição, expandido para mostrar níveis de hierarquia.

¹⁷ Recupera e adiciona recursos criados em um repositório ArchivesSpace. No painel ArchivesSpace, os arquivos transferidos para o *Archivematica* podem ser associados aos recursos criados no ArchivesSpace e os SIPs podem ser enviados para o *Ingest* (ARTEFACTUAL, 201-c).

Figura 24 - Painel Archivespace¹⁸



Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/archivesspace_search.png.

As opções na parte superior do painel ArchivesSpace permitem adicionar e alterar um recurso existente. Observa-se, com essas nuances, as funções que tal painel se descreça referente à descrição arquivística, como os níveis de hierárquia, assim como a ideia de esquema de arranjo e metadados.

d) Guia Ingerir (*Ingest*)

Os objetos digitais na Guia Ingerir (Figura 25), são empacotados em SIPs e executados por vários microsserviços, incluindo a normalização, empacotamento em um AIP e a geração de um DIP.

¹⁸ Mais informações sobre o ArchivesSpace podem ser encontradas em: <http://archivesspace.org/>.

Figura 25 - Guia Ingerir (Ingest)

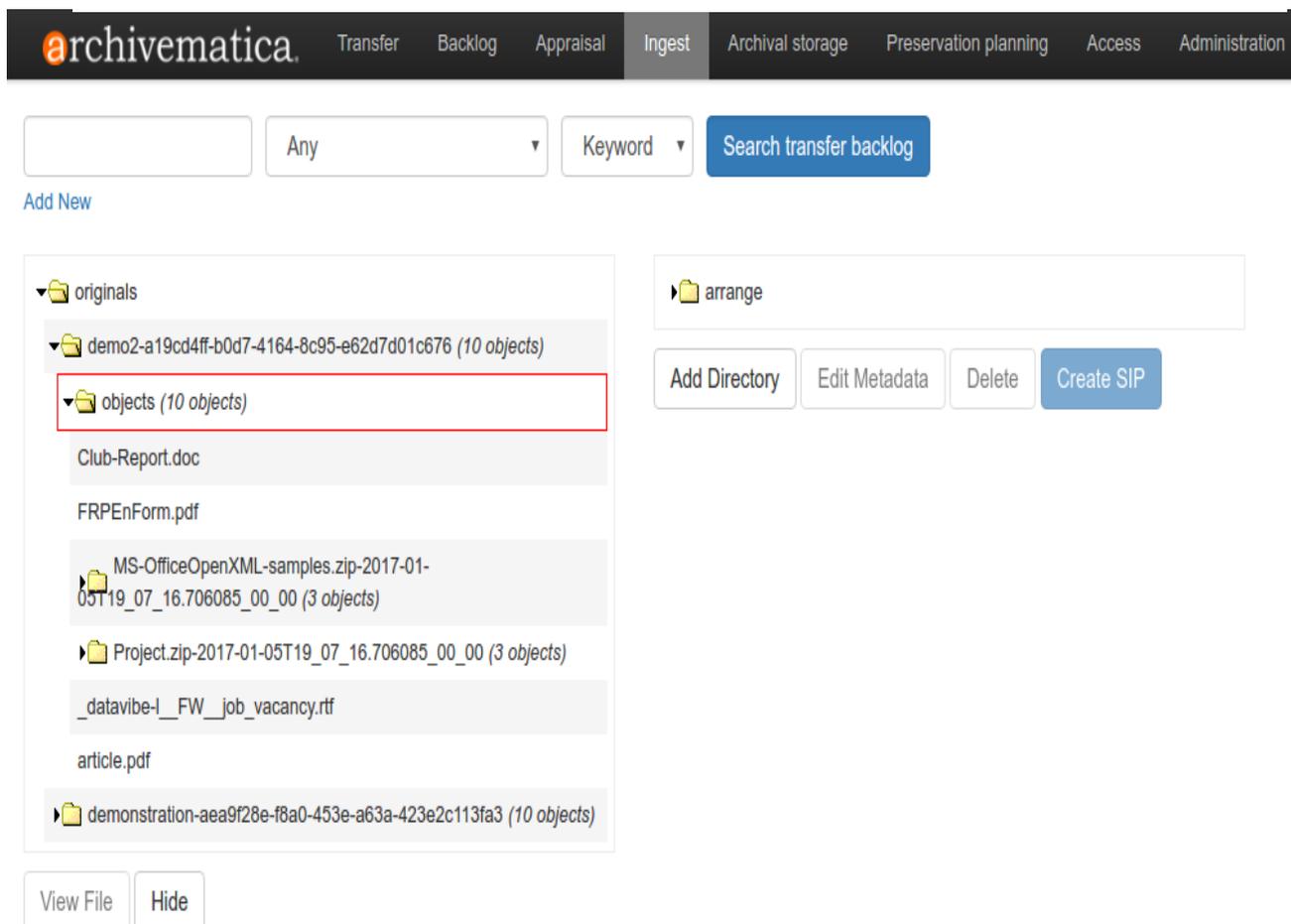
The screenshot shows the Archivematica Ingest interface. At the top, there is a navigation bar with the Archivematica logo and menu items: Transfer, Backlog, Appraisal, Ingest (highlighted with a red notification icon), Archival storage, Preservation planning, Access, Administration, and a user profile dropdown (demo@example.com). Below the navigation bar is a search bar with a text input field, a dropdown menu set to 'Any', another dropdown menu set to 'Keyword', and a blue 'Search transfer backlog' button. Below the search bar is an 'Add New' link. The main content area shows two directory selection boxes: 'originals' and 'arrange'. Below these are buttons for 'View File', 'Hide', 'Add Directory', 'Edit Metadata', 'Delete', and 'Create SIP'. A table titled 'Submission Information Package' is displayed with columns for 'Submission Information Package', 'UUID', and 'Ingest start time'. The table contains a single row for 'sample-series' with UUID '2ec75602-bd2c-4839-8510-20e5434054ed' and start time '2017-01-05 12:05'. Below the table, a 'Micro-service: Normalize' section is expanded, showing a list of jobs with their status and an 'Actions' dropdown menu. The 'Actions' menu is open, showing options: 'Normalize for preservation and access', 'Normalize for preservation', 'Reject SIP', 'Normalize service files for access', 'Do not normalize', 'Normalize manually', and 'Normalize for access'.

Submission Information Package	UUID	Ingest start time
sample-series	2ec75602-bd2c-4839-8510-20e5434054ed	2017-01-05 12:05
▶ Micro-service: Normalize		
Job: Normalize [?]	Awaiting decision	⚙️ Actions
Job: Resume after normalization file identification tool selected.	Completed successfully	⚙️ Actions
Job: Identify file format	Completed successfully	⚙️ - Normalize for preservation and access
Job: Select pre-normalize file format identification command	Completed successfully	⚙️ - Normalize for preservation
Job: Move to select file ID tool	Completed successfully	⚙️ - Reject SIP
Job: Grant normalization options for no pre-existing DIP	Completed successfully	⚙️ - Normalize service files for access
Job: Set remove preservation and access normalized files to renormalize link.	Completed successfully	⚙️ - Do not normalize
Job: Check for Access directory	Completed successfully	⚙️ - Normalize manually
Job: Check for Service directory	Completed successfully	⚙️ - Normalize for access
Job: Identify manually normalized files	Completed successfully	⚙️
▶ Micro-service: Clean up names		
▶ Micro-service: Remove cache files		
▶ Micro-service: Include default SIP processingMCP.xml		
▶ Micro-service: Rename SIP directory with SIP UUID		

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/Normalize1.png.

É possível organizar um SIP a partir do *Backlog*. Para isso, deve-se recuperar o conteúdo da Transferência, usando barras de busca na Lista de pendências, no topo da guia Ingerir (Figura 26), para encontrar a(s) transferência(s) e/ou objeto(s) que se deseja fazer a admissão.

Figura 26 - Organizando um SIP na guia Ingerir



Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/Ingest-panels1.png.

Pode-se clicar no subdiretório desejado, pertencente ao diretório “Arranjo” (Arrange) para selecionar e, em seguida, “Criar um SIP” (Create SIP). O *Archivematica* irá confirmar que se deseja criar um SIP, e depois de receber a confirmação, os próximos passos elencarão os microsserviços para criar AIPs e DIPs como selecionados pelo usuário.

Os metadados podem ser adicionados antes ou depois da etapa de normalização. O *Archivematica* irá disponibilizar um lembrete, se desejado, durante o processamento dos microsserviços de metadados (COSTA *et al*, 2016).

A normalização (Figura 27), é o processo de converter objetos digitais admitidos para formatos de preservação ou de acesso. Os objetos originais são sempre mantidos juntamente com suas versões normalizadas. Nesta etapa, o SIP aparecerá no painel representado pelo ícone de um sino ao lado dele.

Figura 27 - Etapa de normalização

The screenshot shows the Archivematica web interface. At the top, there are browser tabs for 'Archivematica FPR Server', 'Archivematica Dashboard', and 'Archivematica Dashboard'. The address bar shows 'localhost/ingest/'. Below the browser tabs is the Archivematica navigation bar with links for 'Transfer', 'Ingest', 'Archival storage', 'Preservation Planning', 'Access', and 'Administration'. The 'Ingest' link is highlighted with a red '2'. Below the navigation bar is a search bar with a dropdown menu set to 'Any' and a 'Keyword' dropdown. A search button 'Search transfer backlog' and a 'Show files?' checkbox are also present. Below the search bar is an 'Add New' link. The main content area shows a table of submission information. The table has columns for 'Submission Information Package', 'UUID', and 'Ingest start time'. The first row is for a submission named 'Test_files' with UUID 'c374bcb4-a40e-47cd-9dd2-8095b3a5d5e9' and ingest start time '2013-10-10 13:48'. Below this row is a section for 'Micro-service: Normalize'. This section contains a list of jobs. The first job is 'Job: Normalize [?]' with a status of 'Awaiting decision'. An 'Actions' dropdown menu is open for this job, showing several options: 'Normalize for preservation and access', 'Normalize for preservation', 'Reject SIP', 'Normalize service files for access', 'Do not normalize', 'Normalize manually', and 'Normalize for access'. The 'Normalize for preservation and access' option is highlighted in red. Below the 'Job: Normalize [?]' row are several other jobs, all with a status of 'Completed successfully'. These jobs include 'Job: Resume after normalization file identification tool selected.', 'Job: Identify file format', 'Job: Select pre-normalize file format identification command', 'Job: Move to select file ID tool', 'Job: Set resume link after tool selected.', 'Job: Find options to normalize as', 'Job: Move to workFlowDecisions-createDip directory', 'Job: Grant normalization options for no pre-existing DIP', 'Job: Set remove preservation and access normalized files to renormalize link.', 'Job: Check for Access directory', 'Job: Check for Service directory', and 'Job: Identify manually normalized files'. Below these jobs are three more micro-services: 'Clean up names', 'Remove cache files', and 'Include default SIP processingMCP.xml'. At the bottom of the table, there is a partial row for 'localhost -service: Rename SIP directory with SIP UUID'.

Submission Information Package	UUID	Ingest start time
Test_files	c374bcb4-a40e-47cd-9dd2-8095b3a5d5e9	2013-10-10 13:48
Micro-service: Normalize		
Job: Normalize [?]		Awaiting decision
Job: Resume after normalization file identification tool selected.		Completed successfully
Job: Identify file format		Completed successfully
Job: Select pre-normalize file format identification command		Completed successfully
Job: Move to select file ID tool		Completed successfully
Job: Set resume link after tool selected.		Completed successfully
Job: Find options to normalize as		Completed successfully
Job: Move to workFlowDecisions-createDip directory		Completed successfully
Job: Grant normalization options for no pre-existing DIP		Completed successfully
Job: Set remove preservation and access normalized files to renormalize link.		Completed successfully
Job: Check for Access directory		Completed successfully
Job: Check for Service directory		Completed successfully
Job: Identify manually normalized files		Completed successfully
Micro-service: Clean up names		
Micro-service: Remove cache files		
Micro-service: Include default SIP processingMCP.xml		
localhost -service: Rename SIP directory with SIP UUID		

Fonte: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/NormPresAccess1.png>.

Quando o SIP atingir “Normalizar”¹⁹, várias opções aparecerão no *Archivematica*. Deve-se, então, atribuir a opção mais adequada ao fluxo do trabalho. Suas funcionalidades criam cópias para preservação de objetos digitais que serão utilizadas para gerar um DIP (COSTA *et al*, 2016).

Depois desta etapa ser aprovada, o SIP é executado através de microsserviços, incluindo o processamento da documentação: submissão, geração do arquivo METS, indexação, geração do DIP e acondicionamento do AIP (Figura 28).

¹⁹ Todas as características e funcionalidades estão descritas em seu manual: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/user-manual/ingest/manual-normalization/#manual-norm>.

Figura 28 - Execução dos microsserviços na guia Ingerir

The screenshot shows the Archivematica dashboard with the 'Ingest' tab selected. The main content area displays a list of jobs and their associated micro-services. A job titled 'Job: Store AIP (review) [?]' is highlighted in green, indicating it is 'Completed successfully'. An 'Actions' menu is open over this job, showing options: 'Reject AIP' and 'Store AIP'. The 'Store AIP' option is highlighted in orange. The dashboard includes a navigation bar with tabs for Transfer, Ingest, Archival storage, Preservation Planning, Access, and Administration. The URL in the browser is localhost/ingest/.

Job	Status	Actions
Job: Identify manually normalized files	Completed successfully	Actions
Micro-service: Clean up names		
Micro-service: Remove cache files		
Micro-service: Include default SIP processingMCP.xml		
Micro-service: Rename SIP directory with SIP UUID		
Micro-service: Verify transfer compliance		
Micro-service: Verify SIP compliance		
Micro-service: Approve SIP creation		
Sample_series	2c5fedbf-b302-4939-8f8c-10f3ae5f79dd	2013-10-10 13:13
Micro-service: Store AIP		
Job: Store AIP (review) [?]	Awaiting decision	Actions
Job: Move to the store AIP approval directory	Completed successfully	Actions
Micro-service: Prepare AIP		
Micro-service: Upload DIP		
Job: Upload DIP	Awaiting decision	Actions
Micro-service: Prepare DIP		
Micro-service: Process metadata directory		
Micro-service: Process submission documentation		
Micro-service: Normalize		
Micro-service: Clean up names		
Micro-service: Remove cache files		
Micro-service: Include default SIP processingMCP.xml		
Micro-service: Rename SIP directory with SIP UUID		
Micro-service: Verify transfer compliance		
Micro-service: Verify SIP compliance		

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/StoreAIPUpDIP1.png.

No menu “Ação” (Actions), selecione o “Fornecimento AIP” (Store AIP) para movê-lo em direção ao Serviço de Armazenamento. É possível armazenar um AIP em qualquer número de diretórios. Os *uploads* dos DIPs no Archivematica estão suportados e configurados preferencialmente para o AtoM.

e) Guia de Armazenamento Arquivístico (*Archival Storage*)

Neste painel, o pacote AIP é movido para o seu repositório de armazenamento. Com a admissão concluída e após a seleção de um local para armazenamento, o *Archivematica* comprime e fecha o AIP, movendo para essa guia.

O nome AIP é composto dos seguintes itens: o nome da Transferência original, se nenhum novo nome foi atribuído ao SIP criado a partir da transferência; e um UUID atribuído durante a formação SIP (Figura 29).

Figura 29 - Guia Armazenamento Arquivístico

Archival storage / Search

Any Keyword Show files? Show AICs? Search archival storage

Add New

Browse archival storage

Total size: 28.50 MB | Files indexed: 10

AIP	Size	UUID	Date stored	Status	Actions
images	28.50 MB	754d849a-e5f2-49e0-a87e-3ed6b0695f5a	2016-12-20 14:03	Stored	View

Found 1 entries. Showing 1 to 1.

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/ArchStorTab1.png.

Dessa forma, esses pacotes são divididos em subdiretórios (UUID) para armazenamento e recuperação. A consistência desse painel está em uma tabela com informações sobre os AIPs armazenados (nome, tamanho, UUID, data). Para esses arquivos, pode-se fazer cópias se for preciso, através da atribuição de seu nome.

Os conjuntos conjuntos de dados e coleções de objetos digitais (grandes e heterogêneos), podem ser divididos em vários AIPs. Estes são agredados em um AIC, ou coletados como informações de um arquivo, definido pelo modelo de referência SAAI como pacote de informação de arquivo, cuja informação de conteúdo é uma agregação de outros pacotes de informação de arquivo (COSTA *et al*, 2016). A Figura 30 mostra a exclusão de uma AIP:

Figura 30 - Exclusão de um AIP

The screenshot shows the Archivematica web interface. At the top, there is a navigation bar with the Archivematica logo and several menu items: Transfer, Backlog, Appraisal, Ingest, Archival storage (selected), Preservation planning, Access, Administration, and a user profile dropdown (demo@example.com). Below the navigation bar, the breadcrumb path is 'Archival storage / test'. The main content area displays the details for an Archival Information Package (AIP) named 'test'. The details include: UUID (7a5cb1be-1d86-498d-9e79-631d851e11c4), Size (20.26 MB), Date stored (2017-01-04 16:54), Status (Stored), Location (with a Download button and a file path [...]/test-7a5cb1be-1d86-498d-9e79-631d851e11c4.7z), and Pointer file (with a View button). Below the details, there is an 'Actions' section with buttons for Upload DIP, Re-ingest, and Delete. The 'Delete' button is active, and a confirmation dialog is displayed. The dialog asks 'Are you sure that you want to delete this AIP' and provides instructions: 'Choosing to delete this package will send a request to your Archival Storage administrator. If the administrator approves the request, your package will be deleted from your Archival Storage and your index will be updated. If the administrator denies the request, the package will remain in storage and the administrator will contact you.' The dialog also includes a field for the UUID (7a5cb1be-1d86-498d-9e79-631d851e11c4) and a text area for the reason for deletion, which contains the text 'File used for testing purposes'. At the bottom of the dialog, there is a red 'Delete' button.

Fonte: <https://www.archivematica.org/pt-br/docs/archivematica-1.8/images/ReasonDelete.png>.

Para a exclusão de AIP (Figura 30), tal solicitação demanda um motivo, enviando o pedido para o administrador de arquivo (COSTA *et al*, 2016). Se o administrador aprovar o pedido, o seu AIP será excluído do seu armazenamento e seu índice será atualizado. Se negar, o AIP permanecerá no armazenamento e o administrador deverá contatá-lo.

f) Guia de Acesso (*Acess*)

No processo de Ingestão, o acesso às cópias de objetos digitais é gerado e empacotado em um DIP. A guia Acesso (Figura 31), mostrará em uma tabela todos os DIPs carregados no AtoM ou no Binder, incluindo o URL, o AIP associado, a data e o status do *upload* e a opção de excluir da guia Acesso.

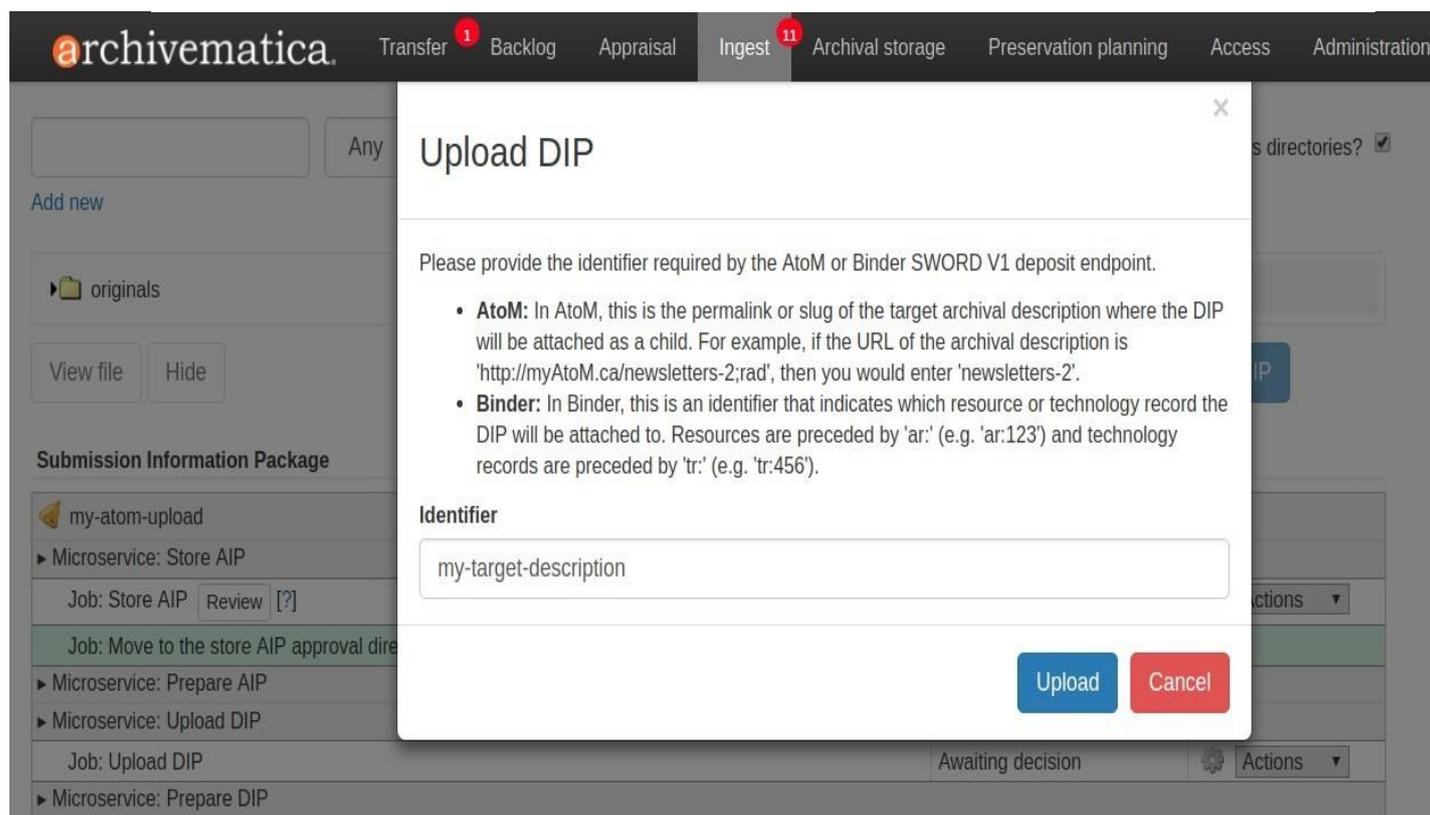
Figura 31 - Guia Acesso

AIP	URL DIP	data de upload	Status de upload
-----	---------	----------------	------------------

Fonte: <https://sandbox.archivematica.org/access/>.

O usuário carrega o DIP no sistema de acesso, ou o armazena para uso futuro. Quando o DIP é preparado, o *Archivematica* move o DIP para o diretório de *uploadDIP*, apresentando no painel as opções de fazer o *upload* DIP para o usuário. Depois que um DIP é armazenado, ele é movido para o diretório *uploadDIPs*, que pode ser esvaziado na guia Administração. O usuário deve criar a descrição em AtoM (ou outro sistema de acesso) antes de carregar o DIP. O usuário terá de indicar parte do URL da descrição ou uma coleção de destino, a fim de enviar o DIP para o local apropriado na descrição (COSTA *et al*, 2016, p. 81, grifo nosso).

O AtoM é o sistema de acesso padrão do *Archivematica*. Subsiste também fluxos de trabalho de integração para o ArchivesSpace (como já exposto na guia Avaliação), o Archivist's Toolkit, o CONTENTdm e o Binder.

Figura 32 - Página do *upload* DIP/AtoM

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/atom-dip-upload.png.

A partir do exposto, pode-se criar a descrição de destino no AtoM antes de fazer o *upload* do DIP. Os objetos digitais que forem enviados são adicionados como filhos da descrição de destino e, por padrão, recebem o nível de descrição do item, sendo possível organizar sua Transferência na guia Avaliação ou Ingerir, além de adicionar metadados descritivos (COSTA et al, 2016).

Antes de qualquer movimentação, nesse sentido, para carregar de fato os DIPs no AtoM, deve-se ter uma instância em funcionamento para interoperar com o *Archivematica*.

g) Guia Plano de Preservação (*Preservation Planning*)

A atribuição dada a este plano remonta a normalização para formatos de preservação e acesso mediante a admissão. As cópias de preservação são adicionadas ao AIP e as de acesso utilizadas para gerar um DIP em *upload* ao sistema de acesso. Os arquivos originais são mantidos para permitir ações no futuro, tais como a própria normalização ou emulação.

Os formatos, regras e comandos padrão são instalados e acessados por meio da guia *Preservation Planning*. [...]. As políticas de formato padrão do *Archivematica* são mantidas pela *Artefactual*. As políticas de formato serão alteradas conforme os padrões, práticas e ferramentas da comunidade evoluírem. Se você quiser sugerir uma política de formato padrão diferente, considere a possibilidade de abrir um problema no repositório de problemas do *Archivematica*. (ARTEFACTUAL, [201-b], p. 1, grifo nosso, tradução nossa).

Formatos como documentos do Microsoft Word não são necessariamente o melhor para preservação, mas ainda são um dos mais presentes (COSTA *et al*, 2016). Nestes casos, o padrão do *Archivematica* é mantê-los em seus formatos originais; avaliação de risco de formato e as informações recebidas de usuários ao longo do tempo permitirão que a *Artefactual Systems* expanda os padrões continuamente para incluir mais formatos através do FPR (Registro da Política de Formatos). Tais políticas precisam estarem alinhadas a TRAC e de acordo com ISO 16363:2012.

O FPR é responsável pelo planejamento e gerenciamento das políticas de formato dos arquivos, contém regras de normalização e extração, que diferem dependendo do arquivo. E portanto, indica ações, ferramentas e configurações para aplicar a um objeto digital, por exemplo, a conversão para o formato de preservação, a conversão para o formato de acesso e extração de formatos de pacotes (COSTA *et al*, 2017; AGUIAR, 2018).

Nesse viés, torna-se importante que as instituições estabeleçam práticas locais que incluam o monitoramento do ambiente de preservação digital para ajudar a informar regras de normalização ao longo do tempo, à medida que os padrões e ferramentas evoluírem.

No Quadro 15 encontram-se algumas definições sobre os elementos do Plano de Preservação de acordo com os preceitos *Archivematica*:

Quadro 15 - Definições dos elementos do Plano de Preservação *Archivematica*

<p>Formato: maneira padronizada pela qual a informação é codificada para armazenamento em um meio digital. O <i>Archivematica</i> é capaz de reconhecer diversos formatos de arquivos, extraindo informações do PRONOM, um registro técnico de formatos de arquivos mantido pelo National Archives of the United Kingdom (UK).</p>
<p>Finalidade: refere-se à função que determinada ferramenta, regra ou comando cumpre em um fluxo de trabalho do <i>Archivematica</i>, dependendo do contexto é implementado. Por exemplo, uma regra para normalizar um arquivo JPG para TIFF usando a conversão de ferramenta pode ter a finalidade Preservação, o que significa que o objetivo dessa regra é criar o TIFF para preservação a longo prazo.</p>
<p>Ferramenta: o <i>Archivematica</i> funciona como um <i>wrapper</i> (um código) para muitas ferramentas de <i>software</i> livre usadas para executar ações de preservação. Estes incluem ferramentas específicas de preservação digital como o FITS, usadas para extrair metadados técnicos de arquivos, bem como ferramentas para lidar com diferentes formatos de arquivos como o <i>Inkscape</i>, que é um programa de design usado para lidar com imagens vetoriais. Ferramentas, no título do registro da política Formato. As ferramentas são chamadas por comandos e atuam em formatos.</p>
<p>Comandos: formatar comandos de política são <i>scripts</i> ou instruções de linha de comando que controlam como uma ferramenta é executada. Um comando é criado para uma finalidade específica, como identificação de arquivo. Por exemplo, durante o processo de transferência, você pode informar ao <i>Archivematica</i> para identificar seus arquivos usando o Siegfried.²⁰ Isso inicia o comando "Identificar usando <i>Siegfried</i>", que executa a ferramenta de identificação de arquivos <i>Siegfried</i> incluídos na transferência. Este comando será executado uma vez em todos os arquivos da transferência.</p>

Fonte: Artefactual ([201-b], grifo nosso).

As regras de identificação (Figura 33) permitem definir a relação entre a saída criada por uma ferramenta de identificação e um dos formatos que existem no FPR. Isto deve ser feito para que o formato seja rastreado pelo *Archivematica* internamente, e para que possa ser utilizado pela normalização mais tarde (COSTA *et al*, 2016).

²⁰ Ferramenta identificadora de formato baseada em PRONOM (COSTA *et al*, 2016).

Figura 33 - Guia de Plano de Preservação

archivematica. Transfer Backlog Appraisal Ingest Archival storage Preservation planning Access Administration support ▾

Identification commands / List

Identification commands

Create new command

Show 10 entries Search:

Command	Type	Tool	Mode	Enabled	Actions
Identify by File Extension	Python script	File Extension version 0.1	ext	Yes	View Replace Disable
Identify using Fido	Python script	Fido version 1	PUID	No	View Replace Enable
Identify using Fido	Python script	Fido version 1.3.7	PUID	Yes	View Replace Disable
Identify using Siegfried	Python script	Siegfried version 1.7.6	PUID	No	View Replace Enable
Identify using Siegfried	Python script	Siegfried version 1.7.6	PUID	Yes	View Replace Disable

Showing 1 to 5 of 5 entries Previous 1 Next

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/identification-commands.png.

Nesse cenário, as regras de política de formato (Figura 34) associam comandos a formatos de arquivo. Como os comandos, essas regras são criadas para uma finalidade específica, conforme a normalização de arquivos. Analogamente, normalizar arquivos para preservação, poderá instruir o *Archivematica* a usar o comando Transcodificação para jpg com conversão para todos os arquivos JPG, enquanto os arquivos SVG serão convertidos para PDFs usando o comando transcodificação para pdf com o *inkscape*. Dessa forma, ferramentas, regras e comandos trabalham em conjunto para implementar políticas de preservação

Figura 34 - Regras de políticas de formato

Format policy rules / List

Format policy rules

Formats [Create new rule](#)

Groups Show entries Search:

	Purpose	Format	Command	Success	Enabled	Actions
Identification	Access	Generic JPEG	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Tools	Access	1987a	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Rules	Access	1989a	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Commands	Access	Adobe Photoshop	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Format policy registry	Access	Bitmap 3.0 NT	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Tools	Access	Canon CIFF raw image data	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Characterization	Access	Canon CR2 Raw	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Rules	Access	Digital Negative Format (DNG)	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Commands	Access	Digital Negative Format (DNG)	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Event Detail	Access	Digital Negative Format (DNG)	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Rules	Access	Digital Negative Format (DNG)	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable
Commands	Access	Digital Negative Format (DNG)	Transcoding to jpg with convert	0 out of 0	Yes	View Replace Disable

Showing 1 to 10 of 344 entries Previous 2 3 4 5 ... 35 Next

Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/normalization-rules.png.

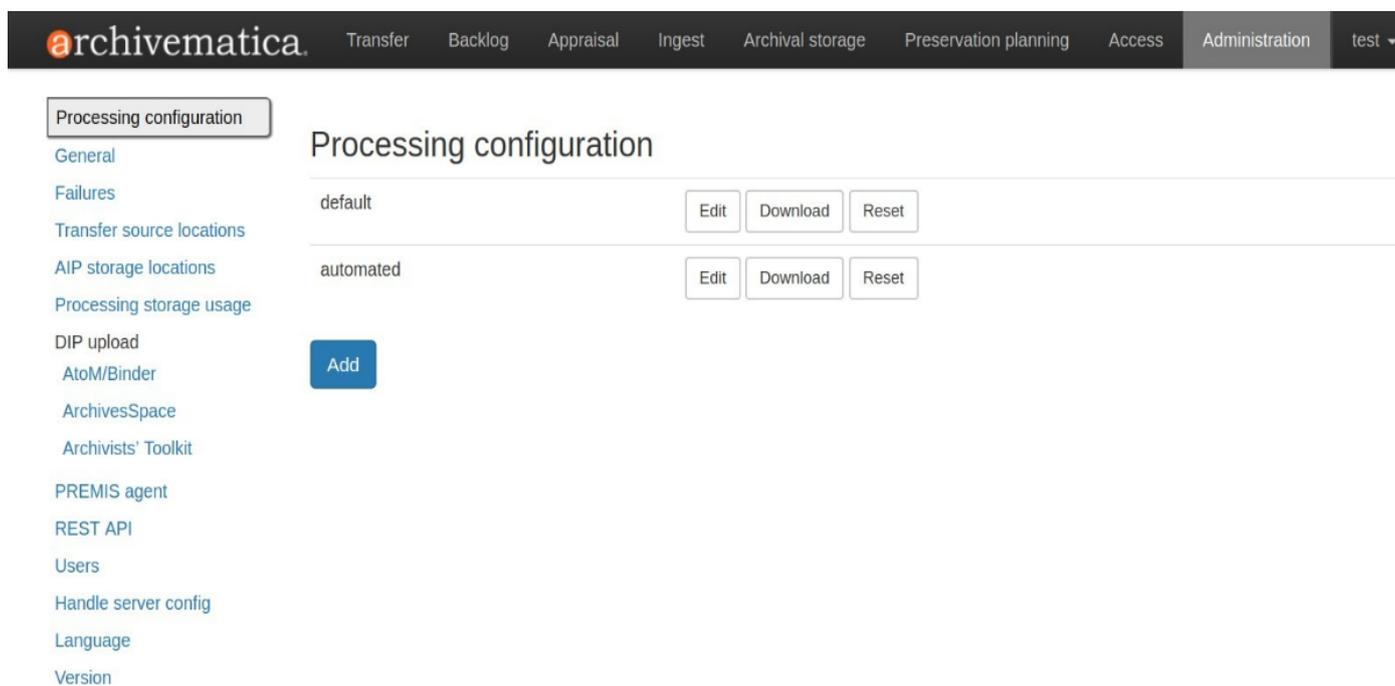
A política de formato padrão do *Archivematica* é mantido pela Artefactual Systemas e suas atualizações podem ser feitas a qualquer momento. As FPRs permitem que comandos possam ser associados a tipos de arquivos específicos. Por exemplo, no caso das regras de normalização, que lhe permitam configurar o comando que usa *ImageMagick* para criar miniaturas a serem executados em .gif e .jpeg, selecionando um comando diferente para ser executado.

h) Guia Administrador (*Administration*)

Esta guia descreve a administração (Figura 35) do *Archivematica*, permitindo configurar os componentes do aplicativo e gerenciar os usuários, uma espécie de painel de controle. A configuração de processamento deste guia permite ajustar os pontos de decisão do trabalho apresentados pelo *Archivematica* durante a Transferência. Conforme Costa *et al* (2016), existem dois tipos de processamento disponíveis na instalação:

1. Padrão - usado para qualquer transferência iniciada manualmente, a partir da guia Transferência, automaticamente por meio das ferramentas de automação ou em qualquer outro contexto em que nenhuma outra configuração de processamento seja especificada.
2. Automatizada - usada para transferências iniciadas automaticamente, a partir de ambientes Jisc RDSS ²¹(pasa usuários usuário Jisc, caso não seja, pode-se excluir esta configuração).

Figura 35 - Guia Administração



Fonte: https://www.archivematica.org/pt-br/docs/archivematica-1.8/_images/processing-config1.png.

É possível editar, por exemplo, um arquivo de configuração de processamento, como mostrado na opção “Editar” (Edit) da Figura acima. Pode-se também, criar outras configurações usando o botão “Adicionar” (Add). Os usuários criam várias configurações para diferentes tipos de conteúdo - para material audiovisual, para imagens, para registros textuais, etc. (COSTA *et al*, 2016).

Do lado esquerdo da imagem vê-se todos os elementos dispostos na guia Administração: Geral; Falhas; Transferir locais de origem; Locais de armazenamento AIP; Processando o uso de armazenamento; *Upload* DIP; AtoM/Binder; ArchiveSpace; *Kit* de Ferramentas dos Arquivistas; Agente PREMIS; API REST; Lidar com configuração do servidor; Língua; Versão. As configurações disponíveis nesta guia são:

²¹ Funciona como serviço de compartilhamento de dados para repositório de código aberto.

1. Selecione se quer ou não enviar transferências para a quarentena (sim/não) e decidir quanto tempo devem permanecer ali.
2. Aprove a normalização, o envio do AIP para armazenamento, e faça o *upload* do DIP sem interromper o *workflow* no painel.
3. Armazene o AIP sem interromper o *workflow* no painel.
4. Selecione se o *Archivematica* deve ou não transcrever arquivos.
5. Selecione se o *Archivematica* deve ou não gerar um relatório de estrutura de Transferência.
6. Selecione se *Archivematica* deve criar um SIP de transferência e continuar o processamento ou enviar a transferência para o *Backlog*.
7. Selecione se deseja extrair pacotes, manter ou excluir os objetos extraídos.
8. Selecione as opções de normalização.
9. Selecione se o *Archivematica* deve ou não lembrá-lo de adicionar metadados no ponto apropriado durante a Admissão.
10. Selecione se quer ou não examinar o conteúdo.
11. Selecione qual ferramenta de identificação, formato e comando para ser executado em ambos/qualquer Transferência ou Admissão como base para sua normalização.
12. Selecione qual ferramenta de identificação, formato e comando de iniciar para a documentação, apresentação e metadados.
13. Selecione se os pacotes devem ser apagados após a extração.
14. Selecione entre 7z usando lzma e 7zip usando bzip ou algoritmos bzip2 paralelos para compressão AIP.
15. Selecione os níveis de compressão:
 - 1 - *fastest mode*
 - 3 - *fast compression mode*
 - 5 - *normal compression mode*
 - 7 - *maximum compression*
 - 9 - *ultra compression*
16. Selecione um local de armazenamento de arquivo em que você vai enviar consistentemente seus AIPs e DIPS (se você estiver armazenando DIPS em vez de enviá-los imediatamente a um sistema de acesso.) (COSTA *et al*, 2016).

Nesta descrição apresentou-se um cenário sintetizado das funcionalidades técnicas do *Archivematica*. Teve-se o intuito de expor as principais atribuições dentro de cada guia/painel

deste software e, não pormenorizá-la em sua totalidade; tal finalidade se encontram *Guia do Usuário Archivematica*, assim como o de *Instalação* em seu site oficial.