

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
DEPARTAMENTO DA CIÊNCIA DA INFORMAÇÃO  
CURSO DE ARQUIVOLOGIA

DANIEL FERREIRA DA SILVA

**Repositórios de objetos arquivísticos confiáveis:  
o *software* RODA e perspectivas para a  
preservação da informação digital**

Porto Alegre – 2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
DEPARTAMENTO DA CIÊNCIA DA INFORMAÇÃO  
CURSO DE ARQUIVOLOGIA

DANIEL FERREIRA DA SILVA

**Repositórios de objetos arquivísticos confiáveis:  
o *software* RODA e perspectivas para a  
preservação da informação digital**

MONOGRAFIA APRESENTADA NO CURSO DE  
ARQUIVOLOGIA DA DA UNIVERSIDADE FEDERAL DO  
RIO GRANDE DO SUL COMO REQUISITO PARCIAL  
PARA OBTENÇÃO DO TÍTULO DE BACHAREL EM  
ARQUIVOLOGIA.

ORIENTADOR: RAFAEL PORT DA ROCHA

Porto Alegre – 2014

**Universidade Federal do Rio Grande do Sul**

Reitor: Carlos Alexandre Netto

**Vice-reitor:** Rui Vicente Oppermann

**Faculdade de Biblioteconomia e Comunicação**

Diretora: Ana Maria Mielniczuk de Moura

Vice-diretora: André Iriburi Rodrigues

**Departamento da ciência da informação**

Chefe: Maria do Rocio Fontoura Teixeira

Chefe-substituto: Valdir Morigi

**Dados internacional na catalogação (CIP)**

S586 Silva, Daniel Ferreira da

Repositórios de objetos arquivísticos confiáveis: o *software* RODA e perspectivas para a preservação da informação digital. / Daniel Ferreira da Silva. – Porto Alegre: 2014.

79 f.

Orientador: Rafael Port da Rocha

Trabalho de Conclusão de Curso (Graduação em Arquivologia) – Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, Curso de Arquivologia, Porto Alegre, 2014.

1. Arquivologia 2. Repositórios Digitais 3. Preservação Digital I. Título

CDU: 651.54

Bibliotecária responsável: Nalin Ferreira CRB 10/2186

**Departamento da ciência da informação**

Rua Ramiro Barcelos, 2507

CEP: 90035-007 – Porto Alegre-RS

TEL: 51 3308-5143

E-mail: dci@ufrgs.br

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
DEPARTAMENTO DA CIÊNCIA DA INFORMAÇÃO  
CURSO DE ARQUIVOLOGIA**

**DANIEL FERREIRA DA SILVA**

**Repositórios de objetos arquivísticos confiáveis: o *software*  
RODA e perspectivas para a preservação da informação digital**

MONOGRAFIA APRESENTADA NO CURSO DE  
ARQUIVOLOGIA DA UNIVERSIDADE FEDERAL DO RIO  
GRANDE DO SUL COMO REQUISITO PARCIAL PARA  
OBTENÇÃO DO TÍTULO DE BACHAREL EM  
ARQUIVOLOGIA.

ORIENTADOR: RAFAEL PORT DA ROCHA

**BANCA EXAMINADORA**

---

DR. RAFAEL PORT DA ROCHA

---

DR. MOISÉS ROCKEMBACH

---

MARLISE GIOVANAZ

Porto Alegre – 2014

*A Universidade Federal do Rio Grande do Sul  
por propiciar ensino de qualidade e gratuito.  
Aos meus Familiares, professores e amigos!*

*Não se encontra o espaço,  
sempre é necessário construí-lo.*

*Gaston Bachelard*

**Resumo**

Investiga, conceitua e analisa repositórios de objetos arquivísticos confiáveis, com destaque para a plataforma RODA. Contextualiza normas arquivísticas brasileiras e recomendações internacionais para gestão e preservação digital em repositórios arquivísticos de objetos confiáveis. Investiga recursos do ambiente RODA que podem apoiar a construção de um repositórios de objetos arquivísticos confiáveis, através de estudos existentes, enfocando a gestão de documentos arquivísticos e o gerenciamento de documento digital confiável.

Palavras-chaves: RODA, Repositórios Confiáveis, Digital, Preservação, Metadados, Gestão

**Abstract**

The work investigates, defines and analyzes trusted digital repositories, focusing the RODA platform. It contextualizes Brazilian and international standards and recommendations for management and digital preservation of objects in trusted digital archives. It explores existing works to investigate features of RODA that can support the construction of a repository of reliable archival objects, focusing on the management of archival documents and trusted digital objects.

Keywords: RODA, Repositories Reliable, Digital Preservation, Metadata Management

## LISTA DE FIGURAS

Figura 1 Modelo OAIS para preservação digital	27
Figura 1.1 Modelo OAIS para preservação digital	31
Figura: 2 Exemplo de um esquema de EAD	43
Figura 3: Modelo conceitual de metadados em PREMIS	45
Figura 4: Organograma básico das principais interações do RODA	57
Figura 5 Interfaces de admissão de uma série de pacotes de arquivos SIP (RODA)	59
Figura 6: Estrutura conceitual de submissão de pacotes de informação	60
Figura 7 Interfaces de admissão de pacotes de arquivos SIP transformados em AIP (RODA)	63
Figura 8: Modelo molecular do RODA	65

## LISTA DE QUADROS

Quadro 1: Funções das entidades no modelo OAIS	27
Quadro 2: Requisitos para preservação no modelo OAIS	28
Quadro 3 - Principais ações no modelo funcional OAIS	30
Quadro 4: Funções do SIP, AIP e DIP	32
Quadro 5: Requisitos para implementação de repositórios	33
Quadro 6: Tipos de Metadados	38
Quadro 7: Entidades envolvidas no modelo PREMIS	46
Quadro 8: Roteiro metodológico	49
Quadro 9: Empacotamento SIP	61
Quadro 10: Principais ações no modelo molecular do RODA	66

## **LISTA DE SIGLAS E ABREVIATURAS**

AIP – Archival Information Package

CCSDS – The Consultive Committee for Space Data System

CONARQ – Conselho Nacional de Arquivos

DIP – Dissemination Information Package

EAD – Encoder Archivist Description

IBICT – Instituto Brasileiro de Ciência e Tecnologia

ISAD (G) General International Standard Archival Description

ISO – International Organization for Standardization

METS – Metadata Encoding and Transmission Standard

NISO – National Information Standards Organization

OAIS – Open Archival Information System

OCLC – Online Computer Library Center

PREMIS – Preservation Metadata: Implementation Strategies

RLG – Research Library Group

SINAR – Sistema Nacional de Arquivos

SIP – Submission Information Package

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>14</b>
<b>2 REVISÃO TEÓRICA</b>	<b>17</b>
<b>2.1 Preservação Digital</b>	<b>17</b>
<b>2.2 Repositórios Digitais</b>	<b>22</b>
<i>2.2.1 Repositórios digitais confiáveis</i>	<b>23</b>
<i>2.2.2 Repositórios digitais de documentos arquivísticos confiáveis</i>	<b>25</b>
<i>2.2.3 O Modelo de referência OAIS</i>	<b>26</b>
<i>2.2.4 Requisitos para um Repositório Digital Confiável</i>	<b>33</b>
<i>2.2.5 CONARQ e a implementação de políticas de preservação digital</i>	<b>35</b>
<b>2.3 Metadados</b>	<b>37</b>
<i>2.3.1 Tipos de metadados: descritivo, estrutural, administrativo, técnico e para preservação</i>	<b>38</b>
<i>2.3.2 Padrões de esquemas de metadados</i>	<b>40</b>
<i>2.3.2.1 Estrutural: METS</i>	<b>40</b>
<i>2.3.2.2 Descritivos: EAD (Encoder Archivist Description)</i>	<b>42</b>
<i>2.3.2.3 Preservação: PREMIS Data Dictionary</i>	<b>44</b>
<b>3 METODOLOGIA</b>	<b>49</b>
<b>4 ANÁLISE DE DADOS</b>	<b>54</b>
<b>4.1 Principais características do RODA</b>	<b>54</b>
<i>4.1.2 O RODA</i>	<b>55</b>
<i>4.1.2.1 O RODA e sua arquitetura</i>	<b>56</b>

4.1.2.2 <i>SIP dentro do RODA</i>	58
4.1.2.3 <i>Estrutura de um Pacote de Informação (SIP)</i>	61
<b>4.2 Gestão de documentos arquivísticos no RODA</b>	<b>67</b>
<b>4.3 Gerenciamento de documento digital confiável</b>	<b>68</b>
<b>4.4 Captura de documentos digitais no RODA</b>	<b>70</b>
<b>4.5 Criação do pacote de arquivamento no RODA</b>	<b>70</b>
<b>4.6 Preservação do AIP no RODA</b>	<b>71</b>
<b>4.7 Acesso no RODA</b>	<b>72</b>
<b>5 CONSIDERAÇÕES FINAIS</b>	<b>74</b>
<b>6 REFERÊNCIAS</b>	<b>76</b>

O gerenciamento de informações atualmente está muito ligado a informática. Tanto para o âmbito administrativo como pessoal. A evolução rápida e transformadora que as tecnologias da informação vêm sofrendo no decorrer dos anos, e questões ligadas ao mercado, criam uma competitividade desenfreada, e muitos componentes tecnológicos entram em desuso rapidamente, assim, como criam-se novas tecnologias tanto de *hardware* como de *software*. Causando mudanças muito rápidas em termos de compatibilidade. Nesse contexto que surge a preocupação com a preservação da informação, e o manejo do conhecimento humano, principalmente relacionada a informação digital.

As principais motivações para esse trabalho foram relacionadas a:

- Necessidades de preservação a longo prazo em repositórios digitais;
- Necessidades por repositórios de objetos arquivísticos confiáveis;
- Surgimento de normas e modelos para implementação de repositórios de objetos arquivísticos confiáveis (OAIS, PREMIS, ISO 16363 de 2012, Resolução 39 de 2014 do CONARQ).

Neste trabalho foram abordados parâmetros relacionados a preservação digital através de estudos em repositórios arquivísticos de objetos confiáveis. É possível estabelecer relações existentes entre esses repositórios disponíveis na *web*, como ferramentas de gerenciamento da informação e suas aplicações na organização e acesso de coleções e fundos arquivísticos. Considerando as questões normativas, técnicas e principalmente científicas que a preservação da informação se torna fundamental no processo de implementação e administração desses repositórios.

Nesse contexto esse trabalho visa a destacar um *software* específico, desenvolvido em Código Aberto<sup>1</sup> (*Open Source*) e disponível para toda

---

<sup>1</sup> O termo código aberto ou *open source* surge na década de 1990 com o surgimento de uma tendência em relação aos programadores de sistemas, principalmente de *softwares* e aplicativos para *web*. Essa tendência queria que fossem abertos os códigos de programação de grandes aplicativos na época, exemplo o Netscape. A partir daí surgem muitas comunidades desenvolvedoras de aplicativos para sistemas em código aberto. Uma onda que se estende até hoje principalmente em usuários das linguagens *Debian* e *Linux*. (<http://opensource.org/history>)

comunidade na *web*, chamado RODA<sup>2</sup>. Através disso, foi possível investigar o que é o RODA e para que serve? Conceituar o que são repositórios, e em que o *software* RODA pode contribuir para a construção de um repositório de objetos arquivísticos confiáveis.

Também foi possível compreender através das normas criadas para a arquivologia para gestão de repositórios, assim como, verificar sua contextualização com esses sistemas desenvolvidos para gerenciar essas informações. E também, o que são essas exigências e como instaurá-las nesses repositórios para proteger a informação digital e torná-la acessível de modo que a informação também possa ser preservada a longo prazo.

Nesse sentido é possível destacar normas arquivísticas vigentes para descrição da informação (ISAD-G e EAD), também o modelo de referência de preservação e gestão digital (OAIS) e os padrões internacionais de metadados de preservação digital (PREMIS) e estruturais (METS). Assim como, sua aplicação no processo de admissão gerência e acesso da informação nesses repositórios.

Através dessas exigências, este trabalho teve como objetivo investigar recursos do Software RODA que podem apoiar a construção de um repositório de objetos arquivísticos confiáveis.

### **Objetivo geral**

- Investigar recursos do ambiente RODA que podem apoiar a construção de um repositórios de objetos arquivísticos confiáveis.

### **Objetivos específicos**

- Analisar as principais características, através de estudos existentes, do *software* RODA;
- Verificar os recursos que o RODA oferece para gestão de documentos arquivísticos;

---

<sup>2</sup> RODA é um repositório digital completo que fornece funcionalidade para todas as principais unidades do modelo de referência OAIS. RODA é capaz de ingerir, gestão e que dá acesso aos vários tipos de conteúdos digitais produzidos por grandes empresas ou organismos públicos. RODA é baseado em tecnologias open-source e é apoiado por normas existentes, como o OAIS, METS, EAD e PREMIS. (<http://www.roda-community.org>)

- Verificar os recursos que o RODA oferece para gerenciamento de documento digital confiável, com relação a admissão (captura de documentos digitais, criação do pacote de arquivamento), armazenamento e acesso.

Esse trabalho foi estruturado da seguinte forma. O capítulo 2 apresenta uma revisão teórica em estudos relacionados a repositórios digitais, principalmente em relação a implementação de repositórios de objetos confiáveis, destacando a norma OAIS. Percorrendo nas diretrizes e aconselhamentos do CONARQ (Conselho Nacional de Arquivos), na qual destaca os principais requisitos e parâmetros necessários através de diretrizes para a implementação de repositórios de objetos confiáveis no Brasil. E levanta, também os principais atuantes envolvidos no processo de preservação da informação digital. Analisando a parte de gestão, a importância das normas arquivísticas aliado a importância dos metadados para preservação digital. Assim como a conceituação do que é repositório, no intuito de criar um panorama geral das questões relacionadas a implementação de repositórios de objetos confiáveis.

O capítulo 3 trata da metodologia empregada para checar as informações relacionadas aos estudos relevantes de repositórios de objetos confiáveis, através de um roteiro que propõe os principais passos para investigar os recursos do Software RODA para a construção de um repositório confiável.

O Capítulo 4 trata da análise dos dados e de conteúdo, esse capítulo vai analisar as informações e estabelecer as relações entre requisitos teóricos e as experiências práticas relatadas nos estudos e demonstradas através destes. Concomitantemente é analisada a arquitetura conceitual dos elementos integrantes do RODA e suas relações com o trato da informação digital. Usando os trabalhos científicos e artigos produzidos sobre o RODA e analisando sua arquitetura e funcionamento para gestão em fundos arquivísticos.

## **2 REVISÃO TEÓRICA**

Através desse trabalho foi possível criar um entendimento mais aprofundado sobre repositórios de objetos arquivísticos confiáveis, em destaque o RODA, e suas relações com as comunidades desenvolvedoras e de usuários. Através de bibliografia consultada e referências que se contextualizam e investigam a aplicação das normas aos *softwares* e plataformas desenvolvidas e disponíveis. Foi possível demonstrar um laboratório conceitual da execução desses repositórios disponíveis em textos e artigos sobre a plataforma RODA.

Através da revisão teórica é que foi possível definir, conceituar e destacar os principais estudos referentes a repositórios digitais existentes. Também revisar os princípios arquivísticos e a suas aplicações em fundos e coleções disponibilizados na *web* através desses repositórios. Simultaneamente, analisar as diretrizes e resoluções estabelecidas no Brasil e internacionalmente, voltadas para a gestão de repositórios de objetos confiáveis e preservação digital. Também foi possível destacar as conexões com os estudos produzidos até então sobre a plataforma RODA verificando a coerência teórica aplicada a prática em repositórios digitais de objetos confiáveis.

### **2.1 Preservação Digital**

A preservação digital ainda é um assunto com muitas lacunas e com muita carência de estudos. Assim tornando-se um campo amplo para desenvolver-se e aprimorar-se. Em decorrência da rápida transformação tecnológica que vemos acontecer em pouco tempo. Sendo o maior desafio o gerenciamento de coleções e fundos arquivísticos em ambientes digitais, assim como a preservação a longo prazo desses objetos ou arquivos digitais.

Segundo Ferreira (2006),

Um objeto digital pode ser definido como todo e qualquer objeto de informação que possa ser representado através de uma sequência de dígitos binários. Esta definição é suficientemente lata para acomodar tanto, informação nascida num contexto tecnológico digital (objetos nato-digitais), como informação digital obtida a partir de suportes analógicos (objetos digitalizados) (p. 21).

Para compreender preservação digital é preciso saber que dados digitais e objetos autênticos requerem de *hardware e softwares para poderem ser exibidos*. O *hardware* é composto pelos componentes eletrônicos que são a parte física do computador. O *software* compõe um sistema operacional que gerencia informações podendo desempenhar uma série de tarefas específicas, como por exemplo o manejo e a visualização da informação. Com o decorrer do avanço das tecnologias da informação esses componentes foram diminuindo o tamanho físico (circuitos eletrônicos, placas, gabinete etc.) e aumentando o espaço de armazenamento digital e lógico. Com isso foram sendo agregados muitos recursos aos computadores. Em detrimento da velocidade em que essas mudanças foram ocorrendo é que é necessário pensar em preservação de informação. Principalmente considerando os problemas de migração de sistema operacional. Preservação Digital pode ser definido como:

Planejamento, alocação de recursos e aplicação de métodos de preservação e tecnologias necessárias para que a informação digital de valor contínuo permaneça acessível e utilizável por longo prazo, considerando-se neste caso longo prazo, o tempo suficiente para preocupar-se com os impactos de mudanças tecnológicas. A preservação digital aplica-se tanto a documentos "nato-digitais" quanto a documentos convertidos do formato convencional para o formato digital. (HEDSTROM, 1997/1998).

Nesse contexto pode-se afirmar que a preservação digital e a tecnologia da informação dependem de elementos tecnológicos.

Para que um ser humano seja capaz de decifrar um objeto digital, há um conjunto de transformações que têm necessariamente de ocorrer. Um objeto digital começa por ser um objeto físico, i.e. um conjunto de símbolos ou sinais inscritos num suporte físico (e.g. disco rígido, CD, DVD, disquete) (FERREIRA, 2006, p. 22).

Essa premissa nos remete a conclusão que Ferreira (2006) faz: o *hardware* assume aqui a responsabilidade de transformar os símbolos inscritos no suporte físico num conjunto de dados que o *software* será capaz de manipular (p.22). Para que ocorra preservação dessa informação produzida em diversos âmbitos e processada através do computador entre *hardware* e *software* é preciso atentar para isso:

Para que a preservação de um objeto digital seja possível, é necessário assegurar que todos os níveis de abstração (...) se encontram acessíveis e interpretáveis. Se a cadeia de interpretação que permite elevar um objeto digital desde o seu nível físico até ao nível conceitual for rompida, a comunicação deixa de ser possível e o objeto perder-se-á para sempre (FERREIRA, 2006, p. 23).

Por se tratar de um tema atual e ainda muito pouco explorado é que a preservação digital deve ser tratada com cuidado. Tendo em vista o avanço informacional e a disponibilização da informação através de tecnologias de baixo custo e acessíveis. A preservação da informação é a preocupação em manter a autenticidade dos elementos informacionais (imagem, texto, áudio, audiovisual e etc.)

Algumas comunidades vêm desenvolvendo estudos em gestão e preservação digital. Os *softwares* livres em código aberto, são comunidades

que têm um papel central no desenvolvimento de ferramentas gerenciais, pois, podem ser customizados conforme a necessidade da instituição. E além de serem de distribuição gratuita são ferramentas muito versáteis. Capazes de agregar segurança e confiabilidade ao uso e gerenciamento da informação.

Foi possível destacar nesse trabalho os modelos de referência e as normas internacionais, assim como as práticas em repositórios digitais. As premissas arquivísticas são fundamentais para que essas implementações possam gerir um repositório de objetos confiáveis.

A dificuldade fundamental da preservação digital advém da natureza dos próprios objetos que busca preservar. Diferentemente dos formatos tradicionais, os objetos digitais são acessíveis somente através de combinações específicas de componentes de *hardware*, *software*, mídia e pessoal técnico. Até bem pouco tempo atrás, a atenção de profissionais da informação concentrava-se, apenas, na longevidade do suporte físico onde a informação era armazenada. Esse posicionamento não é suficiente no mundo digital. Mesmo nas melhores condições de armazenamento, as mídias digitais podem ter sua vida interrompida pela falta ou inadequação de qualquer um dos demais componentes (THOMAZ; SOARES, 2004).

Além dessa preocupação em manter dados intactos e acessíveis é preciso também que se disponibilizem normas acessíveis e atualizadas capazes de autenticar e gerar valor probatório ao objeto digital. Através disso é possível destacar algumas referências que são fundamentais para compreender a preservação digital em sentido amplo, e sua importância em planejamentos estratégicos de gestão em ambientes digitais.

Segundo Lusenet (2001), para compreender o conjunto de problemas associados à preservação digital é preciso, principalmente e antes de mais nada, conscientizar-se de que: 1) as mídias são suportes

transitórios que prestam sua função somente por um período limitado de tempo e que a transferência para novas mídias é absolutamente necessária; 2) o software e o hardware tornam-se obsoletos em questão de anos, ao invés de décadas, e que embora as versões sucessivas de programas possam ser compatíveis, os fabricantes de software normalmente não garantem a compatibilidade por um longo período; e 3) o software proprietário é problemático não somente porque é protegido e o código fonte não está disponível mas, também, porque normalmente está documentado de forma inadequada tornando a conversão de dados muito mais complexa (THOMAZ; SOARES, 2004).

Através disso foi possível deduzir que para haver preservação digital é preciso saber como essa informação foi produzida e se os parâmetros pré-definidos poderão ser processados no *hardware* para tornar-se manipuláveis através do *software*: "a metainformação de preservação é responsável por reunir, junto do material custodiado, informação detalhada sobre a sua proveniência, autenticidade, atividades de preservação, ambiente tecnológico e condicionantes legais (FERREIRA, 2006, p. 54)".

No que diz respeito à proveniência a metainformação de preservação procura descrever a história custodial dos materiais, o caminho percorrido por estes materiais desde a sua criação até à sua incorporação no repositório (FERREIRA, 2006, p. 54).

Daí a importância de agregar os preceitos da arquivologia aos repositórios digitais. Podendo assim gerir com mais qualidade, podendo também gerar autenticidade e proveniência de toda documentação produzida e acumulada de forma digital.

Esta assume também a responsabilidade de

garantir a autenticidade dos materiais. Para tal, agrega um conjunto de metainformação que descreve detalhadamente todo o tipo de atividades desenvolvidas no interior do repositório (FERREIRA, 2006, p. 54).

Esse caminho em que a informação em forma de arquivo digital percorre, desde sua ingestão até seu acesso, dentro de um repositório, deve estar de acordo com o modelos e requisitos, mínimos usados internacionalmente, pois são padrões que protegem a informação de uma maneira em que ela não se perderá futuramente. E poderá ter valor probatório assim como ser um objeto autêntico digital. Evitando extravios da informação e mal uso de documentos digitais.

Com isso é possível destacar a relevância da preservação digital e suas implicações. Para que exista uma confiabilidade em objetos autênticos digitais são necessários diversos cuidados.

Capacidade de manter a integridade e a acessibilidade da informação digital por longo prazo. Esta preservação da integridade e acessibilidade não se limita, apenas, a proteger a informação digital contra o acesso não autorizado mas, também, contra o uso inadequado resultante da má interpretação ou má representação da informação por parte dos sistemas computacionais. Percebe-se, aqui, o aspecto da inseparabilidade entre as atividades de preservação e acesso do mundo digital. (RLG/CPA, 1996).

Nesse contexto que os planejamentos em gestão de documentação digital devem possuir estratégias eficazes no mapeamento da informação, ou seja, a entidade arquivística gestora deve ter políticas voltadas para a preservação digital.

## **2.2 Repositórios Digitais**

Segundo informações contidas no site do IBICT (Instituto Brasileiro de

Informação em Ciência e Tecnologia) Repositórios Digitais são:

(...) são bases de dados *online* que reúnem de maneira organizada a produção científica de uma instituição ou área temática. Os RDs armazenam arquivos de diversos formatos. Ainda, resultam em uma série de benefícios tanto para os pesquisadores quanto às instituições ou sociedades científicas, proporcionam maior visibilidade aos resultados de pesquisas e possibilitam a preservação da memória científica de sua instituição. Os RDs podem ser institucionais ou temáticos. Os repositórios institucionais lidam com a produção científica de uma determinada instituição. Os repositórios temáticos com a produção científica de uma determinada área, sem limites institucionais (<http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais> acessado em 05-08-2014).

Daí a importância desses ambientes digitais na manipulação do conhecimento humano e a necessidade de preservar a informação, visto que a tecnologia em um curto espaço de tempo tem disponibilizado a maioria das produções científicas e o conhecimento, através destes portais na Internet. Podendo ser acessadas de qualquer terminal de computador conectado a *Web*.

Conforme a resolução 39 do CONARQ, repositório digital é:

(...) um ambiente de armazenamento e gerenciamento de materiais digitais. Esse ambiente constitui-se de uma solução informatizada em que os materiais são capturados, armazenados, preservados e acessados. Um repositório digital é, então, um complexo que apoia o gerenciamento dos materiais digitais, pelo tempo que for necessário, e é formado por elementos de hardware, software e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos (CONARQ, 2014, p. 12).

### 2.2.1 *Repositórios digitais confiáveis*

A produção de arquivos digitais em todas as instâncias é uma realidade e o armazenamento dessa documentação se torna cada vez mais vulnerável devido aos avanços tecnológicos relacionados a informática. Existem diversos setores da sociedade que estão em processo de implementação desses repositórios, na intenção de proteger os seus dados e agregar confiabilidade nas informações produzidas.

Um repositório digital confiável é um repositório digital que é capaz de manter autênticos os materiais digitais, de preservá-los e prover acesso a eles pelo tempo necessário. Para cumprir essa missão, segundo o relatório “Trusted Digital Repositories: attributes and responsibilities” (RLG/OCLC, 2002)

Os repositórios digitais confiáveis entram como uma ferramenta eficaz para o gerenciamento dessa informação “repositório digital confiável é aquele que tem como missão oferecer, à sua comunidade-alvo, acesso confiável e de longo prazo aos recursos digitais por ele gerenciados, agora e no futuro” (RLG/OCLC, 2002, p. 5).

Funções em que o repositório deve exercer para se tornar confiável:

- aceitar, em nome de seus depositantes, a responsabilidade pela manutenção dos materiais digitais;
- dispor de uma estrutura organizacional que apoie não somente a viabilidade de longo prazo dos próprios repositórios, mas também dos materiais digitais sob sua responsabilidade;
- demonstrar sustentabilidade econômica e transparência administrativa;
- projetar seus sistemas de acordo com convenções e padrões comumente aceitos, no sentido de assegurar, de forma contínua, a gestão, o acesso e a segurança dos

materiais depositados (RLG/OCLC. 2007, *apud* CONARQ, 2014).

Existe a norma ISO 16363: 2012 que estabelece diretrizes para avaliar e certificar repositórios confiáveis. Também usada para implementação desses repositórios capazes de gerar certificação e autenticidade em relação a informação. Com base nela que a Resolução 39 aconselha o uso de normas específicas como por exemplo o OIAS (Open Information Archival Information System).

ISO 16363 utiliza linguagem e os conceitos dos OIAS, modelo de referência, que permite a avaliação e certificação de um repositório como sendo um repositório digital confiável (WITT, *et al*, 2012).

É pertinente ressaltar que as diretrizes traçadas pelo CONARQ atendem a ISO 16363 de 2012 onde abrange todos os tipos de materiais digitais, inclusive os documentos arquivísticos.

### *2.2.2 Repositórios digitais de documentos arquivísticos confiáveis*

Um repositório digital arquivísticos confiável pode ser definido como:

(...) é um repositório digital que armazena e gerencia esses documentos, seja nas fases corrente e intermediária, seja na fase permanente. Como tal, esse repositório 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. (CONARQ, 2014, p. 12).

Um repositório digital arquivísticos confiável contém os preceitos básicos da arquivologia, podendo gerenciar coleções e fundos arquivísticos em ambientes digitais, agregando as premissas da arquivologia, "repositório digital confiável de documentos arquivísticos deve ser capaz de atender aos procedimentos arquivísticos e aos requisitos de um repositório digital confiável (CONARQ, 2014, p. 13)."

### 2.2.3 O Modelo de referência OAIS

O OAIS (Open Archival Information System) foi criado pelo *Consultative Comitee for Space Data System* e a definição da sua aplicação é:

O modelo OAIS pode ser aplicado para qualquer arquivo. É especificamente aplicáveis à organizações com a responsabilidade de tornar a informação disponível para longo prazo. Isso inclui organizações com outras responsabilidades, como o processamento e distribuição de responsabilidades às necessidades programáticas (CCSDS, 2012, p.1-2).

O OAIS “é um modelo conceitual que se propõe a servir de norma para um sistema de gestão em preservação digital. Sendo aplicado principalmente em repositórios de objetos confiáveis”. Os seus objetivos são: identificar os componentes funcionais que devem fazer parte de um arquivo dedicado à preservação da informação digital e descrever as entidades internas e externas desses sistemas, bem como os objetos de informação que são manipulados no seu interior” (Ver Figura 1) (<http://colecções-digitais.wikidot.com/oais>).

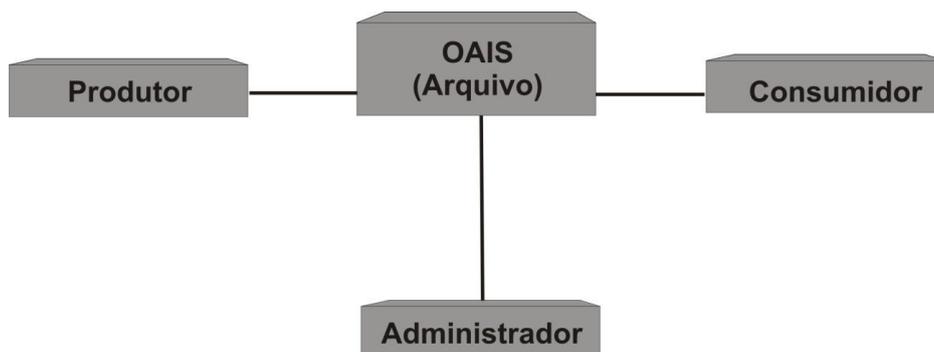
A norma mais importante da área é o *Open Archival Information System – OAIS*<sup>3</sup>, um modelo conceitual descreve as funções de um repositório digital e

---

<sup>3</sup> No Brasil, o modelo OAIS foi traduzido pela ABNT e publicado sob a forma da norma ABNT NBR 15472: 2007, com o título “Sistema Aberto de Arquivamento de Informação – SAAI”.

os metadados necessários para a preservação e o acesso dos materiais digitais gerenciados pelo repositório, que constituem um modelo funcional e um modelo de informação (CONARQ, ano, p. 11).

**Figura 1 Modelo OAIS para preservação digital**



Fonte: (CCSDS, 2012, p.2-2)

Esse esquema nos mostra as principais entidades envolvidas no ambiente de um repositórios digital segundo recomendações do modelo OAIS. Cada entidade envolvida no processo de gestão do repositório terá sua função dentro da administração do objeto digital. É possível destacar as principais funções das entidades envolvidas na administração da informação em ambiente OAIS, como mostra a Quadro 1:

**Quadro 1: Entidades no modelo OAIS**

<b>Produtores, Administradores e Consumidores</b>
Produtor é o papel desempenhado por essas pessoas, ou sistemas de clientes, que fornecem a informação a ser preservada.
Administrador é o papel desempenhado por aqueles que definem a políticas globais como um componente em um domínio mais amplo de políticas, por exemplo, como parte de uma organização maior.
Consumidor é o papel desempenhado por essas pessoas, ou sistemas de clientes, que interagem com serviços OAIS para encontrar e adquirir informações preservado

de interesse. Uma classe especial de consumidores é a Comunidade Designada. A Comunidade Designado é o conjunto de consumidores de que deve ser capaz de compreender a informação preservada.

Fonte: (CCSDS, 2012, p.2-2)

Podemos definir que cada entidade envolvida executará uma função dentro do repositório, através do modelo OAIS, e como essas relações se processarão, a interoperabilidade dos modelos de metadados, e os procedimentos necessários para dar valor probatório ao documento, para que não seja modificado ou alienado dentro repositório digital. Essas relações entre as entidades envolvidas se processarão através procedimentos aplicados ao arquivo digital (Quadro 2), levando em consideração sua trajetória dentro do ambiente digital, desde sua ingestão e ações aplicados no objeto digital, sua administração até seu acesso.

### **Quadro 2: Requisitos para preservação no modelo OAIS**

Uma instituição que pretende assumir a responsabilidade de preservar informação digital por longo prazo precisa, como primeiro passo, observar um conjunto mínimo de requisitos (...)

\* Fixar os limites do objeto a ser preservado: embora a natureza multimídia e hipertextual dos objetos digitais seja bastante vantajosa do ponto de vista da navegação, para fins de preservação é necessário definir, claramente, quais elementos serão efetivamente mantidos.

\* Preservar a presença física: a presença física representa o(s) arquivo(s) físico(s), i.e., a camada primitiva de suporte da informação a ser representada; refere-se, portanto, ao(s) arquivo(s) de computador, às séries de 0's e 1's que são a base para o significado de um objeto digital.

\* Preservar o conteúdo: refere-se a manter a capacidade de acessar o conteúdo em

<p>seu nível mais baixo, como um arquivo texto em ASCII<sup>4</sup>, independente do estabelecimento de variações de fontes e características de leiaute.</p>
<p>* Preservar a apresentação: o conteúdo é apresentado visualmente através da aplicação de fontes de diferentes formatos e tamanhos, uso de espaço em branco, colunas, margens, cabeçalhos, rodapés, paginação e assim por diante. Em alguns tipos de documentos digitais (p.ex., formatos padrão SGML e alguns formatos PDF) as especificações de apresentação ficam separadas do conteúdo.</p>
<p>* Preservar a funcionalidade: objetos digitais podem conter componentes multimídia (i.e., texto, gráficos, áudio e vídeo integrados), existir em formato hipertexto (i.e., podem desviar dinamicamente para outros pontos do próprio documento ou para outro documento), conter conteúdo dinâmico (i.e., gerado automaticamente a partir de bancos de dados) ou ter funções de navegação (i.e., barras de ferramentas, pesquisa a palavra-chave ou tabelas interativas de conteúdo).</p>
<p>* Preservar a autenticidade: é necessário confiar que o objeto acessado é exatamente aquele que se procura e que as possíveis transformações pelas quais passou, para manter sua acessibilidade, preservaram sua forma original.</p>
<p>* Localizar e rastrear o objeto digital ao longo do tempo: imediatamente após a sua criação, os objetos digitais tornam-se passíveis de serem alterados, copiados ou movimentados. Em qualquer referência ao objeto digital, é necessário localizá-lo na edição ou versão correta.</p>
<p>* Preservar a proveniência: identificar a origem de um objeto e detalhar seu histórico ajudam a confirmar sua autenticidade e integridade.</p>
<p>* Preservar o contexto: os objetos digitais são definidos por suas dependências de <i>hardware</i> e <i>software</i>, seus modos de distribuição e relacionamentos com outros</p>

---

<sup>4</sup> ASCII, do inglês *American Standard Code for Information Interchange*, é o Código Padrão Americano para o Intercâmbio de Informação. Trata-se de um código binário (cadeias de bits: 0s e 1s) que codifica um conjunto de 128 sinais: 96 sinais gráficos (letras do alfabeto latino, sinais de pontuação e sinais matemáticos) e 32 sinais de controle. Cada código binário possui 8 bits (= 1 byte): 7 bits para o propósito de codificação e 1 bit de paridade (detecção de erro). A codificação ASCII é usada para representar textos em computadores, equipamentos de comunicação, entre outros dispositivos que trabalham com texto. Desenvolvida a partir de 1960, grande parte das codificações de caracteres modernas a herdaram como base. Os sinais não-imprimíveis, conhecidos como caracteres de controle, são amplamente utilizados em dispositivos de comunicação e afetam o processamento do texto. O código ASCII é muito utilizado para conversão de Código Binário para Letras do alfabeto Maiúsculas ou minúsculas. (<http://pt.wikipedia.org/wiki/ASCII> acessado em 15/10/2014.)

objetos digitais.
-------------------

Fonte: THOMAZ; SOARES, 2004

Nesse contexto é possível identificar padrões fundamentais na trajetória da informação digital dentro de um repositório confiável através do modelo OAIS para gestão. Reconhecendo as entidades atuantes: produtor, administrador e consumidor. Assim como os modelos de metadados empregados e os procedimentos técnicos que ocorrerão na trajetória desses arquivos de fundo arquivístico.

O modelo de referência OAIS constituiu um ponto de partida para a discussão em torno da necessidade de criar um conjunto de elementos de metainformação capazes de dar suporte às atividades relacionadas com a preservação digital (...). Desde então, diversas instituições têm vindo a propor dicionários de metainformação que refletem as necessidades individuais dos projetos em que estão ou estiveram envolvidas (FERREIRA, 2006, p.55).

O modelo funcional nos traz novos elementos que se relacionam com as entidades envolvidas (Quadro 3). É composto pelas seguintes atividades (Ver Figura 1.1)

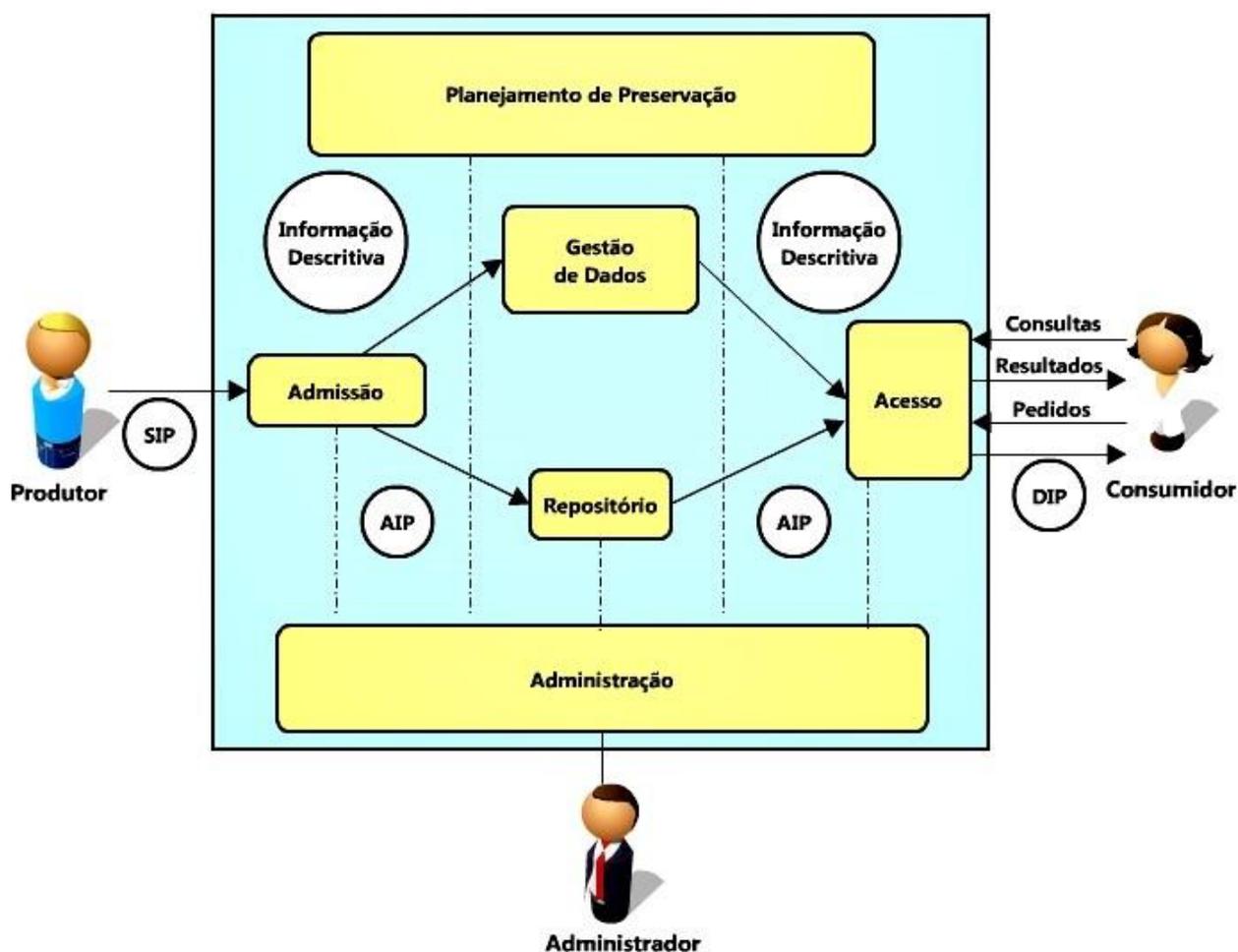
### Quadro 3 - Principais ações no modelo funcional OAIS

1) Ingestão: verifica integridade/qualidade da informação, produz informação descritiva
2) Acesso: descoberta, localização e entrega ao consumidor
3) Planejamento da preservação: políticas e ações de preservação (monitoramento, migração, referescamento,..)
4) Gestão dos dados: armazena e gerencia informação descritiva e do sistema

- 5) Armazenamento: armazenamento seguro para os itens depositados (mídias, recuperação após desastre,...)
- 6) Administração do sistema: operações de rotina o arquivo

Fonte: OAIS

Figura 1.1: Modelo OAIS para preservação digital



Fonte: (CONARQ, 2014, p. 23)

A importância da geração de pacotes fechados inter-relacionados (SIP, AIP e DIP) e informacionais capazes de descrever, estruturar e prevenir a

autenticidade do objeto digital ao ingerir o documento no repositório.

SIP (Submission Information Package), AIP (Archival Information Package) e DIP (Dissemination Information Package) são elementos denominados de pacotes acrescidos de multiníveis de informações. E são gerados a partir da admissão da informação (na forma de SIP), armazenados na forma AIP, até o seu acesso e disseminação (na forma de DIP).

O SIP (Submissão Pacote de Informação) é que o pacote que é enviado para um OAIS por um Produtor. Sua forma e conteúdo detalhado normalmente são negociados entre o produtor e o OAIS. As relações entre SIP e AIP pode ser complexa; bem como simples relações em que um SIP gera um AIP, outras possibilidades incluem: uma AIP sendo produzido a partir de várias SIP produzidos em diferentes momentos por um produtor ou por muitos produtores; um SIP resultando num número de AIP; e muitos SIP a partir de uma ou mais fontes ser desagregado e recombinações em diferentes maneiras de produzir muitos AIP. Mesmo no primeiro caso, o OAIS pode ter que realizar uma série de transformações na SIP o pacote de informações estará sempre presente de alguma forma. (CCSDS, 2012, p.2-7)

Um procedimento que descreve e agrega a um objeto digital multiníveis de informações sobre ele mesmo e que será armazenado dentro de um pacote que poderá produzir muitos níveis de descrição até o ponto de se transformar em outra entidade podendo também gerar relações múltiplas e de vários níveis, tudo isso no intuito de rastrear os passos da informação dentro dos repositórios.

#### **Quadro 4: Funções do SIP, AIP e DIP**

*SIP (Submission Information Package)*

Pacote de informação submetido ao repositório, constituído por uma representação e os metadados associados pelo produtor. (Conceito extraído do modelo OAIS)

*AIP (Archival Information Package)* Pacote de informação preservado pelo repositório, gerado a partir do *SIP*. É constituído pela representação e pelos metadados que lhes foram associados pelo produtor e pelo repositório ao longo do tempo. (Conceito extraído do modelo OAIS).

*DIP (Dissemination Information Package)*

Pacote de informação gerado a partir do *AIP*, entregue ao usuário do repositório na sequência de uma operação de pesquisa sobre o catálogo do repositório. (Conceito extraído do modelo OAIS).

Fonte: (ARQUIVO NACIONAL, p.9)

#### 2.2.4 Requisitos para um Repositório Digital Confiável

Segundo o CONARQ (2014,) os requisitos para um Repositório Confiável estão organizados em três conjuntos: infraestrutura organizacional; gerenciamento do documento digital; e tecnologia, infraestrutura técnica e segurança conforme Quadro 5 a seguir:

#### **Quadro 5: Requisitos para implementação de repositórios**

##### **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 auto avaliação de seu funcionamento e renovar sua certificação; e
- estar comprometido em notificar as entidades certificadoras sobre as mudanças operacionais que afetarão seu *status* de certificação (no caso de repositórios já certificados).

#### 2.4.2 CONARQ e a implementação de políticas de preservação digital

O CONARQ (Conselho Nacional de Arquivos) é um órgão que tem como competência estabelecer os principais parâmetros e requisitos relacionados a implementação de repositórios de objetos autênticos digitais. Assim como norteador das principais políticas de preservação a serem aplicadas em gestão de arquivos. Também servindo de órgão aconselhador das principais entidades responsáveis por essas implementações.

Art. 1º Recomendar aos órgãos e entidades integrantes do Sistema Nacional de Arquivos – SINAR, que tiverem por finalidade a transferência ou o recolhimento de documentos arquivísticos em formato digital, e de forma a garantir a integridade, a autenticidade, a confidencialidade, a disponibilidade e a preservação desses documentos, a adoção das Diretrizes para a Implementação de Repositórios Digitais Confiáveis de Documentos Arquivísticos, anexas a esta Resolução (CONARQ, 2014, p.3).

CONARQ estabelece programas e aconselhamentos em preservação de documentos digitais seguindo os princípios da arquivologia tradicional:

A preservação dos documentos arquivísticos digitais, nas fases corrente, intermediária e permanente, deve estar associada a um repositório digital confiável. Os arquivos devem dispor de repositórios digitais confiáveis para a gestão, a preservação e o acesso de documentos digitais (CONARQ, 2014, p.3).

Essas diretrizes são os primeiros passos e orientação para a implementação de repositórios de objetos autênticos disponibilizada como

premissas para sua melhor implementação e orientações nesse sentido sendo seu objetivo principal:

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 (CONARQ, 2014, p.3).

Em relação a implantação de padrões tecnológicos o CONARQ afirma:

Desde a década de 1990, a comunidade internacional tem desenvolvido iniciativas no sentido de orientar a modelagem e implementação de repositórios digitais, e de apontar os requisitos para atribuir confiabilidade a esses repositórios. A implantação de um repositório digital confiável é fundamental para assegurar a preservação, o acesso e a autenticidade de longo prazo dos materiais digitais (CONARQ, 2014, p.11).

É através dessa resolução que o CONARQ estabelece os principais requisitos para implementação desses repositórios em entidades integrantes ao Sistema Nacional de Arquivos (SINAR<sup>5</sup>).

---

<sup>5</sup> São integrantes do SINAR:

- Arquivo Nacional;
- arquivos do Poder Executivo Federal;
- arquivos do Poder Legislativo Federal;
- arquivos do Poder Judiciário Federal;
- arquivos estaduais dos poderes Executivo, Legislativo e Judiciário;
- arquivos do Distrito Federal dos poderes Executivo, Legislativo e Judiciário; e
- arquivos municipais dos poderes Executivo e Legislativo.

Podem, ainda, integrar o SINAR pessoas físicas e jurídicas de direito privado detentoras de arquivos, mediante convênio com um órgão central. (CONARQ, 2014).

A preservação dos documentos arquivísticos digitais, nas fases corrente, intermediária e permanente, deve estar associada a um repositório digital confiável. Os arquivos devem dispor de repositórios digitais confiáveis para a gestão, a preservação e o acesso de documentos digitais (CONARQ, 2014, p.7).

## 2.3 Metadados

Segundo o Dicionário de Terminologia Arquivística metadados significa "Dados estruturados e codificados, que descrevem e permitem acessar, gerenciar, compreender e/ou preservar outros dados ao longo do tempo" (2005, p. 115).

Metadados descrevem os recursos da *web* com a finalidade de facilitar a sua descoberta, localização e utilização. Motores de busca, ao utilizarem estes metadados, proporcionam consultas bem mais precisas, envolvendo não somente palavras, mas propriedades descritas, como o autor do recurso, o formato do recurso, a data do recurso etc. (ROCHA, 2004, p. 5).

Os metadados são maneiras de classificar a informação, em todos os sentidos e atribuições. Podendo nela se estender em longas definições descritivas até dados básicos como por exemplo a data, o autor e o assunto. Existem diversos padrões desenvolvidos de modelos de metadados. "Metadados é fundamental para garantir que os recursos vão sobreviver e continuar a ser acessível para o futuro (NISO, 2004, p. 3)". Esses padrões surgem para melhorar e proteger a informação.

O termo é usado metadados de forma diferente em diferentes comunidades. Alguns usam para se referir a máquina informação compreensível, enquanto outros usam apenas para registros que descrever recursos eletrônicos no ambiente da biblioteca, metadados é comumente utilizado para qualquer forma de esquema de descrição de recursos, aplicável a qualquer tipo de objeto, arte, digital ou não-digital

biblioteca tradicional catalogação é uma forma de metadados (NISO, 2004, p. 3).

Esses recursos que os metadados acrescentam ao arquivo nada mais é do que a inserção de informações sobre o próprio arquivo, na qual tem a função de servir como orientação e ajudar na busca e na recuperação da informação, além de dar valor probatório a informação, a partir de seu suporte, independente de compatibilidades entre *hardware* e *software*.

Os metadados também tem a função de servir na hora da organização da informação e seus níveis de arquivamento. Quanto mais informações tiver o metadados, mais recursos informacionais serão disponíveis em relação ao arquivo ou objeto digital.

Além de descoberta de recursos, os metadados podem ajudar a organizar os recursos eletrônicos, facilitar a interoperabilidade e integração de recursos de legado, fornece a identificação digital, e suporte de arquivamento e preservação (NISO, 2004, p.4).

### 2.3.1 Tipos de metadados

De acordo com a *National Information Standards Organization* (NISO) é possível distinguir alguns tipos fundamentais de Metadados como mostra a tabela a seguir:

**Quadro 6: Tipos de Metadados**

<p>• <b>Metadados descritivos</b> descreve um recurso para fins tais como descoberta e identificação ele pode incluir elementos tais como título, resumo, autor e palavras-chave.</p>
<p>• <b>Metadados estruturais</b> indica como objetos compostos são colocados em conjunto, por exemplo, como páginas são organizadas para formar capítulos.</p>
<p>• <b>Metadados administrativo</b> fornece informações para ajudar gerenciar um recurso, como quando e como ela foi criada, arquivo tipo e outro técnico informação, e quem pode acessá-lo.</p>

- **Metadados para preservação**, que contém informação necessária arquivar e preservar a recurso. Os metadados podem descrever recursos em qualquer nível de agregação. Pode descrever um conjunto, um único recurso, ou um componente de um maior recurso (por exemplo, uma fotografia num artigo). Assim como catalogadores tomar decisões sobre se um registro do catálogo deve ser criado para um conjunto de volumes ou para cada volume em particular no conjunto, de modo que o criador metadados toma decisões semelhantes. Os metadados também pode ser utilizada para a descrição em qualquer nível do modelo de informações estabelecidas na IFLA (Federação Internacional de Associações e Instituições Bibliotecárias) Requisitos Funcionais para Registros Bibliográficos: obra, expressão, manifestação ou item. Por exemplo, um registro de metadados poderia descrever um relatório, uma edição especial do relatório, ou uma cópia específica do que a edição do relatório.

Fonte: (NISO, 2004)

Cada metadados atende a necessidade em que for direcionado. Tendo sempre como objetivo oferecer informação sobre a informação, seja ela de áudio, visual, escrita e etc, a função principal do metadados é oferecer dados referentes ao elemento informacional. Daí surgem questões ligadas a propriedade intelectual, proveniência, data da criação do objeto entre outras descrições consideradas importantes e de valor probatório e "podem ajudar a organizar os recursos eletrônicos, facilitar a interoperabilidade (...) fornece a identificação digital, e suporte de arquivamento e preservação (NISO, 2004, p. 4).

Os metadados podem ser incorporados num objeto digital ou podem ser armazenados separadamente. Os metadados são muitas vezes incorporados em documentos HTML e nos cabeçalhos de arquivos de imagem. Armazenamento de metadados com o objeto que descreve garante os metadados que a informação não será perdida, evitando problemas de ligação entre dados e metadados, e ajuda a garantir que os metadados do objeto serão atualizados constantemente (NISO, 2004, p.3).

No caso da informação digital o metadados tem funções múltiplas e se torna fundamental para o funcionamento de um repositório digital considerando os mecanismos de busca de informação dentro da Internet como de qualquer outro banco de dados.

### 2.3.2 *Padrões de esquemas de metadados*

Existem três padrões de esquemas de metadados recomendados para a implementação de repositórios de objetos arquivísticos confiáveis. São eles: O METS, EAD e PREMIS.

#### 2.3.2.1 *Estrutural: METS*

O METS é um esquema de metadados estrutural. E tem como função principal criar ações de codificação para objetos digitais dentro de uma biblioteca digital através de um cabeçalho que descreve as principais informações em relação ao objeto digital. Sendo usado especificamente para coleções em ambientes digitais

O esquema METS oferece um mecanismo flexível para codificar metadados descritivos, administrativos e estruturais para um objeto de uma biblioteca digital, e para exprimir as ligações complexas entre estas várias formas de metadados. Assim o METS oferece uma norma útil para a troca de objetos digitais entre repositórios. Adicionalmente, o METS oferece a possibilidade de associar um objeto digital com comportamentos ou serviços. A discussão anterior descreve as principais funcionalidades do esquema, mas uma examinada mais detalhada do esquema e da sua documentação é necessária para compreender todo o alcance das suas capacidades

([http://www.loc.gov/standards/mets/METSOverview.v2\\_port.html](http://www.loc.gov/standards/mets/METSOverview.v2_port.html) acessado em 21/11/2014).

Sendo usado especificamente para coleções em ambientes digitais e por ser escrito no padrão da web o XML<sup>6</sup>, que possui um vasto campo de compatibilidade.

A seção do mapa estrutural de um documento METS define uma estrutura hierárquica que pode ser apresentada a utilizadores do objeto da biblioteca digital para lhes permitir navegar nele. O elemento <structMap> codifica esta hierarquia como série de elementos <div> encaixados. Cada <div> contém informação em atributos que especifica que tipo de divisão é, e também pode conter múltiplos apontadores METS (<mptr>) e elementos apontadores de ficheiros (<fptr>) para identificar o conteúdo correspondente a esse <div>. Apontadores METS especificam outros documentos METS como contendo a informação relevante para o <div> que os contém. Isto pode ser útil quando se codifica grandes coleções de material (e.g., todos os números de um jornal científico) para manter o tamanho de cada ficheiro METS relativamente pequeno. Apontadores para ficheiros especificam ficheiros (ou, em alguns casos, tantos grupos de ficheiros ou localizações específicas dentro de um ficheiro) dentro da secção <fileSec> do documento METS atual que corresponde à porção na hierarquia representada pelo <div>.

(<http://www.loc.gov/standards/mets/METSOverview.v2.html> acessado 21/11/2014).

---

<sup>6</sup> **XML** (*eXtensible Markup Language*) é uma recomendação da W3C para gerar linguagens de marcação para necessidades especiais. É um dos subtipos da SGML (acrônimo de *Standard Generalized Markup Language* ou *Linguagem Padronizada de Marcação Genérica*) capaz de descrever diversos tipos de dados. Seu propósito principal é a facilidade de compartilhamento de informações através da internet.

Entre linguagens baseadas em XML incluem-se XHTML (formato para páginas Web), RDF, SDMX, SMIL, MathML (formato para expressões matemáticas), NCL, XBRL, XSIL. A principal característica do XML, de criar uma infraestrutura única para diversas linguagens, é que linguagens desconhecidas e de pouco uso também podem ser definidas sem maior trabalho e sem necessidade de ser submetidas aos comitês de padronização (<http://tools.ietf.org/html/rfc3023#section-3.2>).

O METS também oferece ao objeto digital informações importantíssimas sobre sua origem e ainda cria muitas maneiras de se inter-relacionar com outros objetos dentro de uma estrutura digital através de um esquema universal dentro da rede. E sendo fundamental para o reconhecimento do objeto digital, assim como para sua disseminação dentro do mesmo contexto.

O METS (*Metadata Encoding & Transmission Standard*) trata-se de uma norma que permite associar metainformação descritiva, administrativa e estrutural sobre objetos digitais. Um documento METS é constituído por sete secções principais:

Cabeçalho METS – O cabeçalho METS contém metadados que descrevem o documento METS em si, incluindo informação sobre o criador, editor, etc.

(<http://www.loc.gov/standards/mets/METSOverview.v2.html> acessado em 21/11/2014).

#### 2.3.2.2 *Descritivo: EAD (Encoder Archivist Description)*

O EAD (Encoder Archivist Description) é considerado um padrão de metadados "emergente utilizado internacionalmente em um número crescente de arquivos e manuscritos bibliotecas para codificar os dados que descrevem registros corporativos e documentos pessoais" (PITTI,1999, <http://www.dlib.org/dlib/november99/11pitti.html>).

EAD é atualmente administrada e mantida conjuntamente pela Sociedade Americana de Arquivistas e a Biblioteca do Congresso dos Estados Unidos. Os desenvolvedores estão atualmente a estudar formas de internacionalizar a administração e manutenção de EAD para refletir e representam a base crescente de usuários (PITTI,1999,

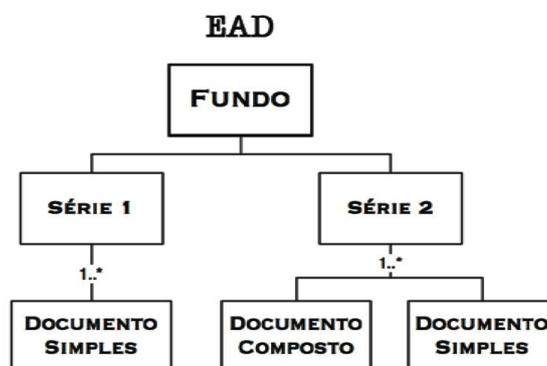
<http://www.dlib.org/dlib/november99/11pitti.html> acessado em 19/09/2014).

O EAD surge como uma maneira eficaz em relação a preservação digital. Tendo em vista que é uma descrição arquivística sofisticada que atende as exigências de troca de *hardware* independente de *software* segundo artigo publicado:

Padrões de codificação hardware e independente de software oferecem a única garantia razoável de informações duradouras. A fim de explorar descrição arquivística totalmente em um ambiente de computador, os arquivistas precisam representar com fidelidade e precisão sua natureza intelectual e de conteúdo (PITTI,1999).

O EAD também pode ser conceituado como um metadados descritivo com base na ISAD (G) que acrescenta no objeto os principais campos descritivos capazes de agregar muitos níveis de informação. Criando campos possíveis de se relacionarem dentro das linguagens compatíveis com a maioria das documentações dispostas na rede, independente do sistema. Abrindo um leque na compatibilidade e nas relações com outras entidades. Podendo ser arranjado dentro de fundos arquivísticos complexos e grandes em repositórios digitais (Figura 2).

**Figura 2 – Exemplo de um esquema de EAD**



Fonte (<http://www.apbad.pt/Downloads/congresso9/COM37.pdf>)

Todos esses elementos podem ser arranjados de modo a obter uma forma hierárquica. "Cada um destes elementos é caracterizado por um identificador único e um nível de descrição que pode assumir um dos seguintes valores" (BARBEDO, *et. al*, 2009).

- Fundo – o conjunto de todos os documentos, independentemente da sua forma ou formato, organicamente criados e/ou acumulados por certa pessoa, família, ou instituição no decurso das suas actividades ou funções;
- Série – conjunto de documentos que dizem respeito a uma actividade funcional;
- Documento composto – uma unidade organizada de documentos agrupados para uso do criador ou no processo de organização arquivística por serem relativos a um mesmo assunto ou actividade;
- Documento simples – a unidade elementar de um arquivo, i.e. um documento intelectualmente indivisível (BARBEDO, *et. al*, 2009).

### 2.3.2.3 *Preservação: PREMIS Data Dictionary*

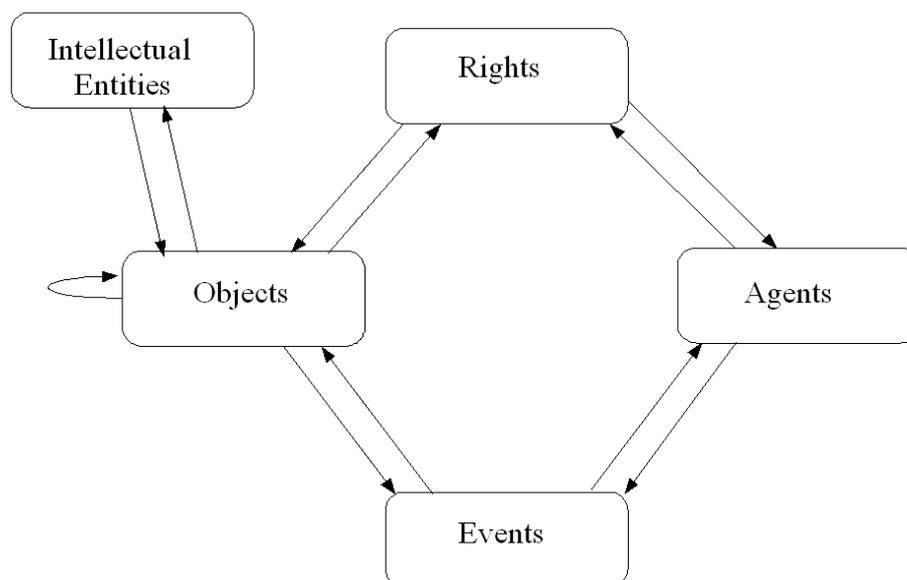
O PREMIS é considerado um metadados voltado para preservação digital. Que possui um vasto vocabulário agregado aos seus recursos. Capaz de identificar muitas entidades e muitos níveis de organização. Não apenas em listagem hierárquica ou ascendente, mas também na questão semântica e de sentido. Criando outras relações entre as informações do arquivo. Complexificando e criando mais alternativas na hora do acesso e recuperação da informação. O PREMIS é capaz de versátil função dentro de um repositório de objetos confiáveis como nos mostra a seguir:

Funções de manutenção da viabilidade, compreensibilidade, autenticidade e identidade num contexto de preservação. Metadados de preservação

abrange, assim, uma série de categorias tipicamente utilizadas para diferenciar tipos de metadados: administrativas (incluindo direitos e permissões), técnicos e estruturais. Foi dada especial atenção à documentação de proveniência digitais (a história de um objeto) e à documentação dos relacionamentos, especialmente os relacionamentos entre os objetos diferentes dentro do depósito de preservação (PREMIS, 2012, p.3).

Esses padrões de metadados criam as condições necessárias para dar autenticidade ao objeto digital. Pois ele identifica as principais característica dos objetos "as entidades estão representadas por caixas; relações entre as entidades são representadas por setas. A direção da seta indica a direção da relação de ligação, uma vez que é registado na preservação de metadados" (PREMIS, 2012, p. 3.). A Figura 3 ilustra essa premissa.

**Figura 3: Modelo conceitual de metadados em PREMIS**



Fonte: PREMIS 2012, p.3

O modelo conceitual cria as principais relações dentro de um metadados voltado para a preservação digital criando possíveis relações entre os envolvidos no processo "Por exemplo, a seta a apontar para a entidade de

Direitos para a entidade Agentes significa que os metadados associado com a entidade Direitos inclui uma unidade de gravação de informação semântica acerca da relação com um Agente " (PREMIS, 2012, p.3).

A seta apontando da entidade Objetos de volta para si mesmo indica que as unidades semânticas definidas no Dicionário de Dados apoiar a gravação de relações entre objetos. Nenhuma outra entidade no modelo de dados suporta relacionamentos desse tipo; em outras palavras, enquanto os objetos podem estar relacionados a outros objetos, eventos podem não estar relacionados a outros eventos, agentes não podem ser relacionados a outros agentes, e assim por diante " (PREMIS, 2012, p.3).

Através da Quadro 6 a seguir é possível definir as ações de cada entidade dentro do modelo conceitual em PREMIS

#### Quadro 7: Entidades envolvidas no modelo PREMIS

• <i>Object</i> – ou Objecto Digital, é a unidade discreta de informação no formato digital
• <i>Intellectual entity</i> – ou entidade intelectual, é um conjunto coerente de conteúdo, que pode ser razoavelmente descrito como uma unidade indissociável de informação (e.g. um livro, uma imagem, uma base de dados). Uma entidade intelectual pode conter outras entidades intelectuais no seu interior, por exemplo um livro pode conter uma imagem.
• <i>Agent</i> – ou agente, trata-se de uma pessoa, organização ou aplicação de software que participou num evento de preservação durante o tempo de vida de um Objeto.
• <i>Event</i> – ou evento, é uma ação que envolve pelo menos um objeto ou um agente conhecido pelo repositório.
• <i>Rights</i> -, ou direitos, é um conjunto de um ou mais direitos ou permissões relativas a um Objeto e/ou Agente.

Fonte: (PREMIS, 2012)

O PREMIS consegue agregar ao objeto digital uma espécie de vocabulário capaz de distinguir as principais informações dentro do próprio objeto digital, sendo um metadados de preservação que também descreve o

objeto dentro de suas características arquivísticas necessárias para poder pensar em catalogação ou organização do conhecimento arquivístico, assim como as questões vinculadas com propriedade intelectuais entre outras informações relevantes para agregar veracidade ao arquivo.

A OCLC/RLG constituiu um segundo grupo de trabalho designado PREMIS (PREservation Metadata: Implementation Strategies) com o objetivo de continuar o desenvolvimento deste esquema de metainformação. O grupo de trabalho foi constituído por um comité internacional com mais de trinta especialistas em preservação digital. Deste trabalho resultou o Dicionário de Dados PREMIS, um documento que identifica e descreve um conjunto básico de elementos de metainformação de suporte à preservação digital, bem como um conjunto de recomendações quanto à forma como estes deverão ser utilizados no contexto de um arquivo digital (Ferreira, 2006, p. 55)

Também é relevante afirmar que o projeto PREMIS criado pela comunidade OCLC/RLG vem adquirindo novos recursos e incrementos para agregar mais incrementos informacionais e sendo aprimorado constantemente.

Em 2003 a OCLC (*Online Computer Library Center*) e a RLG (*Research Libraries Group*) estabeleceram o grupo de trabalho designado *PREservation Metadata: Implementation Strategies* (PREMIS). Em Maio de 2005 o mesmo grupo apresentou o seu relatório final, o *Data Dictionary for Preservation Metadata*, que define o esquema de metainformação que passamos a apresentar. O esquema está organizado segundo um modelo simples (BARBEDO, E-pub acessado em 18/11/2014 <https://repositorium.sdum.uminho.pt/bitstream/1822/6330/1/artigo.pdf>)

Esses tipos de comunidades existem em diversas áreas de atuação no caso do PREMIS, criado pela OCLC (*Online Computer Library Center*) e a RLG (*Research Libraries Group*) nos dá um visão de como esses setores da sociedade atuam no aprimoramento e incremento desses recursos tecnológicos aplicados a informação digital.

### 3 METODOLOGIA

Através de um roteiro de investigação que foram analisados os principais requisitos para saber se o RODA atende a comunidade arquivística através das principais recomendações para implementação de um repositório de objetos confiáveis (Ver roteiro Quadro 8).

**Quadro 8: Roteiro metodológico**

<b>OBJETIVOS</b>	<b>ROTEIRO</b>
<b>Analisar as principais características do RODA: pesquisa bibliográfica em publicações sobre o RODA e manuais (objetivo 1)</b>	Pesquisa bibliográfica em publicações sobre o RODA e manuais
<b>Verificar se o RODA oferece recursos para gestão de documentos arquivísticos (objetivo 2)</b>	Verificar se o RODA oferece ferramentas para administração de plano de classificação  Verificar se o RODA oferece descrição multinível usando padrões arquivísticos e de descrição (ISAD-G, NOBRADE, EAD)
<b>Verificar se o RODA oferece recursos para gerenciamento de documento digital confiável (objetivo 3)</b>	Verificar se o RODA atende aos princípios de representação da informação e informação descritiva em pacotes SIP, AIP e DIP no gerenciamento da informação  Verificar se o RODA padrões de metadados estruturais e de preservação digital

<p><b>Admissão – captura de documentos digitais (SIP)</b></p> <p><b>(objetivo 3)</b></p>	<p>Especificar claramente a informação que deve estar associada (o que compõe o SIP); ter mecanismos para autenticar a origem dos documentos admitidos; ter procedimentos para verificar a integridade do <i>SIP</i>; ter o controle físico dos arquivos do <i>SIP</i>; fornecer ao produtor/depositante relatórios do andamento dos procedimentos durante todo o processo de admissão; ter registros de todas as ações e processos administrativos que ocorrem durante o processo</p>
<p><b>Admissão – criação do pacote de arquivamento (geração do AIP)</b></p> <p><b>(objetivo 3)</b></p>	<p>Descrever como os <i>AIPs</i> são construídos a partir dos <i>SIPs</i>; Atribuir aos <i>AIPs</i>, identificadores que sejam únicos, persistentes e visíveis aos gestores e auditores; registrar metadados de preservação associados aos documentos admitidos; verificar a completude e a correção de cada <i>AIP</i> no momento em que é gerado; documentar todas as ações relevantes à preservação dos documentos e que estão relacionadas à criação do <i>AIP</i>.</p>
<p><b>Armazenamento – Preservação do AIP</b></p> <p><b>(objetivo 3)</b></p>	<p>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);</p> <p>Monitoramento constante da integridade dos <i>AIPs</i>; registro de todas as ações de preservação realizadas nos <i>AIPs</i>.</p>

<p><b>Acesso – DIP (objetivo 3)</b></p>	<p>Concessão de acesso a cada <i>AIP</i>, para os usuários autorizados e da forma devida;          Existe processo que gera o DIP;          Garantia da autenticidade dos <i>DIPs</i>, por meio da entrega de cópias autênticas dos originais ou da viabilidade de rastreamento auditável da relação entre o <i>DIP</i></p>
---	---

### REFERENCIAL TEÓRICO USADO NA ELABORAÇÃO DO ROTEIRO

ARQUIVO NACIONAL – *AN Digital: políticas em preservação digital*. Edição Eletrônica Rio de Janeiro, 2012

ARQUIVO NACIONAL – *Dicionário brasileiro de terminologia arquivística*. Rio de Janeiro, Arquivo Nacional, 2005.

SHELLEMBERG, T.R. *Arquivos modernos: princípios e técnicas*. FGV, SP, 2004.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ) RESOLUÇÃO Nº 39, DE 29 DE ABRIL DE 2014 – *Diretrizes para a implementação de repositórios digitais confiáveis de documentos arquivísticos*, Rio de Janeiro, 2014.

THE SOCIETY OF AMERICAN ARCHIVISTS. *Encoded Archival Description Tag Library (EAD): Version 2002*, Chicago, 2002

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). *NOBRADE – Norma brasileira de descrição arquivística*. Rio de Janeiro, 2006.

<http://www.loc.gov/standards/mets/METSOverview.v2.html>

CCSDS. Recommendation for Space Data System Practices, MAGENTA BOOK, REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS): RECOMMENDED PRACTICE CCSDS 650.0-M-2, June 2012

CONSELHO NACIONAL DE ARQUIVOS. *e-ARQ Brasil: Modelo de Requisitos*

para Sistemas Informatizados de Gestão Arquivística de Documentos Adotada pelo Conselho Nacional de Arquivos em dezembro de 2009. Rio de Janeiro, 2009.

FERREIRA, Miguel. *Introdução a preservação digital: conceitos, estratégias e actuais consensos*. Edição Eletrônica, Portugal, 2006.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). *Understanding Metadata*, USA, 2004. National Information Standards Organization (NISO). *Understanding*.

#### **FONTE PARA COLETA DE DADOS E ANÁLISE DO SOFTWARE**

BARBEDO, Francisco, CORUJO, Luís; CASTRO, Faria; RAMALHO; José Carlos; FERREIRA Miguel. *RODA: Repositório de Objectos Digitais Autênticos, Portugal*.

CASTRO, R., et al. Meet RODA, a full-fledged digital repository for long-term preservation. Communication at International Conference on Preservation of Digital Objects (iPres), Singapore, 2011.

FERREIRA, M., et al. *SCAPE: Report on compliance validation*. SCAPE project, May 2014. *Metadata*, USA, 2004.

PITTI, Daniel V. *Encoded Archival Description An Introduction and Overview*. In: D-Lib Magazine, V. 5 N. 11, 1999,. Institute for Advanced Technology in the Humanities, University of Virginia.

SAYÃO, Luis Fernando. *Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos*. Ponto de Acesso, Salvador, V.4, n.3, p. 68-94, dez 2010

THOMAZ, Katia P. ; SOARES Antonio José. A preservação digital e o modelo de referência Open Archival Information System (OAIS). In: *DataGramaZero* –

*Revista de Ciência da Informação – v.5 n.1 fev. 04.*

FARIA, Luís; CASTRO, Rui; BARBEDO, Francisco; HENRIQUES, Cecília;  
CORUJO, Luís; FERREIRA, Miguel; RAMALHO, José Carlos. RODA  
*Repository of Authentic Digital Objects*, 2009 Atlanta, Georgia, USA

Fonte: Elaborado pelo Autor.

Com base nesse roteiro que foi possível comparar os dados e, relacioná-los com os estudos produzidos sobre repositórios arquivísticos de objetos confiáveis, assim como compreender as principais normas arquivísticas aplicadas a ambientes digitais. Traçando objetivos para checar os principais parâmetros relacionados a implementação desses ambientes digitais através da plataforma RODA.

## 4 ANÁLISE DE DADOS

### 4.1 Principais características do RODA

O RODA pode ser considerado um projeto extenso que ainda está em desenvolvimento, que oferece os recursos e implementos necessários para uma gestão arquivística em ambientes digitais.

O projeto contempla ainda alguns objetivos secundários, nomeadamente; a criação ou identificação de modelos viáveis de financiamento para suportar o Arquivo Digital; a identificação e selecção dos esquemas de metainformação e a definição de uma teoria de classificações e propriedades significativas para cada uma das classes de objetos consideradas (<http://www.apbad.pt/Downloads/congresso9/COM37.pdf>.

Um repositório digital com as características do RODA tem necessariamente de ser suportado por um conjunto de esquemas de metainformação capazes de assegurar a realização de actividades elementares como a gestão de objectos digitais, facilitar a sua localização ou garantir a conservação do valor probatório (<http://www.apbad.pt/Downloads/congresso9/COM37.pdf> acessado em 29/07/2014).

O RODA é um projeto que se propõe a criar as possibilidades de preservação da informação digital, através de um software que é capaz de agregar em si as principais referências em gestão e preservação para fundos arquivísticos. Foi possível analisar os principais requerimentos em relação a estrutura e os elementos que compõem o RODA.

É neste contexto que se desenvolve o projeto RODA (Repositório de Objetos Digitais Autênticos), um projecto que visa desenvolver e promover uma solução tecnológica, ultimada na construção de um protótipo de repositório digital capaz de incorporar, descrever e dar acesso a todo o tipo de informação digital produzida no contexto da Administração Pública. (<http://www.apbad.pt/Downloads/congresso9/COM37.pdf> acessado em 29/07/2014)

#### 4.1.1 O RODA

O RODA é um software em código aberto que congrega as principais recomendações brasileiras e internacionais referente a implementação e gestão de repositórios digitais de objetos autênticos. Por se tratar de um ambiente digital que atende as premissas da comunidade arquivística que é importante ressaltar como que o RODA funciona e suas principais características.

RODA (Repositório de Objetos Digitais Autênticos), um projeto que visa desenvolver e promover uma solução tecnológica, ultimada na construção de um protótipo de repositório digital capaz de incorporar, descrever e dar acesso a todo o tipo de informação digital produzida no contexto da Administração Pública. Procura-se desta forma iniciar um processo sustentado [...]  
(<http://www.apbad.pt/Downloads/congresso9/COM37.pdf> acessado em 29/07/2014).

Para que o RODA seja um protótipo de repositório de objetos autênticos é preciso que atenda as normas e exigências arquivísticas em relação a gestão e preservação digital em repositórios digitais.

O protótipo de arquivo digital será planeado na perspectiva de obter um sistema capaz de assegurar todas as funcionalidades de um arquivo digital constantes da norma OAIS (Open Archival Information System), nomeadamente, a integração (ingestão), a gestão e a disseminação de informação de arquivo. A limitação deste protótipo residirá na restrição de formatos a integrar <http://www.apbad.pt/Downloads/congresso9/COM37.pdf>.

Em relação ao RODA é possível afirmar também que é:

Um repositório digital com as características do RODA tem necessariamente de ser suportado por um conjunto de esquemas de metainformação capazes de assegurar a realização de atividades elementares como a gestão de objetos digitais, facilitar a sua localização ou garantir a conservação do valor probatório (<http://www.apbad.pt/Downloads/congresso9/COM37.pdf> acessado em 29/07/2014).

Contudo, é possível afirmar que essas características do *software* são eficazes em relação a preservação em ambientes digitais, podendo agregar uma série de ferramentas administrativas capazes de gerenciar uma coleção ou um fundo arquivístico.

#### 4.1.2.1 O RODA e sua arquitetura

O projeto RODA foi constituído através de códigos abertos e do Fedora Commons<sup>7</sup>. Logo é um capacitador de ingestão de informação, gestão de

---

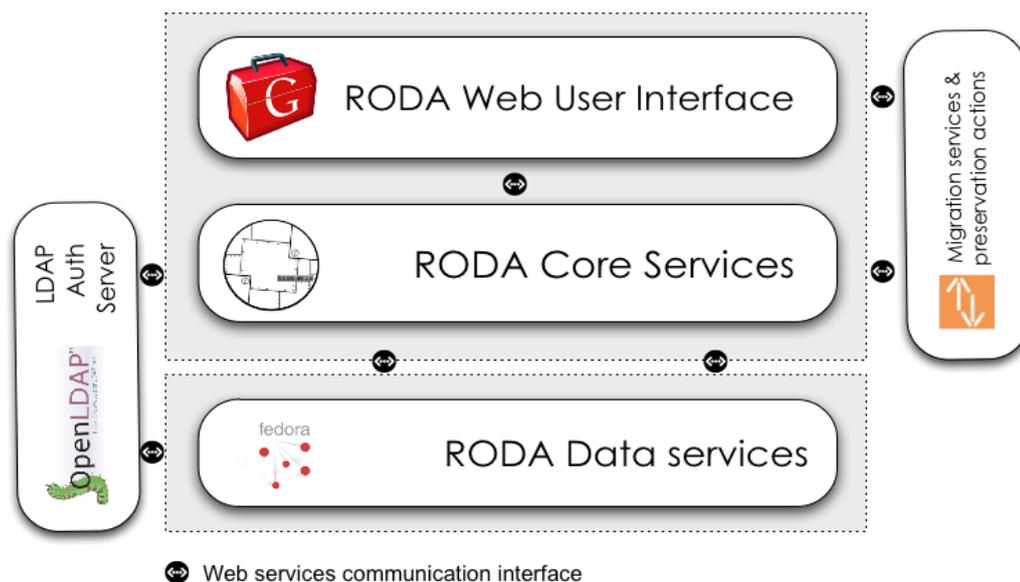
<sup>7</sup> Fedora (ou flexível Objeto Extensível Digital Repository Architecture) é uma arquitetura de gerenciamento de ativos digitais na qual repositórios institucionais, arquivos digitais e sistemas de bibliotecas digitais podem ser construídos. Fedora é a arquitetura subjacente para um repositório digital, e não é uma gestão, indexação, descoberta e aplicação completa entregue. É uma arquitetura modular construído sobre o princípio de que a interoperabilidade e

dados e armazenamento de arquivos. Uma espécie de plataforma com interface pré-definida capaz de ser otimizada conforme as especificidades de cada necessidade.

Sendo capaz de oferecer índices de busca, controle de ingestão de informação e mapeamento e rastreamento dos elementos informacionais. Metadados descritivos capazes de identificar e classificar conforme a orientação de cada objeto nato-digital.

A Figura 4 a seguir nos dá um panorama dos principais aplicativos para o seu funcionamento:

**Figura 4: Organograma básico dos principais aplicativos e interações do RODA**



Fonte: <http://www.roda-community.org/features/>

RODA Core Services são responsáveis pela realização de tarefas mais complexas, como lidar com o fluxo de trabalho de ingerir, consultando o repositório de formas avançadas e no exercício de funções administrativas no repositório. Eles implementam a lógica que suporta

---

extensibilidade são melhor alcançados pela integração de dados, interfaces e mecanismos (ou seja, programas executáveis) [http://en.wikipedia.org/wiki/Fedora\\_Commons](http://en.wikipedia.org/wiki/Fedora_Commons). Acessado em 19/11/2014.

RODA. Este componente é baseado em serviços Web seguros, de modo que qualquer aplicação cliente pode interagir com RODA chamando métodos da Web disponíveis. Isso permite integração de sistemas já existentes na instituição cliente, que se espera para depositar ou consultar a informação disponível no repositório. Além disso, esta AIP permite a criação de novas ferramentas de ingerir que criam automaticamente SIP e depositá-los em RODA. (<http://www.roda-community.org/features/> acessado em 21/10/2014)

Para compreendermos melhor, é possível destacar que o RODA é um pacote de programas e de serviços que através de um conjunto de aplicativos é capaz de oferecer a construção de um repositório de objetos arquivísticos confiáveis

Finalmente, RODA utiliza um componente que são responsáveis pela realização de serviços de migração e ações de preservação. Serviços de migração pode ser executado em diferentes sistemas operacionais. Como um sistema de distribuição do repositório pode ser implementado de muitas maneiras diferentes. Implementações bem-sucedidas de RODA têm sido gerados pela execução de todos os componentes do sistema em um único servidor, bem como pela execução de cada componente em uma máquina diferente. (<http://www.roda-community.org/features/> acessado em 22/10/2014)

#### 4.1.2.2 *SIP (Submission Information Packages) dentro do RODA*

Como o RODA foi desenvolvido com base no modelo de referência OAIS, o seu processo de admissão de informação ocorre através da elaboração do pacote de submissão. A admissão consiste na submissão de SIPs através de fluxos de trabalho. Esses fluxos de arquivos para serem submetidos e gerenciados sem que cause nenhum dano ao arquivo devem ser encapsulados em pacotes denominados de SIP.

O SIP é composto de uma ou mais representações digitais e todos os metadados associados, embalados no interior de um envelope. Produtores poderão tirar proveito de um pequeno aplicativo chamado RODA-nos que lhes permite criar esses pacotes. A estrutura de um SIP suportado pela roda é representada na seguinte figura. A RODA SIP é basicamente um arquivo ZIP comprimido que contém um envelope METS, o conjunto de arquivos que compõem as representações e uma série de registros de metadados. Dentro da SIP deve haver pelo menos um registro de metadados descritivos em formato EAD-Component (<http://www.roda-community.org/features/> acessado em 16/11/2014).

Através disso é possível visualizar esse procedimento dentro repositório RODA através da ingestão ou admissão da informação como nos mostra a Figura 5.

**Figura 5 Interface de admissão de pacotes de arquivos SIP (RODA)<sup>8</sup>**

File	Submission date	Current state	%	Producer
TEST_sf1_c1_sr1_NG.sip	2009-05-11 14:41:47	accepted	100%	ifaria
TEST_sf1_c1_sr1_fox.sip	2009-05-07 12:19:36	accepted	100%	ifaria
TEST_LFARIA_MP48298.sip	2009-03-20 17:50:39	accepted	100%	ifaria
TEST_LFARIA_PNG2.sip	2009-03-20 15:06:14	accepted	100%	ifaria
TEST_LFARIA_PNG492.sip	2009-03-20 15:05:28	accepted	100%	ifaria
TEST_LFARIA_XPM.sip	2009-03-20 11:40:25	accepted	100%	ifaria
TEST_LFARIA_WAV.sip	2009-03-20 11:40:24	accepted	100%	ifaria
TEST_LFARIA_TXT.sip	2009-03-20 11:40:16	accepted	100%	ifaria
TEST_LFARIA_TIFF.sip	2009-03-20 11:40:15	accepted	100%	ifaria
TEST_LFARIA_TGA.sip	2009-03-20 11:40:15	accepted	100%	ifaria
TEST_LFARIA_RTF.sip	2009-03-20 11:40:14	accepted	100%	ifaria
TEST_LFARIA_PDF.sip	2009-03-20 11:40:13	accepted	100%	ifaria
TEST_LFARIA_OGG.sip	2009-03-20 11:40:06	accepted	100%	ifaria
TEST_LFARIA_ODT.sip	2009-03-20 11:40:05	accepted	100%	ifaria
TEST_LFARIA_MP3.sip	2009-03-20 11:40:01	accepted	100%	ifaria
TEST_LFARIA_JPEG.sip	2009-03-20 11:39:28	accepted	100%	ifaria
TEST_LFARIA_ICO.sip	2009-03-20 11:39:25	accepted	100%	ifaria
TEST_LFARIA_GIF.sip	2009-03-20 11:39:24	accepted	100%	ifaria

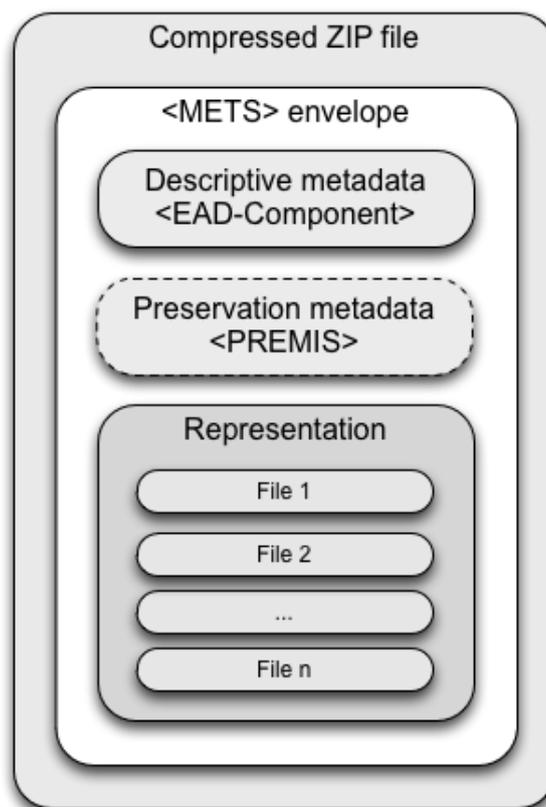
Fonte: <http://www.roda-community.org/features/>

A visualização da figura acima nos dá um panorama de como o RODA, através da sua representatividade gráfica ou interface, ingere pacotes de informação em SIP (Figura 6)

Novas inscrições para o repositório vir na forma de pacotes de Informação de Envio (SIP). Quando o processo de ingerir termina, SIP são transformados em pacotes de informação arquivística (AIP), ou seja, os pacotes atuais que serão mantidas no repositório. Associada à AIP é o, metadados técnicos e preservação estrutural, uma vez que são essenciais para a realização de atividades de preservação (<http://www.roda-community.org/features/> acessado em 22/09/2014).

<sup>8</sup> Ingerir: Processo de adição de objetos ao sistema de armazenamento de um repositório de preservação. No Contexto da OAIS. Ingerir inclui serviços e funções que aceitem pacotes de informação de envio (SIP) de Produtores, e Transforma-los em mais pacotes de Informação de Arquivo (AIP) Para retenção de longo prazo. (PREMIS, 2012).

**Figura 6: Estrutura conceitual de submissão de pacotes de informação**



Fonte: <http://www.roda-community.org/features/> acessado em 10/11/2014

Para haver um empacotamento da informação digital são necessárias diversas ações no intuito de checar a informação e sua autenticidade a partir da ingestão do arquivo no repositório. Esse modelo conceitual ilustra o uso dos principais metadados na inserção ou admissão do objeto digital no repositório.

#### 4.1.2.3 Submissão de um Pacote de Informação (SIP)

Antes do SIP ser totalmente incorporada no repositório são submetidos a uma série de testes para avaliar a sua integridade e conformidade com a política de ingerir. Após descompactar o SIP, o processo de validação realiza o conjunto de tarefas descritas a seguir no Quadro 8:

#### **Quadro 9: Empacotamento SIP**

1. Verificação de vírus: SIP são verificados em busca de vírus. O antivírus está sendo usado sob a superfície para executar esta tarefa.

2. Envelope verificação de sintaxe: Verifique se o envelope METS está bem formado.
3. SIP verificação da integralidade: Verifique se todos os arquivos mencionados no envelope METS existem dentro da SIP.
4. verificação de integridade de arquivos: arquivos incluídos no SIP são acompanhados por uma seqüência de soma de verificação. Esta informação é usada para verificar se algum dos arquivos sofreram a corrupção de algum tipo.
5. Checagem dos metadados descritivos: verificação dos componentes EAD estão inclusos na SIP e se sua sintaxe está correta.
6. Preservação e checagem dos metadados: verificação dos registros PREMIS foram incluídos na SIP e se sua sintaxe está correta.
7. Checagem da representação: Verificação de pelo menos uma representação existe dentro da SIP.
8. Checagem de representação específica: Dependendo do tipo de representação no SIP, uma série de testes mais específicos são realizados para verificar se a representação é completa e formato compatível com a política de ingerir no lugar.
9. Normalização: Representações cujo formato não é compatível com os formatos de preservação definidas pela política de preservação são automaticamente convertidos para o formato correto. A representação original seja mantido pelo repositório por razões diplomáticas.

Fonte: (<http://www.roda community.org/features/> acessado em 20/10/2014)

O SIP é gerado a partir de um procedimento de admissão digital, que é responsável pelas verificações de premissas e parâmetros que são rotinas que estão no sentido da preservação da informação, entre elas agregar todas essas e ações em um único pacote.

Nesse contexto temos que usar um sentido que chama-se a abstração, pois estamos tratando de ambientes digitais e multiníveis informacionais na qual obedecem uma sintaxe específica e uma lógica de programação.

As informações submetidas referentes ao objeto digital se inter-relacionam, e as ações executadas na submissão geram outras informações relacionadas ao objeto, que são adicionadas ao pacote, como por exemplo,

versões derivadas do objeto original e metadados referentes a ações de preservação digital. Pacotes SIP serão adicionados de descritores EAD e METS.

O AIP e o DIP serão derivações de um mesmo pacote de informações (SIP) que têm objetivos diferentes dentro do repositório, armazenamento e disseminação, respectivamente. Daí podemos dizer que as questões de compatibilidade, migração, acesso, armazenamento dependem de multiníveis de relações informacionais capazes de agregar toda essa interoperabilidade relacional em um único objeto.

O SIP é composto por uma ou mais representações digitais e todos os metadados associados, embalados no interior de um envelope de METS. Produtores tirarão proveito de um pequeno aplicativo chamado RODA-nos que lhes permite criar esses pacotes. O SIP é um arquivo ZIP comprimido que está contido em um envelope METS, o conjunto de arquivos que compõem as representações e uma série de registros de metadados. Dentro da SIP deve haver pelo menos um registro de metadados descritivos em formato EAD-Component. (CASTRO, et. al., 2011, p.2)

O SIP, AIP e DIP em uma constatação simplificada são elementos da mesma origem que se inter-relacionam dentro do repositório, contudo tem funções diferenciadas, pois são relações internas que conseqüentemente se tornarão disponíveis para o acesso

Após uma validação com sucesso, o SIP é então desmontado e cada um dos seus componentes é integrado no repositório. Após esse procedimento, a SIP se torna um AIP e está pronto para ser divulgada aos potenciais consumidores que têm autorização para acessar essas informações (CASTRO, et. al., 2011, p.2).

**Figura 7 Interface de admissão de pacotes de arquivos SIP transformados em AIP (RODA)**

The screenshot shows the RODA web interface. At the top, there is a navigation menu with items like 'ABOUT RODA', 'SERVICES', 'POLICY AND TECHNICAL DOCUMENTS', 'R&D', 'CONTACTS', 'ABOUT THE REGISTER', and 'HELP'. Below this is a header with the RODA logo and a user profile for 'ADMIN'. The main content area is titled 'LIST' and shows a table of tasks. The table has columns for 'Name', 'Start date', 'Complete', and 'User'. The tasks listed are:

Name	Start date	Complete	User
Ingest/Create AIP	2014-09-01 18:17:46	2014-09-01 18:17:46	admin
Ingest/Check producer authorization	2014-09-01 18:17:32	2014-09-01 18:17:32	admin
Ingest/Normalize format	2014-09-01 18:17:24	2014-09-01 18:17:24	admin
Ingest/Check SIP syntax	2014-09-01 18:17:21	2014-09-01 18:17:21	admin
Ingest/Virus Check	2014-09-01 18:17:20	2014-09-01 18:17:20	admin
Ingest/Unpack SIPs	2014-09-01 18:16:47	2014-09-01 18:16:47	admin
Ingest/Create AIP	2014-09-01 18:12:46	2014-09-01 18:12:46	admin
Ingest/Check producer authorization	2014-09-01 18:12:32	2014-09-01 18:12:32	admin
Ingest/Normalize format	2014-09-01 18:12:24	2014-09-01 18:12:24	admin
Ingest/Check SIP syntax	2014-09-01 18:12:21	2014-09-01 18:12:21	admin
Ingest/Virus Check	2014-09-01 18:12:20	2014-09-01 18:12:20	admin
Ingest/Unpack SIPs	2014-09-01 18:11:47	2014-09-01 18:11:47	admin
Ingest/Create AIP	2014-09-01 18:07:46	2014-09-01 18:07:46	admin
Ingest/Check producer authorization	2014-09-01 18:07:32	2014-09-01 18:07:32	admin
Ingest/Normalize format	2014-09-01 18:07:24	2014-09-01 18:07:24	admin
Ingest/Check SIP syntax	2014-09-01 18:07:21	2014-09-01 18:07:21	admin
Ingest/Virus Check	2014-09-01 18:07:20	2014-09-01 18:07:20	admin
Ingest/Unpack SIPs	2014-09-01 18:06:47	2014-09-01 18:06:47	admin
Ingest/Create AIP	2014-09-01 18:02:46	2014-09-01 18:02:46	admin
Ingest/Check producer authorization	2014-09-01 18:02:32	2014-09-01 18:02:32	admin
Ingest/Normalize format	2014-09-01 18:02:24	2014-09-01 18:02:24	admin
Ingest/Check SIP syntax	2014-09-01 18:02:21	2014-09-01 18:02:21	admin

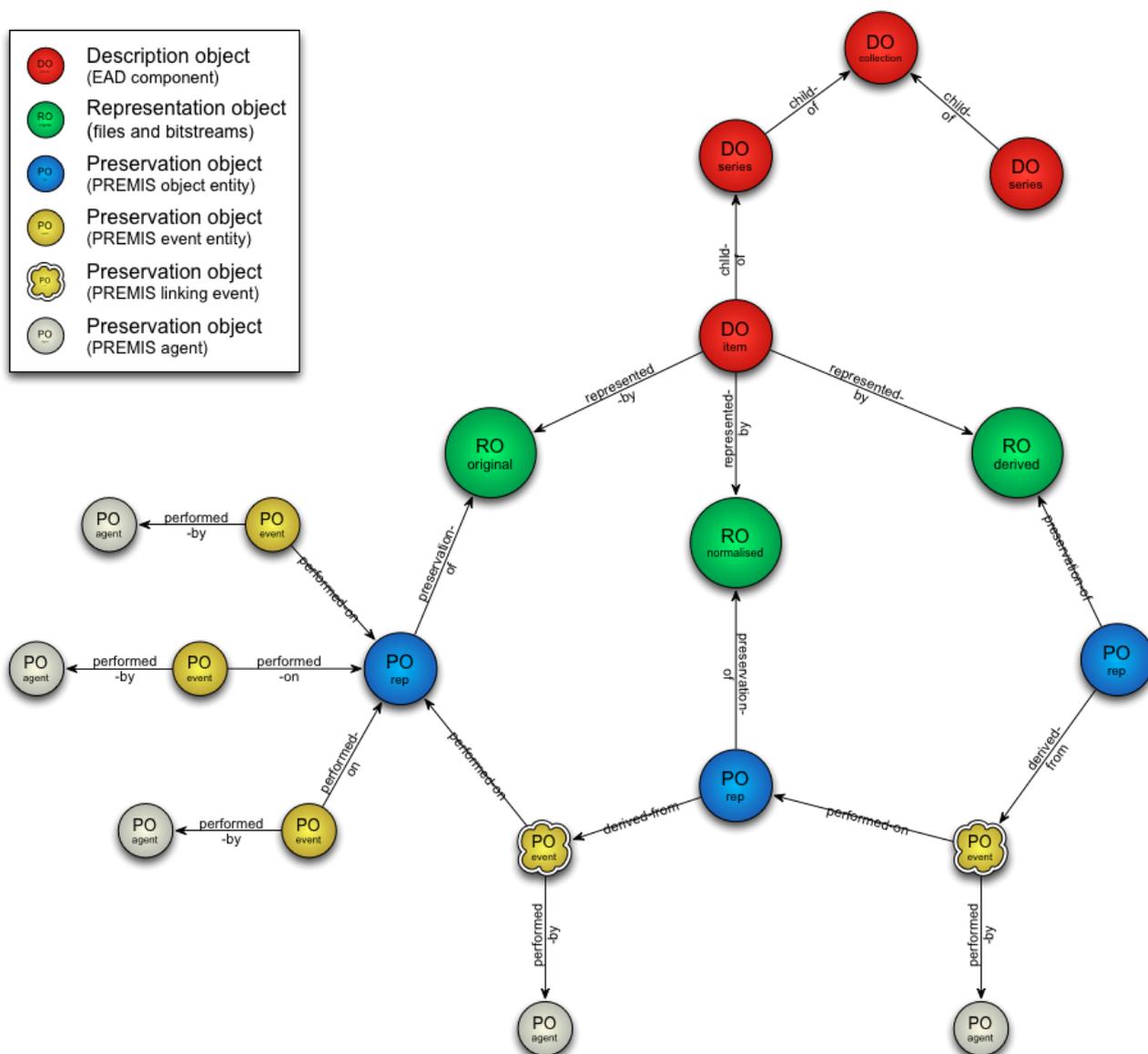
Fonte: (<http://www.roda community.org/features/> acessado em 21/10/2014)

Nessa interface (Figura 7) é possível visualizar a entidade que está ingerindo informação e as ações realizadas na transformação dos pacotes dentro do ambiente RODA.

É através do modelo molecular (fig) do RODA que podemos entender melhor e visualizar as relações e prioridades em um sistema de ingestão de arquivos confiáveis. Nesse sentido essas relações compõem o sistema de informações voltadas para objetos confiáveis. Essas camadas informacionais estão ligadas ao objeto e são parte dos elementos que descrevem o próprio objeto. Assim podendo existir muitos níveis de informação sobre um determinado objeto. Ampliando as possibilidades na hora da pesquisa e busca dos elementos que compõem os objetos e cria as condições ideais para armazenamento e migração de *software*. Por exemplo, a figura 8 ilustra os três principais níveis informacionais são eles, cujas ações são geradoras são apresentadas no Quadro 9:

- descrição do objeto DO (representados pelos nodos em vermelho), envolvendo os níveis da descrição arquivística
- representação do objeto (RO) (caracterizado pelos nodos em verde que representa respectivamente o objeto original, e suas versões normalizada e derivada)
- Descrições de Preservação Digital, envolvendo o objeto em azul, aos eventos realizados sobre o objeto em amarelo e aos agentes que executaram esses eventos em cinza.

Figura 8: Modelo molecular do RODA (Quadro 10)



Fonte: (<http://www.roda-community.org/>).

### Quadro 10: Principais ações no modelo molecular do RODA

Modelo de conteúdo de Roda é orientada para PREMIS. Cada entidade intelectual é descrito por um registo de metadados EAD-componente (DO nós na figura). Esses registros são organizados hierarquicamente, a fim de constituir uma descrição completa de arquivo, mas são mantidos em separado dentro do modelo de conteúdo Fedora Commons. As relações entre EAD-componentes são criados utilizando a própria mecanismo que liga RDF<sup>9</sup> do Fedora

Além disso, cada registro folha (ou seja, um arquivo ou um item) está ligada a um objeto de representação (RO nós na figura), ou seja, um objeto fedora que incorpora todos os arquivos e fluxos de bits que compõem a representação digital. Por fim, cada um desses objetos estão ligados entre si por um conjunto de entidades PREMIS que mantêm informações sobre a proveniência do objeto digital e histórico de eventos (nós PO).

Cada evento preservação que ocorre dentro do repositório é registrado como um novo nó preservação evento (ou seja, os nós de eventos PO na figura). Eventos especiais, como migração de formato, estabelecer relações entre dois nós de preservação-representação. Estes são chamados de ligar eventos. Cada evento preservação é executado por um agente, quer seja um usuário do sistema ou um aplicativo de software acionado automaticamente. O agente que disparou o evento é gravado em nós do agente PO.

Essas relações que o objeto digital é submetido criam as condições necessárias para uma uma gestão com fundamentos arquivísticos em relação

<sup>9</sup> RDF (Resource Description Framework) é um modelo padrão para o intercâmbio de dados na Web. RDF tem características que facilitam a fusão de dados, mesmo se os esquemas diferem subjacentemente, e especificamente suporta a evolução de esquemas ao longo do tempo sem a necessidade de todos os consumidores de dados a ser mudado. RDF estende-se a estrutura de ligação da Web para usar URIs para nomear a relação entre as coisas, bem como as duas extremidades da ligação (isto é normalmente referido como um "triplo"). Usando esse modelo simples, que permite que os dados estruturados e semi-estruturados para ser misturado, exposta, e compartilhados entre aplicações diferentes. Esta estrutura que liga forma, um gráfico marcado dirigido, onde as arestas representam o elo chamado entre dois recursos, representados pelos nós do grafo. Este ponto de vista gráfico é o modelo mental mais fácil possível para RDF e é frequentemente usado em explicações visuais fáceis de entender (<http://www.w3.org/RDF/>).

aos fundos e coleções disponibilizados na *web* ou em rede, pois congrega os procedimentos fundamentais na identificação do objeto digital. O tornando-o confiável e capaz de gerar autenticidade.

## 4.2 Gestão de documentos arquivísticos no RODA

Esta seção avalia recursos que o RODA pode oferecer na construção de um repositório de documentos arquivísticos confiáveis. Tal investigação baseia-se na recomendação brasileiras para repositórios arquivísticos digitais confiáveis do CONARQ. As repostas em relação as confluências entre as recomendações do CONARQ para implementação de repositórios arquivísticos de objetos confiáveis se são atendidas pelo software RODA? Se encontram disponíveis como repostas em quadros temáticos.

Questões a investigar:

- *Verificar se o RODA oferece ferramentas para administração de plano de classificação.*
- *Verificar se o RODA oferece descrição multinível usando padrões arquivísticos e de descrição (ISAD(G), NOBRADE, EAD)*

Resultados encontrados com base na literatura:

Conforme o autor um plano de classificação pode ser criado através de "um aplicativo de desktop multiplataforma que permite que os produtores criem, descrevam, gerencie um plano de classificação que depois pode ser baixado a partir do servidor" (FARIA, et all, 2009, p. 20).
--

O RODA é um repositório digital que incorpora toda a funcionalidade exigida pelo modelo de referência OAIS.
---

O RODA é capaz de incorporar, gerir e dar acesso a vários tipos de material digital produzido no âmbito da atividade de grandes empresas ou organismos públicos. O RODA é baseado em tecnologias <i>open-source</i> e é suportado pelas normas OAIS, EAD, METS e PREMIS.
--

Para além do disposto, o RODA implementa ainda um fluxo de trabalho de
--

ingestão configurável, que não só valida os pacotes depositados, como também permite a avaliação e seleção do material por parte dos profissionais de arquivo.

O módulo de administração permite aos gestores do repositório editar metadados descritivos, lançar ações de preservação (e.g. verificações de integridade, migração de formatos, entre outros), controlar os acessos por parte de utilizadores, consultar estatísticas, logs de acesso, entre muitas outras opções.

Fonte: <http://www.keep.pt/produtos/roda/>

Cabe ressaltar que o metadados em EAD são baseados na norma de descrição internacional ISAD(G) e que a NOBRADE (Norma Brasileira para Descrição) é uma adaptação da ISAD(G) para o Brasil.

O RODA é compatível com normativos como o EAD (metadados descritivos), PREMIS (metadados de preservação), METS (metainformação estrutural) e várias normas ao nível dos metadados técnicos (e.g. NISO Z39.87 no contexto das imagens digitais) (<http://www.keep.pt/produtos/roda/>)

### 4.3 Gerenciamento de documento digital confiável no RODA

#### *Questões a investigar*

– *Verificar se o RODA atende aos princípios de representação da informação e informação descritiva em pacotes SIP, AIP e DIP no gerenciamento da informação*

Resultados encontrados com base na literatura:

Novas inscrições para o repositório na forma de Pacotes de informação (SIP). quando o processo de ingerir termina, SIP são transformados em pacotes de informação de arquivo (AIP), ou seja, os pacotes atuais que serão mantidas no repositório. O AIP é estrutural, técnico e metadados de preservação, tal como eles são essência para a realização de atividades de preservação (CASTRO, 2011)

O RODA tem que permitir que a informação disponibilizada ao utilizador seja essencialmente descritiva e que o ponto de referência para recuperar o AIP e produzir o DIP seja o identificador. Para o gestor a informação obtida tem de permitir identificar, localizar e recuperar todas as componentes que eventualmente constituam o AIP. (BARBEDO, et al, 2009, p.8).

*Questões a investigar:*

– *Verificar se o RODA usa padrões de metadados estruturais e de preservação digital*

Resultados encontrados com base na literatura:

Um repositório digital projetado especialmente para arquivos, com as seguintes características principais:
Preservação a longo prazo e autenticidade
Com base nos padrões (OAIS, Mets, EAD, Premis, etc.)
Confiável Repositórios de Auditoria e Certificação
Interface fácil de navegar
Para os arquivistas, para os produtores, para os consumidores
Em código aberto
RODA é um repositório digital que tem como objetivo principal a preservação e autenticidade do objeto digital. Para isso foi baseada em padrões e especificações internacionais para se tornar um repositório de confiança, por ter alto padrão de segurança e gerenciamento. Interface do usuário também foi uma questão importante, pois foi projetado para arquivistas, para os produtores e para os consumidores.

Fonte: (BARBEDO, et al, 2009, p.3)

#### **4.4 Captura de documentos digitais na admissão no RODA**

Especificar claramente a informação que devem estar associadas (o que compõe o SIP); ter mecanismos para autenticar a origem dos documentos admitidos; ter procedimentos para verificar a integridade do SIP; ter o controle

físico dos arquivos do SIP; fornecer ao produtor/depositante relatórios do andamento dos procedimentos durante todo o processo de admissão; ter registros de todas as ações e processos administrativos que ocorrem durante o processo

Resultados encontrados com base na literatura:

Em seguida, o produtor pode criar um SIP com um arquivo ou item dentro dele. Um painel permite para a inserção dos metadados descritiva e sua validação. Uma representação pode ser criada, escolhendo a classe do objeto. Para cada classe de objeto existe um painel criador diferente. Esta é a criador de imagens de rastreamento. A ordem e estrutura das imagens pode ser definido e uma pré visualização e metadados são apresentados. Depois de o objeto criado e descrito, ele pode ser enviado para o servidor RODA: *on line* pela HTTP ou FTP, ou por correio postal (FARIA, et all, 2009, p. 20).

#### 4.5 Criação do pacote de arquivamento na admissão do RODA

A criação do pacote de arquivamento, segundo o CONARQ envolve descrever como os *AIPs* são construídos a partir dos *SIPs*; Atribuir aos *AIPs*, identificadores que sejam únicos, persistentes e visíveis aos gestores e auditores; registrar metadados de preservação associados aos documentos admitidos; verificar a completude e a correção de cada *AIP* no momento em que é gerado; documentar todas as ações relevantes à preservação dos documentos e que estão relacionadas à criação do *AIP*.

Resultados encontrados com base na literatura:

RODA expõe todas as suas funcionalidades via *Web Services*<sup>10</sup>. Bibliotecas

---

<sup>10</sup> *Web service* é uma solução utilizada na integração de sistemas e na comunicação entre aplicações diferentes. Com esta tecnologia é possível que novas aplicações possam interagir com aquelas que já existem e que sistemas desenvolvidos em plataformas diferentes sejam compatíveis. Os *Web services* são componentes que permitem às aplicações enviar e

Java convenientes estão disponíveis para permitir que os desenvolvedores interagem com RODA por meio de suas APIs Core. Existem várias ferramentas para criar e manipular os SIP e submetê-los ao fluxo de trabalho de ingerir informação através do RODA .

Após uma validação com sucesso, o SIP é então desmontado e cada um dos seus componentes é integrado no repositório. Após esse procedimento, a SIP se torna um AIP e está pronto para ser divulgada aos potenciais consumidores que têm autorização para acessar essas informações.  
<http://repositorium.sdum.uminho.pt/bitstream/1822/18299/1/iPres%202011%20-%20RODA%20short%20paper.pdf>

Fonte: <http://www.roda-community.org/>

#### 4.6 Preservação do AIP no RODA

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);

Monitoramento constante da integridade dos *AIPs*; registro de todas as ações de preservação realizadas nos *AIPs*.

Resultados encontrados com base na literatura:

A migração, data de refrescamento prevista, etc; Notificações periódicas. O RODA deve assegurar que na sequência de um evento será adicionada aos objectos (AIP; Infra estrutura) alvo desse evento; Atualização da PREMIS relativa a eventos. Após um evento de atualização (migração), os AIP resultantes deverão ser confirmados quanto à sua integridade e inteligibilidade sendo para isso sujeitos a processo de validação; Efetuar uma validação no

receber dados em formato XML. Cada aplicação pode ter a sua própria "linguagem", que é traduzida para uma linguagem universal, o formato *XML*. Para as empresas, os *Web services* podem trazer agilidade para os processos e eficiência na comunicação entre cadeias de produção ou de logística. Toda e qualquer comunicação entre sistemas passa a ser dinâmica e principalmente segura, pois não há intervenção humana ([http://pt.wikipedia.org/wiki/Web\\_service](http://pt.wikipedia.org/wiki/Web_service)).

final de qualquer ação de alterar os AIPs (Fluxo de trabalho supracitado) (BARBEDO, *et. al*, 2009).

#### 4.7 Acessos no RODA

Concessão de acesso a cada *AIP*, para os usuários autorizados e da forma devida; Existe processo que gera o *DIP*; Garantia da autenticidade dos *DIPs*, por meio da entrega de cópias autênticas dos originais ou da viabilidade de rastreamento auditável das relações entre o *DIP*

Resultados encontrados com base na literatura:

O RODA tem de ter a capacidade de assegurar que são recuperadas apenas os componentes de um *AIP* necessárias para satisfazer o pedido do utilizador, sem acrescentar ou diminuir informação;

Permitir transformações de formatos entre os *AIP* e os *DIP* o RODA tem que permitir que a informação disponibilizada ao utilizador seja essencialmente descritiva e que o ponto de referência para recuperar o *AIP* e produzir o *DIP* seja o identificador. Para o gestor a informação obtida tem de permitir identificar, localizar e recuperar os componentes que eventualmente constituam o *AIP*. O utilizador(consumidor) pesquisa apenas no EAD e usa o identificador único para recuperar o objecto A certificação do *DIP* tem de respeitar os métodos (Assinatura digital).

A certificação do *DIP* tem de respeitar os métodos legalmente reconhecidos (...) (Assinatura digital); *Assinaturas digitais nos DIPs*

Fonte (BARBEDO, *et. al*, 2009)

As questões relacionadas a certificação, normalmente são designadas para entidades capazes de administrar as demandas causadas pelos avanços tecnológicos. A geração de um *DIP* dentro de um repositório de objetos arquivísticos conviáveis seria a porta de entrada para o usuário. Através do *DIP* é que é possível manejar o objeto digital pronto que será acessado. Com todas ações e procedimentos necessários para que exista confiabilidade. Daí a

importância das entidades governamentais adaptarem seus meios de certificação digital dando acesso e garantindo a informação, de modo que possa gerar valor probatório.

O processo de admissão da informação deve estar em acordo com o modelo OAIS. Assim como novos incrementos aos metadados devem corresponder as necessidades de cada comunidade. O interesse em assegurar um repositório digital deve estar nas políticas de preservação implantadas por setores que desejam gerar autenticidade da informação. Assegurar a proveniência do documento digital, assim como sua disseminação dentro de um contexto de administração de fundos arquivísticos. Sendo eles arquivos correntes, intermediários ou permanentes. Respeitando a idade do documento e sua função dentro de um fundo arquivístico.

Esses padrões gerados para garantir o controle da informação dentro de um repositório criam a inteligibilidade e segurança para poder acrescentar ao objeto digital a confiabilidade necessessária a partir da sua criação, admissão, administração e acesso dentro de um repositório de objetos confiáveis.

## 5 CONSIDERAÇÕES FINAIS

Foi possível através desse trabalho criar uma perspectiva em relação a preocupação com a preservação de objetos digitais. Num sentido amplo e formal capaz de trazer à tona essa necessidade emergente em relação aos repositórios digitais. Sendo possível delimitar as diferenças fundamentais entre os repositórios. Assim, como adentrar nos principais parâmetros relacionados a implementação de repositórios digitais arquivísticos confiáveis. Mais especificamente o *software* RODA. Criado para atender as necessidades emergentes em relação as novas tecnologias e novas formas de acesso a fundos e coleções arquivísticas.

O consumidor é capaz de navegar sobre as coleções disponíveis para visualização ou *download* representações digitais guardados no repositório. Dependendo do tipo do objeto digital, são utilizados diferentes telespectadores ou divulgadores. Por exemplo, documentos de texto são entregues aos consumidores, sem recorrer a qualquer tipo de artefato particulares. Eles são entregues em formato PDF<sup>11</sup>, por isso, o consumidor deve usar o seu aplicativo de visualização de PDF favorito. Os documentos compostos por várias imagens (como obras digitalizadas), por outro lado, são exibidos em aplicações especiais de visualização Web que permitem aos consumidores navegar pelas páginas da representação (<http://www.roda-community.org/>).

---

<sup>11</sup> O PDF (*Portable Document Format*) é um formato de arquivo, desenvolvido pela Adobe Systems em 1993, para representar documentos de maneira independente do aplicativo, do *hardware* e do sistema operacional usados para criá-los. Um arquivo PDF pode descrever documentos que contenham texto, gráficos e imagens num formato independente de dispositivo e resolução. O PDF é um padrão aberto, e qualquer pessoa pode escrever aplicativos que leiam ou escrevam neste padrão. Há aplicativos gratuitos para Microsoft Windows, Apple Macintosh e Linux, alguns deles distribuídos pela própria Adobe e há diversos aplicativos sob licenças livres. PDF pode ser traduzido para português como *formato de documento portátil* ([http://pt.wikipedia.org/wiki/Portable\\_document\\_format](http://pt.wikipedia.org/wiki/Portable_document_format)).

Conjuntamente foi possível compreender o significado e a importância de normas descritivas (EAD e ISAD(g)) e modelos para gestão em repositórios digitais como é o caso do OAIS. Maneiras de procedimentos para classificar os metadados relacionados ao objeto digital. Reconhecendo os modelos de metadados METS e PREMIS. Como se processam as relações dentro do repositório e como que o RODA congrega essas entidades. Também foi possível perceber as relações entre essas normas e resoluções em torno da proteção da informação digital. E as normas arquivísticas aplicadas as principais ferramentas em gestão de objetos digitais.

Também foi possível constatar que o RODA é um projeto em expansão, e que suas melhorias e implementos serão potencialidades geradas pela necessidade de obter confiabilidade dentro de um ambiente digital, e que sua arquitetura obedece uma lógica capaz de agregar muitas informações ao arquivo através dos metadados, e possibilidade de gerir gestão ao gerenciamento da informação, principalmente com fluxos de trabalho pré-definidos para a ingestão/admissão da informação ao repositório. Também capacidade de criar a possibilidade através de seu aplicativos de migração de software e interpretadores da informação digital. Agregando ao objeto digital a potencialidade de ter dados lógicos precisos capazes de serem interpretados independentes do formato. Podendo, também ser convertido, através das metainformações, em uma série de formatos.

Essa incursão no universo do metadado e sua importância no caminho da informação digital, foram significativas para a compreensão de todos requisitos que existem na implementação dessas ferramentas digitais em âmbito administrativo. E a real relevância no futuro da preservação digital.

Contudo, possui uma interface leve e amigável para o manejo dessa informação, podendo atender a expectativa de qualquer setor interessado na administração de grandes coleções e fundos arquivísticos.

## 6 REFERÊNCIAS

ARQUIVO NACIONAL – *AN Digital: políticas em preservação digital*. Edição Eletrônica Rio de Janeiro, 2012

ARQUIVO NACIONAL – *Dicionário brasileiro de terminologia arquivística*. Rio de Janeiro, Arquivo Nacional, 2005.

BARBEDO, Francisco, CORUJO, Luís; CASTRO, Faria; Rui Luís, Ramalho; José Carlos; Miguel Ferreira. *RODA: Repositório de Objectos Digitais Autênticos, Portugal*, <http://www.apbad.pt/Downloads/congresso9/COM37.pdf>.

BELLOTTO, Heloísa Liberali. *Arquivos permanentes: tratamento documental*. FGV, SP, 2004.

CASTRO, Rui; FERREIRA, Miguel; FARIA, Luis. *Meet RODA, a full-fledged digital repository for long-term preservation*. Communication at International Conference on Preservation of Digital Objects (iPres), Singapore, 2011.

CCSDS. Recommendation for Space Data System Practices, MAGENTA BOOK, REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS): RECOMMENDED PRACTICE CCSDS 650.0-M-2, June 2012

CONSELHO NACIONAL DE ARQUIVOS (CONARQ) RESOLUÇÃO Nº 39, DE 29 DE ABRIL DE 2014 – *Diretrizes para a implementação de repositórios digitais confiáveis de documentos arquivísticos*, Rio de Janeiro, 2014.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). *NOBRADE – Norma brasileira de descrição arquivística*. Rio de Janeiro, 2006.

CONSELHO NACIONAL DE ARQUIVOS. *e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos Adotada pelo Conselho Nacional de Arquivos em dezembro de 2009*. Rio de Janeiro, 2009.

FARIA, Luís; CASTRO, Rui; BARBEDO, Francisco; HENRIQUES, Cecília; CORUJO, Luís; FERREIRA, Miguel; RAMALHO, José Carlos. *RODA Repository of Authentic Digital Objects*, 2009 Atlanta, Georgia, USA.

FERREIRA, Miguel. *SCAPE: Report on compliance validation*. SCAPE project, May 2014. *Metadata*, USA, 2004.

FERREIRA, Miguel. *Introdução a preservação digital: conceitos, estratégias e actuais consensos*. Edição Eletrônica, Portugal, 2006.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). *Understanding Metadata*, USA, 2004. National Information Standards Organization (NISO). *Understanding*

PITTI, Daniel. *Encoded Archival Description An Introduction and Overview*. In: D-Lib Magazine, V. 5 N. 11, 1999,. Institute for Advanced Technology in the Humanities, University of Virginia.

PREMIS. *Data Dictionary for Preservation Metadata*, version 2.2, PREMIS Editorial Committee July 2012.

RLG (Research Library Group) and OCLC (Online Computer Library Center). *Trusted Digital Repositories: attributes and responsibilities*. RLG 2002, Mountain View, CA.

ROCHA Rafael Port da. *Metadados, Web Semântica, Categorização Automática: combinando esforços humanos e computacionais para a descoberta e uso dos recursos da web*. In: *Em Questão*, Porto Alegre, v. 10, n. 1, p. 109-121, jan./jun. 2004.

SAYÃO, Luis Fernando. *Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos*. *Ponto de Acesso*, Salvador, V.4, n.3, p. 68-94, dez 2010

SHELLEMBERG, T.R. Arquivos modernos: princípios e técnicas. FGV, SP, 2004.

THE SOCIETY OF AMERICAN ARCHIVISTS. *Encoded Archival Description Tag Library* (EAD): Version 2002, Chicago, 2002

THOMAZ, Katia P.; SOARES Antonio José. A preservação digital e o modelo de referência Open Archival Information System (OAIS). In: *DataGramZero – Revista de Ciência da Informação – v.5 n.1 fev. 04.*

WITT, Michael; KROLL, Matthew; MINOR, David; REILLY; Bernie. *ISO 16363: Trustworthy Digital Repository Certification in Practice*. Purdue University Purdue e-Pubs, 2012.

**Links acessados:**

<http://www.apbad.pt/Downloads/congresso9/COM37.pdf>

[http://www.conarq.arquivonacional.gov.br/media/publicacoes/resol\\_conarq\\_39\\_repositorios.pdf](http://www.conarq.arquivonacional.gov.br/media/publicacoes/resol_conarq_39_repositorios.pdf)

<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>

(<http://www.roda-community.org/>)

<http://www.loc.gov/standards/mets/METSOverview.v2.html>

<http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais>

<http://coleccoes-digitais.wikidot.com/oais>

<http://www.w3.org/RDF/>

[http://pt.wikipedia.org/wiki/Portable\\_document\\_format](http://pt.wikipedia.org/wiki/Portable_document_format)