DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS - RDC-Arq

Conarq
CONSELHO NACIONAL DE ARQUIVOS

Câmara Técnica de Documentos Eletrônicos

RESOLUÇÃO Nº 43, DE 04 DE SETEMBRO DE 2015

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

MINISTÉRIO DA JUSTIÇA

ARQUIVO NACIONAL

CONSELHO NACIONAL DE ARQUIVOS

RESOLUÇÃO Nº 43, DE 04 DE SETEMBRO DE 2015

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

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

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

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

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

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

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

Na página 1, onde se lê: "DIRETRIZES PARA A IMPLEMENTACAO DE REPOSITORIOS DIGITAIS CONFIAVEIS DE DOCUMENTOS ARQUIVISTICOS".

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

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

Nova redação: "DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS

ARQUIVÍSTICOS DIGITAIS CONFIÁVEIS – RDC-Arq".

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

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

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

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

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

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

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

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

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

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

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

JAIME ANTUNES DA SILVA
Presidente do CONARQ

MINISTÉRIO DA JUSTIÇA

ARQUIVO NACIONAL

CONSELHO NACIONAL DE ARQUIVOS

RESOLUÇÃO Nº 39, DE 29 DE ABRIL DE 2014

Estabelece diretrizes para a implementação de repositórios arquivísticos digitais confiáveis para o arquivamento e manutenção de documentos arquivísticos digitais em suas fases corrente, intermediária e permanente, dos órgãos e entidades integrantes do Sistema Nacional de Arquivos – SINAR. [Redação dada pela Resolução nº 43 de 04 de setembro de 2015]

O PRESIDENTE DO CONSELHO NACIONAL DE ARQUIVOS - CONARQ, no uso de suas atribuições, previstas no item IX do art. 23 de seu Regimento Interno, aprovado pela Portaria nº 2.588, do Ministério da Justiça, de 24 de novembro de 2011, em conformidade com a deliberação do Plenário em sua 77ª reunião plenária do CONARQ, realizada no dia 20 de março de 2014,

Considerando que o Conselho Nacional de Arquivos tem por finalidade definir a política nacional de arquivos públicos e privados e exercer orientação normativa visando à gestão documental e à proteção especial aos documentos de arquivo, independente da forma ou do suporte em que a informação está registrada;

Considerando que as organizações públicas e privadas e os cidadãos vêm cada vez mais produzindo documentos arquivísticos exclusivamente em formato digital e que governos, organizações e cidadãos dependem do documento digital como fonte de prova e informação, bem como de garantia de direitos;

Considerando que as instituições arquivísticas devem estabelecer política de preservação e possuir infraestrutura organizacional, bem como requisitos, normas e procedimentos para assegurar que os documentos arquivísticos digitais permaneçam sempre acessíveis, compreensíveis, autênticos e íntegros,

Considerando que a gestão arquivística de documentos, independente da forma ou do suporte adotados, tem por objetivo garantir a produção, a manutenção, a preservação de documentos arquivísticos confiáveis, autênticos e compreensíveis, bem como o acesso a estes;

Considerando a natureza específica dos arquivos digitais, criados e mantidos em ambiente tecnológico de contínua alteração e crescente complexidade, e que não se constituem como entidades físicas convencionais;

Considerando a Carta para a Preservação do Patrimônio Arquivístico Digital do CONARQ, de 6 de julho de 2004, que manifesta a necessidade do, estabelecimento de políticas, procedimentos, sistemas, normas e práticas que levem os produtores de documentos a criar e manter documentos arquivísticos fidedignos, autênticos, preserváveis e acessíveis;

Considerando a Resolução nº 2, de 18 de outubro de 1995, que dispõe sobre as medidas a serem observadas na transferência ou no recolhimento de acervos documentais para instituições arquivísticas públicas;

Considerando a Resolução nº 20, de 16 de julho de 2004, que dispõe sobre a inserção dos documentos digitais em programas de gestão arquivística de documentos dos órgãos e entidades integrantes do Sistema Nacional de Arquivos;

Considerando a Resolução nº 24, de 3 de agosto de 2006, que estabelece diretrizes para a transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas públicas, resolve:

"Art. 1º Aprovar as Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis – RDC-Arq, anexas a esta Resolução, e recomendar sua adoção aos órgãos e entidades integrantes do Sistema Nacional de Arquivos – SINAR, para o arquivamento e manutenção dos documentos arquivísticos em suas fases corrente, intermediária e permanente em formato digital, e de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade e a preservação desses documentos". [Redação dada pela Resolução nº 43 de 04 de setembro de 2015].

Art. 2º Esta Resolução entra em vigor na data de sua publicação.

JAIME ANTUNES DA SILVA

Câmara Técnica de Documentos Eletrônicos

DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVISTICOS DIGITAIS CONFIÁVEIS – RDC-Arq

Rio de Janeiro | 2015

Conarq
CONSELHO NACIONAL DE ARQUIVOS

EQUIPE TÉCNICA DE ELABORAÇÃO DAS DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS ARQUIVISTICOS DIGITAIS CONFIÁVEIS – RDC-Arq

Equipe de redação da Câmara Técnica de Documentos Eletrônicos

Carlos Augusto Silva Ditadi

Claudia Lacombe Rocha

Eloi Juniti Yamaoka

Humberto Celeste Innarelli

João Alberto de Oliveira Lima

Luiz Fernando Sayão

Neire do Rossio Martins

Rosely Curi Rondinelli

Integrantes da Câmara Técnica de Documentos Eletrônicos que participaram deste trabalho

Brenda Couto de Brito Rocco | Arquivo Nacional

Carlos Augusto Silva Ditadi | Arquivo Nacional

Carolina de Oliveira | Arquivo Nacional – a partir de 2012

Claudia Lacombe Rocha | Arquivo Nacional

Daniel Flores | Universidade Federal de Santa Maria

Eloi Juniti Yamaoka | Serviço Federal de Processamento de Dados

Humberto Celeste Innarelli | Universidade Estadual de Campinas

João Alberto de Oliveira Lima | Senado Federal

Luiz Fernando Sayão | Comissão Nacional de Energia Nuclear

Marco Aurélio Rodrigues Braga | Secretaria de Logística e Tecnologia da Informação – a partir de 2013

Margareth da Silva | Universidade Federal Fluminense

Neire do Rossio Martins | Universidade Estadual de Campinas

Rosely Curi Rondinelli | Fundação Casa de Rui Barbosa

Vanderlei Batista dos Santos | Câmara dos Deputados

Colaboração

Andressa Cristiani Piconi | Universidade Estadual de Campinas

Cássia de Paula Moreira Coghi | Universidade Estadual de Campinas

Revisão

José Márcio Batista Rangel

Conarq
CONSELHO NACIONAL DE ARQUIVOS

SUMÁRIO

Conarq
CONSELHO NACIONAL DE ARQUIVOS

I – APRESENTAÇÃO

Os documentos arquivísticos caracterizam-se por registrarem e apoiarem as atividades do órgão ou entidade, servindo de evidência dessas atividades, bem como de fonte de informação para a pesquisa, e para assegurar os direitos dos cidadãos. Assim, é preciso garantir que os documentos sejam acessíveis e permaneçam autênticos em todo o seu ciclo de vida. A produção crescente de documentos arquivísticos em formato digital desafia as organizações produtoras e as instituições de preservação na busca de soluções para a preservação e o acesso de longo prazo. Os documentos digitais sofrem diversas ameaças decorrentes da fragilidade inerente aos objetos digitais, da facilidade de adulteração e da rápida obsolescência tecnológica.

Os documentos arquivísticos digitais em fase corrente e intermediária devem, preferencialmente, ser gerenciados por meio de um Sistema Informatizado de Gestão Arquivística de Documentos – SIGAD, a fim de garantir o controle do ciclo de vida, o cumprimento da destinação prevista e a manutenção da autenticidade e da relação orgânica,[1] características fundamentais desses documentos. Já nessas fases, os produtores precisam tomar cuidados especiais, previstos em um plano de preservação digital, com relação aos documentos digitais que serão mantidos por médio e longo prazos, de forma a garantir sua autenticidade e seu acesso.

A partir da destinação para guarda permanente, ocorre uma alteração na cadeia de custódia, passando a responsabilidade pela preservação dos documentos dos produtores para a instância de guarda. Os documentos digitais em fase permanente são dependentes de um bom sistema informatizado que apoie o tratamento técnico adequado, incluindo arranjo, descrição e acesso, de forma a assegurar a manutenção da autenticidade e da relação orgânica desses documentos.

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.

No contexto internacional, algumas iniciativas indicam a importância do desenvolvimento de repositórios digitais confiáveis como solução para a garantia da autenticidade, da preservação e do acesso de longo prazo. Dentre essas iniciativas, destaca-se a do grupo de trabalho liderado pelo *Research Library Group – RLG* e pelo *Online Computer Library Center – OCLC*.[2] Na perspectiva do grupo de trabalho *RLG*/*OCLC*, um "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).[3]

1 Quando os documentos arquivísticos são produzidos e mantidos dentro de um sistema informatizado (p. ex. sistemas de controle acadêmico em instituições de ensino, sistemas de prontuários médicos, sistemas de controle de ponto), esse sistema deve incorporar as funcionalidades básicas de um SIGAD previstas no e-ARQ Brasil, para assegurar tais objetivos.

2 Desde junho de 2006, o *RLG* e o *OCLC* estão reunidos em uma única organização. Para mais informações, veja o sítio: http://www.oclc.org/

3 Texto no original em inglês: "A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future".

O arquivamento e a preservação digital constituem uma questão complexa que envolve muitas variáveis, compromissos de longa duração e a necessidade de expressivos investimentos em infraestrutura tecnológica, pesquisa e recursos humanos. Diante disso, a formação de consórcios, em determinados casos, pode ser a solução mais viável.

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

I.1 – Objetivo deste documento

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

I.2 – Escopo

Estas diretrizes visam a orientar os órgãos e as entidades integrantes do Sistema Nacional de Arquivos – SINAR na implantação de repositórios digitais confiáveis para documentos arquivísticos digitais.

São integrantes do SINAR:[4]

- 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.

Além de parâmetros tecnológicos e de infraestrutura, as diretrizes aqui apresentadas tratam também de políticas e procedimentos técnicos e administrativos. Os parâmetros indicados atendem às necessidades de repositórios digitais confiáveis para o armazenamento de documentos correntes, intermediários e permanentes.

4 De acordo com o decreto nº 4.073, de 3 de janeiro de 2002, que regulamenta a lei nº 8.159, de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados.

I.3 – Definições[5]

Apresentam-se, aqui, definições importantes no contexto desse documento.

Atualização de suporte
Técnica de migração que consiste em copiar os dados de um suporte para outro, sem mudar sua codificação, para evitar perdas de dados provocadas por deterioração do suporte.

Autenticidade
Credibilidade de um documento enquanto documento, isto é, a qualidade de um documento ser o que diz ser e de que está livre de adulteração ou qualquer outro tipo de corrupção.

Ciclo vital dos documentos
Sucessivas fases por que passam os documentos arquivísticos, de sua produção a guarda permanente ou eliminação.

Confiabilidade
Credibilidade de um documento arquivístico enquanto afirmação de um fato. Existe quando um documento arquivístico pode sustentar o fato ao qual se refere, e é estabelecida pelo exame da completeza, da forma do documento e do grau de controle exercido no seu processo de criação.

Confidencialidade
Propriedade de certos dados ou informações que não podem ser disponibilizadas ou divulgadas sem autorização para pessoas, entidades ou processos.

Conversão
Técnica de migração que pode se configurar de diversas formas, tais como: a) conversão de dados: mudança de um formato para outro; b) conversão de sistema computacional: mudança do modelo de computador e de seus periféricos.

Disponibilidade
Propriedade de estar acessível e utilizável sob demanda por uma entidade autorizada.

Documento arquivístico
Documento produzido (elaborado ou recebido) no curso de uma atividade prática, como instrumento ou resultado dessa atividade, e retido para ação ou referência.

Documento arquivístico digital

5 As definições aqui apresentadas foram baseadas nos glossários dos seguintes documentos:
- "e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos". Câmara Técnica de Documentos Eletrônicos – CTDE. Versão adotada pelo CONARQ em dezembro de 2009;
- Glossário da CTDE/CONARQ.
- Resolução GR – UNICAMP nº 17/2011, de 29/6/2011.
- ABNT 27001/2006 – "Requisitos para sistemas de gestão de segurança da informação".
- "Diretrizes do Preservador – A preservação de documentos arquivísticos digitais: Diretrizes para organizações". Projeto InterPARES 2.
- ISO 14721/2003 – *Reference Model for an Open Archival Information System – OAIS.*

Documento digital reconhecido e tratado como documento arquivístico.

Documento digital
Informação registrada, codificada em dígitos binários, acessível e interpretável por meio de sistema computacional.

Integridade
Estado dos documentos que se encontram completos e não sofreram nenhum tipo de corrupção ou alteração não autorizada nem documentada.

Metadados
Dados estruturados que descrevem e permitem encontrar, gerenciar, compreender e/ou preservar documentos arquivísticos ao longo do tempo.

Migração
Conjunto de procedimentos e técnicas para assegurar a capacidade de os objetos digitais serem acessados face às mudanças tecnológicas. A migração consiste na transferência de um objeto digital: a) de um suporte que está se tornando obsoleto, fisicamente deteriorado ou instável para um suporte mais novo; b) de um formato obsoleto para um formato mais atual ou padronizado; c) de uma plataforma computacional em vias de descontinuidade para outra mais moderna. A migração pode ocorrer por conversão, por atualização ou por reformatação.

Modelo de referência
Uma estrutura conceitual para compreensão dos principais relacionamentos entre as entidades de um ambiente, e para o desenvolvimento de padrões consistentes ou especificações que consolidam esse ambiente. Um modelo de referência é baseado em pequena quantidade de conceitos unificados, e pode ser usado como uma base para aprendizado e explanação de padrões para um não especialista.

Normalização de formatos
Conversão de formatos de arquivo para um elenco gerenciável de formatos apropriados para preservação e acesso.

Preservação digital
Conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo acesso e interpretação dos documentos digitais pelo tempo que for necessário.

Preservador de documentos arquivísticos
Entidade responsável pela custódia física e legal dos documentos do produtor, bem como por sua preservação, isto é, proteger e garantir acesso contínuo aos documentos.

Sistema Informatizado de Gestão Arquivística de Documentos – SIGAD
Conjunto de procedimentos e operações técnicas característico do sistema de gestão arquivística de documentos, processado eletronicamente e aplicável em ambientes digitais ou híbridos, isto é, composto de documentos digitais e não digitais.

II – REPOSITÓRIO DIGITAL CONFIÁVEL DE DOCUMENTOS ARQUIVÍSTICOS – PRINCIPAIS REQUISITOS

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.

A norma mais importante da área é o *Open Archival Information System – OAIS*,[6] um modelo conceitual desenvolvido pelo *Consultive Committee for Space Data Systems – CCSDS*,[7] que resultou na norma ISO 14721:2003. O *OAIS* descreve as funções de um repositório digital e 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.

A preocupação com a confiabilidade dos repositórios digitais foi evidenciada no relatório da *Task Force on Archiving of Digital Information*,[8] uma ação cooperativa do *RLG* e da *Commission on Preservation and Access*, publicado em 1996, no qual se declarou que "um componente crítico da infraestrutura de arquivamento digital é a existência de um número suficiente de instituições confiáveis, que sejam capazes de armazenar, migrar e prover acesso a acervos digitais".[9] O relatório da *Task Force* foi mais além, ao apontar a necessidade de um processo de certificação dos repositórios digitais para atribuir esse caráter de confiabilidade de uma forma mais isenta.

Esse relatório estimulou a colaboração do *RLG*/*OCLC*, iniciada em março de 2000, no sentido de definir as bases conceituais e os principais atributos para um repositório digital confiável. Como resultado desse trabalho, foi publicado, em 2002, um relatório sob o título *Trusted Digital Repositories: Attributes and Responsibilities*.

Em continuidade a esse trabalho, o *RLG* estabeleceu uma parceria com a administração nacional dos arquivos dos Estados Unidos (*National Archives and Records Administration – NARA*), com o objetivo de definir critérios para a certificação de repositórios confiáveis, em sintonia com os resultados apontados no relatório *RLG*/*OCLC*, de 2002, e com o modelo *OAIS*. Assim, foi publicado, em 2007, o documento *Trustworthy Repository Audit & Certification: Criteria and Checklist*, mais conhecido pela sigla *TRAC*,[10] que apresenta um conjunto de critérios e um *checklist* a serem tomados como referência para a certificação de repositórios digitais confiáveis. Esse documento serviu de base para a elaboração da norma ISO 16363: 2012, que lista os critérios que um repositório digital confiável deve atender. Paralelamente a essa iniciativa, encontra-se em fase de desenvolvimento a norma ISO 16919,[11] que estabelece requisitos para entidades certificadoras de repositórios digitais confiáveis.

6 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".

7 Comitê formado pelas maiores agências espaciais do mundo, com o objetivo de oferecer um fórum para discussão de problemas comuns sobre o desenvolvimento e a operação de sistemas de dados espaciais.

8 *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information*. Maio de 1996. Disponível em: http://www.oclc.org/content/dam/research/activities/digpresstudy/final-report.pdf?urlm=161430.

9 Texto no original em inglês: *a critical component of digital archiving infrastructure is the existence of a sufficient number of trusted organizations capable of storing, migrating and providing access to digital collections*.

10 Disponível em: http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf.

11 Disponível em: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57950.

Esses documentos apontam as diretrizes para repositórios digitais confiáveis e fundamentaram a elaboração deste trabalho. Inicialmente, faz-se necessário esclarecer os conceitos de "repositório digital", "repositório arquivístico digital" e "repositório digital confiável".

No contexto deste documento, **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. Tal ambiente tem sido empregado em diversas situações, tais como:

- arquivo corrente e intermediário (em associação com um SIGAD);
- arquivo permanente;
- biblioteca digital;
- acervo de obras de arte digitais;
- depósito legal de material digital; e
- curadoria de dados digitais de pesquisa.

ATENÇÃO: Um repositório digital não se resume a uma solução informatizada para armazenamento (*storage*), que é apenas um componente do repositório.

Um **repositório arquivístico digital** 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.

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 devem:

- 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;
- estabelecer metodologias para avaliação dos sistemas que considerem as expectativas de confiabilidade esperadas pela comunidade;
- considerar, para desempenhar suas responsabilidades de longo prazo, os depositários e os usuários de forma aberta e explícita;
- dispor de políticas, práticas e desempenho que possam ser auditáveis e mensuráveis; e

- observar os seguintes fatores relativos às responsabilidades organizacionais e de curadoria dos repositórios: escopo dos materiais depositados, gerenciamento do ciclo de vida e preservação, atuação junto a uma ampla gama de parceiros, questões legais relacionadas com a propriedade dos materiais armazenados e implicações financeiras.

Uma forma de atestar a confiabilidade de um repositório digital junto à comunidade-alvo dá-se por meio da sua certificação por terceiros. Para esse fim, o *RLG*/*OCLC*, em parceria com o NARA, publicou, em 2007, o documento "Trustworthy Repository Audit & Certification: Criteria and Checklist – *TRAC*".

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

A seguir, serão apresentadas, primeiramente, algumas considerações a respeito dos repositórios digitais de documentos arquivísticos. Num segundo momento, serão abordados os requisitos que um repositório digital deve seguir para que possa ser considerado confiável, com base na norma ISO 16363: 2012, independentemente do tipo de material digital (arquivístico ou não).

II.1 – CONSIDERAÇÕES SOBRE UM REPOSITÓRIO DIGITAL DE DOCUMENTOS ARQUIVÍSTICOS

Responsabilidade pelo repositório
A responsabilidade pelo projeto, implantação e manutenção de um repositório digital de documentos arquivísticos deve ser compartilhada por profissionais de arquivo e de tecnologia da informação, de forma a se cumprirem os requisitos tecnológicos e os procedimentos do tratamento arquivístico.

Tratamento arquivístico
Um repositório digital para documentos arquivísticos tem que ser capaz de organizar e recuperar os documentos, de forma a manter a relação orgânica entre eles. Nesse sentido, deve apoiar a organização hierárquica dos documentos digitais, a partir de um plano de classificação de documentos, e a descrição multinível, de acordo com a norma internacional para descrição arquivística (a "Norma Geral Internacional de Descrição Arquivística – ISAD(G)" e a "Norma Brasileira de Descrição Arquivística – NOBRADE").

Princípios de preservação digital
A preservação digital tem que garantir o acesso de longo prazo a documentos arquivísticos autênticos, o que implica a adoção dos seguintes princípios:

- focar especificamente em documentos arquivísticos, e não em objetos digitais de forma genérica;
- focar em documentos arquivísticos digitais autênticos;
- pressupor que a autenticidade dos documentos arquivísticos digitais está sob ameaça, principalmente no momento da transmissão no espaço (entre pessoas e sistemas) e no tempo (atualização/substituição de *hardware* e *software* usados para armazenar, processar e comunicar os documentos);

- reconhecer que a preservação digital é um processo contínuo, que começa na concepção do documento;
- reconhecer que a autenticidade[12] dos documentos arquivísticos digitais tem por base os procedimentos de gestão e preservação e a confiança tanto no repositório como no órgão responsável pela guarda desses documentos;
- arbitrar o que se considera como documento original, uma vez que a preservação digital implica a necessidade de conversão de formatos e atualização de suportes;
- reconhecer que a elaboração de manuais e os procedimentos de preservação desempenhados pelo repositório digital apoiam a presunção de autenticidade desses documentos;
- reconhecer que o registro, em metadados, das intervenções de preservação em cada documento apoia a presunção de autenticidade desses documentos;
- reconhecer que a autenticidade dos documentos digitais deve ser avaliada e presumida no momento de sua submissão ao repositório.[13]
- reconhecer que o repositório digital é responsável pela manutenção permanente da autenticidade dos documentos a ele submetidos; e
- distinguir claramente a autenticidade e autenticação de documentos, considerando que a primeira é a qualidade de o documento ser verdadeiro, e a segunda é uma declaração dessa qualidade, feita, em um dado momento, por uma pessoa autorizada para tal.

Independência dos repositórios

Um repositório digital deve ter independência; isso significa que seu funcionamento e o acesso aos documentos não podem depender das aplicações que funcionam em conjunto com ele. Por exemplo, em uma aplicação para arquivos correntes e intermediários, deve ser possível acessar os documentos independentemente do SIGAD, isto é, diretamente no repositório, desde que isso seja feito de forma controlada, para não ameaçar a autenticidade dos documentos no repositório. É bom esclarecer que o acesso direto aos documentos no repositório não exclui a necessidade de um SIGAD para apoiar a gestão arquivística.

Interoperabilidade

Um repositório digital deve estar em conformidade com as normas e padrões estabelecidos, de forma a possibilitar níveis de interoperabilidade com outros repositórios digitais e sistemas informatizados que tratam de documentos arquivísticos. Podem ser citados como exemplos dessas normas e padrões: o "Open Archives Initiative Protocol for Metadata Harvesting – OAI-PMH", para coleta de registros de metadados em repositórios digitais; o "Metadata Encoding and Transmission Standard – METS", para a codificação de metadados descritivos, administrativos e estruturais; o "Encoded Archival Description – EAD", para a codificação de metadados descritivos de documentos arquivísticos; e os "Padrões de Interoperabilidade de Governo Eletrônico – e-PING",[14] no caso dos órgãos e entidades do governo federal.

12 Ver Resolução nº 37, de 19 de dezembro de 2012, do CONARQ, que aprova as "Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais".

13 Ver Resolução nº 24, de 3 de agosto de 2006, do CONARQ, que estabelece diretrizes para a transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas públicas.

14 Informações disponíveis em: http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padroes-de-interoperabilidade.

II.2 – REQUISITOS PARA UM REPOSITÓRIO DIGITAL CONFIÁVEL

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

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

II.2.1 – Infraestrutura organizacional

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

a. Governança e viabilidade organizacional

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

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

b. Estrutura organizacional e de pessoal

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

c. Transparência de procedimentos e arcabouço político

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

- definir a comunidade alvo e sua base de conhecimento;
- possuir políticas e definições, acessíveis publicamente, que demonstrem como os requisitos do serviço de preservação serão contemplados;
- possuir políticas, procedimentos e mecanismos de atualização, na medida em que o repositório cresce e a tecnologia e práticas da comunidade evoluem;
- documentar permissões legais – por meio de acordos de custódia, normas de procedimentos e outros – que o isentem de responsabilidade, no caso de alterações passíveis de ocorrer em estratégias de preservação digital;
- fazer o registro histórico das mudanças de procedimentos, de *software* e *hardware*;
- relacionar o registro histórico, acima referido, com as estratégias de preservação digital, e descrever os potenciais efeitos dessas mudanças sobre os documentos digitais;
- demonstrar que está sistematicamente avaliando a satisfação das expectativas dos

produtores e dos usuários, e buscando atendê-las;

- estar comprometido com a definição, coleta, auditoria e fornecimento (sob demanda) de mecanismos de controle da integridade dos documentos digitais sob sua custódia;
- estar comprometido em realizar regularmente uma autoavaliação de seu funcionamento e 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).

d. Sustentabilidade financeira

Um repositório digital confiável deve demonstrar sustentabilidade financeira. Para isso, deve ter um plano de gestão que observe os seguintes aspectos:

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

e. Contratos, licenças e passivos

Os contratos, licenças e passivos firmados pelo repositório devem ser claros e mensuráveis; delinear papéis, responsabilidades, prazos e condições; e ser facilmente acessíveis ou disponíveis aos interessados. Esses contratos, licenças e passivos podem envolver tanto a relação entre o repositório e os produtores de documentos digitais, como a relação entre o repositório e fornecedores de serviços. Esses mesmos instrumentos devem especificar todos os direitos e obrigações do repositório sobre os documentos digitais a ele confiados, em especial no que diz respeito à propriedade intelectual e a restrições de uso.

II.2.2 – Gerenciamento do documento digital

O gerenciamento dos documentos de um repositório digital confiável deve estar de acordo com o modelo de referência *OAIS*, que estabelece a formação de pacotes de informação envolvendo os documentos digitais (informação de conteúdo) e seus metadados (informação de representação).

São três os tipos de pacotes de informação:

- Pacote de informação para submissão (*submission information package – SIP*) – refere-se à admissão dos documentos digitais e seus metadados associados.
- Pacote de informação para arquivamento (*archival information package – AIP*) – refere-se ao acondicionamento e armazenamento dos documentos digitais e seus metadados associados.
- Pacote de informação para disseminação (*dissemination information package – DIP*) – refere-se ao acesso aos documentos digitais e seus metadados associados.

A *TRAC* apresenta os requisitos para gerenciamento do documento no repositório digital, categorizados em seis grupos, com base nas funcionalidades, conforme detalhado a seguir:

a. Admissão: captura de documentos digitais

A admissão consiste na entrada dos documentos e seus metadados no repositório digital. Os requisitos de admissão variam dependendo do tipo de material, do contexto legal e da relação entre o produtor de documento e o repositório. Independentemente dessas variações, pode-se afirmar que a admissão se inicia com o recebimento de um *SIP*, que é convertido em *AIP*, e termina quando um *AIP* está seguro no repositório, incluindo a criação de cópias de segurança.

A seguir, apresentam-se requisitos gerais a serem cumpridos pelo repositório, cuja adequação deve ser avaliada de acordo com a missão e as necessidades de cada repositório:

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

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

b. Admissão: criação do pacote de arquivamento

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

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

- descrever cada classe de informação (texto estruturado, imagem matricial, banco de dados, imagem em movimento e outras) a ser preservada pelo repositório, e como ela está implementada – essa descrição deve apontar os componentes-chave do *AIP*: o documento arquivístico, sua informação de representação (informação estrutural e semântica) e as várias categorias de informação descritiva de preservação (fixidade, proveniência e contexto), e ainda como esses componentes se relacionam;

- descrever minuciosamente as diferentes classes de informação e como os *AIPs* são implementados, nos casos em que a especificidade daquelas classes exigir ações de preservação diferentes (por exemplo, a imagem *TIFF* que é processada por um sistema pode necessitar de ações de preservação diferentes das ações necessárias à imagem *TIFF* que é apresentada para o olho humano);
- descrever como os *AIPs* são construídos a partir dos *SIPs*, ou seja, apontar todas as transformações pelas quais passarão os documentos e os metadados submetidos, e os metadados a serem adicionados no momento da formação do *AIP*;
- ser capaz de demonstrar se os *SIPs* foram aceitos e transformados em um *AIP*s integralmente ou em parte, ou ainda se foram recusados;
- atribuir aos *AIPs*, identificadores que sejam únicos, persistentes e visíveis aos gestores e auditores, de acordo com padrões reconhecidos (por exemplo: Handle System, DOI, URN, PURL);
- no caso de o documento já possuir um identificador único, a ele atribuído no *SIP*, o repositório deverá mantê-lo no *AIP*, ou criar um outro identificador, que deverá ser associado, de maneira persistente, ao do *SIP*;
- ter acesso a ferramentas amplamente reconhecidas para apoiar o monitoramento dos componentes digitais dos documentos, tais como diretórios de formatos de arquivos (ex.: *PRONOM* – base de dados com registro de formatos mantida pelo arquivo nacional do Reino Unido[15]) e registros de outras informações de representação;
- registrar, em um banco de dados local, a informação de representação dos documentos admitidos, quando essa informação não estiver disponível nas ferramentas mencionadas no ponto anterior;
- registrar metadados de preservação associados aos documentos admitidos, de maneira a apoiar sua integridade, localização, legibilidade e proveniência, dentre outros;
- ter procedimentos para testar se os documentos são compreensíveis pela comunidade-alvo e, em caso negativo, adequá-los às necessidades dessa comunidade (ex.: documentos voltados para deficientes visuais);
- verificar a completude e a correção de cada *AIP* no momento em que é gerado, isto é, no momento em que o *SIP* é convertido em *AIP*;
- ter um mecanismo independente para verificar a integridade do conjunto do seu acervo, ou seja, verificar que todos os documentos previstos foram, de fato, admitidos no repositório, justificando possíveis lacunas; e
- documentar todas as ações relevantes à preservação dos documentos e que estão relacionadas à criação do *AIP*.

c. Planejamento da preservação

Um repositório digital deve fazer o planejamento da preservação dos documentos sob sua custódia, a fim de enfrentar os problemas trazidos pela obsolescência tecnológica e fragilidade do suporte. Esse planejamento deve ser feito a partir de uma política de preservação digital, ser bem documentado e incluir:

- estratégias de preservação bem definidas e periodicamente atualizadas, apontando e detalhando cada procedimento a ser adotado, como, por exemplo, a normalização de formatos;

15 Disponível em: http://www.nationalarchives.gov.uk/PRONOM/

- mecanismos para monitoramento e notificação quando alguma informação de representação dos documentos no repositório estiver se tornando obsoleta ou inviável (ex.: um formato de arquivo que esteja entrando em desuso, um suporte que esteja no final de sua vida útil);
- mecanismos de mudanças do plano de preservação como resultado do monitoramento; e
- fornecimento de evidências sobre a eficácia do plano de preservação.

d. Armazenamento e preservação / manutenção do *AIP*

Um repositório deve atender a um conjunto de condições para garantir o bom desempenho da preservação de longo prazo dos *AIPs*. Tais condições são:

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

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

e. Gerenciamento de informação

Uma funcionalidade essencial de um repositório digital confiável é o gerenciamento da informação, aqui entendido como a gestão das informações descritivas (metadados) dos documentos admitidos no repositório. O principal objetivo desses metadados é apoiar o acesso e a recuperação dos documentos, e vão além das informações descritivas mais usuais (autor, título, data), envolvendo outras informações descritivas úteis aos usuários, tais como tamanho do arquivo disponível para *download* ou informação sobre a aplicação necessária para ler o arquivo. O gerenciamento da informação descritiva envolve os seguintes aspectos:

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

f. Gerenciamento de acesso

Todo repositório deve produzir pacotes de disseminação de informação (*DIP*), atendendo aos seguintes requisitos:

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

II.2.3 – Tecnologia, infraestrutura técnica e segurança

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

a. Infraestrutura de sistema

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

- funcionamento do repositório com base num sistema operacional e outros *softwares* de infraestrutura que tenham um bom suporte do mercado e da comunidade de usuários;
- adequação dos processos, do *hardware* e do *software* do sistema de *backup* às necessidades do repositório;

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

b. Tecnologias apropriadas

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

c. Segurança

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

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

III – PADRÕES E NORMAS DE REFERÊNCIA

A seguir, são apresentados documentos de referência para a construção de repositórios arquivísticos digitais confiáveis – RDC-Arq.

Os documentos são bastante variados:

- documentos que definem modelos ou orientam a certificação de repositórios confiáveis;
- definição de metadados, que podem ser utilizados de acordo com o propósito do repositório; e
- codificações, em *XML*, de metadados e de padrões de transmissão.

III.1 – Modelo de referência *OAIS*

FONTES:
Reference Model for an Open Archival Information System (OAIS) – Magenta Book. Issue 2 – CCSDS: junho de 2012.
Space data and information transfer systems – Open archival information system – Reference model: ISO 14721:2012.
Sistemas espaciais de dados e informações – Modelo de referência para um sistema aberto de arquivamento de informação (SAAI): ABNT NBR 15472:2007.

O modelo de referência *OAIS* (*Open Archival Information System*[16]) é uma recomendação internacional desde 2003 (ISO 14721). Trata-se de um modelo conceitual que define um repositório digital, identificando o ambiente, os componentes funcionais, suas interfaces internas e externas, os objetos de dados e informações. No Brasil, foi adaptado e publicado como norma ABNT NBR 15472: 2007, sob o título "Sistema Aberto de Arquivamento de Informação – SAAI".

Um repositório que segue a norma *OAIS* é constituído por pessoas e sistemas com a responsabilidade de preservar a informação e torná-la disponível. O modelo aborda questões fundamentais relativas à preservação de longo prazo de materiais digitais, independentemente da área de aplicação (arquivo, biblioteca, museu etc.).

O ambiente do modelo conta com três entidades externas:

- Produtor – é o papel desempenhado por pessoas ou sistemas que fornecem a informação a ser preservada.
- Administrador – é o papel desempenhado por aqueles que estabelecem as políticas gerais que governam o repositório.
- Consumidor – é o papel desempenhado por pessoas ou sistemas que interagem com os serviços *OAIS* para acessar a informação preservada desejada.

O *OAIS* é composto por dois modelos: o modelo funcional e o modelo de informação. O modelo funcional delineia as funções que precisam ser desempenhadas por um repositório *OAIS*. A figura 1 apresenta os componentes funcionais, os pacotes de informação e as entidades externas de um repositório digital compatível com o *OAIS*.

16 O termo *open* ("aberto") é usado para indicar que o modelo de referência é construído em fóruns abertos, e não que o acesso ao arquivo é irrestrito.

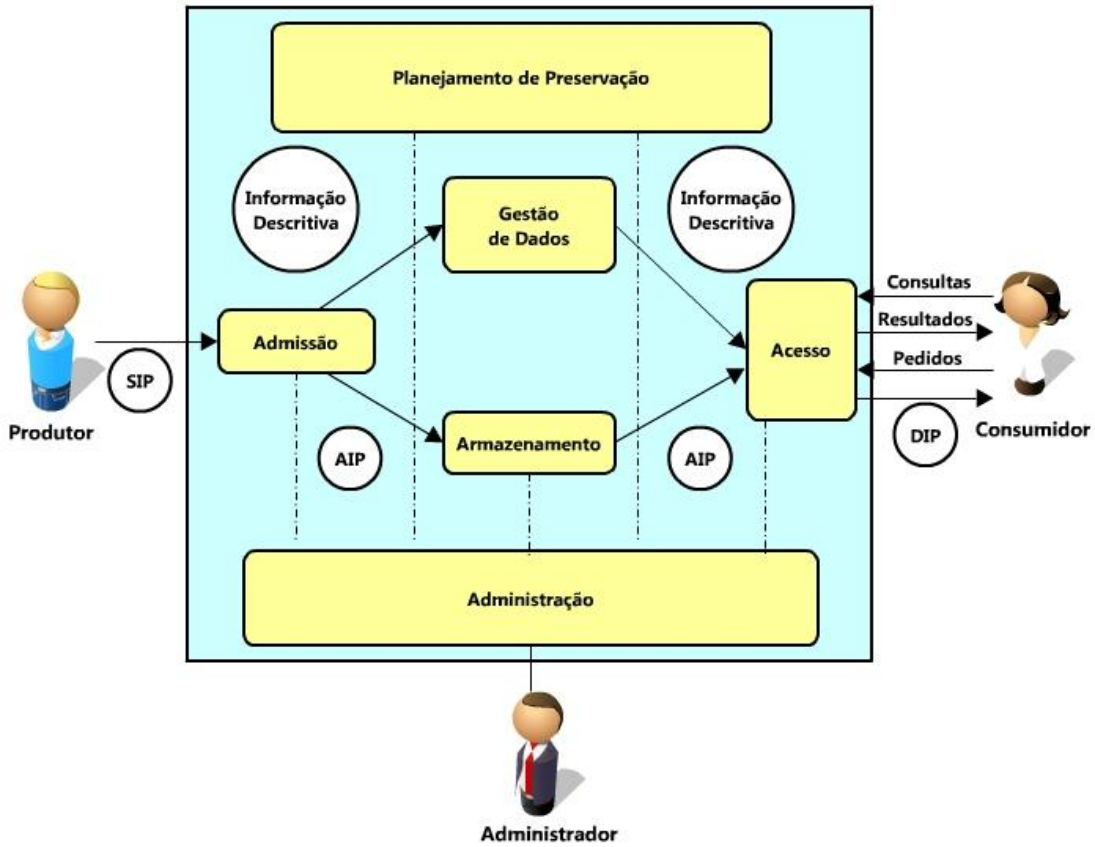

Figura 1 – Entidades funcionais do *OAIS*

Para fins de preservação, o entendimento claro de determinados conceitos é central. Assim, no âmbito do *OAIS*, esses conceitos são:

- **Informação** é qualquer tipo de conhecimento que pode ser intercambiado, sempre representado por algum tipo de dado;
- **Objeto de informação** (figura 2) é resultante do objeto de dado, que é interpretado com o uso da informação de representação; essa informação de representação pode ser decomposta em informação semântica e estrutural, como, por exemplo, um texto em português (informação semântica) codificado no formato *ASCII* (informação estrutural).

Figura 2 – Informação a partir dos dados

O Modelo de Informação do *OAIS* propõe o conceito de **pacote de informação** (figura 3), que é formado pela informação de conteúdo e pela informação de descrição de preservação, encapsuladas e identificadas pela informação de empacotamento. A informação de conteúdo é o objeto de informação (objeto de dado + informação de representação) a ser preservado. A informação de descrição de preservação é a informação necessária para a adequada preservação da informação de conteúdo, e que pode ser categorizada como informação sobre proveniência, referência, fixidade e contexto.

O pacote de informação é associado a outras informações descritivas que vão possibilitar sua localização no repositório.

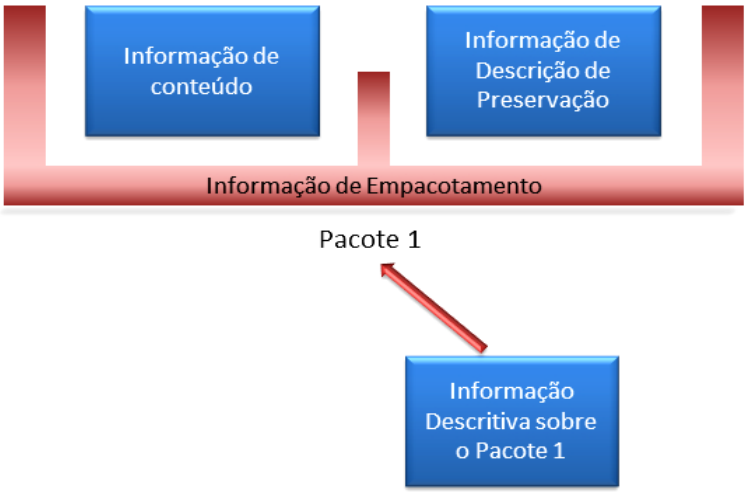

Figura 3 – Conceitos e Relacionamentos do Pacote de Informação

III.2 – Relatório da *Research Library Group* (*RLG*) e da *Online Computer Library Center* (*OCLC*) – Repositórios digitais confiáveis: atributos e responsabilidades

FONTE: Trusted Digital Repositories: Attributes and Responsibilities – RLG-OCLC Report: maio de 2002.

O relatório, publicado em maio de 2002, apresenta uma proposta conjunta para a implementação de repositórios de organizações de ciência e pesquisa, a partir do modelo *OAIS*. O relatório estabeleceu as características essenciais e as responsabilidades para a criação e manutenção de repositórios digitais confiáveis que atendessem aos acervos de instiuições culturais e científicas, garantindo seu acesso a longo prazo, sua integridade e confiabilidade.
De acordo com a *OCLC*, um dos princípios básicos de um repositório digital confiável é o de demonstrar sua capacidade de sustentabilidade, no longo prazo, e de qualificação para o tratamento técnico dos acervos digitais, em diferentes formatos, além de contar com uma infraestrutura tecnológica robusta.

III.3 – Certificação e auditoria de repositórios confiáveis: critérios e *checklist* – *TRAC*

FONTES:
Trustworthy Repositories Audit & Certification: Criteria and Checklist – OCLC, CRL, NARA: fevereiro de 2007.
Space data and information transfer systems – Audit and certification of trustworthy digital repositories: ISO 16363:2012.

O documento apresenta um conjunto de critérios e um *checklist* que são tomados como referência para a certificação de repositórios digitais. Nessa direção, ele oferece ferramentas para auditoria, avaliação e certificação potencial de repositórios; estabelece a documentação exigida para a auditoria; delineia um processo de certificação; e estabelece as metodologias apropriadas para determinar a solidez e a sustentabilidade de repositórios digitais.

III.4 – Requisitos técnicos para entidades de auditoria e certificação de organizações candidatas a serem repositórios digitais confiáveis

FONTES:
Requirements for bodies providing audit and certification of candidate trustworthy digital repositories – Magenta Book – CCSDS: novembro de 2011.
Space data and information transfer systems – Requirements for bodies providing audit and certification of candidate trustworthy digital repositories: ISO/DIS 16919.

É uma recomendação técnica, criada pelo *Consultative Committee for Space Data Systems* (*CCSDS*), que estabelece requisitos para as entidades de auditoria e certificação de repositórios digitais confiáveis. Desde novembro de 2011, encontra-se em fase de desenvolvimento como norma ISO/DIS 16919.

O principal objetivo do documento é definir uma prática sobre a qual devem se basear as operações de uma organização que realiza auditorias para avaliar a confiabilidade de repositórios digitais e fornecer a certificação apropriada. Nesse sentido, apoia o credenciamento de entidades que prestam tal certificação. As exigências contidas nesta norma precisam ser demonstradas, em termos de competência e confiabilidade, por qualquer organização ou organismo de certificação de repositórios digitais.

III.5 – Metadados de preservação – *PREMIS*

FONTE:
PREMIS Data Dictionary for Preservation Metadata – Versão 2.2: julho de 2012.

É uma norma internacional que apresenta um conjunto básico (*core*) de elementos de metadados de preservação para apoiar sistemas que gerenciam objetos digitais. O grupo de trabalho *PREMIS* (*Preservation Metadata: Implementation Strategies*) tem ampla abrangência junto à comunidade dedicada à preservação digital, e seu principal documento de referência é o *PREMIS Data Dictionary for Preservation Metadata*.

Os metadados definidos no *PREMIS Data Dictionary*:

- contribuem para a viabilidade, disponibilidade, clareza, autenticidade e identidade de objetos no contexto da preservação digital;
- representam as informações sobre os documentos digitais que a maioria dos repositórios precisa saber para preservar esses documentos ao longo do tempo;
- prestam especial atenção aos metadados rigorosamente definidos, com base em diretrizes para a criação, gestão e uso, voltados para fluxos de trabalho automatizados; e
- são tecnicamente neutros, ou seja, não assumem o uso, em particular, de quaisquer tecnologias de preservação, estratégias, sistemas de armazenamento, gerenciamento de metadados etc.

O *PREMIS Data Dictionary* também inclui um modelo de esquema em *XML* que permite incorporar o dicionário de dados em sistemas de gestão de objetos digitais.

A norma é mantida pelo *Network Development and MARC Standards Office*, da Biblioteca do Congresso dos EUA (*Library of Congress*).

III.6 – Norma Geral Internacional de Descrição Arquivística – *ISAD(G)*

FONTE:
ISAD (G): Norma Geral Internacional de Descrição Arquivística: segunda edição – CIA: 2000.

É uma norma elaborada no âmbito do Conselho Internacional de Arquivos – CIA, publicada, pela primeira vez, em 1994 e, em segunda edição, em 2000, que estabelece diretrizes gerais para a preparação de descrições arquivísticas. Tem, por objetivos, identificar e explicar o contexto e o conteúdo de documentos de arquivo, a fim de promover o acesso aos mesmos.

III.7 – Norma Brasileira de Descrição Arquivística – NOBRADE

FONTE:
Norma Brasileira de Descrição Arquivística – NOBRADE – CTNDA / CONARQ: 2006.

É uma norma elaborada pela Câmara Técnica de Normalização de Descrição Arquivística do Conselho Nacional de Arquivos – CTNDA/CONARQ, publicada em 2006, em conformidade com a ISAD(G) e a "Norma Internacional de Registro de Autoridade Arquivística para Entidades Coletivas, Pessoas e Famílias – *ISAAR(CPF)*". Consiste na adaptação – e não simplesmente na tradução – das normas internacionais à realidade brasileira, visando a facilitar o acesso e o intercâmbio de informações, em âmbito nacional e internacional, por meio de descrições consistentes, apropriadas e autoexplicativas dos documentos arquivísticos.

III.8 – Metadados do e-ARQ Brasil

FONTE:
Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos – e-ARQ Brasil, v1.1 – CTDE / CONARQ: dezembro de 2009.

O e-ARQ Brasil é o modelo de requisitos para sistemas informatizados de gestão arquivística de documentos, elaborado pela Câmara Técnica de Documentos Eletrônicos do Conselho Nacional de Arquivos (CTDE/CONARQ) e adotado pelo Sistema Nacional de Arquivos – SINAR, por meio das resoluções nº 25, de 27 de abril de 2007, e nº 32, de 17 de maio de 2010, do CONARQ. O objetivo do modelo é orientar a implantação da gestão arquivística de documentos, fornecer especificações técnicas e funcionais e metadados para orientar a aquisição e/ou desenvolvimento de sistemas informatizados, independentemente da plataforma tecnológica em que forem desenvolvidos e/ou implantados.

A parte II da versão 1.1 do e-ARQ Brasil elenca os metadados a serem associados aos documentos, a fim de apoiar a gestão, a preservação e a presunção de autenticidade dos documentos arquivísticos. A especificação dos metadados considera as seguintes entidades: documento, evento de gestão, classe, agente, componente digital e evento de preservação.

O e-ARQ Brasil deve ser levado em consideração para a implementação dos repositórios arquivísticos digitais, já que a integração dos repositórios aos sistemas informatizados de gestão arquivística de documentos (SIGADs) é fundamental para o sucesso das iniciativas de gestão.

III.9 – Protocolo para coleta de metadados – *OAI-PMH*

FONTE:
Open Archives Initiative Protocol for Metadata Harvesting – OAI-PMH, Version 2.0 – Open Archives Initiative: junho de 2002.

É um protocolo para coleta de metadados que permite a interoperabilidade entre repositórios. Está baseado nas normas abertas *HTTP* e *XML*, e visa a facilitar a disseminação eficiente de conteúdo. O *OAI-PMH*[17] não realiza pesquisas em dados, mas possibilita a reunião dos dados num só lugar.

III.10 – Padrão de codificação e transmissão de metadados – *METS*

FONTE:
METS – Metadata Encoding & Transmission Standard.

É um esquema *XML* que permite a codificação e o intercâmbio dos metadados descritivos, administrativos e estruturais relativos a objetos digitais. Trata-se de um padrão para empacotamento que permite organizar, em um único arquivo compactado, tanto os dados quanto os metadados. Atualmente, o *METS* é mantido pela Biblioteca do Congresso dos EUA (*Library of Congress*), que

17 Disponível em: http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm.

mantém um sítio oficial.[18]

Algumas implementações do padrão *OAIS* utilizam o *METS* para estruturar os pacotes *SIP*, *AIP* e *DIP*. Além disso, alguns repositórios digitais utilizam o *METS* para intercâmbio de objetos.

A estrutura de um pacote *METS* é definida por um modelo que detalha os elementos para a estruturação de um objeto digital. Basicamente, um pacote *METS* possui, obrigatoriamente, um cabeçalho (*header*) e até seis seções, que abrangem os metadados descritivos, os metadados administrativos, a lista de arquivos do pacote, seus relacionamentos e comportamento.

O *METS* é neutro em relação aos formatos de metadados encapsulados no pacote digital. Dessa forma, é necessário que se defina como o pacote digital será estruturado, estabelecendo os formatos de metadados que serão utilizados. Essa configuração-padrão é denominada "perfil de aplicação". Como boa prática, deve-se, antes de criar uma novo perfil para um determinado repositório, pesquisar se os perfis existentes atendem aos requisitos desejados.

III.11 – Descrição arquivística codificada – *EAD*

FONTE: EAD – Encoded Archival Description: dezembro de 2002.

Trata-se de uma codificação desenvolvida e utilizada para a descrição de metadados arquivísticos baseados na linguagem de marcação *XML*. O projeto, iniciado na Universidade da Califórnia em 1993, teve como base o padrão *MARC* (*machine-readable cataloging*), dando origem à "EAD.DTD", que foi publicada, em sua versão 1.0, em 1998, e consolidada em dezembro de 2002.[19] A versão vigente atualiza e incorpora metadados relacionados aos padrões de metadados *MARC*, *ISAD(G)* e *Dublin Core*.

A *EAD* permite a descrição, estruturação e interoperabilidade dos metadados arquivísticos referenciais, que, quando associados ao *XML*, possibilitam a decodificação e a apresentação das informações referenciais de forma estruturada aos usuários. O padrão *EAD*, atualmente, é mantido pelo *Network Development and MARC Standards Office* da Biblioteca do Congresso dos EUA, em parceria com a *Society of American Archivists*.

18 Disponível em: http://www.loc.gov/standards/mets.
19 Disponível em: http://www.loc.gov/ead.