

Relato

Modelo Câmara dos Deputados de Preservação Arquivística Digital

The Chamber of Deputies Digital Archival Preservation Model

Marcelo Carneiro da Fontoura | Mestre em Ciência da informação e Arquivista pela Universidade de Brasília, Analista Legislativo, especialização Arquivista, da Câmara dos Deputados, Brasil. E-mail: marcelo.fontoura@camara.leg.br. Currículo Lattes: <http://lattes.cnpq.br/7518546393794780>.

Márcia Ênia Lopes de Sousa | Especialista em Administração de Arquivos e Documentação pela Universidade Federal de Alagoas, graduada em Arquivologia pela Universidade de Brasília, licenciada em Letras pela Universidade Paulista. Analista Legislativo, especialização Arquivista, da Câmara dos Deputados, Brasil. E-mail: marcia.enia@camara.leg.br. Currículo Lattes: <https://lattes.cnpq.br/273271179631946>.

Jose Raymundo Ribeiro Campos Filho | Mestre em Administração de Empresas pela Universidade Federal de Pernambuco, graduado em Administração de Empresas pela Universidade de Pernambuco, doutorando em Administração Pública pelo Instituto Superior de Ciências Sociais e Políticas (ISCSP) da Universidade de Lisboa. Analista Legislativo da Câmara dos Deputados, Brasil. E-mail: joseraymundo.campos@camara.leg.br

Thiago Almeida Rodrigues Borges | Graduado em Arquivologia pela Universidade de Brasília, mestrando em Ciência da Informação pela Universidade de Brasília. Analista Legislativo, especialização Arquivista, da Câmara dos Deputados, Brasil. E-mail: thiago.borges@camara.leg.br.

Darlan Eterno Silvério de Sousa | Especialista em Inteligência Organizacional e Competitiva, graduado em Arquivologia pela Universidade de Brasília. Analista Legislativo, especialização Arquivista, da Câmara dos Deputados, Brasil. E-mail: darlan.sousa@camara.leg.br.

Laila Monaiar | Graduada em Arquivologia pela Universidade de Brasília. Analista Legislativo, especialização Arquivista, da Câmara dos Deputados, Brasil. E-mail: laila.monaiar@camara.leg.br.

Resumo

O Modelo Câmara dos Deputados de Preservação Arquivística Digital (Modelo CD-PAD) foi desenvolvido no âmbito do projeto de revisão e consolidação das ações de preservação digital implementadas na Câmara dos Deputados a partir de 2012. Esse Modelo prevê a aplicação integral do conjunto de práticas arquivísticas (já adotado pela Coordenação de Arquivo ao acervo em suporte físico) à gestão do acervo digital. Objetivo: apresentar uma visão geral do Modelo proposto e do atual estágio de implementação. Metodologia: trata-se de relato de experiência no qual se registram as ações implementadas e se analisam os resultados alcançados. Resultados: foram definidos os padrões de pastas pré-SIP (*Submission Information Package*) e AIP (*Archival Information Package*), ambas consideradas caixas-arquivo digitais e utilizadas para armazenamento e movimentação do acervo digital. Foi implementado o Repositório Institucional de Preservação Digital da Câmara dos Deputados (RPD-CD), previsto em Ato da Mesa n.º 48/2012. Este Modelo apresenta, em detalhes de alto nível, a implementação da gestão de documentos arquivísticos digitais, bem como sua preservação de longo prazo na Câmara dos Deputados. A presente divulgação objetiva contribuir para o desenvolvimento e a implementação de ações tecnológicas inovadoras pelas instituições públicas brasileiras.

Palavras-chave: RD-Arq. Repositório de Preservação Digital. Repositório Arquivístico. Preservação Digital.

Abstract:

The Chamber of Deputies Digital Archival Preservation Model (CD-PAD Model) was developed within the scope of a project aimed at reviewing and consolidating digital preservation actions implemented by the Chamber of Deputies since 2012. This model envisions the comprehensive application of archival practices (already adopted by the Archive Coordination regarding the physical collections) in managing digital holdings. This article intends to provide an overview of the proposed Model and its current implementation status through the experience of the involved employees, documenting the implemented actions and analyzing the achieved results. The CD-PAD Model defines standards for pre-SIP (Submission Information Package) and AIP (Archival Information Package) folders, both considered digital archival boxes used for storing and transferring digital holdings. Additionally, the Institutional Digital Preservation Repository of the Chamber of Deputies (RPD-CD), as stipulated in Resolution number 48/2012, was successfully implemented. This article presents in detail the implementation of digital archival documents management policies and its long-term preservation within the Chamber of Deputies. By disseminating this information, the article seeks to contribute to the development and adoption of innovative technological practices in Brazilian public institutions.

Keywords: DR-Arc. Digital Preservation Repository. Archival Repository. Digital Preservation.

1. Introdução

As atividades de recebimento e tratamento documental do Arquivo da Câmara dos Deputados remontam a 1823, quando a Secretaria da Assembleia foi incumbida da guarda e da gestão dos documentos legislativos e administrativos produzidos durante a realização da *Assembleia Geral Constituinte e Legislativa do Império do Brasil de 1823*.

Além dos acervos arquivísticos próprios do fundo arquivístico da Câmara dos Deputados (produzidos e recebidos a partir da primeira legislatura (1826-1829) e da primeira Assembleia Constituinte Brasileira de 1823), o Arquivo da Câmara mantém sob sua custódia e gestão o acervo documental das Assembleias Constituintes de 1891, 1934, 1946 e 1988. O acervo de todos os fundos equivale atualmente a 110.000 (cento e dez mil) caixas-arquivo, o que corresponde a aproximadamente 11.000 metros lineares e é constituído por documentos dos mais diferentes suportes e formatos: de documentos manuscritos do início do século XIX até os modernos documentos nato-digitais atuais. São 200 anos de história devidamente registrados, tratados, armazenados e gerenciados, conforme os mais rigorosos processos arquivísticos de gestão, descrição e difusão.

Nos últimos 20 anos, a ampliação dos processos de digitalização de acervos físicos e o aumento da produção de documentos nato-digitais trouxeram novos desafios e necessidades técnicas para a gestão de um acervo digital que, atualmente, ocupa 170 *terabytes* e está sob custódia do Arquivo da Câmara, em processo de tratamento e preparação para preservação. A falta de soluções prontas, a necessidade de adaptar modelos e as propostas variadas sobre esse assunto levaram as equipes da Coordenação de Arquivo (COARQ) e da Coordenação de Preservação de Conteúdos Informativos (COBEC) a pesquisar cada passo necessário para ampliar o conhecimento em torno da produção, gestão e preservação arquivística digital de longo prazo.

O trabalho desenvolvido nos últimos dez anos foi baseado na legislação nacional sobre arquivos e no conjunto de Atos da Mesa promulgados pela Mesa Diretora da Câmara dos Deputados em 2012. Tais Atos resultaram de estudos realizados pela direção e pelo corpo técnico do Centro de Documentação e Informação, com a participação de servidores de diversas outras áreas da Casa. O Ato da Mesa n.º 48, de 16 de julho de 2012, que instituiu a Política de Preservação Digital da Câmara dos Deputados, art. 7º, *caput*, determina a criação e a manutenção de repositório institucional — aqui denominado Repositório Institucional de

Preservação Digital da Câmara dos Deputados (RPD-CD) — para a preservação digital, ao mesmo tempo em que incumbe à COARQ, em seu art. 10, a responsabilidade pelo depósito, gestão e consulta aos documentos desse repositório.

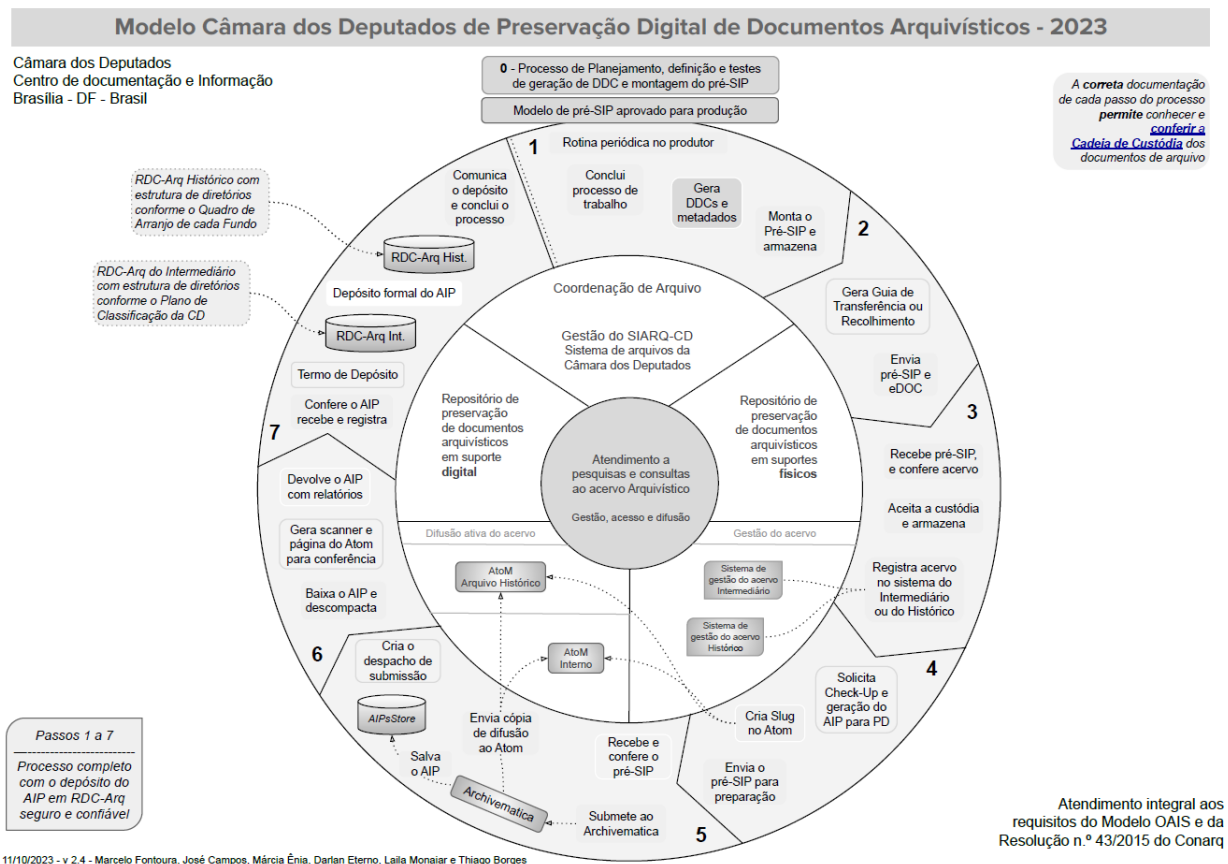
Além dessa legislação citada, o Modelo Câmara dos Deputados de Preservação Arquivística Digital (Modelo CD-PAD) utilizou, para sua constituição, a norma ISO 14721:2012 (traduzida para a ABNT NBR ISO 14721:2021) e as melhores práticas na área de preservação digital.

2. Visão geral do Modelo de Preservação

O Modelo CD-PAD, consolidado em 2023, é um conjunto sequencial de processos, graficamente representado pela figura de um círculo, formado por sete passos, que serão descritos detalhadamente na seção 6. O primeiro desses passos é realizado na fase corrente, quando ocorre a consolidação dos documentos digitais pelo órgão produtor, e o último passo é o efetivo depósito dos documentos arquivísticos no RPD-CD.

A Figura 1 seguinte (Representação Gráfica do Modelo da Câmara dos Deputados de Preservação Digital) ilustra o ciclo seguido pelo Modelo CD-PAD. A etapa zero trata da fase de planejamento, comum a cada vez que um setor produtor de documentos necessita transferir sua documentação digital para a Coordenação de Arquivo. Os passos de um a sete são outras etapas realizadas em diferentes fases do processo de preservação digital.

Figura 1 - Representação Gráfica do Modelo CD-PAD.



Fonte: produção própria.

Este Modelo, por adotar os conceitos constantes na ABNT NBR 14721:2021, utiliza para denominação dos pacotes de informação os seguintes termos: pacote de informação de submissão, pacote de informação de arquivamento e pacote de informação de disseminação. Uma vez que os termos correspondentes em inglês são amplamente utilizados pela comunidade nacional e internacional de preservação digital, utilizam-se seus acrônimos nessa língua. Deste modo, para o pacote de informação de submissão, utiliza-se o acrônimo SIP (*Submission Information Package*); para o pacote de informação de arquivamento, AIP (*Archival Information Package*); para o pacote de informação de disseminação, DIP (*Difusion Information Package*). Além desses pacotes, utiliza-se o termo pré-SIP para o pacote, ou pasta, elaborado a partir do produtor de documentos digitais (arquivo corrente) ou pelo próprio Arquivo, quando o documento está em sua custódia na fase intermediária ou permanente.

Ressalta-se que a organização, a movimentação e o armazenamento dos documentos digitais, em todos os passos do Modelo CD-PAD, pressupõem a utilização de armazenadores digitais (pastas pré-SIP e pastas AIP) que possuem a mesma finalidade e função das caixas-arquivo físicas utilizadas para organização, movimentação e armazenamento dos documentos em suportes físicos.

A premissa de organizar o acervo digital em caixas-arquivo digitais é uma recomendação do *The Consultative Committee for Space Data Systems* (CCSDS) às agências espaciais de diversos países (o Brasil estava representado pelo Departamento de Ciência e Tecnologia Aeroespacial, órgão integrante da Força Área), constante em seu documento *Reference Model for an Open Archival Information System* (OAIS)- CCSDS 650.0-M-2. Essa recomendação está contida no Modelo OAIS, Modelo considerado referência para instituições que tratam da preservação digital.

O Modelo CD-PAD, em complemento ao Modelo OAIS, atribui a denominação de pasta pré-SIP àquela criada para receber os documentos, tão logo sejam produzidos e/ou consolidados no órgão produtor. Essa pasta digital é utilizada para agrupar, armazenar e identificar os conjuntos documentais ou dossiês que serão objeto de transferência ou recolhimento. Isso permite que o arquivamento final na fase corrente e a ação de transferência de custódia com a movimentação dos documentos digitais ocorram quando eles já estão separados e armazenados em caixas-arquivo digitais. Essas caixas possuem conteúdo, registro e identificação inequívocos para o órgão produtor, são registradas na guia pertinente e podem ser conferidas durante o processo de recebimento e aceite pela área de Arquivo.

No momento de criação do pré-SIP, que conterà os documentos arquivísticos, será produzido um arquivo padronizado que contenha os metadados básicos de identificação de cada documento digital. Esses metadados serão posteriormente utilizados para cadastro no sistema de controle de acervo e para a produção da página de descrição arquivística de cada documento no sistema de difusão de acervo. Essa página ficará abaixo da descrição agregadora (criada previamente pela equipe do Arquivo) e permitirá a busca e o acesso à versão de consulta do documento (DIP). A difusão da cópia de acesso é outra característica do Modelo a ser ressaltada, pois ela ocorre automaticamente durante as atividades de verificação e criação da caixa-arquivo digital de guarda de longo prazo (a pasta AIP). Após essas atividades, conferem-se novamente a forma e o conteúdo da pasta AIP, último passo antes do registro em sistema de controle do acervo e a produção do *Termo de Depósito*, o que finaliza o depósito no RPD-CD.

Para essa etapa final de depósito da pasta AIP, o Modelo CD-PAD implementa, de forma inovadora, um Repositório de Preservação Digital que atende integralmente aos requisitos arquivísticos expressos na Resolução n.º 43/2015 do CONARQ, cujas características serão analisadas na seção seguinte.

Todas as ações e as atividades realizadas, para cada conjunto documental do acervo, são registradas individualmente em um processo eletrônico, que é produzido e tramitado no ambiente digital do Sistema

Informatizado de Gestão Arquivística de Documentos (SIGAD) interno da Câmara dos Deputados. Esse registro permite o rastreamento, a qualquer tempo, das ações de criação, destinação, movimentação, mudança e aceite de custódia. Os procedimentos técnicos realizados e o registro citado permitem mapear e comprovar a “cadeia de custódia”¹ relativa a cada documento sob a responsabilidade jurídica (temporária ou definitiva) do Arquivo.

Por fim, e após a conclusão da tramitação do processo eletrônico no SIGAD (com a anexação do Termo de Depósito referente aos documentos que estão na pasta AIP), realiza-se sua consolidação e armazenamento na mesma pasta na qual o AIP foi salvo. Esse é o documento que permitirá identificar a procedência do acervo e todos os atos procedimentais a ele relacionados, desde a sua produção até seu efetivo depósito no RPD-CD, para a guarda e a preservação de longo prazo.

O Modelo CD-PAD (desenvolvido pelo Centro de Documentação e Informação da Câmara dos Deputados (CEDI), acima descrito) orienta os demais setores sobre os procedimentos técnicos necessários para a garantia da preservação digital de longo prazo. Isso possibilita que mais conjuntos documentais sejam analisados e os pré-SIPs sejam adequadamente projetados, testados e modelados. Essa é a resposta das áreas técnicas de gestão arquivística para receber e tratar adequadamente esse acervo, com vistas ao devido processamento e preservação digital de longo prazo.

3. O Repositório Institucional de Preservação Digital da Câmara dos Deputados (RPD-CD)

Em atendimento ao disposto no Ato da Mesa n.º 48/2012 e sob a orientação da Diretoria-Geral (DG), o CEDI, em conjunto com a Diretoria de Tecnologia da Informação e Inovação (DITEC), efetivou, em maio de 2023, a criação do RPD-CD, ação final para a implementação completa do Modelo CD-PAD.

Ainda segundo esse Ato, cabe à COARQ a responsabilidade pelas ações de gestão do RPD-CD e depósito do acervo arquivístico. Sendo assim, conforme estabelecido na Resolução n.º 43/2015 do CONARQ e nas práticas recomendadas no Modelo OAIS, foram criados repositórios arquivísticos de preservação digital com estrutura multinível, o que permite manter a relação orgânica entre os documentos e a efetiva gestão arquivística em longo prazo.

O caráter arquivístico desses repositórios é atribuído em função da estrutura de pastas de diretórios criada para a organização do acervo em cada um deles: Repositório Arquivístico do Arquivo Intermediário e Repositório Arquivístico do Arquivo Permanente.

O Arquivo Intermediário custodia o acervo produzido ou recebido pela Câmara dos Deputados desde 1946 até a atualidade, sendo um fundo aberto. Constitui-se de documentos arquivísticos que cumprem a fase intermediária de guarda. Cumprida essa fase, parte desse acervo seguirá para o Arquivo Permanente, em atendimento à legislação arquivística e a normas internas.

O Arquivo Permanente se compõe de documentos destinados à guarda permanente. Ele custodia os fundos fechados da Câmara de 1826 até 1945 e de cada uma das Assembleias Constituintes de 1823, 1891, 1934, 1946 e 1988 (cada Assembleia constitui um fundo).

O Repositório Arquivístico do Arquivo Intermediário preserva os documentos digitais de guarda intermediária, com longos prazos de guarda. Cita-se o exemplo do dossiê nato-digital da folha de pagamentos, com prazo de guarda de 100 anos. Esse repositório utiliza uma estrutura/árvore de diretórios organizada

¹ A cadeia de custódia é entendida conforme expresso no art. 158-A do Código de Processo Penal Brasileiro e pág. 21 do Dicionário de Terminologia Arquivística, Camargo e Bellotto, 1996.

de acordo com o Plano de Classificação Funcional do Fundo Câmara (disponível no Portal da Câmara dos Deputados), que define níveis principais e seus respectivos subníveis, e é acessível pelo *Explorador de Arquivos do Windows*. Essa estrutura é replicada no sistema de acesso interno (AtoM), exclusivo para usuários cadastrados nas seções transferidoras de documentos e para o Arquivo Intermediário.

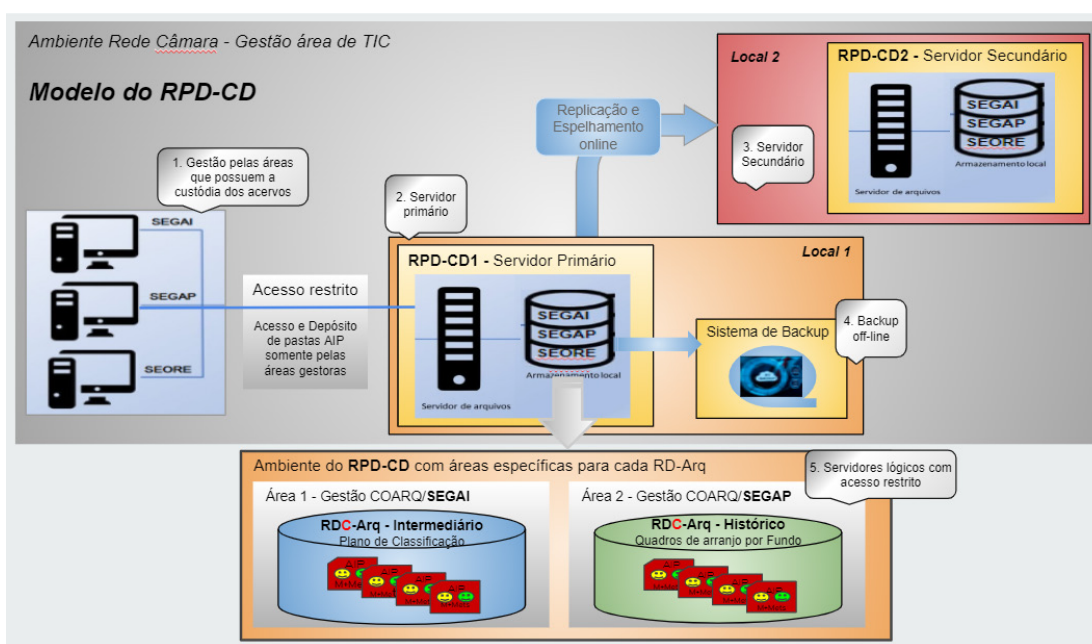
O Repositório Arquivístico do Arquivo Permanente preserva os documentos digitais de guarda permanente, inalienáveis e imprescritíveis. Esse repositório utiliza uma estrutura/árvore de diretórios construída de acordo com o *Quadro de Arranjo* elaborado para cada um dos fundos de arquivo citados anteriormente. Essa estrutura é replicada no sistema de difusão de acervos (AtoM), de acesso público.

4. Características técnicas e estrutura do RPD-CD

O RPD-CD está implementado em servidor de arquivo específico para esse fim, que estabelece regras rígidas de criação e gestão de usuários e possui controle de acesso por trilha de auditoria permanente. Possui espelhamento integral *on-line* em parques tecnológicos distintos, situados em endereços e edificações diferentes, de propriedade da Câmara dos Deputados. Possui plano para recuperação de desastres, com programa de cópia de segurança (*backup*) especialmente modelado para ele.

Nesse servidor, foram criadas duas áreas distintas para armazenamento de acervo arquivístico, conforme pode ser observado na *Figura 2*.

Figura 2 - Diagrama do RPD-CD



Fonte: produção própria.

O RPD-CD é composto por duas áreas:

- Área 1- RD-Arq-int: é o Repositório Arquivístico de Preservação Digital do Arquivo Intermediário. Possui estrutura de diretórios multinível, conforme previsto no Plano de Classificação Funcional da CD, tendo como pasta raiz “\BR-DFCD-INT”.
- Área 2 - RD-Arq-hist: é o Repositório Arquivístico de Preservação Digital do Arquivo Permanente. Constituído por uma estrutura de diretórios multinível, conforme definido nos Quadros de Arranjo de cada fundo custodiado.

O depósito e a recuperação dos documentos digitais podem ser realizados sem a necessidade de uso de um sistema informatizado (por exemplo, um banco de dados) entre o usuário e a pasta AIP que deseja, utilizando-se a simples navegação entre os diretórios, por meio do Explorador de Arquivos do sistema operacional. A estrutura hierárquica de diretórios, pastas e subpastas replica os instrumentos de classificação arquivísticos citados anteriormente, o que permite manter a relação orgânica entre os documentos digitais depositados, realizar a gestão e a destinação de documentos (após o cumprimento dos prazos precaucionais), nos moldes do que preceitua a Resolução n.º 43/2015 do CONARQ.

5. Passos e atividades do Modelo CD-PAD para um pré-SIP testado e aprovado

O Modelo CD-PAD possui um passo inicial específico e voltado à análise de cada conjunto documental, denominado passo zero. A conclusão desse passo ocorre com a aprovação do Plano de Preservação e a assinatura do *Termo de Acordo de Produção de pré-SIP*.

Concluído esse passo, a área produtora de documentos pode dar início ao processo que possui sete passos, descritos a seguir:

1. rotina periódica de criação de documentos no produtor;
2. produção de guia e envio do pré-SIP;
3. recebimento e aceite da custódia pelo Arquivo Intermediário ou pelo Arquivo Permanente;
4. solicitação de checagem e preparação para preservação;
5. realização de checagem, publicação no AtoM e criação do AIP;
6. baixa e conferência do AIP; e
7. registro no sistema de acervo, depósito no RPD-CD e exclusão de cópias intermediárias.

Cada passo será descrito de forma mais detalhada a seguir, mas deve-se observar que, apesar de ser um processo circular, ele não é necessariamente contínuo para cada conjunto de documentos produzidos. Entre alguns passos, deverá ser cumprido o prazo de guarda definido na *Tabela de Temporalidade* (TT) da Câmara dos Deputados, o que resulta em retenção dos documentos em repositórios intermediários.

Os passos seis e sete (conferência e efetivo depósito do AIP no RPD-CD) são implementações inovadoras do Modelo CD-PAD. O passo sete se constitui em ato administrativo formal relativo ao depósito no Repositório Arquivístico Digital (RD-Arq), considerado fundamental na Câmara dos Deputados e marco inicial da fase de preservação digital de longo prazo ou permanente.

Visando ampliar a capacidade de processamento dos pré-SIPs, criação de AIPs e depósitos no RD-Arq, está em fase de desenvolvimento um *Sistema de Automatização de Procedimentos*, cujo projeto será identificado internamente como SAP-RPD. Essa automatização visa também atender às normas e às práticas arquivísticas de registro e documentação dos procedimentos (conforme adotado pela Câmara dos Deputados), o que possibilita a gestão do acervo.

Passo 0 - Planejamento, definição e testes de geração de documento digital e montagem do pré-SIP

Essa atividade (considerada o passo zero) é vital para a execução de todo o processo e tem como foco caracterizar cada conjunto documental a ser tratado pelo Arquivo. Em determinadas situações, o formato, a dimensão, a forma de produção e mesmo a qualidade do original podem indicar soluções e

propostas diferentes de criação de um pré-SIP, o que resulta em um *Plano de Preservação* específico para cada série documental.

Esse processo de trabalho deve ser realizado de forma conjunta entre a equipe do Arquivo, do setor produtor e da equipe de informática. Nesse momento, será planejada a forma de consolidação do documento, a estrutura do pré-SIP e a forma de extração dos metadados. Esse planejamento definirá o padrão de criação do documento e da caixa-arquivo digital utilizada para a movimentação inicial dos documentos digitais do produtor para o Arquivo Intermediário ou Permanente.

Nessa etapa de planejamento, definição e testes deverão ser definidos itens como:

- formação do nome do pré-SIP: Deverá ser único e fazer sentido para o produtor; será utilizado também para a criação do slug no AtoM e para identificar o AIP;
- formação e nome dos documentos originais: Deve possuir um identificador único (apesar de possibilitar nome/título duplicado) e não poderá conter espaços ou caracteres especiais;
- formatos dos arquivos: Na definição do formato e tamanho do documento original, foco da preservação, deverão ser atendidos os requisitos mínimos de qualidade e possibilidade de leitura e/ou migração a longo prazo. Caso não seja possível criar um original em formato de amplo uso e conhecimento, deverá ser prevista a criação de cópia em formato aberto para ampliar a possibilidade de leitura a longo prazo;
- estrutura de pastas e número de arquivos em cada pré-SIP;
- forma de extração/criação e a regra para criação do arquivo `metadata.csv` (valores separados por vírgulas), onde devem ser registrados os metadados de identificação e difusão;
- por último, deverá ser definido se ocorrerá ou não a geração de cópia de baixa resolução para difusão no AtoM. Em caso positivo, deverá ser definido seu formato e revisados os metadados de identificação, pois passam também a ser metadados de difusão.

O arquivo `metadata.csv` utiliza os 15 elementos básicos do padrão Dublin Core na versão 1.1, que são lidos pelo AtoM para a criação da página de descrição arquivística do documento. Caso seja necessária a inclusão de outros metadados no AtoM, esse processo deverá ser feito posteriormente, de forma manual, e em complemento ao preenchimento automático com esses elementos.

Os testes de verificação dos documentos, difusão no AtoM e criação da pasta AIP deverão ser realizados primeiro no ambiente de testes e, se aprovados, serão novamente realizados no ambiente de homologação, antes de serem finalmente liberados para execução contínua no ambiente de produção. Diversos testes anteriores (realizados durante meses pelas equipes envolvidas) evidenciaram a necessidade de criação desses três ambientes distintos, independentes e funcionais, para o uso efetivo e integrado dos *softwares Archivematica* e AtoM.

O envio de descrições para o ambiente de produção do AtoM (sem a realização prévia de experimentos nos ambientes de teste e homologação) poderá resultar na necessidade de exclusão de descrições e arquivos, o que traz o risco de exclusões acidentais ou a manutenção de informações incorretas ou duplicadas no ambiente de produção.

Ao se utilizar, adequadamente, cada um dos ambientes em sua função específica (testes, homologação e produção) e ao se obter a aprovação de um padrão para a criação do pré-SIP, considerar-se-á definida sua estrutura e conteúdo. Citam-se os seguintes exemplos de pré-SIPs padronizados: arquivos da folha de pagamento de um mês/ano; diário da Câmara dos Deputados; captura de vídeo de uma determinada sessão

plenária; arquivos de um Projeto de Lei específico; o produto da digitalização de um códice manuscrito constituído de arquivos de imagens de cada página, em alta resolução, e o respectivo arquivo em PDF de baixa resolução, onde essas imagens estão dispostas na forma de uma publicação única.

O produto final dessa etapa é um *Termo de Acordo de Produção de pré-SIP* (caixa-arquivo digital inicial), no qual constará um conjunto de regras e procedimentos de consolidação de documentos, criação do arquivo de metadados e estrutura do pré-SIP, tudo isso definido de comum acordo e aprovado por um grupo de trabalho constituído pelas áreas de produção de documentos, Arquivo, Preservação e Tecnologia da Informação.

Após a conclusão de todos os testes e a definição do padrão de pré-SIP, o produtor poderá iniciar a produção de pacotes de informação, o que iniciará o ciclo dos sete passos do Modelo CD-PAD.

Passo 1 - Rotina periódica de produção de documentos no produtor

Essa atividade inicial deve levar em conta as definições previamente estabelecidas no *Termo de Acordo de Produção de pré-SIP*. Nesse documento, consta a forma de extração e consolidação do documento de arquivo, a descrição da estrutura do pré-SIP e do arquivo de metadados. Essa atividade inicial será realizada de forma esporádica (diária ou mensalmente), a depender da produção efetiva do conjunto documental gerado. É nesse ponto que ocorre a consolidação dos originais e a produção do arquivo de metadados. Como atividade final, deve ocorrer a montagem e conferência da estrutura e conteúdo do pré-SIP. Esse passo poderá ser realizado com auxílio do SAP-RPD, que visa validar a estrutura e gerar os documentos de controle e aprovação.

Passo 2 - Produção de guia e envio do pré-SIP

Quando cumprido o prazo de guarda do documento digital no Arquivo Corrente, realizar-se-á a devida destinação, conforme a TT em vigor. Nesse momento, o Arquivo poderá ser novamente consultado para a conferência do acervo, validação final do pré-SIP e produção da Guia de Transferência (GT) ou da Guia de Recolhimento (GR) no SIGAD.

Após elaboração no SIGAD da GT ou da GR, será realizado o contato com o Arquivo, que receberá a custódia do acervo para sua efetiva disponibilização. Essa atividade também será auxiliada pelo SAP-RPD.

Passo 3 - Recebimento e aceite da custódia pelo Arquivo Intermediário ou Permanente

O Arquivo Intermediário ou o Permanente, ao ser notificado pelo produtor sobre a formalização da disponibilização do acervo, realizará a conferência da forma e do conteúdo do pré-SIP, manualmente ou com auxílio do SAP-RPD. Caso exista divergência, não se realizará a transferência de custódia e ajustes serão solicitados. Caso o pré-SIP seja aprovado, realizar-se-á sua movimentação para uma área virtual de armazenamento temporário do Arquivo.

Cada Guia e seus respectivos pré-SIPs deverão ser registrados no sistema de controle do acervo, da mesma forma que se faz com o acervo físico, onde cada pré-SIP corresponde a uma caixa-arquivo física.

Da mesma forma, no Arquivo Corrente, é possível que, por razões legais ou procedimentais, exista um prazo entre o recebimento do pré-SIP e a solicitação de verificação e preparação para preservação. Confirmada a existência desse prazo, o acervo ficará armazenado em área específica de tratamento e guarda temporária da área de Arquivo.

Passo 4 - Solicitação de checagem e preparação para preservação

Após o cumprimento do prazo, caso exista, e concluídas as ações de tratamento da documentação presente no pré-SIP, será produzido um documento formal de solicitação de realização dos procedimentos de checagem, difusão e criação da pasta AIP. Esse documento será encaminhado ao Arquivo por meio do SIGAD, para aceite da operação e captura do acervo disponibilizado.

No caso de difusão via AtoM, antes da solicitação de checagem, o Arquivo criará previamente uma página de descrição do item agrupador dos documentos do pré-SIP. Esse procedimento antecede a admissão do pré-SIP no Archivematica (sistema informatizado adotado para a preparação do AIP e do DIP, bem como para o armazenamento temporário desses pacotes), pois é necessário informar o *slug* do AtoM ao qual será vinculada a descrição dos documentos presentes no pré-SIP, em campo próprio desse aplicativo, o que possibilita a automatização do envio do DIP resultante. Dessa forma, a área de Arquivo deverá identificar em qual ponto da estrutura existente no AtoM os arquivos do pré-SIP deverão ser publicados como DIP. Essa vinculação é importante, pois essa mesma estrutura de níveis hierárquicos do AtoM será replicada no RD-Arq, para facilitar a localização e o arquivamento da pasta AIP.

Passo 5 - Realização de checagem, publicação no AtoM e criação do AIP

Ao receber o SIGAD, o Arquivo deverá se certificar de que os documentos listados na Guia estejam adequadamente disponibilizados na área virtual de transferência, para a devida captura e tratamento. Dessa forma, para cada pré-SIP disponibilizado, a conferência deve ser realizada visando garantir a correta passagem de custódia sobre os documentos submetidos ao Archivematica. Ao receber o *Processo de Solicitação de Preparação para Preservação* no SIGAD, a equipe responsável confere as informações constantes em formulário próprio e captura o pré-SIP constante na área citada.

Caso seja encontrado algum problema na preparação do pré-SIP disponibilizado, a área que solicitou o preparo para preservação será notificada para realizar os ajustes necessários. Concluídos os ajustes, o pré-SIP será novamente disponibilizado para nova verificação.

O processo de admissão do pré-SIP se inicia no *software* Archivematica, que realiza diversos microsserviços, em duas etapas. A primeira é identificada como *Transferência* (do inglês *Transfer*), onde se produz o SIP a partir do pré-SIP, caso todos os documentos sejam aprovados. Concluída essa etapa inicial, é realizada a segunda etapa, denominada *Admissão* (do inglês *Ingest*), quando os documentos do SIP são submetidos a outro conjunto de microsserviços, do qual resultam o DIP (pacote publicado no AtoM) e o AIP (pacote de arquivamento), que será baixado para conferência e envio para o RPD-CD.

O processamento realizado pela plataforma Archivematica é constituído por uma sequência de microsserviços de conferência, testes de integridade e validação de cada um dos arquivos originais a serem preservados. Ressalte-se que, durante a realização desses microsserviços, é criado o arquivo METS-UUID-da-Etapa.xml (onde são registrados todos os metadados do processamento citado, bem como as informações sobre os arquivos submetidos), considerado o arquivo de metadados da preparação para preservação. Esse arquivo METS-UUID-da-Etapa.xml é salvo na pasta *"/data"* do AIP e pode ser consultado para análise e verificação de todos os processos realizados, visando atestar a integridade dos documentos preservados.

Segue abaixo a relação dos microsserviços que são realizados durante o processamento técnico de preparação para preservação dos documentos presentes nos pré-SIPs, utilizando-se o sistema Archivematica, versão 1.12.2.

a) Microsserviços executados na etapa *Transferência* (de pré-SIP para SIP).

1. Verificar a conformidade: verifica se a transferência está devidamente estruturada.
2. Renomear com UUID²: atribui um identificador universal exclusivo para essa etapa.
3. Atribuir UUIDs de arquivo e somas de verificação aos objetos: atribui um identificador universal exclusivo e somas de verificação para cada objeto digital.
4. Verificar somas de verificação: verifica todas as somas de verificação dos arquivos incluídos.
5. Produzir documento METS-UUID.xml: cria um arquivo METS do tipo .xml, que captura a ordem original da transferência. Esse arquivo é adicionado ao SIP e seu respectivo AIP.
6. Verificar vírus: verifica existência de vírus ou arquivos de *malware*.
7. Produzir relatório de estrutura: elabora uma árvore de diretórios em .txt da transferência original e o inclui no AIP.
8. Alterar nomes de arquivos: remove caracteres proibidos de pastas e nomes de arquivos como, por exemplo, & (“e” comercial).
9. Identificar formato de arquivo: identifica o tipo e o formato de arquivo e verifica se a extensão do arquivo é compatível com o conteúdo.
10. Caracterizar e extrair metadados: extrai os metadados técnicos incorporados a cada arquivo, para registro no METS.
11. Validar arquivos: valida os formatos de arquivo em relação à especificação do formato contida nos metadados técnicos.
12. Criar SIP ao final da etapa: concluídos os microsserviços acima, os arquivos do pré-SIP e os outros arquivos citados formam o arquivo SIP, armazenado em área de trabalho temporária do Archivematica, até que o responsável pela admissão autorize os próximos microsserviços constantes na etapa *Admissão*.

b) Microsserviços executados na etapa de *Admissão* (de SIP para DIP e AIP).

1. Verificar conformidade do SIP: verifica se o SIP está organizado de acordo com a estrutura de pastas exigida pelo Archivematica.
2. Renomear diretório SIP com o UUID: altera o UUID da etapa anterior do SIP para o UUID que identifica essa etapa. Tanto o UUID da etapa anterior quanto o UUID dessa etapa serão capturados no AIP final, compondo o nome dos arquivos METS.
3. Normalizar: cria, caso seja previsto, as derivadas (cópias) dos objetos digitais em formatos abertos, conforme o plano de preservação.
4. Processar arquivos normalizados manualmente: processa arquivos que foram normalizados antes da transferência ou permite a normalização manual.
5. Verificar políticas de produção de derivadas (cópias): verifica se as derivadas (cópias) de acesso e de leitura de longo prazo (criadas durante a normalização) atenderam ao definido no Registro da Política de Formatos³.
6. Processar documentação de envio: processa qualquer documentação complementar sobre os originais que tenha sido incluída no SIP e a adiciona ao diretório “/objects” do AIP.

² *Unique Universal Identifier* (UUID) ou, em tradução livre, Identificador Universal Único.

³ *Format Policy Registry* (FPR) ou, em tradução livre, Registro da Política de Formatos.

7. Produzir AIP METS: produz o arquivo METS-UUID.xml, que registra todas as ações e microsServiços realizados durante o processamento técnico.
8. Preparar DIP: cria um DIP contendo cópias de qualidade adequada para difusão, miniaturas e uma cópia do arquivo METS-UUID.xml.
9. Preparar AIP: cria um AIP no formato *BagIt*⁴ e o compacta sem perdas; cria também um arquivo de ponteiro e indexa o AIP no *Archival Storage* do Archivematica.
10. Carregar DIP: sobe o DIP para o AtoM.
11. Armazenar DIP: armazena o DIP no *Storage Service*.
12. Armazenar AIP: produz uma cópia do AIP no diretório temporário do servidor e realiza as verificações dos cálculos padronizados do padrão BagIt: *verifyvalid*, *checkpayloadoxum*, *verifycomplete*, *verifypayloadmanifests*, *verifytagmanifests*. Após essas validações, produz os documentos de controle, compacta a pasta AIP com o software 7.zip e o armazena no *Storage Service*.

O AIP e o DIP são armazenados no Archival Storage do Archivematica. Esse local possui uma árvore de diretórios criada a partir de cada AIP, pela quebra do UUID aleatório que lhe foi atribuído durante a segunda etapa. Essa forma de organizar os AIPs/documentos não atende a qualquer requisito arquivístico, sendo uma solução meramente técnica para evitar que todos os AIPs fiquem em uma mesma pasta. Essa forma de armazenamento é considerada, pela Câmara dos Deputados, como temporária e fora dos padrões arquivísticos necessários para guarda em longo prazo, acesso e localização.

Após a conclusão das etapas descritas, produz-se um relatório, que informa todos os microsServiços realizados nos documentos originais que permitiram a devida publicação e criação da descrição arquivística no AtoM, bem como a criação da caixa-arquivo digital no formato BagIt, que armazenará os documentos originais aprovados em todos os testes realizados.

Passo 6 - Baixa e conferência da criação do AIP

Inicia-se a fase de conferência do AIP ao final da realização dos microsServiços e do armazenamento temporário no Storage Service do Archivematica.

Nessa atividade de verificação, o AIP (que é um arquivo compactado) será baixado para uma área de acesso restrito, descompactado e conferido integralmente.

Essa conferência consiste na confirmação de que os processos previstos, acima descritos, foram efetivamente realizados em todos os documentos do pré-SIP. Caso algum erro seja detectado, realiza-se uma perícia em busca do problema e da solução, após a qual se realiza outra conferência.

Tendo o AIP sido considerado conforme e aprovado pela equipe de tratamento técnico, são produzidos dois relatórios: um com o resultado da publicação no AtoM e outro que registra, em planilha de cálculo, o conteúdo integral do AIP com o somatório total de bytes de seus arquivos. Esse somatório poderá ser utilizado, juntamente com as somas de verificação criadas, para permitir a conferência da integridade dos documentos preservados por outras seções ou entidades, em qualquer tempo.

Esses dois documentos, além do despacho de conclusão da verificação, serão inseridos no processo eletrônico do SIGAD, que documenta a gestão do acervo pelo Arquivo. Ato contínuo, esse processo é tramitado de volta para a área que solicitou a preparação e o AIP (que foi conferido para análise, registro em sistema e depósito) é disponibilizado em área virtual de acesso ao Arquivo.

⁴ *BagIt* é uma especificação para empacotar diretórios de arquivos, hierarquicamente, para armazenamento a longo prazo ou para a transferência entre ambientes de armazenamento. Disponível em: http://wiki.ibict.br/index.php/Guia_do_Usu%C3%A1rio_-_Archivematica. Acesso em: 14 nov. 2023.

Passo 7 - Registro no sistema de acervo e depósito no RPD-CD

No passo sete, depósito formal dos AIPs, as normas, os conceitos e as práticas arquivísticas (desenvolvidos internamente e adotados pelo Arquivo da Câmara nos últimos 50 anos) foram implementados com sucesso. Além da criação do RPD-CD, considera-se necessária também a implementação dos requisitos arquivísticos expressos na Resolução n.º 43/2015 do CONARQ, que preconiza a criação de repositórios arquivísticos digitais confiáveis com estrutura hierárquica multinível, que permitam manter a relação orgânica entre os documentos e a efetiva gestão arquivística no longo prazo.

Dessa forma, e dando continuidade ao processo de gestão do acervo (após a verificação e o armazenamento na caixa-arquivo digital no padrão *BagIt*), o Arquivo acessa a área virtual de transferência interna, localiza o AIP e realiza a conferência final, conforme dados constantes no processo eletrônico do SIGAD.

Considerando que todos os processos solicitados foram realizados e devidamente documentados e aprovados, elabora-se o *Termo de Depósito de Documento em Repositório Arquivístico Digital*, no qual o responsável pela custódia do documento se identifica, registra dados sobre o AIP, data e o local exato onde será realizado o armazenamento de longo prazo da caixa-arquivo digital. Ato contínuo, o AIP será movimentado da área de transferência para a pasta específica, criada de acordo com o *Plano de Classificação* ou o *Quadro de Arranjo* do fundo correspondente.

Após conferência do *Termo de Depósito* e a ação de depósito por um segundo servidor público, esse Termo é anexado ao processo eletrônico no SIGAD, o que documenta a ação final de recebimento e armazenamento de longo prazo. Como atividade final, é realizada a exportação do processo eletrônico do SIGAD. Esse processo (consolidado em um arquivo .pdf) é salvo na mesma pasta onde foi depositado o AIP.

Com a conclusão de todos esses passos, considera-se que os documentos digitais foram devidamente produzidos, transferidos, gerenciados no arquivo, verificados, tiveram uma cópia de acesso disponibilizada no AtoM e foram armazenados em caixa-arquivo digital apropriada, o que efetiva o arquivamento de longo prazo em repositório arquivístico digital. O passo sete fecha o círculo do Modelo CD-PAD.

6. Histórico sobre Auditoria e certificação de Repositórios Digitais

A auditoria e a certificação de repositório digital confiável são assuntos tratados há pelo menos 30 anos. Em 1994, formou-se o Grupo de Trabalho constituído pela Comissão de Preservação e Acesso dos Estados Unidos (CPA)⁵ e *The Research Libraries Group* (RLG), o que resultou na publicação do relatório *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*, em 1996.

Posteriormente, em 2002, foi publicado o relatório *Trusted Digital Repositories: Attributes and Responsibilities* (TDR), redigido pelas instituições *The Research Libraries Group* (RLG) e Online Computer Library Center (OCLC).

Em 2003, constituiu-se uma força-tarefa formada pelo RLG e o *National Archives and Records Administration* (NARA), que publicou, em 2007, o relatório *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC), onde consta uma lista de verificação de critérios para auditoria e certificação de repositório digital confiável, o que engloba desde repositórios acadêmicos de preservação institucional até arquivos e bibliotecas nacionais.

Em 2011, o *The Consultative Committee for Space Data Systems* (CCSDS) publicou as *Recommendation for Space Data System Practices: audit and certification of trustworthy digital repositories*, CCSDS 652.0-M-1.

⁵ A Comissão de Preservação e Acesso foi uma organização privada sem fins lucrativos que atuava em nome das bibliotecas, arquivos e universidades dos Estados Unidos para desenvolver e incentivar estratégias colaborativas para preservar e fornecer acesso ao registro humano acumulado.

Segundo o CCSDS, essas recomendações fornecem embasamento para que as agências espaciais possam auditar e certificar seus repositórios digitais quanto à aderência aos requisitos listados.

Ambos os relatórios, TRAC e *Recommendation for Space Data System Practices: audit and certification of trustworthy digital repositories*, CCSDS 652.0-M-1, fundamentam o padrão ISO 16363:2012 (*Space data and information transfer systems. Audit and certification of trustworthy digital repositories*).

No Brasil, o CONARQ publicou a Resolução n.º 43/2015, que aborda as diretrizes para a implementação de repositórios arquivísticos digitais confiáveis, dentre outros assuntos. Esse órgão conceitua e lista critérios para um repositório arquivístico digital confiável, o que acrescentou o arquivístico ao conceito desenvolvido nos documentos citados. Com isso, o CONARQ inclui as técnicas próprias da ciência arquivística a esses repositórios.

Os repositórios digitais aqui descritos (que constituem Modelo CD-PAD) são arquivísticos, uma vez que atendem às normas emanadas pelo CONARQ. Em resumo, eles apresentam estrutura multinível, onde o Repositório Arquivístico do Arquivo Intermediário utiliza uma estrutura/árvore de diretórios organizada de acordo com o *Plano de Classificação Funcional* e o Repositório Arquivístico do Arquivo Permanente utiliza uma estrutura/árvore de diretórios construída de acordo com o *Quadro de Arranjo* elaborado para cada fundo de arquivo custodiado.

Um Grupo de Trabalho Multidisciplinar foi constituído para realizar uma autoauditoria, segundo os critérios listados na ISO 16363:2012 (*Space data and information transfer systems. Audit and certification of trustworthy digital repositories*) e aqueles elencados na Resolução n.º 43/2015 do CONARQ. Essa é uma ação que está em andamento.

Considerações finais

O Modelo CD-PAD está em produção e aperfeiçoamento, uma vez que alguns procedimentos descritos serão automatizados com a utilização de um *Sistema de Automatização de Procedimentos* (SAP-RPD), em desenvolvimento e, também pelo uso e aplicação de Inteligência Artificial (IA).

O Grupo de Trabalho Multidisciplinar realiza autoauditoria, com vistas à obtenção da certificação, segundo os critérios listados na ISO 16363:2012 (*Space data and information transfer systems. Audit and certification of trustworthy digital repositories*) e na Resolução n.º 43/2015 do CONARQ.

Por fim, esse Modelo consolida e apresenta os resultados de experiência aplicada na Câmara dos Deputados resultante da implementação de processos de trabalho e de sistemas informatizados de tratamento, difusão e preservação de documentos digitais nos últimos cinco anos. O relato visa contribuir para a análise, a reflexão teórica e a aplicação prática por todas as entidades públicas brasileiras interessadas em enfrentar os desafios de gestão arquivística, acesso e preservação digital no longo prazo.

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 14721**: sistemas espaciais de transferência e de informação- Sistema Aberto de Arquivamento de Informação (SAAI)- Modelo de referência. Rio de Janeiro: ABNT, 2021. 121p.

BELLOTTO, Heloísa Liberalli; CAMARGO, Ana Maria de Almeida. **Dicionário de Terminologia Arquivística**. São Paulo: AAB-Núcleo Regional de São Paulo, Secretaria de Estado da Cultura, Departamento de Museus e Arquivos, 1996.

BRASIL. Decreto-Lei 2.848, de 7 de dezembro de 1940. Código Penal. **Diário Oficial da União**: seção 1, Rio de Janeiro, 31 dez. 1940. Disponível em: <https://www2.camara.leg.br/legin/fed/declei/1940-1949/decreto-lei-2848-7-dezembro-1940-412868-norma-pe.html>. Acesso em: 14 nov. 2023.

BRASIL. Câmara dos Deputados. Ato da Mesa nº 48, de 16 de julho de 2012c. Institui a Política de Preservação Digital da Câmara dos Deputados. **Diário da Câmara dos Deputados, Suplemento**, p. 71, 17 jul. 2012. Disponível em: <https://www2.camara.leg.br/legin/int/atomes/2012/atodamesa-48-16-julho-2012-773828-publicacaooriginal-137097-cd-mesa.html>. Acesso em: 11 out. 2023.

BRASIL. Ministério da Justiça e Segurança Pública. Conselho Nacional de Arquivos. Resolução nº 43, de 4 de setembro de 2015. Altera a redação da Resolução do CONARQ nº 39, de 29 de abril de 2014, que estabelece diretrizes para a implementação de repositórios digitais confiáveis para a transferência e recolhimento de documentos arquivísticos digitais para instituições arquivísticas dos órgãos e entidades integrantes do Sistema Nacional de Arquivos - SINAR. **Diário oficial da União**: seção 1, Brasília, DF, ano 152, n. 171, p. 34, 8 set. 2015. Disponível em: <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=34&data=08/09/2015>. Acesso em: 8 nov. 2023.

CENTER FOR RESEARCH LIBRARIES (CRL); ONLINE LIBRARY CENTER (OCLC). **Trustworthy Repositories Audit & Certification (TRAC)**: criteria and checklist. Version 1.0. Chicago: Center for Research Libraries; Dublin: Online Computer Library Center (OCLC), c2007. 88 p. Disponível em: https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf. Acesso em: 8 nov. 2023.

COUNCIL OF THE CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). **Audit and Certification of Trustworthy Digital Repositories**: Recommended Practice, Issue 1, CCSDS, 652.0-M-1. Magenta Book. Washington, DC, USA. Council of the Consultative Committee for Space Data Systems, September 2011. 77 p. Disponível em: <https://public.ccsds.org/pubs/652x0m1.pdf>. Acesso em: 14 nov. 2023.

COUNCIL OF THE CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). **Reference Model for an Open Archival Information System (OAIS)**: Recommended Practice, Issue 2, CCSDS, 652.0-M-2. Magenta Book. Washington, DC, USA. Council of the Consultative Committee for Space Data Systems, June 2012. 135 p. Disponível em: <https://public.ccsds.org/pubs/650x0m2.pdf>. Acesso em: 14 nov. 2023.

WATERS, Donald; GARRETT, John. **Preserving Digital Information**. Report of the Task Force on Archiving of Digital Information. Washington, DC, USA. Commission on Preservation and Access, Washington, DC, and The Research Libraries Group, Inc., Mountain View, CA., May 1, 1996. Disponível em: <https://eric.ed.gov/?id=ED395602>. Acesso em: 14 nov. 2023.