



Capítulo 01

# **PRESERVAÇÃO DE DOCUMENTOS EM FORMATO DIGITAL**

**PRESERVATION OF DOCUMENTS IN  
DIGITAL FORMAT**

Miguel Ángel Márdero Arellano<sup>1</sup>



---

<sup>1</sup> Dados do autor: Doutor em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, miguel@ibict.br.



# A informação digital não sobreviverá acidentalmente.

BEAGRIE, 2006

**Resumo:** Examina-se a degradação dos objetos digitais e a demora na aplicação de estratégias para garantir sua preservação, assim como, a importância do emprego de técnicas focadas na manutenção das informações contidas nesses objetos. Neste capítulo são analisadas ações de preservação que em seu conjunto formam sistemas de informação integrados beneficiando a preparação e gestão dos conteúdos e coleções em cadeias de transferências entre repositórios e locais de arquivamento confiáveis. O papel da pesquisa nas áreas da Ciência da Informação é mencionado na especificação dos atributos dos registros e dados digitais que dão suporte à presunção de autenticidade. São apresentadas algumas respostas sistêmicas para a padronização da prática de preservação digital que priorizam o uso da infraestrutura governamental e a aplicação de um financiamento contínuo. A abordagem para os sistemas de preservação integrados e distribuídos em rede é recomendado para o tratamento dos objetos digitais, especificamente a sua vulnerabilidade física e lógica, parcerias e contratos e, no recrutamento de pessoal devidamente qualificado.

**Palavras-Chave:** Preservação digital. Objeto digital. Preservação digital distribuída. Autenticidade.

**Abstract:** It examines the degradation of digital objects and the delay in applying strategies to ensure their preservation, as well as the importance of employing techniques focused on maintaining the information contained in these objects. This chapter analyzes preservation actions that together form integrated information systems benefiting the preparation and management of contents and collections in transfer chains between repositories and reliable archiving locations. The role of research in the areas of Information Science is mentioned in the specification of the attributes of digital records and data that support the presumption of authenticity. Some systemic responses to the standardization of digital preservation practice are presented that prioritize the use of government infrastructure and the application of continuous funding. The approach to integrated and distributed network preservation systems is recommended for the treatment of digital objects, specifically their physical and logical vulnerability, partnerships and contracts, and in the recruitment of suitably qualified personnel.

**Keywords:** Digital preservation. Digital object. Distributed digital preservation. Authenticity.

# ◆ 1 INTRODUÇÃO

Todos os grupos humanos adotam tecnologias para prolongar a prática de crenças e tradições, e objetos produzidos com a tecnologia adotada no momento podem, também, causar mudanças nessas práticas. O governo, a administração pública e privada, a pesquisa científica e tecnológica e a expressão cultural dependem cada vez mais de documentos digitais, não disponíveis em outro formato, para o exercício de suas atividades. A digitalização em massa vem cobrando pesados investimentos em tecnologia da informação, trazendo à tona uma crescente debilidade estrutural dos sistemas eletrônicos de informação, que os incapacita de assegurar a preservação em longo prazo e o acesso contínuo às informações geradas em um contexto de rápido avanço tecnológico. Começam a surgir artefatos digitais que permitirão a identificação das características dos objetos digitais, produtos da necessidade de preservar o significado e a usabilidade da informação codificada neles.

O tema da preservação dos documentos digitais está cada vez mais presente entre os criadores e gerenciadores de coleções digitais. Para aqueles que começam a enfrentar a possibilidade de não poder recuperar informações produzidas há 5 ou 10 anos, é vital contar com estratégias que garantam a persistência desses dados. Para bibliotecários, arquivistas, museólogos e o pessoal de tecnologia da informação, a preservação digital é percebida como um conjunto de atividades desenvolvidas para conservar e prolongar a vida da informação digital, protegendo-a da obsolescência tecnológica, da perda e das falhas de equipamentos. Segundo a Associação Americana de Bibliotecários, “a preservação digital combina políticas, estratégias e ações que garantem o acesso ao conteúdo digital através do tempo” (ALA, 2007). As técnicas de preservação que eram conhecidas pelas bibliotecas e pelos centros de informação se modificaram e encontraram novas práticas em um contexto de rede de computadores, no qual a informação não fica apenas em um só lugar. A preservação da informação em formato digital precisa de um conjunto de práticas técnicas e de gerenciamento que mudam constantemente.

Desde as primeiras definições da preservação digital, a parte de planejamento, alocação de recursos e aplicação de métodos e tecnologias sempre estiveram presentes (HEDSTROM, 1997). Já neste século, seu universo instrumental e o papel dos mecanismos que permitem o armazenamento em repositórios digitais começou a receber mais atenção/ estar em foco (CASTRO, 2018). A preservação dos documentos continua a ser determinada pela capacidade de o objeto informacional servir às utilizações que lhe são imputadas e às suas atribuições, as quais garantem que ele continue a ser satisfatório para utilizações posteriores. Na era da informação digital, a ênfase está na geração e/ou aquisição de material digital, deixando em segundo plano a preservação e o acesso, a longo prazo, aos acervos eletrônicos existentes, quer dizer, a aplicação de estratégias de preservação digital deveria estar presente, pois sem elas não existiria nenhuma garantia de acesso, de confiabilidade e de integridade dos documentos a longo prazo (MÁRDERO ARELLANO; LEITE, 2009).

O domínio da preservação digital se estende para incluir o ciclo de vida completo dos objetos digitais, todos os processos relacionados à ingestão, armazenamento, gerenciamento e fornecimento de acesso ao objeto em questão. Os registros digitais são documentos compostos que consistem em objetos com dependência de intermediários tecnológicos necessários para torná-los visíveis para humanos. São mais facilmente alterados e corrompidos do que materiais analógicos, e as alterações, quando ocorrem, são mais difíceis de detectar. Muitas vezes, dada a complexidade das partes inter-relacionadas que compõem um registro digital, criadores e preservadores acham difícil estabelecer

a identidade e a integridade de um registro, duas características pelas quais a autenticidade é determinada (AUTHENTICITY TASK FORCE, 2002). Esta dificuldade em estabelecer sua autenticidade afeta negativamente o valor dos registros, pois como Ross (2002, p. 7) afirma: “objetos digitais que carecem de autenticidade têm valor limitado como evidência ou como um recurso de informação.”

Para que o objetivo da preservação digital possa ser atingido, os objetos digitais devem ser compreendidos e gerenciados em vários níveis: como um objeto físico, como uma codificação lógica, como objetos conceituais ou possuidores de significado para os humanos e como um conjunto de elementos essenciais, que devem ser preservados para oferecer aos futuros usuários a essência do objeto (UNESCO, 2003).

## ♦ 2 O OBJETO DA PRESERVAÇÃO

Os objetos digitais seriam aqueles que foram criados em computador, podendo ser originais ou versões depois de terem sido convertidas ou digitalizadas. Pela sua natureza multifacetada, a produção e disseminação dos objetos digitais têm sido ampla e geral, assim como as informações do seu contexto registrado e preservado para que, dessa maneira, futuros usuários possam entender o ambiente tecnológico no qual ele foi criado. Para Thibodeau (2002), objeto digital é o objeto de informação, de qualquer tipo de informação ou de qualquer formato que está expresso em forma digital. Esse raciocínio foi corroborado por Ludwing (2010), que apontou para os três diferentes aspectos visados pela preservação digital: a preservação de *bitstream*, que lida com o objeto físico e garante a estabilidade da cadeia bits, conquanto o objeto físico possa mudar; a preservação de conteúdo, que tenta manter o conteúdo estável, conquanto o formato do objeto lógico possa mudar; e a curadoria de dados, que permite complementações e mudanças no objeto conceitual com o propósito de assegurar a usabilidade a longo prazo.

Entre as ações gerenciais e as técnicas exigidas dentro do planejamento da preservação digital, a alocação de recursos e a aplicação de métodos e tecnologias de preservação, combinadas com políticas e estratégias para garantir o acesso a conteúdo reformatado e “nativo digital” (*born-digital*), levam sempre em consideração a obsolescência tecnológica dos objetos digitais, a fragilidade dos suportes e a sua interpretação a longo prazo (MUHANNA, 2018; CONARQ, 2020). A preservação digital não é apenas um problema técnico, e, sim, “um componente de uma ampla agregação de serviços, políticas e partes interessadas interconectados que, juntos, constituem um ambiente digital” (LAVOIE; DEMPSEY, 2004). As condições básicas à preservação digital seriam, então, a adoção desses métodos e dessas tecnologias que integrariam a preservação física, lógica e intelectual dos objetos digitais.

A preservação física dos objetos digitais está centrada no conteúdo armazenado em mídia magnética e em discos óticos. Assim como ocorre com o suporte físico da informação, em papel, a superfície metálica magnetizada se desintegra ou pode se tornar irrecuperável. Existem, ademais, os efeitos da temperatura, da umidade, do nível de poluição do ar e das ameaças biológicas; os danos provocados pelo uso indevido e regular, as catástrofes naturais e a obsolescência tecnológica. Para Duranti (2010), diferentemente dos documentos tradicionais, os digitais podem ser armazenados em diferentes partes do suporte e eles não existem como objetos físicos até o momento em que são gerados intencionalmente, pois os elementos do documento digital podem se manifestar de múltiplas formas. Uma das dificuldades da aplicação prática das técnicas de preservação digital é a capacidade que possuem as ferramentas tecnológicas de tratar os diferentes objetos digitais adicionais de um único objeto digital.

O dado e a mídia que suportam a informação devem possuir um nível de funcionalidade representacional que permita a sua reprodução a qualquer momento que a instituição mantenedora precisar recuperar o dado. Cada mídia pode armazenar uma sequência de bits de forma diferente, segundo as propriedades físicas da mídia. O *bitstream* precisa, então, ser interpretado, uma vez que toda sequência significativa de bits pode representar qualquer coisa. Os objetos digitais são salvos como coleções de bits, representando documentos específicos, significativos apenas para o programa que os criou. A preservação digital não passa tanto pela preservação do suporte físico (como acontece com a preservação tradicional), mas, sim, por garantir que a informação nele contida permaneça acessível e autêntica ao longo do tempo (FERREIRA, 2011).

Já na preservação da natureza lógica dos objetos digitais, são procurados/buscados, na tecnologia, formatos atualizados para inserção dos dados (correio eletrônico, material de áudio e audiovisual, material em rede etc.), novos *software* e hardware que mantenham vigentes seus bits para conservar sua capacidade de leitura. Na publicação digital, a preservação lógica está associada à necessidade de garantir a conversão dos formatos originais que passaram a ser obsoletos ou de custosa manutenção.

No caso da preservação intelectual, o foco são os mecanismos que garantem a integridade e a autenticidade da informação nos documentos eletrônicos. A preservação do objeto físico e lógico não garante a obtenção do objeto conceitual. É necessário preservar uma forma de traduzir o objeto lógico para o conceitual (YAMAOKA; GAUTHIER, 2013, p. 85,). A preservação da informação descritiva dos materiais digitais é de crucial importância, devido, principalmente, à capacidade de o objeto digital ser passível de modificação no seu desenho (leiaute), apresentação ou interação no formato de publicação. Com isso, a perda do conteúdo intelectual original pode ser declarada inaceitável pelo autor.

Em “Teoria dos Objetos Digitais”, Kallinikos, Aaltonen e Marton (2010) argumentam que estes artefatos possuem algumas características próprias, dentre as quais as de serem editáveis, por permitirem modificações contínuas e sistemáticas; interativos, com funções incorporadas para oferecer caminhos aos agentes; abertos, no sentido de serem acessíveis e modificáveis por um programa e distribuídos, pela possibilidade de disponibilização, em diversas fontes e instituições.

Um fator relevante da natureza intelectual dos objetos digitais é que ela envolve a preservação da propriedade intelectual (IPR) – *Intellectual Property Rights*, que tem um significado mais argúvel do que na mídia tradicional. O IPR deve considerar não apenas o conteúdo, mas também qualquer ação relacionada ao *software* (cópias, encapsulamento de conteúdo, emulação de *software*, migração de conteúdo) que envolva atividades que podem infringir permissões específicas daqueles que mantêm os direitos. Para Bullock (1999), o IPR é uma das principais barreiras que interferem na preservação dos objetos digitais e sua adoção pode interferir nas informações contidas, como funcionalidade, aparência e conteúdo. Por esse motivo, Lavoie e Gartner (2005) recomendam documentar os direitos associados aos objetos arquivados, para que os processos de preservação estejam coordenados com as restrições impostas aos objetos.

## ◆ 3 A PESQUISA DA PRESERVAÇÃO DIGITAL

Desde suas primeiras conceptualizações, a preservação digital levou em consideração, além da interpretação dos dados, todas as ações que permitiram sua produção e acesso. A pesquisa da preservação digital investiga a natureza de objetos digitais, incluindo registros e dados e os atributos que podem apoiar a sua confiabilidade, a integridade e a presunção de sua autenticidade (GRÁCIO, 2011). Embora muita pesquisa tenha sido e continua a ser conduzida na proteção da autenticidade, no contexto de requisitos para preservação digital, os meios atuais de avaliação da autenticidade para os profissionais da informação ainda não oferecem medidas quantificáveis, e modelos generalizáveis que podem reduzir o problema. Para Moore (2008), as ações de preservação são uma forma de comunicação com o futuro e o grande desafio é como incorporar novas tecnologias de forma eficaz, ao mesmo tempo em que se resguardem as propriedades de preservação, como autenticidade, integridade e cadeia de custódia.

A pesquisa da preservação tem tido foco na natureza dos objetos digitais, no desenvolvimento de padrões, estruturas e sistemas de repositório, em definir e usar/compartilhar esquemas de metadados, nas tecnologias de preservação e nos formatos de arquivo. Todos eles compartilham um objetivo comum: o de preservar objetos digitais confiáveis (ROGERS, 2015). Novas questões surgem a partir desse objetivo, colocando a preservação digital não apenas como um processo isolado, mas também como componente de um conjunto agregado de serviços, políticas e especialistas que constituem o contexto do ciclo de vida da informação digital.

A preservação digital inicia-se no agrupamento da perspectiva que se tinha dos requisitos associados com as atividades tradicionais nas áreas de conservação e de preservação dos objetos informacionais. Os objetos digitais não podem ser deixados em formatos obsoletos para serem transferidos, depois de longos períodos de negligência, para repositórios digitais. O desafio é muito mais um problema social e institucional do que técnico, porque, principalmente para a preservação digital, depende-se de instituições que passam por mudanças de direção, missão, administração e fontes de financiamento (TERMENS, 2009; SASTRE, 2015).

Ao analisar o ciclo da informação (geração, tratamento, preservação, conservação e, finalmente, difusão da informação), percebe-se que a preservação e a conservação necessitam ser revistas no âmbito da informação digital. Até pouco tempo, entendia-se a preservação no sentido de conservar e prevenir os documentos do risco de deterioração. É importante analisar a preservação a partir da necessidade de assegurar o acesso e recuperação da informação como fundamentos para a pesquisa. Aliado a isso, deve-se considerar o crescente uso dos computadores e a quantidade de informações disponíveis em meio eletrônico (algumas, inclusive, somente neste meio).

Muitos materiais publicados digitalmente são produto de serviços de informação disponibilizados por organizações que adotam alguma infraestrutura tecnológica. Essas instituições levam em consideração aspectos legais e culturais que afetam a oferta desses serviços orientados a atender as necessidades de determinados usuários. Mas, muitas vezes, essa preocupação não é a mesma que têm os produtores das novas tecnologias. É fato que, na preservação de documentos digitais, assim como na de documentos em papel, é necessária a adoção de ferramentas que protejam e garantam a sua manutenção. Essas ferramentas deverão servir para reparar e restaurar registros protegidos, prevendo os danos e reduzindo os riscos dos efeitos naturais (preservação prospectiva), ou para restaurar os documentos já danificados (preservação retrospectiva).

Na área da Ciência da Informação, o uso da tecnologia digital, que toma o lugar dos tradicionais meios de conservação e preservação, trouxe consigo a preocupação com as normas para o uso das técnicas digitais e sua prontidão na tarefa da preservação a longo prazo (CHEPESUIK, 1997). Os especialistas da área que trabalham com informação em formatos digitais elaboram normas necessárias para armazenar e compartilhar, de maneira adequada, esses materiais, assim como buscam a formulação de políticas institucionais de preservação.

Os primeiros estudos desenvolvidos na área da preservação digital tiveram origem no seio da Arquivologia. Entretanto, considerando as diferenças existentes no que tange à preservação de acervos bibliográficos e à preservação de documentos arquivísticos ao longo do tempo, a Biblioteconomia e a Ciência da Informação passaram a dedicar-se, também, a este campo. Isto pelo fato de as bibliotecas, os arquivos e os museus serem responsáveis por manter coleções digitais para uso permanente, protegendo-as de ameaças, ou salvando-as e reparando-as para compensar seus impactos. Os sistemas de bibliotecas e de arquivos atuais envolvem um grande esforço de gerenciamento de objetos digitais, que vai além das tarefas tradicionais dessas instituições patrimoniais (aquisição, seleção, classificação, arquivamento etc.), na construção da interoperabilidade de acervos digitais (arquiteturas, metadados, formatos padrão), o que é possível por meio de sistemas relacionados e desenvolvidos para propósitos e comunidades específicas.

As bibliotecas digitais foram as primeiras a serem consideradas o caminho mais adequado para a preservação dos recursos de informação (HILDRETH, 1996). Posteriormente, os repositórios arquivísticos foram considerados, dentro das instituições, locais confiáveis capazes de armazenar, migrar e dar acesso a coleções digitais. O treinamento de especialistas e técnicos também começou a ser contemplado nos custos das atividades desses repositórios, produzindo melhora na aplicação efetiva das estratégias (CHAPMAN, 2004).

Segundo Becker *et al.* (2009), existe um constante crescimento na oferta de ferramentas, teorias e modelos para a preservação digital. Entretanto, apresenta-se uma carência de estudos que auxiliem os gestores na tomada de decisão a respeito da seleção das ferramentas e sobre como implementar os modelos propostos de forma concreta. A responsabilidade de adotar medidas preventivas e corretivas, tendo como objetivo minimizar a ação do tempo sobre o suporte físico da informação e assegurar sua disponibilidade, foi atribuída às instituições arquivísticas, as quais teriam a capacidade de compreender os limites e os significados dos documentos (autenticidade, capacidade probatória, integridade das informações, contexto de produção, manutenção etc.), dando ênfase às tarefas que essas organizações devem observar para lidar com a guarda permanente dos objetos digitais autênticos.

A gestão dos documentos eletrônicos constitui uma área de pesquisa que tem produzido contribuições relevantes relacionadas com a preservação de documentos digitais de “valor permanente” para instituições públicas. Como componentes do documento digital arquivístico, os metadados requerem os mesmos procedimentos de preservação, visto que o contextualizam e colaboram para a manutenção de suas características arquivísticas (SANTOS, 2012).

## ◆ 4 AS PRÁTICAS DA PRESERVAÇÃO DIGITAL

A preservação digital é a parte mais longa e também a última do ciclo de gerenciamento de objetos digitais. Ela permite o emprego de mecanismos que viabilizam o armazenamento em repositórios de objetos digitais e que garantem a autenticidade e a perenidade dos seus conteúdos. São necessários, além de procedimentos de manutenção e recuperação de dados (no caso de perdas acidentais para resguardar a mídia e seu conteúdo), também estratégias e procedimentos para manter acessibilidade e autenticidade através do tempo, podendo requerer colaboração entre diferentes financiadoras e boa prática de licenciamento, de metadados e de documentação, antes de aplicar ações técnicas.

Nas últimas três décadas, as práticas relacionadas com a preservação digital estão baseadas na ideia de garantir a longevidade dos arquivos, e elas dependem, principalmente, da solução tecnológica adotada e dos custos que as envolvem. Mas, ainda hoje, ainda é evidente a ausência de conhecimento sobre as estratégias de preservação digital e o que isso poderá significar diante da necessidade de garantir a longevidade dos arquivos digitais. Os atuais detentores de acervos digitais percebem a imperiosa necessidade de contar com mecanismos que garantam a preservação de seus documentos em formato digital. Especificamente, essa preocupação parte das comunidades responsáveis pelas bibliotecas e pelos arquivos, nas quais o desenvolvimento de padrões e de mecanismo legais para lidar com arquivos eletrônicos precisa de estratégias metodológicas bem definidas.

As iniciativas se multiplicam e novas soluções são testadas por instituições detentoras de acervos de especial relevância para o desenvolvimento científico e tecnológico de seus países de origem. Muitos projetos e iniciativas têm conseguido estabelecer os benefícios de algumas metodologias e estratégias. A conclusão a que essas iniciativas chegam é que devem ser usados padrões e converterem-se os documentos nos formatos livres, para que eles sejam acessados após a obsolescência dos equipamentos e dos programas informáticos em que foram criados.

Por causa da natureza interdisciplinar, do escopo abrangente e do custo da preservação digital, os projetos são frequentemente realizados por alianças ou consórcios nacionais e internacionais de universidades, bibliotecas e arquivos, agências governamentais, empresas e indústrias. Cada um deles é definido por sua perspectiva epistêmica e propósitos particulares. No entanto, cooperação e colaboração são constantes em toda a comunidade da área. Pesquisadores em várias partes do mundo estão desenvolvendo modelos do que seria a infraestrutura para a preservação, a longo prazo, de informação em formato digital e muitas iniciativas propostas pelos grandes centros de pesquisa continuam sendo utópicas, porém, proporcionam uma base para a discussão sobre a implantação de soluções futuras em relação ao problema. Entre elas, está o uso de padrões abertos que permite seu estudo e sua conversão para novos padrões.

Os principais métodos recomendados para a preservação dos objetos digitais podem ser agrupados em dois tipos: estruturais e operacionais. Os estruturais tratam dos investimentos iniciais por parte das instituições que estão se preparando para implementar algum processo de preservação e que adotam ou adaptam um dos modelos de metadados existentes ou seu próprio esquema. As atividades operacionais são as medidas concretas aplicadas aos objetos digitais, aquelas que englobariam os novos requisitos de preservação, migração de suporte, atualização da mídia (preservação física), conversão dos formatos, emulação (preservação lógica) a preservação do conteúdo (intelectual).



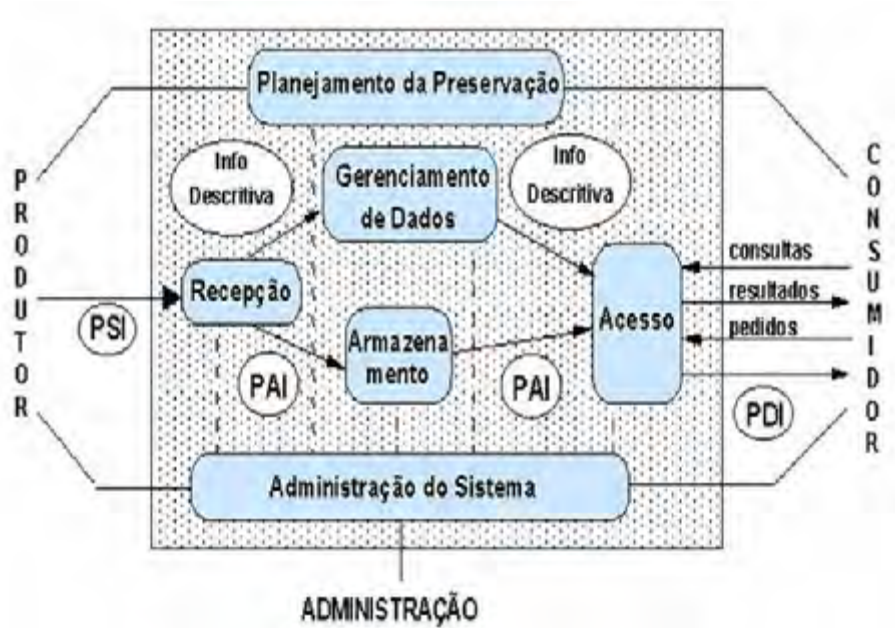
As tarefas da preservação digital são permanentes e devem ser pensadas a longo prazo desde a criação do próprio documento. Por este motivo, é importante tomar como referência padrões internacionais e nacionais existentes, tais como: o modelo OAIS, o sistema de metadados PREMIS e o esquema METS. As coleções que servem de informações estruturadas e padronizadas em metadados tornam-se mais vivíveis em redes de alcance global, o que permite que seus recursos informacionais sejam identificados, localizados, selecionados por critérios de relevância e distinguidos por diferenças e similaridades.

Desde a década de 1990, a importância dos metadados foi reconhecida (HURLEY, 1995) como sendo a chave para a criação de registros digitais confiáveis, parte importante em todas as estratégias estruturais e operacionais de preservação digital, uma vez que elas estão baseadas na conservação de *software* e *hardware*, emulação ou migração, como um meio de garantir a autenticidade, registrar o gerenciamento de direitos e coleções de dados e de interagir com recursos de busca (GILLILAND; MCKEMMISH, 2012). Metadados são afirmações legíveis por máquinas e por humanos a respeito de recursos de informação que permitem o controle físico, intelectual e técnico sobre esses recursos. Os responsáveis pelos acervos digitais criam e anexam e, em seguida, mantêm e preservam metadados, automática e/ou manualmente, por serem elementos ou estruturas de organização da informação que, associados a um objeto digital, classificam, categorizam e descrevem essa informação.

Uma forma especializada de administrar metadados que podem ser usados como um meio de estocar a informação técnica (e que suporta a preservação dos objetos digitais) é utilizar metadados de preservação. Eles visam a apoiar e a facilitar a retenção, a longo prazo, da informação digital (FORMENTON, *et al.*, 2017), codificando e rastreando a identidade e a integridade do material ao longo do tempo por meio de mudanças tecnológicas. Para definir estruturas padronizadas de metadados e determinar a sua utilização nos sistemas eletrônicos de gestão de informação, com o propósito de gerir a preservação e a acessibilidade dos documentos digitais, foi necessária a adoção de um modelo de alto nível para sistemas de preservação digital, o Open Archival Information System (OAIS). Ele é referência em sistemas de preservação digital, abordando todos os aspectos da preservação de longo prazo da informação digital: ingestão, armazenamento em arquivo, gerenciamento de dados, acesso, disseminação e migração para novas mídias e formas.

Desenvolvido em 2002 pelo Comitê Consultivo para Sistemas de Dados Espaciais, o OAIS é hoje uma norma ISO aprovada (ISO 14721: 2003) e passou por várias revisões, a mais recente em 2012 (CCSDS 2012). Um arquivo nos termos do OAIS significa considerar os arquivos como organizações de pessoas e sistemas, que aceitaram a responsabilidade de preservar a informação e torná-la disponível, além de melhorar a comunicação e a produtividade entre diferentes comunidades. No modelo, atuam quatro entidades: produtores, consumidores, administração e o arquivo propriamente dito. No esquema do OAIS, está incluído um modelo de informação para inserção dos metadados de preservação (Figura 1), a seguir.

Figura 1 - Modelo funcional OAIS



Fonte: ABNT (2007).

As entidades funcionais gerenciam o fluxo de informação entre as entidades que fazem parte do ambiente OAIS e identificam os componentes funcionais dos arquivos que estão relacionados com a preservação dos objetos digitais. Cada entidade é responsável por algum processo do sistema no ambiente OAIS, ou seja, o produtor pode ser considerado a pessoa ou sistema cliente que fornece a informação; o consumidor é a pessoa ou sistema cliente que quer obter o material arquivado ou a versão transformada do mesmo; a comunidade-alvo refere-se aos usuários potenciais (MÁRDERO ARELLANO, 2008).

Porém, algumas decisões independem do Modelo OAIS. Como é o caso do tipo de material a ser armazenado, sobre o local do armazenamento das informações descritivas, de relacionamento entre objetos e se outros componentes do gerenciamento devem ou não ser incluídos. O modelo apresenta desafios para as instituições, assim como pode ser considerado uma estratégia de preservação de longo prazo da informação digital. Ele provê conceitos importantes e a adequação necessária dos componentes que podem chegar a ser desenvolvidos no futuro. Em uma adaptação do OAIS especificamente direcionado para documentos arquivísticos digitais, as funções do programa de preservação de documentos arquivísticos digitais devem seguir o modelo de Cadeia de Preservação do InterPARES 2 (2008).

O Dicionário PREMIS (LIBRARY OF CONGRESS, 2008) é usado para definir o conjunto básico de unidades semânticas que os repositórios digitais devem conhecer ao realizar funções de preservação. Para ser considerado um repositório de preservação, é preciso ser ou estar em conformidade com a “responsabilidade exclusiva ou como parte de múltiplas responsabilidades que realiza a preservação de longo prazo de objetos digitais sob sua custódia”. As tarefas de preservação podem variar de um repositório para outro, mas elas devem garantir a viabilidade e a sustentabilidade dos objetos digitais. Além disso, devem assegurar, também, que objetos digitais no repositório não sejam alterados sem nenhuma documentação prévia. Os principais usos do PREMIS são direcionados ao *design* do repositório e à sua avaliação e troca de pacotes de informação entre repositórios. Também, muitos dos metadados PREMIS são projetados para serem fornecidos por um aplicativo de *software*.

Já o formato padronizado para transmissão de metadados, projetado para envio de arquivos, imagens e objetos multimídia em um repositório digital que utiliza a estrutura XML é o Metadata Encoding & Transmission Standard METS (LIBRARY OF CONGRESS, 2007). O padrão é mantido no Gabinete de Desenvolvimento de Redes e Padrões MARC da Biblioteca do Congresso, e consiste na codificação de metadados necessários para o gerenciamento de objetos digitais e na troca entre repositórios digitais.

Em 2014, no âmbito do projeto SCAPE, da União Europeia, ficou amplamente documentada a tendência de as instituições limitarem-se em elaborar políticas de preservação digital, não avançando para as políticas de procedimento (SIERMAN *et al.*, 2014). A expectativa é que as instituições possam definir políticas de gerenciamento que garantam a confiabilidade; definam regras; apliquem as políticas; estabeleçam capacidades que implementem as funções de preservação necessárias; e definam metadados de preservação, para, assim, capturar informações reais sobre a aplicação das funções de preservação. Dessa forma, os metadados de preservação poderão ser consultados para analisar se os critérios de avaliação foram satisfeitos, uma vez que muitos repositórios fazem uso indevido do modelo OAIS, tendo em vista que, nem sempre, eles foram auditados e certificados para verificação de tal conformidade (LAVOIE, 2014).

A comunidade de preservação vem desenvolvendo critérios de avaliação para validar a confiabilidade dos repositórios digitais. Um repositório capaz de autenticar os critérios de avaliação pode ser considerado confiável e, portanto, seria um ambiente razoável para a preservação de dados em longo prazo. Desde 2007, quando um conjunto inicial de critérios de avaliação foi proposto pelo Research Libraries Group and the National Archives and Records Administration (RLG; NARA, 2007), gestores de sistemas de preservação reconhecem que um ambiente de preservação deve ser quantificado em termos de operações de preservação reais (que são realizadas) e de políticas de gestão (que controlam a execução dos processos de preservação).

A importância da automação de todo o processo de preservação digital é resultado da crescente produção de informações a serem preservadas pelas instituições em seus repositórios digitais. Para Térmens e Ribera (2009), os sistemas de preservação digital só podem funcionar com eficiência (ou seja, com custos razoáveis e confiabilidade fidedigna), se puderem normalizar completamente seus processos e automatizar os procedimentos de ingestão. Dessa maneira, os repositórios digitais poderão integrar ferramentas de preservação digital capazes de realizar as tarefas relacionadas à identificação, validação, caracterização, arquivamento e acesso de longo prazo a documentos em formato digital.

## ◆ 5 CONCLUSÃO

Devido à rápida degradação física, a obsolescência tecnológica, complexidade e os custos envolvidos na preservação, os objetos digitais estão em uma posição de alto risco, tendo em vista a fragilidade e demora na aplicação de estratégias para garantir preservação em um longo período de tempo. A vantagem está no fato de, no meio digital, a destruição do suporte não representar perda de informação, desde que tenham sido empregadas técnicas de preservação neste objeto digital. A preservação digital não é sobre um sistema particular, ou uma série de ações de preservação. Trata-se de preparar conteúdo e coleções em uma grande cadeia de transferências (OWENS, 2018).

Os gestores de repositórios digitais ainda possuem pouca ou nenhuma experiência em relação às práticas de preservação digital. Enquanto isso, a pesquisa continua a investigar a natureza dos objetos digitais, incluindo registros, dados, e os atributos que podem apoiar a presunção da autenticidade. Também continua a procura por respostas sistêmicas que possam ser padronizadas na prática, como infraestrutura governamental e financeira. É imperiosa a necessidade de preservar sistemas complexos de forma integrada e distribuída em rede.

## ◆ REFERÊNCIAS

ABNT. Associação Brasileira de Normas Técnicas. **NBR 15472**: modelo de referência para um sistema aberto de arquivamento de informação. Rio de Janeiro: ABNT, 2007.

AMERICAN LIBRARY ASSOCIATION. **Definitions of digital preservation**. Chicago: American Library Association. 2007. Disponível em: <http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/defdigpres0408.pdf>. Acesso em: 1 ago. 2021.

AUTHENTICITY TASK FORCE. Appendix 2 – requirements for assessing and maintaining the authenticity of electronic records. In: L. DURANTI, L. (ed.). **The long-term preservation of authentic electronic records**: Findings of the InterPARES project. Vancouver, BC: InterPARES, 2002. Disponível em: [http://www.interpares.org/display\\_file.cfm?doc=ipl\\_authenticity\\_requirements.pdf](http://www.interpares.org/display_file.cfm?doc=ipl_authenticity_requirements.pdf) Acesso em: 1 ago. 2021.

BEAGRIE, N. Digital curation for science, digital libraries, and individuals. **International Journal of Digital Curation**, [s.l.], v. 1, n. 1, 2006. Disponível em: <http://www.ijdc.net/article/view/6> Acesso em: 1 ago. 2021.

BECKER, C. *et al.* Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. **International Journal on Digital Libraries**, [s.l.], v. 10, n. 4, p. 133-157, 2009. Disponível em: <https://doi-org.ez74.periodicos.capes.gov.br/10.1007/s00799-009-0057-1>. Acesso em: 28 jul. 2021.

BULLOCK, A. **Preservation of digital information**: issues and current status. Ottawa: National Library of Canada, 1999. Disponível em: <http://www.referenciasarquivisticas.fci.unb.br:8080/jspui/handle/123456789/3537>. Acesso em: 28 jul. 2005.

CASTRO, R. M. Objetos digitais: da maleabilidade do não-finito à uma ontologia. **Pragmatizes**: Revista Latino-Americana de Estudos em Cultura, [s.l.], n. 16, p. 31-43, 3 jun. 2019.

CHAPMAN, S. Counting the costs of digital preservation: is repository storage affordable? **Journal of Digital Information**, [s.l.], v. 4, n. 2, 2004. Disponível em: <https://journals.tdl.org/jodi/index.php/jodi/article/view/jodi-113>. Acesso em: 30 jul. 2021.

CHEPESUIK, R. The future is here: America's libraries go digital. **American Libraries**, v. 2, n. 1, p. 47-49, 1997.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). Câmara Técnica de Documentos Eletrônicos (CTDE). **Glossário Documentos Arquivísticos Digitais**. 2020. 8ª versão. Disponível em: [http://antigo.conarq.gov.br/images/ctde/Glossario/glosctde\\_2020\\_08\\_07.pdf](http://antigo.conarq.gov.br/images/ctde/Glossario/glosctde_2020_08_07.pdf) Acesso em: 20 mai. 2021.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). **Reference Model for an Open Archival Information System (OAIS): recommended Practice**, issue 2. Washington, D.C.: CCSDS/NASA, 2012. 135 p. Disponível em: <http://www.ccsds.org>. Acesso em: 20 jul. 2021.

DURANTI, L. Concept and principles for the management of electronic records, or records management theory is archival diplomatics. **Record Management Journal**, [s.l.], v. 20, n. 1, p. 78-95, 2010.

FERREIRA, C. A. S. **Preservação da informação digital**: uma perspectiva orientada para as bibliotecas. 2011. 155f. Dissertação (Mestrado em Informação, Comunicação e Novos Media) – Universidade de Coimbra, Coimbra, 2011.

FORMENTON, D. *et al.* Os padrões de metadados como recursos tecnológicos para a garantia da preservação digital. **Biblios**, Pittsburgh, n. 68, p. 82-95, 2017. Disponível em: <http://www.scielo.org.pe/pdf/biblios/n68/a06n68.pdf> Acesso em: 20 jul. 2020.

GILLILAND, A.; MCKEMMISH, S. Recordkeeping Metadata, the Archival Multivers, and Societal Grand Challenges. *In*: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 2012, p. 106-113. **Proceedings [...]** Kuching, Sarawak, Malaysia. Disponível em: <http://dcevents.dublincore.org/IntConf/dc-2012/paper/view/108/66>. Acesso em: 25 jul. 2021.

GRÁCIO, J. C. A. **Preservação digital na gestão da informação**: um modelo processual para as instituições de ensino superior. 2011. 223 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2011.

HEDSTROM, M. Digital preservation: a time bomb for digital libraries. **Computers and the humanities**, v. 31, n. 3, p. 189-202, 1997. Disponível em: [https://www.researchgate.net/publication/30843346\\_Digital\\_Preservation\\_A\\_Time\\_Bomb\\_for\\_Digital\\_Libraries](https://www.researchgate.net/publication/30843346_Digital_Preservation_A_Time_Bomb_for_Digital_Libraries). Acesso em: 11 jul. 2021.

HILDRETH, C. R. Preserving what we really want to access, the message, not the medium: challenges and opportunities in the digital age. *In*: HELAL, A. H.; WEISS, J. W. (ed.). INTERNATIONAL ESSEN SYMPOSIUM, 18., 1995, Essen. **Proceeding [...]** Essen: Germany University, 1996. p. 76-95.

HURLEY, C. Ambient Functions: Abandoned Children to Zoos. **Archivaria**, v. 40. 1995. Disponível em: <https://archivaria.ca/index.php/archivaria/article/view/12095>. Acesso em: 31 jul. 2021.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 14721:2012**: Space data and information transfer systems: Open archival information system (OAIS). Reference model. Genebra: ISO, 2012.

INTERPARES 2 PROJECT (The). **Experiential, Interactive and Dynamic Records**. 2008. Disponível em: <http://www.interpares.org/ip2/book.cfm> . Acesso em: 18 jun. 2021.

KALLINIKOS, J.; AALTONEN, A.; MARTON, A. A theory of digital objects. **First Monday**, [s.l.], v. 15, n. 6, 2010. Disponível em: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3033/2564>

Acesso em: 1 ago. 2021.

LAVOIE, B. **The Open Archival Information System (OAIS) Reference Model**: introductory guide. 2. ed. York: DPC, 2014. Disponível em: <http://dx.doi.org/10.7207/twr14-02>. Acesso em: 1. Ago. 2021.

LAVOIE, B; DEMPSEY, L. Thirteen ways of looking at - digital preservation. **D-Lib Magazine**, v. 10, n. 7/8, jul./ago., 2004. Disponível em: <http://www.dlib.org/dlib/july04/lavoie/07lavoie.html> Acesso em: 28 jun. 2021.

LAVOIE, B.; GARTNER, R. **Preservation Metadata**. Dublin, Ohio: OCLC, September 2005. Disponível em: <https://www.dpconline.org/docs/technology-watch-reports/894-dpctw13-03/file>. Acesso em: 30 jul. 2021.

LIBRARY OF CONGRESS. **Metadata Encoding and Transmission Standard**. 2007. Disponível em: <http://www.loc.gov/standards/mets> Acesso em: 1 ago, 2021.

- LIBRARY OF CONGRESS. **Data dictionary for preservation metadata:** PREMIS. Version 2.2. (2012). Disponível em: <http://www.loc.gov/standards/premis/v2/premis-dd-2-2.pdf> Acesso em: 28 jul. 2021.
- LUDWIG, J. **About the complexity of a digital preservation theory and different types of complex digital objects.** Germany: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2010. Disponível em: <http://www.dagstuhl.de/Materials/Files/10/10291/10291.LudwigJens.ExtAbstract.pdf>. Acesso em: 22 mai. 2021.
- MÁRDERO ARELLANO, M. A. **Crítérios para a preservação digital da informação científica.** 2008. 354 f. Tese (Doutorado) – Universidade de Brasília, Departamento de Ciência da Informação, 2008. Disponível em: <http://eprints.rclis.org/12649/>. Acesso em: 28 jul. 2021.
- MÁRDERO ARELLANO, M. A.; LEITE, F. C. L. Acesso aberto à informação científica e o problema da preservação digital. **Biblios**, Brasília, n. 35, 2009. Disponível em: <http://repositorio.unb.br/handle/10482/4937> Acesso em: 1 ago. 2021.
- MOORE, R. Towards a Theory of Digital Preservation. **International Journal of Digital Curation**, [s.l.], v. 3, n. 1, 2008. Disponível em: <http://www.ijdc.net/article/view/63> Acesso em: 30 jul. 2021.
- MUHANNA, E. What Does “Born Digital” Mean? **International Journal of Middle East Studies**, [s.l.], v. 50, n. 1, Feb. 2018, p. 110–112.
- OWENS, T. **The theory and craft of digital preservation.** Baltimore, Maryland : Johns Hopkins University Press, 2018.
- RESEARCH LIBRARIES GROUP & NATIONAL ARCHIVES AND RECORDS ADMINISTRATION. **Trustworthy repositories audit & certification:** criteria and checklist. 2007. Disponível em: [https://www.crl.edu/sites/default/files/d6/attachments/pages/trac\\_0.pdf](https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf) Acesso em: 1 ago. 2021.
- ROGERS, C. **Virtual authenticity:** authenticity of digital records from theory to practice. Electronic Theses and Dissertations, Vancouver: University of British Columbia, 2015. Disponível em: <http://dx.doi.org/10.14288/1.0166169>. Acesso em: 28 jul. 2021.
- ROSS, S. Position paper on integrity and authenticity of digital cultural heritage objects. In: **DigiCULT: Integrity and Authenticity of Digital Cultural Heritage Objects**, August, 2002. p. 7–8. Disponível em: <https://www.digicult.info/downloads/html/7/7.html> . Acesso em: 24 jul. 2021.
- SANTOS, V. B. dos. Preservação de documentos arquivísticos digitais. **Ciência da Informação**, Brasília, v. 41, n. 1, 2014. Disponível em: <http://revista.ibict.br/ciinf/article/view/1357>. Acesso em 28 jul. 2021.
- SASTRE NATIVIDAD, G. **Preservación y conservación de documentos digitales** [en línea]. In: ArchivPost. Salamanca: Asociación de Archiveros de Castilla y León, 2015. Disponível em: <http://www.acal.es/index.php/archivpost-a-fondo>. Acesso em: 29 jul. 2021.
- SIERMAN, B. *et al.* **Catalogue of preservation policy elements.** 2014. Disponível em: [http://scape-project.eu/wp-content/uploads/2014/02/SCAPE\\_D13.2\\_KB\\_V1.0.pdf](http://scape-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf) Acesso em: 1 ago. 2021.
- TÉRMENS, M. Investigación y desarrollo en preservación digital: un balance internacional. **El Profesional de la Información**, v. 18, n. 6, p. 613–624, 2009. Disponível em: <http://www.elprofesionaldelainformacion.com/contenidos/2009/noviembre/04.html>. Acesso em: 1 ago. 2021.

TÉRMENS, M; RIBERA, M. El control de los formatos en la preservación digital. *In*: JORNADAS ESPAÑOLAS DE DOCUMENTACIÓN, 11, 2009, Zaragoza. **Anais [...]** Zaragoza: FESABID, 2009. p. 139-146.

THIBODEAU, K. Overview of technological approaches to digital preservation and challenges in coming years the state of digital preservation: an international perspective. **Proceedings [...]** Washington: CLIR and Library of Congress, 2002. Disponível em: <http://www.clir.org/pubs/reports/pub107/pub107.pdf#page=10>. Acesso em: 22 maio 2021.

UNESCO. **Directrices para la preservación del patrimonio digital**. 2003. Disponível em: <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>. Acesso em: 1 ago. 2021.

YAMAOKA, E. J.; GAUTHIER, F. O. Objetos digitais: em busca da precisão conceitual. **Informação & Informação**, v. 18, n. 2, p. 77-97, 2013. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/16162/13077>. Acesso em: 26 jul. 2021.

### COMO CITAR:

MÁRDERO ARELLANO, Miguel Ángel. Preservação de documentos em formato digital. *In*: BRAGA, Tiago Emmanuel Nunes; MÁRDERO ARELLANO, Miguel Ángel. (org.). **Hipátia**: modelo de preservação para repositórios arquivísticos digitais confiáveis. Brasília: Ibict, 2022. p. 10-25. (Informação, Tecnologia e Inovação, v. 1). DOI 10.22477/9786589167501.cap1 .