

{ Preservação Digital

*Universidade de Brasília
Faculdade de Ciência da Informação
Disciplina: Conservação e Restauração de Documentos
Profª Lillian Alvares*

Situação

- ⋞ A **obsolescência tecnológica** e a...
- ⋞ ... **fragilidade das mídias** de armazenamento está...
- ⋞ ... causando **perda de documentos digitais** e...
- ⋞ ... poderão levar a **perda da memória digital**.

Situação

- ⊆ O problema da **preservação digital** está na...
- ⊆ *ausência de métodos comprovados* para ...
 - ⊆ ... **assegurar a existência** e a ...
 - ⊆ **continuidade do acesso** à informação digital...
 - ⊆ ... utilizando as **ferramentas tecnológicas disponíveis**.

Preservação Digital

- ⋞ Conjunto de ações **gerenciais e técnicas** exigidas para...
- ⋞ ... **superar as mudanças tecnológicas** e a fragilidade dos suportes...
- ⋞ .. **garantindo o acesso e a interpretação** de documentos digitais...
- ⋞ pelo **tempo** que for necessário

(Glossário da Câmara Técnica de Documentos Eletrônicos do Conselho Nacional de Arquivos)

- ⋈ É a atividade de **garantir o acesso à informação** em meio digital, mantendo a sua integridade e autenticidade.
- ⋈ Esta informação deverá ser interpretada por uma **plataforma tecnológica do futuro que será diferente da que foi utilizada no momento de sua criação.**

⊆ É o **planejamento, alocação** de recursos e ...

⊆ ... aplicação de **métodos e tecnologias...**

⊆ ... para assegurar que a informação digital de valor contínuo permaneça...

⊆ **acessível e utilizável.**

{ Certezas

Século XXI

- ⋞ É um dos grandes desafios do século XXI, pois há possibilidade real de comprometimento:
 - ⋞ Do acesso à informação
 - ⋞ Da preservação do conhecimento para futuras gerações.

Responsabilidade

- ⌞ Muitas informações e coleções digitais importantes estão sendo **construídas e publicadas diretamente** em meio digital.
 - ⌞ Essa preocupação deve envolver tanto os **produtores quanto detentores** dessa informação.

Ausência de Estratégias

- ⋞ Percebeu-se a **ausência de conhecimento** sobre...
- ⋞ ... as **estratégias de preservação** digital.

Padrões para Informação Digital

- ⋞ Devem ser usados **padrões** para que a informação digital...
- ⋞ seja acessada após a **obsolescência dos equipamentos** e ...
- ⋞ ... **programas de computador** em que foram criados.

Tradicionais Meios de Preservação

- ⋞ O uso da tecnologia digital está...
- ⋞ ...**tomando lugar** dos tradicionais meios de preservação...
- ⋞ ... como a **microfilmagem**.

Requisitos de Preservação Digital

- ⋞ Preservar o conteúdo.
- ⋞ Preservar a autenticidade.
- ⋞ Preservar a proveniência.
- ⋞ Preservar a apresentação.
- ⋞ Preservar a funcionalidade.
- ⋞ Preservar o contexto.
- ⋞ Preservar a presença física.

Vulnerabilidade da Informação Digital

- ⊆ O documento digital possui uma série de especificidades, entre elas:
 - ⊆ **Vulnerabilidade**: facilidade de alterar, reformatar e falsificar sem deixar vestígios...
 - ⊆ ... o que poderá comprometer a sua:
 - ⊆ *Integridade*
 - ⊆ *Autenticidade*
 - ⊆ *Capacidade probatória*
 - ⊆ *Acesso em longo prazo.*

Políticas de Preservação

∞ Tendo em vistas estas questões...

⊆ ... será necessária a **interferência de políticas de preservação digital...**

⊆ ... pois uso inadequado das ferramentas tecnológicas poderá...

⊆ ... **colocar em risco a integridade dos acervos arquivísticos.**

Objetos Digitais Autênticos

- ⋞ A perspectiva arquivística parte da compreensão de...
 - ⋞ *autenticidade, capacidade probatória, integridade das informações, contexto de produção, manutenção...*
 - ⋞ ... **dando ênfase ao entendimento de que instituições arquivísticas são responsáveis pela guarda permanente desses documentos e devem observar para lidar com objetos digitais autênticos.**

Arquivos, Bibliotecas e Museus

⋞ Arquivos, bibliotecas e museus...

⋞ ... assumem neste contexto um papel determinante...

⋞ ... responsabilizando-se pela preservação e longevidade desses artefatos.

Formatos Obsoletos

⋞ Os objetos digitais não podem ser deixados em **formatos obsoletos** para serem transferidos depois de longos períodos de negligência.

⋞ Obsolescência do *Hardware*

⋞ Obsolescência do *Software*

Preservação Física, Lógica e Intelectual

⋞ As condições básicas à **preservação digital** seriam, então, a adoção de métodos e tecnologias que integrariam a:

⋞ Preservação **Física**

⋞ Preservação **Lógica**

⋞ Preservação **Intelectual**

dos objetos digitais.

{ Preservação Digital

- ∫ Preservação Física: Diz respeito **aos meios de armazenamento** e suporte.
- ∫ Preservação Lógica: Diz respeito aos **formatos, sua interpretação por hardwares e softwares** e a necessidade de atualização para a continuidade da leitura e interpretação.
- ∫ Preservação Intelectual: a preservação intelectual refere-se ao **conteúdo intelectual.**

{ Estratégias

Modelos, Políticas e Variedades

- ⋞ A implementação de estratégias de preservação digital **deve seguir uma política** previamente definida, para então **proceder às atividades** de manutenção dos objetos digitais.
- ⋞ Quanto maior a **variedade de estratégias** de preservação, **mais objetos digitais** serão contemplados

Estratégias de Preservação

⋞ Preservação Tecnológica

⋞ Emulação

⋞ Encapsulamento

⋞ Migração

Refreshamento

{ Primeira Estratégia: Preservação Tecnológica

- ∫ Baseada na criação de museus tecnológicos que **mantêm equipamentos e software obsoletos**, de forma que os documentos digitais possam ser **processados no seu ambiente original**.
- ∫ É uma solução de curto prazo, de **transição**.
- ∫ Preserva e mantém todo o hardware e software utilizados na concepção do objeto digital em sua forma original. **Preserva o objeto lógico original**.

{ Primeira Estratégia: Preservação Tecnológica

∫ Problemas:

- ∫ **Dependência** de hardware, software e profissionais específicos tornará o acervo vulnerável
- ∫ **Alto custo operacional**, tornando-se **inviável** em longo prazo.
- ∫ O **acesso ao público** externo ao acervo pode ser tornar extremamente **restrito** ou mesmo inviável.
- ∫ Os **ciclos de obsolescência** cada vez mais **acelerados** tornam as **peças de reposição** e os **profissionais qualificados** para a manutenção cada vez mais caros e escassos.

{ Segunda Estratégia: Emulação

- ∫ A emulação **simula plataformas de hardware e/ou software obsoletas**, possibilitando a **interpretação dos objetos digitais em sua concepção original**.
- ∫ Recriam o comportamento do hardware e dos softwares obsoletos em computadores novos.
- ∫ As estratégias de emulação são particularmente relevantes para preservar tanto a **aparência, quanto os recursos dos objetos digitais**.
- ∫ **Preserva o objeto digital original**.

{ Segunda Estratégia: Emulação

∫ Problemas

- ∫ O **emulador deverá contemplar atualizações**, pois poderá sofrer **obsolescência** como qualquer outro software. Ou ainda, pode necessitar **de emular o próprio emulador**, aumentando demasiadamente o grau de complexidade da preservação.
- ∫ Necessidade de **desenvolver um emulador ou aquisição de licenças de uso**. Estes são fatores que além de encarecer o plano de preservação **poderão torná-lo inviável na preservação em longo prazo**. *O uso de emulador proprietário, a dependência de um emulador específico e a sua obsolescência resultarão em instabilidades ao plano de preservação, podendo impossibilitar a recuperação dos documentos.*

{ Terceira Estratégia: Encapsulamento

- ∫ Consiste num **pacote de metadados anexados ao arquivo a ser preservado** permitindo futuro desenvolvimento de suas funcionalidades. Pode incluir no pacote **além de metadados, software e arquivos específicos constituintes do recurso digital, que possibilite seu acesso no futuro.**
- ∫ Permite manter o **objeto digital original inalterado até o acesso** a este material seja requerido. Quando solicitado, o objeto digital original deverá ser tratado para tornar o acesso possível.

{ Terceira Estratégia: Encapsulamento

∫ Problemas:

- ∫ Recuperar os objetos digitais em **software proprietário e de código fechado** implica em uma série de **restrições com relação ao acesso** a sua estrutura interna fator que pode inviabilizar o ciclo de preservação.
- ∫ O encapsulamento aplicado **isoladamente não tem como foco recuperar documentos digitais de forma imediata**, sua aplicação tem por objetivo reunir toda a informação necessária para a sua recuperação no futuro.

{ Quarta Estratégia: Migração

- ∫ Migração periódica **de um patamar tecnológico** em vias de se tornar **obsoleto** e/ou de se degradar fisicamente **para outro mais atualizado**.
- ∫ Voltada para a preservação do **objeto conceitual**.
- ∫ A migração parte do princípio de converter ou atualizar os formatos de arquivo considerados **antigos para formatos atuais**. Consiste na transferência periódica de material digital de uma dada configuração de hardware/software para uma outra, ou **de uma geração de tecnologia para outra subsequente**.

{ **Refrescamento**

- ∫ Trata-se da transferência de objetos digitais contidos em um determinado **suporte físico de armazenamento**, o qual é considerado obsoleto, para outro suporte considerado atual.
- ∫ O refrescamento está concentrado na preservação do **objeto físico, o suporte de armazenamento**.
- ∫ *Se o suporte físico se deteriorar ou se tornar obsoleto haverá risco do seu conteúdo ser perdido de forma definitiva.*

{ Convergência entre as Estratégias

Sincronia e Dependência

{ Preservação de Tecnologia & Emulação & Encapsulamento

- ∞ Preservação de tecnologia é uma estratégia fundamental para compreender o contexto tecnológico no qual o documento digital está inserido.
- ∫ É possível extrair informações relevantes para o desenvolvimento de emuladores
- ∫ É possível extrair informações relevantes para o encapsulamento.

{ Preservação de Tecnologia & Migração & Encapsulamento

- ∫ Preserva o objeto digital original, juntamente com seu contexto de hardware e software **enquanto pesquisa-se pela migração considerada ideal.**
- ∫ Preservar a tecnologia esperando a migração ideal deve ser aplicada somente **em curto prazo**, para longos períodos de tempo recomenda-se o Encapsulamento

{ Migração & Refreshamento

- ∞ Os documentos são extraídos do suporte, em seguida são migrados e finalmente são gravados no novo suporte, incorporando assim, refreshamento e migração.
- ∞ Devido à obsolescência tecnológica e a própria fragilidade dos suportes, o refreshamento deverá ser uma estratégia sempre presente nos acervos a fim de evitar a perda física dos documentos digitais.

{ Metadados de Preservação

Metadados

- ⋞ Metadados são um conjunto dados que identificam determinado documento.
- ⋞ São representações das informações documentadas.
- ⋞ Informação estruturada que descreve, explica, localiza, ou ainda possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar.

Metadados de Preservação

- ⋞ Parte essencial das estratégias de preservação digital.
- ⋞ Todas as estratégias de preservação dependem de vários tipos de dados que informem sobre histórico, características técnicas, estruturas, dependências e alterações sofridas pelo objeto digital, entre outros.
- ⋞ Viabilizam o pleno acesso e permitem a recriação e a interpretação da estrutura e do conteúdo da informação digital ao longo do tempo.

5 Categorias de Informação de Metadados

{ 1. Proveniência

∫ Devem registrar informações sobre a...

∫ ... **história do objeto** desde sua origem...

∫ traçando a sua **cadeia de custódia** e de **propriedade**.

5 Categorias de Informação de Metadados

{ 2. Autenticidade

∫ Devem incluir informações suficientes para...

∫ validar que o objeto é de fato o que diz ser e...

∫ ... que não sofreu alterações, intencionais ou não, não documentadas.

5 Categorias de Informação de Metadados

{ 3. Atividades de Preservação

∫ Devem documentar as **ações tomadas ao longo do tempo** para...

∫ ... **preservar** o objeto digital e ...

∫as **consequências dessas ações sobre aparência,
usabilidade e funcionalidades** do objeto.

5 Categorias de Informação de Metadados

{ 4. Ambiente Técnico

∫ Devem descrever as dependências técnicas necessárias...

∫ ... para a apresentação e uso dos objetos digitais...

∫ .. tais como hardware, sistema operacional e software de aplicação.

5 Categorias de Informação de Metadados

{ 5. Gestão de Direitos

∫ Devem registrar todos os itens relacionados às...

∫ ... questões de **propriedade intelectual** que **limitem as ações de preservação...**

∫ ... **de disseminação e uso** por parte de usuários de hoje e do futuro.

{ Padrões de Preservação Digital

- ⊆ BRTF-SPDA: Blue Ribbon Task Force on Sustainable Digital Preservation and Access
- ⊆ NISO: Z39.87 Data Dictionary - Technical Metadata for Digital Still Images
- ⊆ **OAIS: Open Archival Information System Reference Model**
- ⊆ PAIMAS: Producer-Archives Interface Methodology Abstract Standard
- ⊆ PREMIS: Preservation Metadata Implementation Strategies
- ⊆ TDR : Trusted Digital Repository Checklist (ISO 16363)
- ⊆ TRAC: Trustworthy Repositories Audit & Certification

OAIS

Open Archival Information System (OAIS) ou

Sistema Aberto para Arquivamento de Informações (SAAI)

- ⋞ Orienta um sistema de arquivo para a **preservação e manutenção do acesso a informações digitais por longo prazo**
- ⋞ Descreve as especificações em termos de metadados necessários para um repositório.
- ⋞ Estabelece **responsabilidades** que uma organização deve distribuir para operar como um arquivo de acesso livre.

- ⋮ Modelo genérico ISO 14721:2002 (ABNT NBR 15472), *aplicável a qualquer contexto de preservação digital.*

- ⋮ A norma descreve um ambiente digital **aberto, interoperável** e com garantias de **confiabilidade**, que se autodefine como uma organização de pessoas e sistemas que aceitaram a responsabilidade de preservar a informação e torná-la disponível para uma comunidade-alvo.

⋆ Organizações que têm a responsabilidade de tornar a informação disponível a longo prazo:

⋆ *Arquivos, bibliotecas, centros de informação*, entre outros, ...

⋆ ... bem como organizações que *adquirem tais informações*.

- ⌞ O modelo OAIS desdobra-se em três partes distintas:
 - ⌞ **Produtor**: pessoas ou sistemas que fornecem a informação a ser submetida e preservada.
 - ⌞ **Administrador**: entidade responsável pelo estabelecimento de políticas e regras gerais do arquivo, como a escolha do formato a ser adotado como padrão para acesso e preservação de longo prazo.
 - ⌞ **Consumidor**: pessoas ou sistemas que irão interagir com os serviços seja por meio de um repositório ou sistema de gerenciamento arquivístico.

{ Plataformas Tecnológicas Disponíveis

⊆ ARCHIVEMATICA

⊆ DAITTS (Dark Archive in the Sunshine State)

⊆ DPSP (Digital Preservation Software Platform)

⊆ DURACLOUD

⊆ LOCKSS

⊆ PRESERVICA

⊆ RODA (Repository of Authentic Digital Objects)

{ Archivemática

⋞ Software livre para criação de repositórios digitais, desenvolvido em código aberto e com acesso ao código fonte.

⋞ Artefactual Systems

⋞ Programa Memória do Mundo da Unesco

⋞ Arquivo Municipal da Cidade de Vancouver, Canadá

⋞ Biblioteca da Universidade de British Columbia, Canadá.

⋞ Arquivo Central de Rockefeller, Canadá.

⋞ Arquivo da Universidade Simon Fraser, Canadá.

- ⋞ Contempla **preservação de longo prazo...**
- ⋞ ... evitando a **obsolescência tecnológica** e...
- ⋞ ... garantindo que os documentos digitais **permaneçam autênticos, acessíveis e utilizáveis** ao longo dos anos.

⋞ O repositório do Archivematica aceita **múltiplos formatos** de arquivos, incluindo:

⋞ Documentos textuais

⋞ Músicas

⋞ Vídeos

⋞ Imagens

⋞ Plantas

⋞ Entre outros.

⋞ Segue o padrão **OAIS**.

⋞ O Archivematica utiliza os padrões de **metadados**:

⋞ Mets, **Premis**, **Dublin Core**, BagIt

⋞ e *outros padrões e práticas* para fornecer **pacotes de arquivamento confiáveis, autênticos, confiáveis e interoperáveis** para realizar o armazenamento em qualquer repositório escolhido.

- ⌞ O grande marco do Archivematica é oferecer aos Arquivistas e aos Bibliotecários **confiança para fazerem as preservações digitais de hoje de forma segura e prática.**
- ⌞ O Archivematica utiliza basicamente **três etapas** distintas dentro do repositório:
 - ⌞ SIP | Pacote de informação de submissão
 - ⌞ AIP | Pacote de Informação de Arquivamento
 - ⌞ DIP | Pacote de informação de difusão

{ Pacote de Informação de Submissão (SIP)

- ∫ Primeira etapa após a instalação do repositório
- ∫ O produtor fará a **submissão dos pacotes de informação**, transferindo os arquivos para o repositório.
- ∫ Após a submissão dos pacotes, é possível **inserir os metadados** que auxiliarão na posterior recuperação destas informações submetidas.
- ∫ Após a transferência, os pacotes serão submetidos ao *ingest*, quando são realizadas a **conversão de formatos antes de serem arquivados**.

{ Pacote de Informação de Submissão (SIP)

∫ Metadados

1. Proveniência
2. Autenticidade
3. Atividades de Preservação
4. Ambiente Técnico
5. Gestão de Direitos

{ Pacote de Informação de Arquivamento (AIP)

∫ Segunda etapa

∫ **Arquivamento do pacote de informação** submetido anteriormente.

∫ *Para documentos com algum grau de sigilo o serviço acaba aqui, não será realizada a terceira etapa para a difusão dos documentos digitais.*

{ Pacote de Informação de Difusão (DIP)

- ∫ Terceira e última etapa ocorre a **disseminação da informação** ostensiva, por meio do software **ICA AtoM** (Acesso à Memória), software para descrição e difusão de documentos arquivísticos de caráter permanente.
- ∫ Trata-se de um software para gestão de conteúdo baseado em padrões para descrição arquivística e nas normas internacionais de Arquivos do Conselho Internacional de Arquivos.

{ Pacote de Informação de Difusão (DIP)

- ∫ Nesta etapa será feita a descrição dos documentos digitais, de acordo com as normas internacionais do ICA, que abrange:
 - ∫ ISAD(G): Norma Geral Internacional de **Descrição Arquivística**.
 - ∫ ISAAR (CPF): Norma Internacional de **Registro de Autoridade Arquivística** para Entidades Coletivas, Pessoas e Famílias.
 - ∫ ISDIAH: Norma Internacional para **Descrição de Instituições**.
 - ∫ ISDF: Norma Internacional para **Descrição de Funções**.

{ Identificadores Persistentes

- ↳ Identificadores persistentes ou permanentes são...
- ↳ recursos digitais que **permanecem independentemente da localização do objeto digital** isto é, mesmo que se mude a URL.
- ↳ O identificador persistente assegura que **mesmo quando um documento é movido os links para ele permaneçam** efetivamente acionáveis
- ↳ Estão relacionados diretamente com os **padrões de preservação digital**, sendo um componente importante da infraestrutura dos repositórios digitais.

- ⋞ Identificador de Objeto Digital (DOI).
- ⋞ É composto de números e letras, e atribuído ao objeto digital para que este seja unicamente identificado na Internet.
- ⋞ Oferece identificação unívoca da propriedade intelectual de **livros, artigos, periódicos e até imagens** encontrados na Internet, associando a cada objeto seus dados básicos e sua origem.

↳ Originou-se de iniciativa conjunta de três associações comerciais na indústria editorial:

↳ International Publishers Association.

↳ International Association of Scientific, Technical and Medical Publishers.

↳ Association of American Publishers.

↳ A CrossRef é uma associação de editores e instituições que publicam na internet e que necessitam registrar seu conteúdo e metadados de forma única e persistente. Ela é o principal agente de atribuições de DOI no mundo.

Conclusão

{ **International Research on Permanent Authentic Records in Electronic Systems (InterPARES)**

- ‡ Grupo de Pesquisa Internacional sobre **Documentos Arquivísticos Autênticos Permanentes em Sistemas Eletrônicos**.
- ‡ Coordenado pela Universidade de British Columbia, **Canadá**.
- ‡ O resultado dos estudos desenvolvidos no projeto InterPARES têm gerado novos conhecimentos e referenciais para a preservação em longo prazo para que os documentos arquivísticos se mantenham autênticos e inalterados, de acordo com os princípios arquivísticos.

InterPares

{ Resultados do InterPares

- ∫ Requisitos conceituais para avaliar a autenticidade dos documentos arquivísticos digitais.
- ∫ Modelos de processos de avaliação e de preservação para documentos arquivísticos digitais autênticos.
- ∫ Base de dados de terminologia.
- ∫ Modelos conceituais de preservação.
- ∫ Diretrizes para produção e preservação de documentos digitais autênticos incluindo Diretrizes do Produtor e Diretrizes do Preservador.

CONARQ

{ Carta para a Preservação do Patrimônio Arquivístico Digital

- ∫ Considerando o **patrimônio arquivístico digital**
- ∫ Considerando que este patrimônio arquivístico digital se encontra em **perigo de desaparecimento**
- ∫ Considerando que a Carta para a Preservação do Patrimônio Digital da **Unesco**

{ Alerta para os problemas

- ∫ Dependência da informação digital
- ∫ Rápida **obsolescência** da tecnologia digital
- ∫ Incapacidade dos atuais sistemas eletrônicos de informação em **assegurar a preservação** de longo prazo
- ∫ **Fragilidade** intrínseca do armazenamento digital
- ∫ Complexidade e **custos da preservação** digital

{ Implementação de ações

- ∫ Elaboração de **estratégias e políticas**.
- ∫ Estabelecimento de **normas**.
- ∫ Promoção do **conhecimento**.

Referências

ARELLANO, M. A. Preservação de documentos digitais. *Ciência da Informação*, Brasília, v. 33, n. 2, 2004.

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

SANTOS, H. M. dos; FLORES, D. Estratégias de preservação digital para documentos arquivísticos: uma breve reflexão. *Cadernos BAD*, n. 1, p. 87-101, 2015.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, v. 15, n. 30, p.1-31, 2010.