

UNIDADE 3

PRESERVAÇÃO DIGITAL

3.1 OBJETIVO GERAL

Apresentar os aspectos de preservação dos objetos da biblioteca digital e algumas de suas principais iniciativas.

3.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) avaliar os objetivos da preservação digital;
 - b) identificar suas principais técnicas e métodos;
 - c) analisar algumas de suas principais iniciativas.
-

3.3 INTRODUÇÃO

Como estudamos na Unidade 1, atualmente existe uma grande facilidade em criar objetos digitais e, a partir daí, formar coleções é um passo que também é facilitado pela tecnologia. *Márdero Arellano* (2004, p. 16) alerta que “[...] muitas coleções digitais importantes estão sendo construídas fora das bibliotecas por diferentes organizações, ou sendo publicadas diretamente na Internet”. Entretanto, se, por um lado, a tecnologia atua como um elemento facilitador da publicação e do acesso, por outro, oferece desafios, dentre os quais podemos citar:

- fragilidade das mídias;
- multiplicidade de tipos de conteúdo e um número crescente de formatos de arquivo;
- grandes volumes de dados digitais;
- falta de clareza sobre procedimentos e responsabilidades.

Um dos primeiros relatos desses problemas aparece em 1995, no artigo de *Jeff Ruthenberg* no periódico *Scientific American* (v. 272, n. 1, p. 42-47), intitulado “Ensuring the Longevity of Digital Documents”, em que o autor relata sua preocupação com a substituição dramática do documento em papel pelo digital, e a possibilidade da perda desses documentos digitais, devido a sua fragilidade. A partir daí, vê-se um reconhecimento gradual do problema, entre profissionais da informação, e algumas iniciativas começam a surgir.

3.4 PRESERVAÇÃO DO OBJETO DIGITAL

Para a maioria dos bibliotecários, os desafios associados à preservação digital são imensos:

- incapacidade de determinar por onde começar a ação;
- falta de conhecimentos especializados;
- falta de ferramentas confiáveis e simples de utilizar;
- incapacidade de determinar custos; entre outros.

No entanto, como bem observa *Ferreira* (2011), o contexto digital não deve ser entendido pelas bibliotecas como uma ameaça, mas, sim, como uma oportunidade de redefinição, tirando partido das novas tecnologias em prol de uma melhor distribuição e acesso da informação, sem deixar, no entanto, de identificar os desafios que o objeto digital apresenta para sua preservação.

O objetivo da preservação digital é garantir o acesso e a fidedignidade das informações originais em qualquer ambiente tecnológico disponível, a qualquer tempo (CONARQ, 2004). Para isso, é necessário planejar as

Preservação digital

De acordo com o *Glossário de Documentos Arquivísticos Digitais do Conarq*, a preservação digital pode ser definida como um:

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, 2014, p. 29).

Pode-se entendê-la como a capacidade de garantir que a informação digital permaneça acessível e com qualidade de autenticidade para que possa, no futuro, ser interpretada numa plataforma tecnológica diferente daquela utilizada em sua criação (FERREIRA, 2006).



ações, prever a alocação de recursos e a aplicação de métodos e tecnologias de preservação necessárias (HEDSTROM, 1998). Cabe destacar que, como sabemos, o objeto digital possui uma complexidade diferente daquela do documento convencional, sendo assim, preservar um documento digital é bem mais difícil. Se no convencional nossa preocupação está voltada para o suporte, no digital essa é apenas uma das preocupações, estando a dificuldade maior em preservar o conteúdo em si (FERREIRA, 2011).

Considerando todos os aspectos envolvidos, uma questão desafiadora no ambiente digital é como prover o acesso permanente ao objeto e sua apresentação adequada, dadas as mudanças tecnológicas que têm ocorrido e que irão continuar a ocorrer (HODGE; FRANGAKIS, 2004).



Curiosidade

Hoje, o “MovieCD” (produzido na década de 1990) pode ser considerado o primeiro formato de vídeo para computadores oficialmente desaparecido, já que seu *codec* “MotionPixels” não foi comprado por nenhuma empresa ou teve seu código-fonte disponibilizado. De acordo com *Pereira* (2012, p. 101),

Nada resta de seu *codec*, a página do desenvolvedor e do produto desapareceu, e [...] os poucos filmes que sobraram não podem ser reproduzidos pelos computadores atuais, nem por qualquer outro aparelho reproduzidor (PEREIRA, 2012, p. 101).

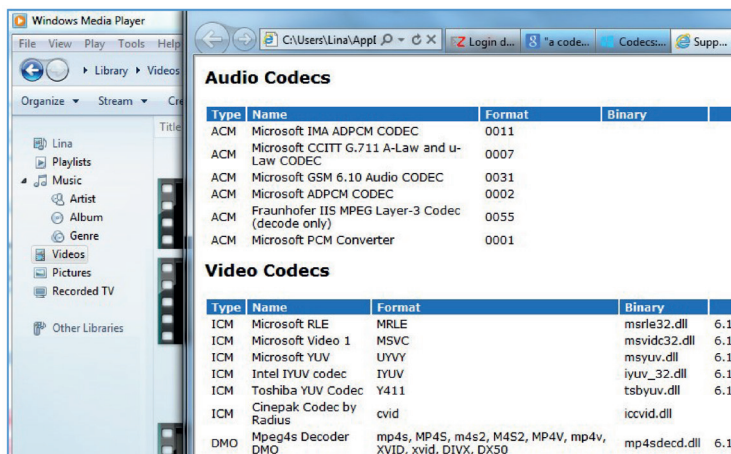


Explicativo

Os áudios e vídeos em formato digital costumam ser “empacotados” dentro de formatos que funcionam como embalagens que contêm todas as informações necessárias para apresentar o conteúdo; por exemplo, o fluxo de vídeo em si, o áudio, as legendas, informações sobre seções e sincronismo, entre outras.

Como arquivos de áudio e vídeo costumam ser muito grandes, sua compactação torna-se imperativa, em especial se consideramos seu uso via internet. Sendo assim, existem *softwares* específicos para compactar e descompactar esse tipo de conteúdo antes de eles serem “encapsulados” em sua forma final, ou seja, o arquivo que você baixa e assiste (ex.: .mp4, .avi etc.). Esses *softwares* são denominados *codecs*, e são usados para comprimir ou descomprimir um arquivo de mídia digital (áudio, vídeo). Muitas vezes você nem percebe sua existência, pois muitos dos *codecs* costumam vir pré-instalados nos *softwares* usados para reproduzir arquivos de mídia digital, por exemplo, o *Windows Media Player*. Veja, na imagem a

seguir, alguns *codecs* de áudio e vídeo que já vieram instalados em uma versão desse *software*.



Fonte: produção da própria autora.

3.5 MAS O QUE DEVEMOS PRESERVAR?

Como já falamos nas unidades anteriores, o objeto digital é composto de uma parte lógica e de uma parte física, que estão sujeitas à obsolescência das tecnologias, tanto de *software* quanto de *hardware*.

Dessa forma, quando pensamos em preservação de um documento digital, devemos levar em conta seu nível conceitual além do lógico e do físico. Ou seja, deve haver a preocupação com a integridade e a autenticidade da informação, com os formatos de arquivo e com as mídias, que se tornam obsoletas com o tempo.

Algumas pessoas podem pensar que as cópias de segurança (*backups*) são suficientes para preservar o documento digital. Mas repare que elas apenas reproduzem, em um meio físico, cópias do documento digital. Ora, se o meio físico depende de um equipamento para ser lido e se esse equipamento se torna obsoleto, então, nesse caso, a cópia não resolve. Por exemplo, os primeiros disquetes que surgiram foram os de 8 polegadas. Você sabe onde encontrar, hoje, um computador com uma unidade de leitura desses disquetes? Então, mesmo que o conteúdo esteja intacto dentro do disquete, se não encontrarmos o equipamento adequado para a leitura da mídia, de nada terá adiantado guardar a cópia de segurança. Outro problema é que a mídia se degrada com o tempo, ocasionando erro na leitura. Caso, por exemplo, você tenha encontrado um equipamento para ler o disquete de 8 polegadas (veja um na Figura 10), existe uma chance grande de seu conteúdo estar degradado e,

portanto, irrecuperável. Mas o problema não para por aí. Se você tem algum arquivo gravado em um meio tão antigo (os primeiros datam do final da década de 1960), provavelmente esse arquivo foi criado em um *software* bem antigo e sua especificação de formato de arquivo não está disponível, nem havendo mais nenhum *software* capaz de abri-lo. Então, mesmo que você encontre um equipamento para ler o disquete e o conteúdo não esteja corrompido, ainda assim é bem provável que não seja possível ler esse arquivo de modo que faça sentido para seu usuário. Nesse caso, estamos falando da obsolescência do *software*.

Figura 10 – Disquetes de diferentes tamanhos



Fonte: Wikimedia.³

Mas não podemos esquecer a importância da preservação intelectual, que diz respeito ao nível conceitual. De acordo com *Márdero Arellano* (2004), ela é mais crítica no caso dos documentos digitais do que nos convencionais, devido, principalmente, à capacidade de o objeto digital ser passível de modificação em sua forma de apresentação ou formato de publicação. Por exemplo, margens ou fontes podem ser mudadas, alterando a paginação. Essas mudanças, dependendo de como afetam o documento original, podem ser consideradas inaceitáveis pelo autor.

Considerando esses fatores, *Innarelli* (2007, p. 4) propõe dez mandamentos de boas práticas para a preservação digital:

1. Manterás uma política de preservação.
2. Não dependerás de *hardware* específico.
3. Não dependerás de *software* específico.
4. Não confiarás em sistemas gerenciadores como única forma de acesso ao documento digital.
5. Migrarás teus documentos de suporte e formato periodicamente.
6. Replicarás os documentos em locais fisicamente separados.
7. Não confiarás cegamente no suporte de armazenamento.
8. Não deixarás de fazer *backup* e cópia de segurança.
9. Não preservarás lixo digital.
10. Garantirás a autenticidade dos documentos digitais.

³ WIKIMEDIA COMMONS. George Chernilevsky. **Três dos mais populares formatos de disquete: 8, 5¼ e 3½ polegadas**. Disponível em: <https://pt.wikipedia.org/wiki/Disquete#/media/File:Floppy_disk_2009_G1.jpg>. Acesso em: 2 dez. 2018.

A adoção de uma política de preservação é importante para definir papéis, responsabilidades, bem como as práticas que devem ser adotadas, destacando-se que, conforme *Boeres e Márdero Arellano* (2005) mencionam:

A responsabilidade de que a informação científica e tecnológica seja preservada para benefício das futuras gerações está nas mãos dos criadores desses materiais (BOERES; MÁRDERO ARELLANO, 2005, p. 9).

Ainda, de acordo com esses autores, uma política de preservação digital implica observar aspectos como:

- a compatibilidade de *hardware* e *software*;
- a migração dos dados para prevenir a obsolescência tecnológica;
- a integridade do conteúdo intelectual a ser preservado;
- a análise dos custos envolvidos;
- o desenvolvimento e a formalização de uma política de seleção do que será preservado e, ainda,
- as questões de direito autoral (BOERES; MÁRDERO ARELLANO, 2005, p. 10).

Os autores destacam, também, princípios que devem ser seguidos em políticas de preservação digital, citando o contexto das instituições de ensino superior, tais como:

- a colaboração entre organizações e pessoas envolvidas na criação da política;
- as características dos *softwares* que vão abrigar os objetos digitais;
- a definição clara dos papéis e responsabilidades dos envolvidos no processo;
- o escopo das estratégias e métodos de preservação, que deverão estar contemplados em todo o ciclo de vida da informação digital (BOERES; MÁRDERO ARELLANO, 2005).

Dessa forma, para uma abordagem efetiva, devem-se contemplar vários aspectos organizacionais, tecnológicos, legais (GRACIO et al., 2013).

De acordo com *Gracio et al.* (2013), os aspectos organizacionais estão ligados aos objetivos da instituição, que deve estabelecer atos administrativos para apoiar uma política de preservação digital, definir responsabilidades, uma equipe multidisciplinar (dadas as diferentes facetas do objeto digital) e prover os recursos financeiros necessários.

Do ponto de vista legal, deve-se ter a preocupação de observar as leis e garantir os direitos autorais, quando for o caso.

Já do ponto de vista técnico, a preocupação deve ser: fazer a seleção do que se vai preservar e até quando; definir os modelos, padrões e iniciativas existentes que podem ser aproveitadas; definir uma infraestrutura tecnológica adequada, incluindo os repositórios digitais e, ainda, definir estratégias de preservação digital, mecanismos para garantir



a autenticidade dos objetos digitais e os suportes para eles, ou seja, as mídias e os metadados que vão descrever esses objetos.

A independência de *software* e *hardware* aponta para uma política de uso de *hardwares* que são compatíveis com várias tecnologias e de *softwares* não proprietários e abertos.

Guardar documentos digitais confiando, exclusivamente, em sistemas gerenciadores pode acarretar perda, no caso de esse sistema ser descontinuado. É importante ter acesso ao documento de outra forma também. Por exemplo, em uma estrutura de pastas independente.

Como vimos, a mídia se degrada ao longo do tempo. Dessa forma, em algum momento (dependendo da durabilidade estimada da mídia e da obsolescência dos equipamentos utilizados para acessá-la), deve-se migrar ou copiar os documentos digitais contidos em um determinado suporte para outra mídia mais atual. Quanto aos formatos, eles também tendem a se tornar obsoletos com o tempo. Assim, é importante monitorar essa obsolescência e, quando necessário, converter os documentos digitais armazenados no formato obsoleto para outro formato compatível.

Outro importante aspecto da preservação digital são as cópias de segurança (além dos *backups*). Entretanto, essas cópias devem ser armazenadas em locais fisicamente distantes dos *backups* originais, uma vez que ocorrências como incêndios, por exemplo, podem comprometer a integridade física das mídias em que os documentos estão armazenados. Se estiverem todas as cópias em um mesmo local, ambas ficarão perdidas. Cabe destacar que as cópias de segurança devem ser feitas periodicamente, de acordo com boas práticas para tal, e preferencialmente com ferramentas adequadas e pessoal capacitado.

Deve-se, ainda, estabelecer medidas para avaliar a integridade estimada das mídias e garantir que suas condições de armazenamento estejam adequadas, uma vez que, como mencionamos, elas se degradam com o tempo.

Como o volume dos documentos digitais tende a ser muito grande e crescente, é importante definir o que, de fato, deve ser preservado. Não é só a questão da quantidade de espaço disponível, mas também o tempo gasto para *backup*, para acesso, descrição, entre outras atividades.

Por fim, é importante que o documento se mantenha autêntico e íntegro ao longo do tempo, preservando sua forma e seu conteúdo originais.



3.5.1 Atividade

Explique o que você entende por preservação digital e por que ela é mais complexa do que a preservação de documentos convencionais.

Resposta comentada

O objetivo da preservação digital é garantir que o acesso e a interpretação de documentos digitais sejam possíveis em longo prazo, mesmo quando a tecnologia de *hardware* e *software* relacionada ao objeto digital for diferente. A preservação de documentos digitais é mais complexa do que a de documentos convencionais, pois envolve mais do que preservar o suporte físico: deve também preservar os aspectos intelectuais e lógicos.

Para atacar os problemas da preservação digital, vamos abordar três aspectos importantes relacionados: técnicas e métodos, iniciativas e metadados.

3.6 TÉCNICAS E MÉTODOS DE PRESERVAÇÃO DIGITAL

Márdero Arellano (2012) separa os métodos de preservação digital em dois grupos:

- operacionais;
- estruturais.

Os métodos operacionais dizem respeito a ações tomadas sobre os objetos digitais e consistem na aplicação de técnicas (que veremos a seguir) para a preservação desses objetos. Os métodos estruturais, por outro lado, dizem respeito a decisões ou providências tomadas em nível institucional para apoiar o processo de preservação digital, como a adoção de padrões, normas e metadados que veremos nesta Unidade.

Dentre as técnicas de preservação digital comumente mencionadas na literatura, temos:

- emulação;
- encapsulação (ou encapsulamento);
- preservação da tecnologia;
- migração/conversão;
- refrescamento (ou transferência).

A emulação consiste em tecnologias de *software* que reproduzem as condições necessárias de um ambiente original onde o *software* era executado, de modo que ele possa ser executado em um ambiente



diferente. Por exemplo, existem *softwares* que emulam jogos antigos que rodavam em ambientes hoje obsoletos, como o console de jogos Atari, de modo que eles possam ser jogados em ambiente Windows. Mas a emulação também pode ser usada para executar programas em ambientes diferentes, mas atuais. Por exemplo, é possível emular um ambiente de um *smartphone* Android em uma máquina *desktop* Windows, de modo que aplicativos que só rodariam originalmente em celulares Android possam ser executados nesse tipo convencional de máquina.



Multimídia

Veja alguns exemplos de emuladores de Atari e de Android nos sítios a seguir:

<<http://www.tecmundo.com.br/android/85518-5-programas-emular-android-computador.htm>>;

<<http://www.atari.com.br/emulacao/>>.

Embora, em alguns casos, a emulação seja considerada fundamental para preservar determinados recursos, ela possui uma complexidade e um custo que nem sempre a tornam viável (SANTOS; FLORES, 2014), além de muitas vezes apresentar problemas quanto ao grau de precisão do emulador em relação às características do ambiente do *software* original. Lembrando, ainda, que, como o emulador é um *software*, ele também pode vir a sofrer de sua própria obsolescência (SANTOS; FLORES, 2014).

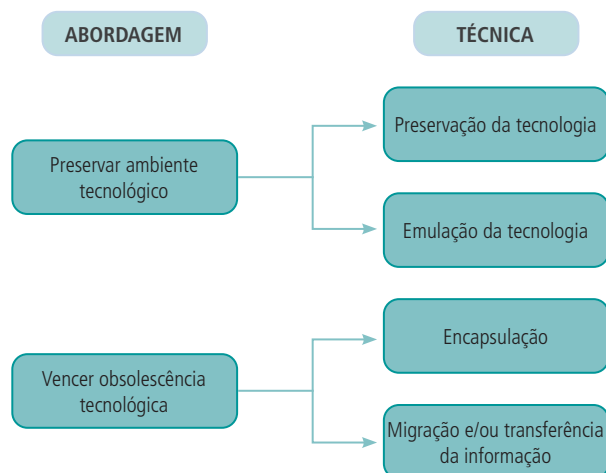
A encapsulação consiste em reunir, em conjunto com o recurso digital, o que quer que seja necessário para manter o acesso a ele. Ou seja, mantém o conteúdo digital original encapsulado em um envelope que contém instruções de *software* para recuperar, exibir e processar esse conteúdo.

Por sua vez, a migração/conversão consiste em transformar o documento digital de um formato para outro (ex.: doc para pdf-a). Nesse aspecto, destacamos a importância da adoção de formatos abertos, em especial os que são mantidos por organizações de padronização internacionais, como a ISO. Exemplos de formatos que se enquadram nesses critérios são o pdf-a e o odf.

A transferência, ou refrescamento, atua na mídia, tratando-se da mudança do suporte físico do documento digital para um suporte físico mais atual. Por exemplo, trocar o documento de um CD para outro mais recente, ou trocá-lo de CD para HD externo.

Essas abordagens podem ser agrupadas também em relação à estratégia adotada para a preservação, ou seja, se o foco é preservar o ambiente tecnológico ou vencer a obsolescência em si (LEE et al., 2002). A Figura 11 ilustra esse aspecto.

Figura 11 – Técnicas de preservação digital por tipo de abordagem



Fonte: adaptado de Lee et al. (2002).



3.6.1 Atividade

Associe a técnica de preservação ao nível físico, lógico ou conceitual, de acordo com o foco a que ela se destina.

Refreshamento: _____

Emulação: _____

Encapsulamento: _____

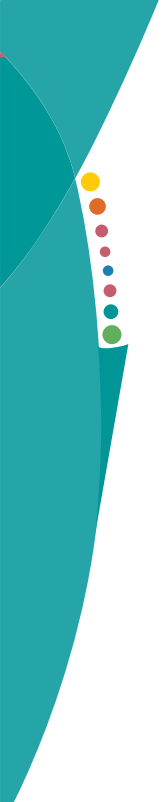
Migração/conversão: _____

Resposta comentada

O refreshamento atua no nível físico, pois trata da mídia; a emulação e o encapsulamento, no nível lógico, pois estão voltadas para a preservação da tecnologia usada para acessar o objeto digital; já a migração/conversão atua no nível conceitual, pois, ao migrar um documento ou converter seu formato, estamos preocupados com a preservação do patrimônio intelectual.

3.7 INICIATIVAS DE PRESERVAÇÃO DIGITAL

Como podemos observar, preservação digital é um procedimento complexo, com muitas facetas diferentes, tendo originado várias propostas



para dar conta do problema. Algumas das mais relevantes são: o modelo OAIS, o projeto InterPares e a Carta para a Preservação do Patrimônio Digital, da UNESCO, que serão abordadas a seguir, entre outras.

Iniciativas da UNESCO e do Conarq

A Carta para a Preservação do Patrimônio Digital (UNESCO, 2003) foi publicada em 15/10/2003 pela UNESCO, refletindo a preocupação com a questão da preservação do patrimônio digital mundial. Ela define 12 princípios que devem ser adotados de modo a prevenir a perda desse patrimônio para as gerações atuais e futuras. Ainda, indica um conjunto de medidas com o objetivo de elaborar estratégias, selecionar e proteger o que deve ser preservado.

Mais recentemente, dando seguimento às preocupações sobre o assunto e com o propósito de obter soluções efetivas para o problema, o Diretor Geral da UNESCO fez uma convocação para uma conferência internacional, intitulada “A memória do mundo na era digital: digitalização e preservação” (UNESCOVANC, 2012), que ocorreu em setembro de 2012 em Vancouver, no Canadá. Esse encontro teve como resultado uma série de recomendações para promover a preservação digital em nível mundial, envolvendo os estados membros da UNESCO, as organizações profissionais do setor de patrimônio cultural e do setor privado, cada qual com um conjunto específico de recomendações, mas todas com o mesmo intuito: promover a conscientização de que é necessário tomar medidas para evitar a perda do patrimônio digital que é de interesse público.

Seguindo essa linha de atuação, O Conselho Nacional de Arquivos (CONARQ), em julho de 2004, no Rio de Janeiro, aprovou a Carta para a Preservação do Patrimônio Arquivístico Digital (CONARQ, 2004). Por meio dela, convoca as instituições públicas e privadas a desenvolverem estratégias, políticas, normas e ações concretas para assegurar a preservação das informações digitais produzidas e armazenadas, assim como para compartilhar os conhecimentos adquiridos na área de preservação.

Dentre as iniciativas de preservação digital oriundas da área da Arquivologia, podemos citar, ainda, o “Modelo de Requisitos para Sistemas Informatizados de Gestão de Processos e Documentos da Justiça Federal” (MOREQ-JUS, 2009), que contém diretrizes e metadados para a preservação digital, que são aplicáveis também a documentos digitais não arquivísticos.

Projeto InterPARES

O projeto InterPARES (*International Research on Permanent Authentic Records on Electronic Systems*) é uma iniciativa acadêmica de pesquisa de preservação digital destinada a desenvolver soluções para a preservação de documentos arquivísticos digitais autênticos (INTERPARES, 2009), embora muitas de suas recomendações possam ser aplicadas a qualquer tipo de documento digital.

O projeto é hospedado na *School of Library, Archival and Information Studies*, na *Universidade de British Columbia*, em Vancouver, Canadá, sendo coordenado, atualmente, pela pesquisadora brasileira *Luciana Duranti*. Ele conta com a participação de pesquisadores de diversas áreas (Arquivologia, Biblioteconomia, Tecnologia da Informação, Informática,

Direito e outros), vindos da Europa, Ásia, África e Américas, sendo que o Brasil está presente no projeto, via participação do Arquivo Nacional.

É desenvolvido em fases, tendo se iniciado em 1998. Sua quarta fase (2013-2018) previa “[...] explorar questões relacionadas a registros digitais e dados confiados à Internet” (INTERPARES, 2009, n.p.), com o foco em desenvolver políticas, procedimentos, padrões etc., de âmbito local, nacional e internacional para a preservação da memória digital em longo prazo.

Modelo OAIS

Com a preocupação da preservação digital em mente e tendo como pano de fundo a preservação de dados digitais gerados em missões espaciais norte-americanas, o Comitê Consultivo para Sistemas de Dados Espaciais (CCSDS) da *Nasa*, contando com contribuições da comunidade científica internacional, criou o modelo de referência OAIS (*Open Archival Information System*) (LEE, 2010). Ele se tornou um padrão para a concepção de repositórios digitais confiáveis.

O modelo de referência OAIS especifica os principais aspectos que devem ser contemplados em iniciativas de preservação digital, abarcando seus processos, assim como os dados e metadados necessários. Entretanto, seu escopo também abrange o arquivamento de objetos físicos, além dos digitais.

De acordo com *Saramago* (2004, p. 3), o modelo OAIS “[...] descreve um enquadramento conceitual para um repositório digital genérico, aberto a todas as comunidades com garantias de confiabilidade”. Nesse sentido, ele ajuda os fabricantes de *softwares* de repositórios digitais a construir seus produtos de acordo com diretrizes que vão auxiliar na gestão e na preservação dos objetos digitais.

O OAIS se tornou um padrão ISO em 2003 (ISO 14721:2003) e especifica um modelo funcional e um modelo de informação. O modelo funcional descreve as atividades específicas realizadas por um repositório, tais como arquivamento e acesso. Já o modelo de informação descreve os tipos de informação que são trocadas e gerenciadas, bem como os metadados associados à preservação de um objeto digital ao longo do tempo.

Existem repositórios ou ferramentas de preservação que estão em conformidade ou são influenciados pelo modelo OAIS. Dentre eles, destacamos o DSpace (TANSLEY; BASS; SMITH, 2003), traduzido para o português e disponibilizado pelo IBICT (<www.ibict.br/dspace>):

O modelo de referência OAIS, ao detalhar sobre os tipos de informação necessária para a preservação dos objetos digitais (Informação Descritiva de Preservação), constitui, de certo modo, o ponto de partida para a reflexão sobre que metadados são necessários para assegurar a preservação e acesso dos objetos digitais em longo prazo (FERREIRA, 2011, p. 78).

Os metadados são muito importantes para a preservação digital, pois descrevem não só as características do objeto digital, mas também as ações que foram tomadas sobre ele ao longo de sua vida, como veremos a seguir.

Modelo de referência

Um modelo de referência é uma “[...] representação genérica abstrata para a compreensão de entidades de uma dada área e as relações significantes entre essas entidades” (CAMARINHA-MATOS; AFSARMANESH, 2006, p. 197).



3.7.1 Metadados de preservação digital

De acordo com *Ferreira* (2011), metadados de preservação são metadados descritivos, estruturais e administrativos que apoiam a preservação em longo prazo dos materiais digitais.

Márdero Arellano (2008, p. 75) os define como:

[...] aqueles relacionados com o conteúdo do documento, seu contexto e estrutura no momento de sua criação, assim como das mudanças acontecidas em todo seu ciclo de vida. Essa perspectiva inclui sua estrutura física, técnica e lógica (MÁRDERO ARELLANO, 2008, p. 75).

De acordo com *Sayão* (2010), metadados descritivos descrevem um recurso para fins de descoberta e identificação. Exemplos deles são: título, autor, palavras-chave, entre outros. Já os metadados estruturais descrevem como os objetos digitais complexos são estruturados; por exemplo, como as páginas digitalizadas de um livro são organizadas e ordenadas para formar um capítulo. Por fim, os metadados administrativos são informações relacionadas à gestão do ciclo de vida dos objetos digitais. Alguns exemplos são: dados sobre a data de criação, proveniência, dependências técnicas de *software* e *hardware*, versão do formato de arquivo, direitos legais, ações e consequências de ações de preservação adotadas sobre o objeto ao longo de sua vida.

Os metadados servem para descrever tanto a parte conceitual quanto a lógica e a física dos documentos digitais. Eles são agregados aos documentos pela ação humana ou, automaticamente, pelos *softwares* usados para criá-lo ou modificá-lo. No último caso, constituem uma camada de informação separada do conteúdo do documento (MOREIRA, 2012).

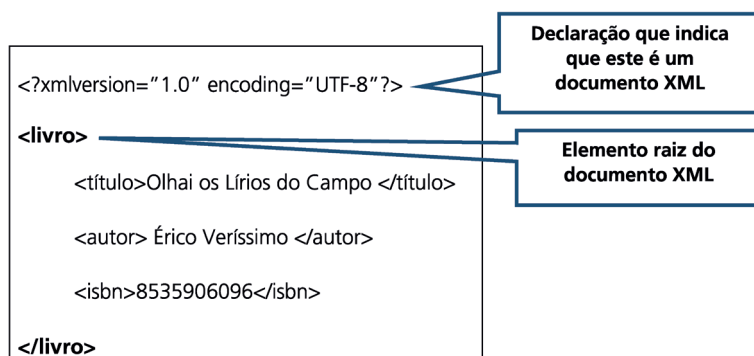
Como iniciativas importantes de metadados de preservação, podemos citar a PREMIS (*Preservation Metadata: Implementation Strategies*) (PREMIS, 2005) e a METS (*Metadata Encoding Transmission Standard*) (METS, 2016). Essas iniciativas costumam representar seus metadados no padrão **XML**.



Explicativo

XML

XML é uma linguagem de marcação padronizada pelo W3C que permite o armazenamento de dados de forma semiestruturada e, ainda, a criação livre de *tags* (rótulos para os dados). Um documento XML é de natureza textual, estruturado a partir de um elemento raiz, de forma hierárquica, formando uma “árvore” com seus elementos. Um exemplo de um documento XML simples é ilustrado na figura a seguir.



Fonte: produção da própria autora.

No documento de exemplo mostrado na figura, além do elemento raiz, que não possui conteúdo, observamos outros três que possuem: o título, o autor e o ISBN. Cada um desses elementos é iniciado por uma *tag* de abertura, que é delimitada pelos sinais de menor "<" e maior ">", e terminado por uma *tag* de fechamento, a qual é precedida por uma barra "/" após o sinal de menor "<". Entre as *tags* de abertura e fechamento, está o conteúdo do elemento. Um documento XML pode possuir outros detalhes. Se quiser saber mais, existe um bom tutorial sobre essa linguagem em: <<http://www.clem.ufba.br/tuts/xml/c03.htm>>.

3.7.2 Metadados PREMIS

De acordo com Sayão (2010, p. 21), a PREMIS surgiu com o objetivo de "detalhar os aspectos práticos de implementação dos metadados de preservação no contexto de sistemas de preservação digital", ou de colocar em ação os conceitos definidos pelo modelo de informação do OAIS. Ainda segundo o autor, ela teve como base o consenso extraído de experiências de várias instituições, tanto públicas como privadas, e da expertise de profissionais de bibliotecas, museus, arquivos etc. de vários países.

Como fruto da iniciativa, foi gerado um relatório, que contém um dicionário de dados (PREMIS *Data Dictionary*), que se constitui em um guia abrangente de metadados para apoiar a preservação digital de longo prazo (SAYÃO, 2010).

Um exemplo de metadados PREMIS é ilustrado a seguir, em que estão descritas propriedades relacionadas ao conteúdo de um documento. Essas propriedades são: o número de caracteres (*character count*), o número de palavras (*word count*), o número de linhas (*line count*) e o número de parágrafos (*paragraph count*). Tais metadados podem ser úteis para comparar se o documento mantém esses aspectos depois de um processo de migração.

Figura 12 – Exemplo de metadados PREMIS

```
-<premis:significantProperties>
  <premis:significantPropertiesType>character count</premis:significantPropertiesType>
  <premis:significantPropertiesValue>4063</premis:significantPropertiesValue>
</premis:significantProperties>
- <premis:significantProperties>
  <premis:significantPropertiesType>word count</premis:significantPropertiesType>
  <premis:significantPropertiesValue>712</premis:significantPropertiesValue>
</premis:significantProperties>
- <premis:significantProperties>
  <premis:significantPropertiesType>line count</premis:significantPropertiesType>
  <premis:significantPropertiesValue>33</premis:significantPropertiesValue>
</premis:significantProperties>
- <premis:significantProperties>
  <premis:significantPropertiesType>paragraph count</premis:significantPropertiesType>
  <premis:significantPropertiesValue>9</premis:significantPropertiesValue>
</premis:significantProperties>
```

Fonte: adaptado de *Metadados PREMIS* (2017).

3.7.3 Metadados METS (*Metadata Encoding Transmission Standard*)

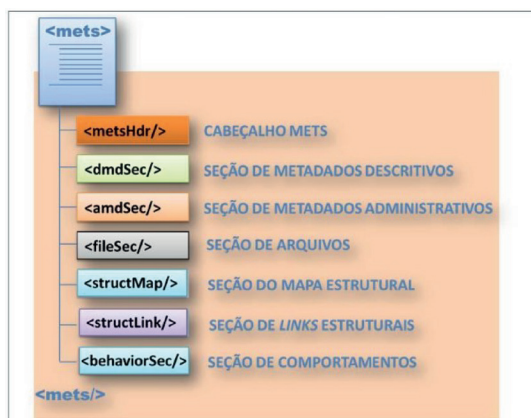
O METS é um esquema padrão para codificar metadados descritivos, administrativos e estruturais, relacionados a objetos em uma biblioteca digital. Ele é expresso na linguagem XML. O padrão METS (2016) é mantido pela *Library Of Congress* (Biblioteca do Congresso) dos Estados Unidos da América.

De acordo com *Sayão* (2007, p. 38):

O METS foi desenvolvido para preencher a necessidade de uma estrutura de dados padronizada para descrever objetos digitais complexos no contexto de uma biblioteca digital (NISO). O METS [...] provê a codificação e o encapsulamento de metadados descritivos, administrativos e estruturais necessários para a recuperação, apresentação, gestão, e preservação de objetos digitais dentro de um repositório e ainda para o intercâmbio entre repositórios (SAYÃO, 2007, p. 38).

De acordo com o sítio oficial do METS (2016), um documento METS é estruturado em sete seções, ilustradas na Figura 13, a seguir.

Figura 13 – Estrutura de um documento METS



Fonte: *Sales* (2014), com base em METS.

Um exemplo de metadados METS é ilustrado na Figura 14, em que se descreve um documento (denominado página1), que possui um *link* para outro documento (denominado página2). No exemplo, pode-se observar a seção *structLink*, usada para estabelecer a ligação entre essas duas páginas.

Figura 14 – Exemplo de metadados METS

Seção **StructLink** abaixo sendo usada para descrever *links* entre páginas:

```
<structLink>  
<smLinkxlink:from='página1' xlink:to='página2'/>  
</structLink>
```

Fonte: produção da própria autora.



3.7.4 Atividade

Imagine um cenário onde você é funcionário de uma instituição que possui uma biblioteca digital que não tem iniciativas concretas de preservação de seu acervo digital. Escreva uma carta para seu chefe explicando a gravidade da situação e sugerindo medidas que devem ser adotadas.

Resposta comentada

A carta deve explicar o risco de perda que o material digital apresenta e sugerir iniciativas que contemplem: a definição de processos, normas, políticas, mecanismos de gestão, e também a questão da definição de metadados, que são fundamentais para apoiar os diversos aspectos envolvidos na preservação do contexto e das características dos objetos digitais.

Semestre

7

3.8 CONCLUSÃO

A produção de objetos digitais tem evoluído de maneira crescente e não se pode mais imaginar o mundo sem documentos dessa natureza. Entretanto, esse tipo de documento é exposto a uma fragilidade que é mais complexa do que a do documento convencional: a obsolescência tecnológica, que possui várias facetas que vão além do suporte físico. Nesse cenário, surge a preocupação com a preservação digital desses documentos, que visa assegurar sua acessibilidade e integridade em longo prazo. A preservação digital engloba um leque variado de aspectos, que atuam em diferentes níveis de especificidade. Por um lado, mais global, existem iniciativas que expressam preocupações em definir políticas, normas, modelos, estratégias de gestão. Por outro, mais específico, existe a preocupação com as técnicas que devem ser adotadas para a preservação do objeto em si e metadados que devem ser usados para descrevê-lo e, assim, ajudar a preservar seu contexto e suas características. Estar ciente da necessidade de se pensar a preservação digital e entender seus objetivos deve ser o papel do profissional da informação. Da mesma forma, é importante saber o que tem sido feito nesse sentido e aprender com as experiências já conduzidas, de modo a poupar esforços.

RESUMO

A seguir uma relação dos tópicos mais importantes que foram apresentados na Unidade 3.

Preservação do objeto digital

A preservação digital visa garantir o acesso e a interpretação de documentos digitais em longo prazo, mesmo quando a tecnologia de *hardware* e *software* relacionada ao objeto digital for diferente. Ela envolve mais do que preservar o suporte físico: deve também preservar os aspectos intelectuais e lógicos e, para isso, deve-se pensar em estratégias globais para o planejamento dessas iniciativas, suas técnicas e métodos.

Técnicas e métodos de preservação digital

Dentre as várias técnicas e métodos de preservação digital, destacamos:

- o refrescamento, que atua no nível físico, pois trata da mídia;
- a emulação e o encapsulamento, que atuam no nível lógico, pois estão voltados para a preservação da tecnologia usada para acessar o objeto digital;
- a migração/conversão, que atua no nível conceitual, pois, ao migrar um documento ou convertê-lo de formato, estamos preocupados com a preservação do patrimônio intelectual.

Iniciativas de preservação digital

Dentre as várias iniciativas de preservação digital, destacamos:

- a) a Carta para a Preservação do Patrimônio Digital, publicada em 15/10/2003, pela UNESCO;
- b) a Carta para a Preservação do Patrimônio Arquivístico Digital, elaborada pelo Conarq, no Brasil, seguindo o exemplo da UNESCO;
- c) o projeto InterPARES (*International Research on Permanent Authentic Records on Electronic Systems*), iniciativa acadêmica de pesquisa de preservação digital, voltada para desenvolver soluções para a preservação de documentos arquivísticos digitais autênticos;
- d) o modelo OAIS, que especifica os principais aspectos que devem ser contemplados em iniciativas de preservação digital, abarcando seus processos, assim como os dados e metadados necessários. O modelo descreve uma proposta conceitual para um repositório digital genérico, aberto, interoperável e confiável;
- e) metadados de preservação digital, que servem para descrever tanto a parte conceitual quanto a lógica e a física dos documentos digitais. Como iniciativas importantes de metadados de preservação, podemos citar a PREMIS (*Preservation Metadata: Implementation Strategies*) e a METS (*Metadata Encoding Transmission Standard*). De forma complementar, elas definem uma série de metadados que vão ajudar na questão da preservação digital.

