

Preservação de documentos arquivísticos digitais: reflexões sobre o uso de padrões abertos nos acervos

Henrique Machado dos Santos^{*}
Daniel Flores^{*}

Artículo recibido:
1 de enero de 2016
Artículo aceptado:
3 de mayo de 2016

RESUMEN

Este artículo tiene como objetivo llevar a cabo una reflexión sobre el uso de estándares abiertos dentro de las colecciones, específicamente con respecto al software, formatos de archivo y sistemas de gestión y conservación. Así, se busca explorar los requisitos que deben ser considerados en la producción y conservación de los documentos para colocarlos en una política de preservación. Desde un punto de vista metodológico, este estudio es una investigación aplicada que lleva a cabo una revisión bibliográfica de los materiales publicados anteriormente en la que los datos se analizaron cualitativamente. Se inicia con un análisis que va

* Universidade Federal de Santa Maria, Brasil
henrique.hms.br@gmail.com danielflores@ufsm.br

de lo general a lo específico a partir del ámbito de la preservación digital, a través de las estrategias y políticas para llegar a la importancia del uso de estándares abiertos.

Palabras clave: Estándares abiertos; Documento digital; Preservación digital; Acceso a largo plazo

Preservation of digital archival documents: reflections on the use of open standards in collections

Henrique Machado-dos-Santos, Daniel Flores

ABSTRACT

This paper provides a reflection on the use of open standards within collections, specifically with regard to software, file formats and management and preservation systems, exploring requirements of a preservation policy for the production and preservation of documents. The applied methodology deployed in this research entails a literature review from which data is gathered and subjected to qualitative analysis, which moves from the general to the specific, i.e., from examining the scope of digital preservation to the strategies and policies in order to argue the importance of using open standards.

Keywords: Open standards; Digital document; Digital preservation; Access in long-term

INTRODUÇÃO

Os constantes e desenfreados avanços das tecnologias da informação e comunicação vêm modificando os hábitos da sociedade contemporânea, principalmente no que tange a troca de informações e conhecimentos. Esta forte demanda por inovação não foi diferente dentro dos limites epistemológicos e pragmáticos da Arquivologia: o documento antes produzido somente em suportes tradicionais como, por exemplo, o papel, começou a também ser produzido em meio digital, através dos computadores e registrado em suportes eletrônicos. Desta forma, a relação orgânica do documento

digital com os demais documentos passou a denominá-lo como documento arquivístico digital.

A informática se incorporou na gestão e na disseminação, tanto dos documentos tradicionais, quanto dos documentos digitais (Bellotto, 2006). Além disso, uma parcela significativa dos documentos arquivísticos está sendo produzida e armazenada em meio digital (Conselho Nacional de Arquivos do Brasil, 2004a; Conselho Nacional de Arquivos do Brasil, 2004b; Innarelli, 2011; Interpares 2 Project, 2002-2007; Santos e Flores, 2015a; Thomaz, 2005 e 2006). Os fatos apresentados montam um cenário de renovação, onde as teorias e as práticas arquivísticas passam por um processo de reformulação para que possam contemplar a natureza digital dos documentos.

Neste contexto, documentos de interesse para a sociedade começaram a ser produzidos em meio digital e a sua preservação ganhou relevância no âmbito da Arquivologia e da Ciência da Informação. Desta forma, tornou-se necessária a implementação de estratégias e políticas de preservação digital para garantir o acesso contínuo em longo prazo à documentos arquivísticos digitais íntegros e autênticos.

Este artigo tem por objetivo realizar uma reflexão sobre o uso de padrões abertos nos acervos, perpassando pela sua utilização no âmbito dos *softwares*, formatos de arquivo e sistemas de gestão e preservação. Pretende-se ir além da implementação de estratégias de preservação digital, muito bem fundamentadas e já aprofundadas em diversos trabalhos (Baggio e Flores, 2013; Corrêa, 2010; Ferreira, 2006; Grácio e Fadel, 2010; Hedstrom, 2001; Heminger e Robertson, 2000; Lopes, 2008; Márdero Arellano, 2008; Rothenberg, 1999; Santos e Flores, 2014 e 2015b; Santos, 2005; Saramago, 2002; Thibodeau, 2002; Thomaz, 2004). Assim, será possível explorar os requisitos que devem ser considerados na produção e na preservação dos documentos, e então situá-los nas definições de uma política de preservação.

Do ponto de vista metodológico, este estudo parte de uma pesquisa de natureza aplicada, na qual os dados são analisados de forma qualitativa. Realiza-se um levantamento bibliográfico de materiais previamente publicados como livros, teses, *sites* da *Internet* e artigos de periódicos científicos indexadas no *Google Scholar* (Gil, 2010; Luna, 1997; Silva e Menezes, 2005). Desta forma, parte-se de uma análise que perpassa do geral para o específico: iniciando no escopo da preservação digital, passando pelas estratégias e políticas até chegar à importância do uso de padrões abertos para a preservação em longo prazo.

PRESERVAÇÃO DIGITAL: CONTEXTO, CONCEITO E FINALIDADE

Desde que a informação começou a ser registrada pelas sociedades, tanto na forma de desenho quanto na forma escrita, surgiram novos suportes para registrar a memória, seja ela coletiva ou individual. A partir desse momento, o registro de novas informações e conhecimentos tornou-se mais fácil, além de permanecerem fixados por muito mais tempo. E foi por intermédio desses suportes, que grandes acontecimentos políticos, sociais e culturais das mais diversas civilizações, foram preservados.

Entretanto, muitos desses registros vindos de sociedades passadas se perderam ao longo do tempo, ou estão se deteriorando em virtude da despreocupação ou inexistência de métodos adequados de preservação. Destaca-se que o mesmo que aconteceu com os documentos de gerações passadas, se reflete nos dias de hoje com os documentos digitais. Não existem métodos plenamente eficazes para a preservação em longo prazo, os documentos digitais são muito frágeis e isto acrescenta uma série de incertezas quanto a sua longevidade.

Após a adesão das ferramentas de tecnologia da informação e do próprio documento digital pela sociedade, não se questionou a sua permanência neste cenário. Isto fez com que todas as inovações posteriores fossem aceitas *a priori* sem qualquer análise, isto porque não se pensou em diferenciar as tecnologias mais “simples” das mais “complexas”. Tal fato torna a preservação digital um grande dilema, pois não se tem controle dos avanços da tecnologia, e mesmo assim, documentos digitais continuam sendo produzidos. (Santos e Flores, 2015a: 47)

A preservação digital surge com o objetivo de garantir a integridade, a autenticidade e o acesso em longo prazo aos documentos, de modo que a informação contida neles possa ser interpretada corretamente. O termo “*Digital Preservation*” então traduzido como “Preservação Digital”, está sendo amplamente utilizado desde a década de 90 (Thomaz, 2004), abrange tanto os documentos gerados diretamente em meio digital, quanto àqueles os quais foram convertidos para um formato digital. Dentre suas atividades, podem ser destacados o planejamento, a alocação de recursos e a aplicação de métodos e tecnologias consideradas necessárias para garantir o acesso contínuo em longo prazo à informação digital (Hedstrom, 1998).

Atualmente, os problemas relacionados à preservação de documentos digitais vêm crescendo rapidamente propagando seus efeitos sobre a memória contemporânea. A fragilidade intrínseca do meio digital e a consequente degradação física dos suportes é uma das principais ameaças que circundam os

documentos digitais. Entretanto, preservar apenas o suporte onde estão armazenados os documentos digitais, não será suficiente para garantir o acesso em longo prazo, é preciso ir muito além, definindo políticas que contemplem também o *hardware* e o *software*. “Não é exagero afirmar que os próprios responsáveis pela preservação, como bibliotecas e arquivos, estão com dificuldades para se adaptar a esta nova realidade. Conceitos necessitam ser reformulados e novas teorias precisam ser criadas para contemplar o lado digital da documentação”. (Santos e Flores, 2015a: 46).

Outro fator que dificulta a preservação dos documentos digitais são os formatos de arquivo utilizado, pois estes formatos são afetados pela obsolescência em nível de *software*. Desta forma, o armazenamento de documentos em formatos de arquivo obsoletos coloca em risco a garantia de acesso a estes, bem como a suas qualidades de integridade e autenticidade. É preciso questionar como o *software* decodificador dos documentos permanecerá estável ao longo do tempo, e se haverá suporte ao produto.

DOCUMENTO DIGITAL: ESTRUTURA LÓGICA E INTERAÇÃO

O documento digital pode ser claramente separado em *hardware*, *software* e suporte, entretanto, há especificidades para além desta divisão. A seguir, a *Figura 1* permite detalhar as especificidades do *software*, apresentando os atributos oriundos da estrutura lógica dos documentos digitais.

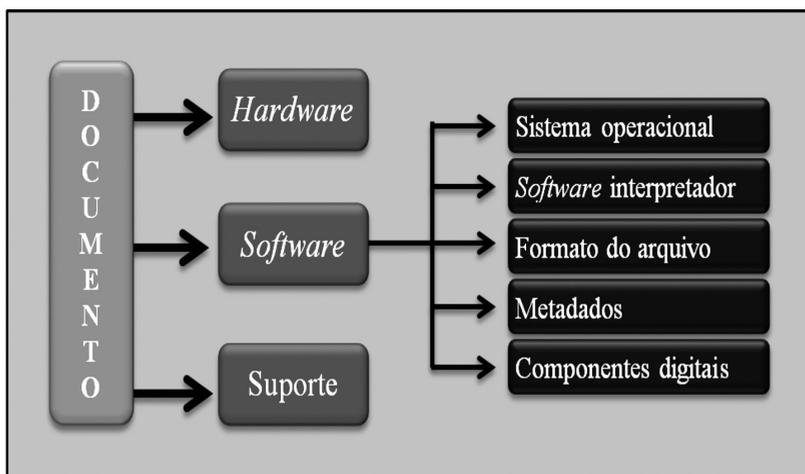


Figura 1. Componentes lógicos do documento digital
Fonte: elaborado pelos autores

Conforme a *Figura 1*, o documento digital é representado através de uma estrutura composta por *hardware*, *software* e suporte. A parte referente ao *software* compreende o sistema operacional, o *software* interpretador, o formato de arquivo, os metadados e os demais componentes digitais.

O sistema operacional compreende o “todo maior”, é a plataforma de *software* onde estão inseridos os mais diversos aplicativos que corroboram para o funcionamento do sistema. Nesta plataforma o usuário do computador pode inserir outros aplicativos que auxiliem suas atividades cotidianas ou lazer.

O *software* interpretador está inserido em um sistema operacional, sendo que sua função é decodificar o conteúdo dos documentos digitais. Como exemplos de *softwares* interpretadores, podem-se citar os editores de texto, os *browsers* e os *players* de música. Estes são os responsáveis pela correta leitura do que está codificado.

O formato de arquivo vem a ser a estrutura dotada de regras lógicas que codifica o conteúdo de documento digital. Cada formato tem suas especificidades e pode ser gerado por diferentes *softwares* ou a partir de um único. Desta forma, questões como a qualidade, riqueza de detalhes e tamanho lógico variam conforme o formato de arquivo. Existem diversos formatos de arquivo, sejam eles para áudio (mp3, wma, wav, acc), vídeo (avi, flv, mov, mkv), texto (doc, pdf, odt, rtf), imagem (jpg, png, tif, gif), entre outros (exe, html, css, accdb).

Metadados na linguagem popular da Arquivologia, Ciência da Informação e Ciência da Computação, significa “dados sobre dados”, ou seja, são informações que descrevem um objeto digital. A informação dos metadados auxilia na compreensão do contexto tecnológico, do histórico de alterações e demais informações que forem relevantes para a preservação dos documentos digitais.

Os componentes digitais correspondem todos os demais objetos digitais que corroboram para a correta representação do documento digital. Pode-se citar a critério de exemplo, uma página web, esta possui um arquivo principal em formato “html” e diversos outros objetos, como áudios, vídeos e imagens, que são referenciados por meio deste objeto principal. Da mesma forma, os metadados do documento poderão estar armazenados em um objeto digital separado em formato “xml”. Há casos, porém, onde o documento digital não possuirá componentes externos, por exemplo, arquivos em formatos “pdf”, “doc”, “jpg”, os quais apresentem todo o seu conteúdo sem a necessidade de outros componentes digitais.

Percebe-se que cada atributo lógico desempenha uma função específica e que não pode ser substituído por outro atributo. Desta forma, a manutenção da integridade e da autenticidade do documento digital, bem como a sua

correta representação, depende da conformidade destes elementos com os requisitos definidos *a priori*. A *Figura 2* apresenta um detalhamento destes elementos, sintetizando as suas relações diretas e indiretas em um contexto tecnológico genérico.

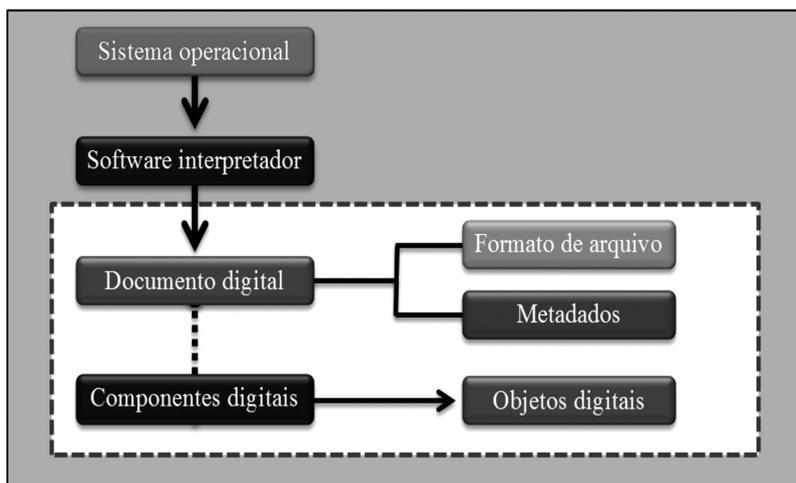


Figura 2. Interação entre componentes lógicos

Fonte: Elaborado pelos autores

Neste contexto, o “sistema operacional” é a plataforma que suporta os mais diversos softwares, sem ele não é possível instalar nenhum aplicativo, nem acessar quaisquer documentos. Tem-se o sistema operacional como uma base para o funcionamento dos demais atributos de *software*.

O “*software* interpretador” é instalado sobre uma plataforma de “sistema operacional”, e então, é possível criar, editar e acessar o “documento digital”. Percebe-se assim, que há uma relação muito estreita entre o documento digital e o *software* interpretador. Observando o documento digital, há uma relação indissociável estabelecida com o “formato de arquivo” e com os “metadados”. Todo documento digital é codificado e armazenado em um respectivo formato de arquivo, e por sua vez, possui um conjunto de metadados que fornecem informações básicas de sua criação. O documento digital, o formato de arquivo e os metadados são criados e estruturados por um *software*, logo, determinados *softwares* criam documentos em formatos de arquivo específicos, e utilizam um conjunto de metadados padronizado.

Paralelamente aos atributos apresentados, podem existir ainda os “componentes digitais” que são fundamentais para correta representação do

documento digital. Dentre estes componentes, podem-se incluir os mais diversos “objetos digitais”, nos mais diversos formatos de arquivo. No momento da decodificação, o *software* interpretador buscará informações no documento digital e nos seus respectivos componentes digitais. Vale lembrar que estes componentes digitais nem sempre serão necessários para a plena representação do documento, ou seja, existem documentos digitais “simples” e “compostos”.

POLÍTICAS DE PRESERVAÇÃO DIGITAL: NECESSIDADES, DEFINIÇÕES E IMPLICAÇÕES

Por muito tempo, os documentos arquivísticos digitais foram produzidos de forma desenfreada sem que houvesse qualquer preocupação com a sua preservação. A “novidade” do meio digital alicerçada pela sua praticidade de criação, edição e disseminação, levou ao pensamento de que a sua preservação estaria garantida, e que este, seria mais confiável que os documentos em suportes tradicionais como o papel. Esta hipótese foi rapidamente descartada em virtude das perdas decorrentes da fragilidade dos suportes eletrônicos e da própria obsolescência tecnológica. Além disso, constatou-se que a sua facilidade de edição era também um ponto fraco, pois não deixava vestígios e assim comprometia a autenticidade dos registros.

Entretanto, quando as comunidades da Arquivologia e da Ciência da Informação perceberam o fato da vulnerabilidade dos documentos digitais, muito já tinha se perdido, além de haver outros registros sob o risco de perda. Neste sentido ocorreu uma exaustiva busca por tecnologias que solucionassem o problema, tão ameaçador para a memória das sociedades contemporâneas. Não foram encontradas soluções definitivas, talvez essa seja a questão mais indagadora da preservação digital até aqui. O que pode ser feito é definir um conjunto de regras e recomendações que facilitem o processo de preservação. Dentre estas questões, podem ser destacadas algumas iniciativas como a implementação de padrões de metadados que auxiliem na presunção de autenticidade; a adoção de padrões abertos que facilitem a recuperação no futuro; a implementação de sistemas informatizados e estratégias de preservação digital, de modo que essas atividades venham a gerar a confiabilidade necessária ao acervo.

Inicialmente as políticas de preservação devem descrever os requisitos diplomáticos de forma fixa, conteúdo estável, variabilidade limitada. Em seguida definem-se os procedimentos para adoção de estratégias e softwares a serem

utilizados, implementação de repositórios digitais, escolha dos formatos de arquivo recomendados para preservação e adoção dos padrões de metadados. Durante o planejamento da preservação, a tecnologia deve estar hierarquicamente abaixo das políticas institucionais, com isto, entende-se que os sistemas informatizados deverão estar em conformidade com os requisitos de preservação em longo prazo, definidos previamente, garantindo integridade, autenticidade e confiabilidade aos documentos armazenados. (Santos e Flores, 2015c: 205)

A preservação de documentos arquivísticos digitais dependerá da definição de políticas organizacionais, isto é fato, a prevenção será a melhor forma de preservar o patrimônio digital. Após definir uma política de preservação, devem-se escolher as estratégias que farão parte do plano de preservação digital, sempre considerando os requisitos diplomáticos relacionados.

A política de preservação deverá informar claramente as estratégias adotadas, informando qual nível de abstração do objeto se pretende preservar (físico, lógico ou conceitual). Além disso, deve considerar o contexto social, econômico e organizacional em que se insere (Ferreira, 2006). O contexto local e temporal em que os documentos digitais estão inseridos poderá influenciar no processo de preservação, alterando as prioridades da organização, fazendo com que a preservação ganhe ou perca relevância. Deve-se destacar que este contexto é facilmente mutável em virtude de acontecimentos cotidianos que podem mudar a “escala de valores” da sociedade.

Basicamente, uma política de preservação digital deverá ser composta por um conjunto de normas, procedimentos e estratégias com a finalidade de garantir a integridade, a autenticidade do documento digital em longo prazo. Simultaneamente, é preciso estabelecer como estas políticas de preservação serão implementadas, a fim de gerar a confiança do público-alvo. Para que isto aconteça, Ferreira comenta que uma política de preservação digital deve envolver, além do arquivo, todos os setores da organização que dependem dos documentos digitais, sejam unidades de gestão documental, administrativas e de tecnologia da informação. Desta forma, deve-se proceder a avaliação e seleção de materiais, a identificação do padrão de metadados adequado, a definição de estratégias de preservação para cada formato de arquivo, a criação de planos de sucessão em caso de encerramento das atividades da organização (Ferreira, 2006).

A política organizacional de preservação digital deverá estabelecer claramente os procedimentos para a preservação física, lógica e intelectual dos documentos. Sendo assim, deverá contemplar as especificidades dos mais diversos formatos de arquivo presentes no acervo.

ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL: ESCOPO, PADRÕES ABERTOS E METADADOS

As estratégias de preservação digital compreendem o conjunto de objetivos e métodos para realizar a manutenção dos documentos digitais em longo prazo, contemplando os seus respectivos atributos relacionados, sendo possível garantir a autenticidade (Webb, 2003). Podem ser divididas em dois grupos: as de ordem estrutural e operacional. As estratégias estruturais são os investimentos iniciais como, por exemplo, a definição de normas, padrões e infraestrutura. As estratégias operacionais compreendem as atividades para a preservação física, lógica e conceitual dos documentos digitais, como por exemplo, migração, emulação, preservação de tecnologia e encapsulamento (Márdero Arellano, 2004; Thomaz, 2004).

Preservar um acervo arquivístico demanda a implementação de mais de uma estratégia operacional. Nesta etapa, o uso de padrões abertos fará uma diferença significativa, seja na questão dos direitos autorais, ou mesmo na compreensão do funcionamento dos *softwares* e conhecimento da estrutura dos formatos de arquivo. A escolha de padrões abertos com licenças de uso claramente definidas simplificará o processo de preservação no que tange a utilização de plataformas de *hardware* e *software*, além de reduzir os custos com licenças.

Destaca-se ainda, que durante a implementação das estratégias operacionais, devem-se adotar padrões de metadados que documentem informações relevantes a sua preservação como o histórico de alterações (Saramago, 2004), bem como o contexto tecnológico utilizado durante a criação do documento.

Os metadados são fundamentais para qualquer estratégia operacional, em especial para as estratégias de migração, as quais dependem fortemente deste registro. Da mesma forma, é recomendável usar padrões abertos para metadados de modo que estes possam ser atualizados para versões futuras (Márdero Arellano, 2008). Através da adoção de padrões de metadados que documentam toda e qualquer alteração, será possível verificar a integridade e a autenticidade dos documentos digitais.

Os metadados, enquanto parte integrante dos documentos arquivísticos digitais, auxiliam na análise diplomática, sendo possível estabelecer métodos que garantam a fidedignidade dos documentos (Rondinelli, 2005). Desta forma, os sistemas de gestão e preservação devem preservar os metadados de todos os objetos digitais que integrem os documentos arquivísticos digitais, objetivando a comprovação da fidedignidade.

PADRÕES ABERTOS: UM NOVO HORIZONTE PARA A PRESERVAÇÃO DIGITAL

Dentre as atividades de preservação digital, compreendem-se ações como a identificação e a avaliação de estratégias, *softwares*, suportes, padrões de metadados e formatos de arquivo escolhidos. Todas estas definições devem ser fundamentadas em uma política de preservação digital definida *a priori*, pois esta será a base para realizar a manutenção dos documentos, visando o acesso contínuo em longo prazo com garantia de fidedignidade.

A definição e a utilização de padrões abertos é altamente recomendável, e assim, deve-se considerar a sua ampla aceitação para criar, armazenar, transmitir e acessar documentos digitais (Conselho Nacional de Arquivos do Brasil, 2004a). De maneira genérica, os formatos abertos podem ser entendidos como os *softwares* ou formatos de arquivos criados por estes, os quais possuem suas especificações publicadas e disponibilizadas através de uma licença de uso gratuita. Assim, os formatos abertos podem ser amplamente utilizados e livres de *royalties* (Interpares 2 Project, 2002-2007). Neste sentido, é preferível que tanto o *software* que cria os documentos quanto o formato de arquivo criado por este possuam código aberto. Tais fatos facilitarão as atividades de preservação caso seja necessário conhecer especificidades do código fonte, seja para alterar as funcionalidades ou mesmo recriar o *software* ou o formato de arquivo.

Dentre a variedade de *softwares* e formatos de arquivo disponíveis, devem ser escolhidos aqueles que tenham uma expectativa que possibilite o acesso em longo prazo. Neste sentido, entende-se que os formatos abertos utilizados em larga escala oferecem maior confiança em virtude de sua sólida disseminação na sociedade. Além disso, em caso de obsolescência tecnológica, podem ser recriados para o novo contexto tecnológico sem que haja restrições decorrentes do direito de propriedade intelectual.

A adesão de formatos abertos deve contemplar os mais diversos componentes do plano de preservação digital, abrangendo assim, emuladores, conversores, padrões de metadados, sistemas informatizados de gestão e preservação, e os demais *softwares* utilizados no acervo. O uso de *softwares* proprietários torna o acevo dependente do suporte oferecido pelos desenvolvedores, além de restringir a possibilidade de reconstrução em caso de obsolescência. Desta forma, a descontinuidade de um determinado produto implica na migração para outra tecnologia considerada estável.

O uso de padrões fechados se configura como um empecilho para a preservação digital, pois este transforma o documento em uma “caixa preta”, não havendo nenhum controle por parte do acervo. Logo, o uso de padrões abertos minimiza o risco de obsolescência tecnológica, bem como reduz o

risco de perda dos documentos digitais no caso de uma migração de tecnologia (Innarelli, 2007), a qual é considerada uma ação de risco para a integridade e autenticidade dos documentos arquivísticos digitais.

Há diversos estudos e iniciativas surgindo no mundo todo, confirmando a necessidade de usar padrões abertos, além de substituir os *softwares* e migrar os documentos de formato fechado para formatos abertos, pois o seu uso permite o estudo e a conversão para novos padrões (Márdero Arellano, 2004). Desta forma, o processo de migração/conversão torna-se mais próximo do idealizado pela preservação.

Dentre os estudos de preservação digital, o Arquivo da Web Portuguesa (AWP) apresenta uma proposta relevante quanto aos formatos de arquivo. Trata-se uma lista apresentando diversos formatos de arquivo, os quais são classificados conforme uma escala de recomendação (alta, média e baixa). Esta lista compreende formatos de arquivo para texto, imagem, áudio, vídeo e outros.

Um formato de arquivo seja adequado à preservação deve ser livre de qualquer restrição de uso; ter o código aberto e documentado; ser amplamente utilizado; ser multiplataforma; preferencialmente sem compressões ou de compressão sem perdas; ser uma norma emitida por um organismo oficial (AWP, 2007).

Já um formato com características não recomendadas para preservação é aquele que possui direito proprietário, de código fechado; pouco difundido; criado e interpretado por poucas plataformas de *software* de código fechado; que possui compressão com perdas; composto por elementos internos como, por exemplo, macros (AWP, 2007).

Os critérios utilizados pelo AWP são relevantes para a escolha dos formatos de arquivo. Embora as recomendações sejam para a preservação de conteúdo de *sites*, podem ser facilmente adotadas para preservar documentos pessoais e institucionais. As recomendações de formatos de arquivo para a preservação do AWP (2007) convergem com estudos relevantes como os de Márdero Arellano (2004), Innarelli (2007) e do projeto Interpares (Interpares 2 Project, 2002-2007).

De maneira geral, constata-se que a preservação em longo prazo necessita ser mediada por padrões abertos, que sejam usados em larga escala, preferencialmente sem compressão, sendo acessíveis em diversas plataformas e se possível integrem uma norma reconhecida. Sendo assim, é preciso que os administradores de sistemas informatizados de preservação solicitem padrões abertos aos produtores de conteúdos.

O USO DE PADRÕES ABERTOS: DA PRODUÇÃO AOS SISTEMAS DE GESTÃO E PRESERVAÇÃO

As atividades de preservação digital devem envolver todo o ciclo de vida dos documentos arquivístico, isto abrange sua produção, tramitação e sua guarda em longo prazo. Desta forma, as atividades de preservação digital devem ser implementadas nas três idades documentais: corrente, intermediária e permanente.

Tendo em vista a preservação de documentos arquivísticos digitais, as fases corrente e intermediária deverão ser gerenciadas por um Sistema Informatizado de Gestão Arquivística de Documentos (SIGAD), este sistema deve possuir ferramentas que realizem a manutenção da integridade e da autenticidade dos documentos digitais, garantindo assim sua fidedignidade. Preferencialmente, o SIGAD deverá ser um sistema integrado com o repositório digital, assim, após cumprirem o valor imediato de sua função administrativa, os documentos de fase corrente e intermediária do SIGAD serão recolhidos ao repositório digital, onde serão preservados em longo prazo devido ao seu valor mediato.

Neste ponto, chama-se a atenção para a necessidade de garantir a interoperabilidade entre o SIGAD e o repositório. Estas tecnologias precisam estar em sincronia para cumprirem as funções atribuídas na política de preservação digital. Além do já exposto, deve-se atentar para o fato de que os documentos arquivísticos digitais em fase corrente e intermediária possuirão um determinado período de permanência no SIGAD antes de serem recolhidos ao repositório. A questão é que este período de permanência no SIGAD poderá ser suficiente para que ocorra algum problema de obsolescência tecnológica em virtude da rápida evolução das tecnologias da informação. Desta forma, em casos de uma documentação com muitas especificidades, podem-se implementar repositórios digitais em fase corrente/intermediária. Esta é uma questão muito peculiar dos documentos, um caso raro, mas não é descartável e por isto, as políticas de preservação devem contemplar todas as especificidades necessárias as documentos.

Inicialmente, os documentos digitais não podem ser armazenados em formatos obsoletos para que somente após longos períodos de tempo sejam recolhidos aos repositórios digitais. Isto se justifica porque há o risco de não haver compatibilidade retrospectiva, isto é, da tecnologia futura não recuperar as informações de formatos produzidos no passado. Este é um entrave que afeta tanto os produtores quanto as organizações detentoras de conteúdos digitais (Márdero Arellano, 2008). A tecnologia está evoluindo de forma acelerada, gerando ciclos de obsolescência cada vez menores (Thomaz e Soares, 2004). Este fato

reforça a necessidade de monitorar as tendências e contemplar as especificidades dos documentos digitais para evitar perdas.

A seguir na *Figura 3*, Flores (2014) apresenta um esquema que sintetiza o uso do SIGAD e do repositório digital, abordados anteriormente. Dentre estas, há algumas recomendações e padrões amplamente difundidos pela comunidade de preservação digital como o modelo *Open Archival Information System* (OAIS). O autor realça importância do uso de repositórios digitais confiáveis durante as três idades documentais: corrente, intermediária e permanente. A implementação de repositórios digitais é abordada como um ponto opcional nas fases corrente e intermediária, tornando-se elementar na idade permanente.

Há ainda peculiaridades brasileiras como o “Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos” (e-ARQ Brasil), o qual “[...] estabelece requisitos mínimos para um Sistema Informatizado de Gestão Arquivística de Documentos (SIGAD), independentemente da plataforma tecnológica em que for desenvolvido e/ou implantado” (Conselho Nacional de Arquivos do Brasil, 2011: 9). Em linhas gerais a e-ARQ Brasil especifica os requisitos para desenvolver o SIGAD, enquanto o OAIS define os requisitos para o repositório digital.

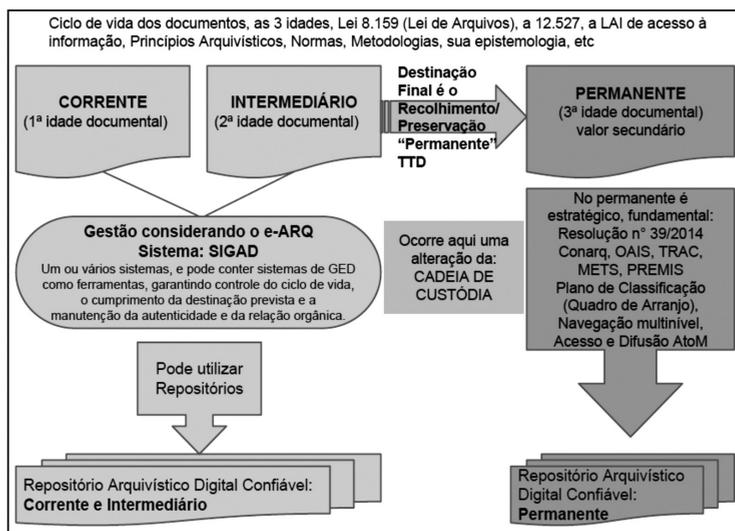


Figura 3. Ciclo de vida dos documentos

Fonte: Flores, 2014: slide 17

Além do uso de tecnologias adequadas para gestão e preservação, chama-se atenção para o uso de padrões abertos no que tange aos formatos de arquivo. Produzir documentos digitais em formatos recomendados para a preservação em longo prazo é primordial para o sucesso das atividades de qualquer acervo.

Os padrões para formatos de arquivo precisam ser planejados e implementados com o objetivo de possibilitar a gestão e o acesso aos documentos de guarda permanente com garantia de integridade e autenticidade (Innarelli, 2007 e 2012). Assim, a produção documental mediada por padrões abertos para *softwares*, formatos de arquivo e sistemas informatizados de gestão e preservação, diminui os riscos de obsolescência, perda ou corrupção de dados.

Paralelamente aos investimentos em tecnologia e a definição de políticas de preservação, é importante que o acervo divulgue para a sua comunidade-alvo quais são as normas e as metodologias de custódia adotadas. Esta é uma estratégia para que a comunidade-alvo julgue os métodos como confiáveis ou não.

CONCLUSÃO

Este artigo realizou uma reflexão sobre a necessidade do uso de formatos abertos nos acervos arquivísticos. O embasamento teórico utilizado se constituiu basicamente de autores das áreas da Arquivologia e da Ciência da Informação, destacando as complexidade e especificidades dos documentos digitais.

Dentre as constatações mais salientes, pode-se destacar a importância de estabelecer uma política de preservação *a priori*, contemplando o uso de padrões abertos tanto para *softwares*, formatos de arquivo e sistemas de gestão e preservação. A partir desta base é possível um conjunto de boas práticas para a preservação em longo prazo. O uso de sistemas informatizados e demais tecnologias de *hardware* e *software* deve ser fundamentado e regido por políticas que contemplem as características as quais se querem preservar em longo prazo.

O uso de formatos abertos, no que tange a *softwares*, formatos de arquivo e sistemas informatizados, proporciona maior liberdade seja para conhecer as especificidades do código fonte, e assim modificar suas funcionalidades ou recriá-los em caso de obsolescência. Da mesma forma, estas tecnologias precisam ser estáveis e permitir futuras atualizações, para não gerar a dependência de *softwares* específicos.

Por fim, a questão dos formatos abertos deve ser pensada antes mesmo de produzir os documentos digitais, desta forma, podem-se escolher formatos abertos amplamente utilizados, sem compressão e que sejam considerados como padrões de preservação. Destaca-se que toda a atividade de preservação digital realizada a *priori* da produção dos documentos, diminui os gastos com licenças de *softwares*, bem como os futuros investimentos em estratégias para atualizar formatos.

REFERÊNCIAS

- AWP (Arquivo da Web Portuguesa). 2007. *Formatos adequados para preservação*. Portugal: Lisboa. Acesso em: 04 nov. 2014. <http://sobre.arquivo.pt/colaboracoes/recomendacoes-para-autores-de-sitios-web/formatos-adequados-para-preservacao>
- Baggio, Claudia Carmem e Daniel Flores. 2013. “Documentos digitais: preservação e estratégias”. *Biblos: Revista do Instituto de Ciências Humanas e da Informação* (Rio Grande) 27 (1) (jan./jun.): 11-24. Acesso em: 10 abr. 2015. <http://www.seer.furg.br/biblos/article/view/2654/2395>
- Bellotto, Heloísa Liberalli. 2006. *Arquivos permanentes: tratamento documental*, 4. ed. Rio de Janeiro: Fundação Getúlio Vargas.
- Conselho Nacional de Arquivos. Câmara Técnica de documentos eletrônicos. 2004a. *Carta para a Preservação do Patrimônio Arquivístico Digital*. Rio de Janeiro: Arquivo Nacional. Acesso em: 10 ago. 2014. <http://www.conarq.arquivonacional.gov.br/Media/publicacoes/cartapreservpatrimarqdigitalconarq2004.pdf>
- Conselho Nacional de Arquivos. Câmara Técnica de documentos eletrônicos. 2004b. *Gestão Arquivística de Documentos Eletrônicos*. Rio de Janeiro: Arquivo Nacional. Acesso em: 09 jul. 2014. <http://pt.scribd.com/doc/37174068/Gestao-Arquivistica-de-Documents-Eletronicos-CONARQ-Por-Claudia-Rocha>
- Conselho Nacional de Arquivos. Câmara Técnica de documentos eletrônicos. 2011. *e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos*. Rio de Janeiro: Arquivo Nacional. Acesso em: 05 ago. 2014. http://www.conarq.arquivonacional.gov.br/media/publicacoes/earq/conarq_earqbrasil_model_requisitos_2009.pdf
- Corrêa, Amarílis Montagnolli Gomes. 2010. “Preservação digital: autenticidade e integridade de documentos em bibliotecas digitais de teses e dissertações”. Dissertação, Mestrado em Ciência da Informação, Universidade de São Paulo, São Paulo. Acesso em: 03 jul. 2014. <http://www.teses.usp.br/teses/disponiveis/27/27151/tde-05112010-105831/pt-br.php>
- Ferreira, Miguel. 2006. *Introdução à preservação digital: conceitos, estratégias e atuais consensos*. Portugal: Escola de Engenharia da Universidade do Minho. Acesso em: 02 ago. 2014. <https://repositorium.sdum.uminho.pt/bits-tream/1822/5820/1/livro.pdf>

- Flores, Daniel. 2014. “Inovação em Arquivos: a formação e a pesquisa em arquivologia diante da complexidade e especificidade do documento arquivístico digital, da gênese à preservação”, em *Encontro Internacional de Arquivos* (EIA), Portugal: Évora. Acesso em: 05 nov. 2014. http://eiarquivos2014.weebly.com/uploads/2/6/0/6/26068081/p1_3_df.pdf
- Gil, Antonio Carlos. 2010. *Como elaborar projetos de pesquisa*, 4. ed. São Paulo: Atlas.
- Grácio, José Carlos Abbud e Bárbara Fadel. 2010. “Estratégias de preservação digital”, em *Gestão, mediação e uso da informação*, Marta Valentim (org.), 59-83. São Paulo: Editora UNESP, Cultura Acadêmica. Acesso em: 01 dez. 2014. <http://books.scielo.org>
- Hedstrom, Margaret. 1998. “Digital preservation: a time bomb for digital libraries”. *Computer and the humanities* (Netherlands) 31: 189-202. Acesso em: 01 out. 2014. http://deepblue.lib.umich.edu/bitstream/2027.42/42573/1/10579_2004_Article_153071.pdf
- Hedstrom, Margaret. 2001. *Digital Preservation: Problems and Prospects*. University of Michigan, USA. Acesso em: 10 jan. 2015. http://www.dl.slis.tsukuba.ac.jp/DL-journal/No_20/1-hedstrom/1-hedstrom.html
- Heminger, Alan R. e Steven B. Robertson. 2000. “The Digital Rosetta Stone: a model for maintaining long-term access to static digital documents”. *Communications of AIS* 3, artigo 2 (jan.). Acesso em: 24 set. 2014. <http://dl.acm.org/citation.cfm?id=374497&dl=ACM&coll=DL&CFTOK>
- Innarelli, Humberto Celeste. 2012. *Instrumenta 2: Preservação de Documentos Digitais*. Associação dos Arquivistas de São Paulo. São Paulo: ARQ-SP.
- Innarelli, Humberto Celeste. 2011. “Preservação digital: a influência da gestão dos documentos digitais na preservação da informação e da cultura”. *RDBCi: Revista Digital de Biblioteconomia e Ciência da Informação* (Campinas) 8 (2) (jan./jun.): 72-87. Acesso em: 07 jul. 2014. <http://www.sbu.unicamp.br/seer/ojs/index.php/rbc/article/view/487>
- Innarelli, Humberto Celeste. 2007. “Preservação digital e seus dez mandamentos”, em *Arquivística: temas contemporâneos, classificação, preservação digital, gestão do conhecimento*, Vanderlei Batista Santos, org., 21-75. Distrito Federal: SENAC.
- Interpares 2 Project. 2002-2007. *Diretrizes do Produtor. A elaboração e a manutenção de materiais digitais: diretrizes para indivíduos*. TEAM Brasil. Tradução: Arquivo Nacional e Câmara dos Deputados. Acesso em: 09 ago. 2014. http://www.interpares.org/ip2/display_file.cfm?doc=ip2_creator_guidelines_booklet--portuguese.pdf
- Lopes, Vitor. 2008. *Preservação Digital*. Portugal: Universidade do Minho, Guimarães. Acesso em: 28 ago. 2012. http://www.vitorlopes.com/Trabalhos/Preservacao_Digital-Vitor_Lopes.pdf
- Luna, Sergio Vasconcelos de. 1997. *Planejamento de pesquisa: uma introdução*. São Paulo: EDUC.
- Márdero Arellano, Miguel Ángel. 2008. “Critérios para a preservação digital da informação científica”. Tese, Doutorado em Ciência da Informação, Universidade Federal de Brasília, Departamento de Ciência da Informação. Acesso em: 15 jun. 2014. http://btdt.bce.unb.br/tesdesimplificado/tde_busca/arquivo.php?codArquivo=4547

- Márdero Arellano, Miguel Ángel. 2004. "Preservação de documentos digitais". *Ciência da Informação* (Brasília) 33 (2) (maio/ago.): 15-27. Acesso em: 25 jul. 2014. <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/305/1452>
- Rondinelli, Rosely Curi. 2005. *Gerenciamento arquivístico de documentos Eletrônicos: uma abordagem teórica da diplomática arquivística contemporânea*, 4. ed. Rio de Janeiro: Editora FGV.
- Rothenberg, Jeff. 1999. "Avoiding technological quicksand: finding a viable technical foundation for digital preservation", em *Commission on Preservation and Access and Council on Library and Information Resources: a report to the Council on Library and Information Resources*. EUA: Washington, DC. Acesso em: 24 set. 2014. <http://www.clir.org/pubs/reports/rothenberg/pub77.pdf>
- Santos, Henrique Machado dos e Daniel Flores. 2015a. "As vulnerabilidades dos documentos digitais: Obsolescência tecnológica e ausência de políticas e práticas de preservação digital". *Biblios: Revista de Bibliotecología y Ciencias de la Información* (Lima) 59: 45-54. , 2015a. Acesso em: 04 jul. 2015. <http://biblios.pitt.edu/ojs/index.php/biblios/article/view/215>
- Santos, Henrique Machado dos e Daniel Flores. 2015b. "Preservação de documentos arquivísticos digitais autênticos: reflexões e perspectivas". *Acervo* (Rio de Janeiro) 28 (1) (jan./jun.): 241-253. Acesso em: 01 jul. 2015. <http://www.revistaacervo.an.gov.br/seer/index.php/info/article/view/740>
- Santos, Henrique Machado dos e Daniel Flores. 2015c. "Repositórios digitais confiáveis para documentos arquivísticos: ponderações sobre a preservação em longo prazo". *Perspectivas em Ciência da Informação* (Belo Horizonte) 20 (2) (abr./jun.): 198-218. Acesso em 03 jul. 2015. <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2341>
- Santos, Henrique Machado dos e Daniel Flores. 2014. "Estratégias de preservação digital em arquivos: garantia de autenticidade e acesso contínuo em longo prazo", em Congresso Nacional de Arquivologia, 6., Santa Maria. *Anais eletrônicos...* Santa Maria: AARS. Acesso em: 21 mar. 2015. https://www.academia.edu/17741035/ESTRAT%C3%89GIAS_DE_PRESERVA%C3%87%C3%83O_DIGITAL_EM_ARQUIVOS_GARANTIA_DE_AUTENTICIDADE_E_ACESSO_CONT%C3%8DNUO_EM_LONGO_PRAZO
- Santos, Vanderlei Batista dos. 2005. *Gestão de documentos eletrônicos: uma visão arquivística*, 2. ed. rev. aum. Brasília: ABARQ.
- Saramago, Maria de Lurdes. 2004, "Metadados para preservação digital e aplicação do modelo OAIS", em Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas 08. Estoril. *Anais eletrônicos...* Acesso em: 04 jul. 2014. <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640>
- Saramago, Maria de Lurdes. 2002. "Preservação digital a longo prazo: boas práticas e estratégias". *Cadernos BAD* (Lisboa) 2: 54-68. Acesso em: 10 fev. 2015. <http://www.bad.pt/publicacoes/index.php/cadernos/article/view/866>
- Silva, Edna Lúcia da e Estera Muszkat Menezes. 2005. *Metodologia da pesquisa e elaboração de dissertação*, 4. ed. rev. Florianópolis: UFSC. Acesso em: 13 jun. 2014. https://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf

- Thibodeau, Kenneth. 2002. “Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years: The State of Digital Preservation: An International Perspective2”, em *Anais eletrônicos...* Washington: CLIR and Library of Congress. Acesso em: 19 ago. 2014. <http://www.clir.org/pubs/reports/pub107/thibodeau.html>
- Thomaz, Kátia de Pádua. 2004. “A preservação de documentos eletrônicos de caráter arquivístico: novos desafios, velhos problemas”. Tese, Doutorado em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Acesso em: 28 jul. 2014. http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/VALA-68ZRKF/doutorado___katia_de_padua_thomaz.pdf
- Thomaz, Kátia de Pádua. 2005. “Documentos eletrônicos de caráter arquivístico: fatores condicionantes da preservação”. *Perspectivas em Ciência da Informação* (Belo Horizonte)10 (1) (jan./jun.): 34-53. Acesso em: 07 set. 2014. <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/301>
- Thomaz, Kátia de Pádua. 2006. “Gestão e preservação de documentos eletrônicos de arquivo: revisão de literatura – parte 2”. *Arquivística.net* (Rio de Janeiro) 2 (1) (jan./jun.): 114-131. Acesso em: 09 set. 2014. www.brapci.ufpr.br/download.php?dd0=6733
- Thomaz, Kátia de Pádua e José Antônio Soares. 2004. “A preservação digital e o modelo de referência Open Archival Information System (OAIS)”. *DataGramZero* (Rio de Janeiro) 5 (1) (fev.). Acesso em: 10 jun. 2014. www.dgz.org.br/fev04/Art_01.htm
- Webb, Colin. 2003. *Guidelines for the Preservation of Digital Heritage*. Preparado pela Biblioteca Nacional da Austrália para a Divisão de Sociedade de Informação, UNESCO, relatório n. CI-2003/WS/3. Acesso em: 17 Ago. 2014. <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>

Para citar este texto:

- Machado dos Santos, Henrique y Daniel Flores. 2018. “Preservação de documentos arquivísticos digitais: reflexões sobre o uso de padrões abertos nos acervos”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (74): 35-53.
<http://dx.doi.org/10.22201/iibi.24488321xe.2018.74.57905>