

REFLEXÕES SOBRE A PRESERVAÇÃO DE DOCUMENTOS ARQUIVÍSTICOS DIGITAIS: ESTRATÉGIAS E REPOSITÓRIOS

REFLECTIONS ON THE PRESERVATION OF DIGITAL ARCHIVAL DOCUMENTS: STRATEGIES AND REPOSITORIES

HENRIQUE MACHADO DOS SANTOS¹ E DANIEL FLORES²

Recebido em: 04/12/2014

Aprovado em: 13/01/2016

RESUMO

Os constantes avanços da Tecnologia da Informação contribuíram para a disseminação de suas ferramentas na sociedade contemporânea. Tais ferramentas também foram introduzidas no campo arquivístico, de modo que documentos de valor histórico e sociocultural começaram a ser produzidos apenas em meio digital, tornando necessária, assim, a sua preservação em longo prazo. Tendo isso em vista, este artigo aborda a implementação de estratégias e repositórios para a preservação em longo prazo dos documentos arquivísticos digitais. Com relação à metodologia, pode-se definir esta pesquisa como bibliográfica, de natureza aplicada e com abordagem qualitativa. Baseando-se no levantamento de materiais já publicados, discutiu-se a aplicação e o conceito das estratégias de preservação digital em documentos arquivísticos. O referencial teórico utilizado pertence principalmente aos campos da Arquivologia e da Ciência da Informação, estabelecendo relações interdisciplinares com outras áreas.

Palavras-chave: Tecnologia da Informação; Documento arquivístico digital; Preservação digital; Arquivologia.

ABSTRACT

The constant advances in information technology have contributed to the spread of their tools in contemporary society. Such tools have also been introduced in archival field, so that documents value history, social and cultural started to be produced just digitally, necessitating thus preservation in long-term. With this in mind, this article discusses the implementation of strategies and repositories for long-term preservation of digital archival documents. Regarding the methodology, you can set this as a bibliographical research of applied nature and with qualitative approach. Based on the survey of material already published, was discussed the application and the concept of digital preservation strategies in archival documents. The theoretical framework used to belong mainly to the fields of Archival Science and Information Science, establishing interdisciplinary relations with other areas.

Keywords: Information Technology; Digital archival documents; Digital preservation; Archival Science.

¹ Bacharel em Arquivologia pela Universidade Federal de Santa Maria (UFSM), Brasil. E-mail: henrique.hms.br@gmail.com.

² Doutor em Ciência da Informação pela Universidade Federal do Rio de Janeiro (UFRJ), Brasil. Atualmente é Professor Adjunto do Departamento de Documentação e do Mestrado em Patrimônio Cultural da Universidade Federal de Santa Maria (UFSM). E-mail: dlfloresbr@gmail.com.

1 Introdução

O advento da Tecnologia da Informação acarretou mudanças significativas nos hábitos do público em geral. Conforme Rondinelli (2005),

No mundo do trabalho, por exemplo, os profissionais da informação foram profundamente atingidos e, entre eles, os arquivistas. Tal informação se fundamenta no fato de que o avanço tecnológico mudou radicalmente os mecanismos de registro e de comunicação de informação nas instituições e, conseqüentemente, seus arquivos também mudaram. Ora considerando que os arquivos se constituem no principal objeto da arquivologia, fica evidente o impacto da informática sobre esse campo do conhecimento (p. 24).

Diante disso, o documento arquivístico adquiriu novo suporte e características próprias que foram proporcionadas pela dinâmica do meio digital. Dentre essas características, pode-se destacar a “sua produção, classificação, avaliação, descrição e comprovação de autenticidade” (SOUSA, 2007, p. 81).

Diversos estudos mostram a fragilidade dos documentos digitais, os quais possuem uma série de complexidades e especificidades, como vulnerabilidade em relação à corrupção de dados, podendo ser alterados, reformatados e falsificados sem deixar vestígios. Esses fatores podem comprometer a sua autenticidade e o seu acesso no futuro (CONARQ, 2004; 2011; 2012; CORRÊA, 2010; FERREIRA, 2006; INNARELLI, 2006; INTERPARES, 2007).

Devido à complexidade da natureza dos documentos digitais, faz-se necessária a interferência de profissionais especializados para definir políticas de preservação digital (MÁRDERO ARELLANO, 2004; INNARELLI, 2006; FERREIRA, 2006). As políticas de preservação são elementares, pois o uso inadequado das ferramentas de Tecnologia da Informação poderá ser

responsável “pelo desaparecimento de registros e, conseqüentemente, colocar em risco a integridade dos acervos arquivísticos” (PAES, 2004, p. 158).

Desse modo, este artigo tem como objetivo identificar as estratégias de preservação digital para documentos arquivísticos, visado à sua guarda permanente em repositórios digitais, com a finalidade de preservar e garantir que esses documentos possam ser acessados em longo prazo.

2 Referencial teórico

As estratégias de preservação digital devem ser precedidas de uma definição das políticas institucionais que nortearão as atividades no acervo. Da mesma forma, após a escolha das estratégias, estas devem ser implementadas seguindo uma metodologia própria para cada especificidade do acervo. Conforme apresenta Lopes (2008),

Para o sucesso da preservação digital é importante proceder-se à implementação da mesma, tendo como base uma estratégia de preservação que garanta a integridade dos objetos digitais independentemente dos seus formatos e suportes, bem como a evolução dos mesmos (p. 16).

Dessa forma, deve-se proceder à identificação das estratégias de preservação e verificar sua aplicabilidade ao acervo.

2.1 Estratégias de preservação digital

A finalidade das estratégias de preservação digital é realizar a “manutenção” dos documentos arquivísticos digitais, evitando a depreciação do registro no que tange tanto ao seu conteúdo quanto aos seus metadados. A esse respeito, ressalta-se que os metadados são essenciais para a comprovação da autenticidade documental, a qual é um princípio arquivístico elementar. Tendo em vista que ocorreram transformações

no documento arquivístico, observa-se que a análise diplomática também teve de ser adaptada para a realidade digital e que os metadados se constituem como elemento fundamental que devem ser preservados durante a implementação de estratégias de preservação para garantir a autenticidade dos documentos arquivísticos.

Ressalta-se que a implementação dessas estratégias busca minimizar e até mesmo evitar os problemas causados pela obsolescência tecnológica, tais como as dificuldades de leitura. Dessa forma, essas estratégias possibilitarão uma representação fidedigna dos do-

cumentos em um contexto tecnológico futuro, diferente do qual tais documentos foram originados, sendo assim uma ação em longo prazo.

As estratégias de preservação digital podem ser classificadas em três níveis, cada um com suas prioridades: físico, lógico e conceitual. Neste artigo, serão abordadas as seguintes estratégias: refrescamento, encapsulamento, emulação, preservação de tecnologia, pedra de Rosetta digital e migração (Figura 1). Tais estratégias, assim como as não contempladas neste estudo, possuem especificidades no que se refere à sua aplicação e ao seu nível de preservação.

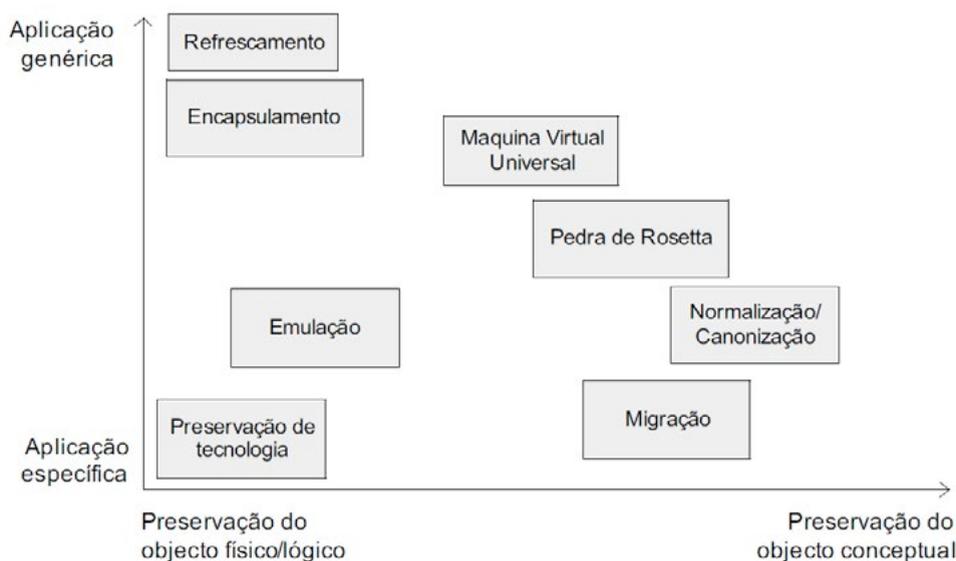


Figura 1 - Distribuição das estratégias.
Fonte: Thibodeau (2002 apud FERREIRA, 2006, p. 32).

2.1.1 Preservação do nível físico

A preservação do nível físico consiste em realizar esforços para garantir a integridade do suporte em que os documentos digitais estão armazenados. A única estratégia selecionada neste estudo para esse nível foi o refrescamento, também conhecido como rejuvenescimento ou migração de suporte.

O refrescamento consiste em transferir os objetos digitais contidos

em um determinado suporte físico de armazenamento, considerado antigo, para um suporte atual. É realizado antes que o suporte de armazenamento antigo se deteriore ou se torne obsoleto. A atividade de recopiar dados de um suporte físico para outro será uma atividade necessária sempre que o formato selecionado se tornar obsoleto (INTERPARES, 2007). Além disso, “a confiabilidade também deve ser verificada de forma periódica, pois o frágil

suporte digital sofre degradação física, o que pode colocar em risco a documentação digital” (INNARELLI, 2007, p. 57). Dessa forma, o refrescamento de suporte será uma atividade constante conforme a evolução das tecnologias de armazenamento.

2.1.2 Preservação do nível lógico

A preservação do objeto lógico consiste em priorizar a manutenção dos objetos digitais em sua forma original, mantendo o documento e os seus respectivos componentes digitais intactos. Dentre as estratégias de nível lógico que foram selecionadas neste estudo, estão a preservação de tecnologia, a emulação e o encapsulamento.

A preservação de tecnologia consiste na preservação e manutenção contínua de todo o *hardware* e *software* utilizados para a concepção do objeto digital em sua forma original. Fundamenta-se na criação de “museus tecnológicos”, que são as plataformas consideradas ideais para a correta representação dos objetos digitais.

Em sua implementação, além de salvar o *software* capaz de interpretar corretamente os objetos digitais, deve-se manter o *hardware* da máquina em correto funcionamento (MÁRDERO ARELLANO, 2008). Essa estratégia evita qualquer perda de informação, mantendo o objeto lógico livre de alterações em sua cadeia de *bits* e permitindo a obtenção de um alto grau de fidelidade aos objetos digitais.

Assim como a preservação de tecnologia, as estratégias de emulação partem do princípio de preservar o objeto lógico em seu formato original (FERREIRA, 2006), mantendo a integridade sobre o funcionamento e as características de todos os componentes digitais do documento. A emulação visa simular plataformas de *hardware* e *software* obsoletas, possibilitando a recuperação dos documentos digitais. As estratégias de emulação mantêm os do-

documentos inalterados em seu estado original, ou seja, a cadeia de *bits* permanece a mesma. Através das estratégias de emulação, é possível recuperar diversas plataformas obsoletas e possibilitar, assim, um ambiente para a preservação e o acesso aos documentos digitais.

Além da preservação de tecnologia e da emulação, pode-se recorrer ao encapsulamento para a preservação do nível lógico dos objetos digitais. O encapsulamento consiste em reunir as informações referentes aos suportes de armazenamento, isto é, a descrição do contexto tecnológico de *hardware* e *software* necessário para a correta interpretação dos objetos digitais. Dentre essas informações, estão as aplicações utilizadas durante o ciclo de vida dos objetos digitais, inclusive o *software* usado na sua criação (MÁRDERO ARELLANO, 2004; SARAMAGO, 2004). Dessa forma, são preservados os objetos digitais e todas as informações necessárias para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores (DIGITAL PRESERVATION TESTBED, 2001 apud FERREIRA, 2006), possibilitando a recuperação dos documentos digitais.

Em casos mais específicos, o encapsulamento pode ser utilizado para documentos textuais, utilizando o PDF/A1 conhecido como “PDF arquivístico”. Trata-se de um formato de arquivo estandardizado que garante a forma fixa e o conteúdo estável, preconizada pela diplomática contemporânea. O PDF/A1 permite anexar, por exemplo, as fontes que são utilizadas no documento textual.

2.1.3 Preservação do nível conceitual

Dentre as estratégias de nível conceitual, foram selecionadas neste estudo a pedra de Rosetta digital e a migração. Essas estratégias estão direcionadas para a preservação do nível conceitual, ou seja, do documento manifestado.

A pedra de Rosetta digital consiste na reprodução de documentos digitais em suportes analógicos. Segundo Thibodeau (2002), sua aplicação consiste, por exemplo, em imprimir em papel um conjunto de documentos textuais juntamente com a sua representação binária. Dessa forma, há uma expectativa de que no futuro seja possível interpretar as regras necessárias e migrar os objetos para novos formatos. Este processo teria como base o método de análise e comparação dos documentos impressos com a sua representação binária (THIBODEAU, 2002). É preciso escolher uma mídia durável no que se refere à qualidade, além de proporcionar o acondicionamento e o ambiente de conservação adequado.

Já as estratégias de migração estão centradas na preservação do objeto conceitual, isto é, no modo como o documento está sendo apresentado, independente da forma. Consistem, assim, em converter ou atualizar os formatos de arquivo considerados antigos para formatos atuais. Conforme a definição do *Task Force on Archiving of Digital Information* (1996 apud FERREIRA, 2006, p. 36), a migração é a “transferência periódica de material digital de uma dada configuração de *hardware/software* para uma outra, ou de uma geração de tecnologia para outra subsequente”. Com as estratégias de migração, há a possibilidade de os objetos digitais criados em um contexto tecnológico do passado continuarem sendo acessados e interpretados pelas tecnologias atuais.

2.2 Repositórios digitais

Com a intenção de abordar os repositórios digitais, deve-se esclarecer que “a preservação digital compreende os mecanismos que permitem o armazenamento em repositórios de dados digitais que garantiriam a perenidade dos seus conteúdos” (MARDERO ARELLANO, 2004, p. 17). Ou seja, o repositório digital faz parte do grupo

de tecnologias a serem implementadas para a preservação de documentos em longo prazo.

Dessa forma, Ferreira (2006) comenta a importância dos repositórios no contexto da preservação digital:

Não obstante, para além da definição de uma política de preservação e do estabelecimento de estratégias de preservação adequadas, é fundamental adoptar um sistema de arquivo digital, i.e. um repositório capaz de albergar os objectos, bem como facilitar a implementação dessas políticas e respectivas estratégias de preservação (FERREIRA, 2006, p. 67).

Os repositórios digitais são, assim, ferramentas de auxílio para a preservação de documentos em longo prazo. Desse modo, é possível implementar parte das políticas institucionais no próprio repositório. A esse respeito, Márdero Arellano (2008) destaca o uso de repositórios que contemplem normas de preservação digital:

Especificamente a preservação digital está sendo observada como o resultado do uso de normas nos processos de arquivamento dos documentos digitais. O uso de repositórios digitais que adotam essas normas está sendo apontado como uma melhoria na qualidade da produção da informação digital a ser preservada pelas organizações (p. 48).

Dentre essas normas, pode-se citar o modelo de referência *Open Archival Information System* (OAIS) apresentado na Figura 2, que é um arquivo digital (repositório) com o objetivo de preservar informações e garantir seu acesso para uma comunidade específica. A documentação armazenada no OAIS é considerada de preservação em longo prazo independentemente da longevidade do próprio modelo. Nesse contexto, considera-se longo prazo como o tempo suficiente para se preocupar com os impactos da evolução das tecnologias (CCSDS, 2002; 2012).

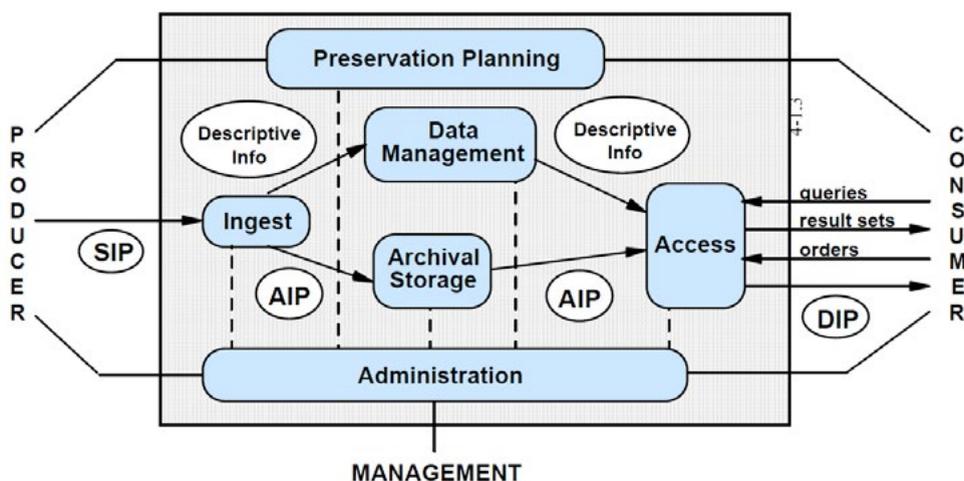


Figura 2 - Modelo OAIS.
Fonte: CCSDS, 2002, p. 4-1.

Nesse modelo conceitual, existem três agentes: produtor (*Producer*), administrador (*Management*) e consumidor (*Consumer*). O produtor faz a submissão de um pacote contendo documentos e informações de descrição (SIP) para a entidade de ingestão (*Ingest*). Após a aceitação do pacote e a inserção da informação de descrição (*Descriptive info*), o SIP transforma-se em um pacote de informação de armazenamento (AIP).

O pacote AIP é então dividido, de modo que os metadados são extraídos e armazenados na entidade gestão de metadados (*Data management*). Logo os objetos digitais são armazenados na entidade do repositório de arquivos (*Archival storage*), onde serão aplicadas as estratégias de preservação digital definidas na entidade do plano de preservação (*Preservation planning*). Após o armazenamento do AIP nas entidades de gestão de metadados e repositório de arquivos, pode-se gerar o pacote de informação de disseminação (DIP). O DIP possui características diferenciadas, pois não está focado diretamente nas atividades de preservação como o SIP e o AIP. O DIP contém formatos de arquivo “mais leves”, próprios para o acesso do público consumidor, formatos estes que são definidos na entidade

do plano de preservação. A entrega dos documentos digitais é realizada pela entidade de acesso (*Access*), que é o meio pelo qual os consumidores realizam consultas (*Queries*). A entidade de acesso retorna resultados para o DIP, que recebe os metadados e é entregue ao consumidor. Todo o processamento realizado no interior do repositório é monitorado pela entidade de administração (*Administration*), a qual poderá gerenciar todo o sistema.

3 Método

Este estudo é fruto de uma pesquisa de natureza aplicada, pois “objetiva gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos” (SILVA & MENEZES, 2005, p. 20). Sua abordagem é qualitativa, pois, conforme Silva e Menezes, a pesquisa qualitativa considera que:

há uma relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números. [...] Não requer o uso de métodos e técnicas estatísticas. O ambiente natural é a fonte direta para coleta de dados e o pesquisador é o instrumento-chave. É descritiva (2005, p. 20).

Os procedimentos técnicos são de ordem bibliográfica, partindo da leitura de materiais já publicados e resultando em uma revisão de literatura. Dentre os materiais coletados na pesquisa bibliográfica, destacam-se: livros, teses, dissertações, diretrizes do Conselho Nacional de Arquivos (Conarq), recomendações do *Consultative Committee for Space Data Systems (CCSDS)* e artigos publicados em periódicos científicos.

4 Apresentação e discussão dos resultados

Ao definir uma política para a preservação de documentos arquivísticos digitais, procede-se à execução das estratégias apropriadas para cada objeto digital. A escolha das estratégias deverá ter como base as propriedades significativas dos documentos que se deseja preservar. Como a formatação, um determinado recurso de interatividade, poderá ser significativa para alguns documentos e para outros não, procede-se a uma avaliação inicial destas propriedades e, após, define-se que ação será tomada, ou seja, qual estratégia será mais indicada.

Um segundo ponto a ser observado, que é fundamental, consiste na qualidade “arquivística” dos documentos digitais. A esse respeito, estudos sobre diplomática contemporânea apresentam determinados requisitos que um documento digital deve manter para ser considerado arquivístico.

Rondinelli (2005; 2013) realça questões como a relação orgânica, as pessoas envolvidas no processo de criação do documento (autor, destinatário, escritor, criador e originador), a forma documental e o contexto tecnológico no qual o documento foi criado. Paralelamente aos conceitos de forma fixa e conteúdo estável, há a variabilidade limitada, a qual permite que o documento arquivístico tenha variação, elemento este que deve ser controlado pelo seu

produtor. Nesse sentido, é fundamental que tais conceitos sejam contemplados durante a implementação de estratégias e repositórios digitais.

4.1 Estratégias de preservação digital: vantagens e desvantagens

Tendo em vista as estratégias pesquisadas, identificou-se uma série de vantagens e desvantagens para cada uma. Da mesma forma, as especificidades de cada formato de arquivo, juntamente com as propriedades significativas que se deseja manter, serão fatores a serem considerados antes da aplicação de qualquer estratégia.

Com o surgimento da preocupação em preservar os documentos digitais,

[...] as práticas relacionadas com a preservação digital estavam baseadas na ideia de garantir a longevidade dos arquivos, mas essa preocupação agora está centralizada na ausência de conhecimento sobre as estratégias de preservação digital e o que isso poderá significar na necessidade de garantir a longevidade dos arquivos digitais (MÁRDERO ARELLANO, 2004, p. 15).

Essa falta de conhecimento vem reforçando a importância de se pesquisar as estratégias de preservação digital, identificando suas vantagens e desvantagens.

Nesse contexto, observa-se que o refrescamento deverá estar sincronizado com outras estratégias, uma vez que sua aplicação isolada oferece apenas preservação física, não contemplando os níveis lógico e conceitual dos objetos digitais. Uma constante verificação da integridade dos suportes físicos e o seu refrescamento periódico são atividades vitais no contexto da preservação digital (FERREIRA, 2006). Desse modo, quando for identificado que o suporte está se tornando obsoleto, será necessário selecionar um novo suporte (INNARELLI, 2007),

já que, quando o documento digital é danificado fisicamente, ele é perdido, fazendo do refrescamento é um pré-requisito para qualquer acervo digital.

A preservação de tecnologia apresenta vantagens como o alto grau de fidedignidade que proporciona aos documentos. Porém, a sua implementação em longos períodos de tempo se tornará inviável, pois tanto as peças de reposição quanto os profissionais qualificados serão raros e os custos de manutenção elevados.

A preservação do *hardware* e do *software* envolve algumas ações específicas, tais como: o armazenamento das cadeias de bits em uma mídia digital estável e a preservação dessa mídia; copiar ou atualizar os dados para uma nova mídia e preservar sua integridade e também, preservar a aplicação original, os programas e as plataformas para processar essas aplicações. Em conclusão, a aplicação desta metodologia tem sentido quando o objeto digital foi criado em formato proprietário e em um *software* obsoleto (MÁRDERO ARELLANO, 2004, p. 71).

Preservar todo o *hardware*, o *software* e os objetos digitais de maneira estável torna-se inviável em longos períodos de tempo. Conforme Grácio e Fadel (2010),

Com os constantes avanços das TIC, sobretudo do *hardware* e do *software* utilizados, esse tipo de preservação, entretanto, apresenta problemas, pois é inevitável que qualquer plataforma tecnológica, mesmo a mais popular, acabe se tornando obsoleta e deixando de existir. Dessa forma não será possível no futuro recuperar a informação digital preservada (p. 73).

De maneira geral, entende-se que a preservação de tecnologia deverá ser utilizada em curtos períodos de tempo, pois sua aplicação em longo prazo poderá impossibilitar a recuperação dos documentos digitais.

As estratégias de emulação, por sua vez, possuem um grau de fidedignidade semelhante ao da preservação de tecnologia. Apresentam, além disso, algumas vantagens, pois a emulação elimina a dependência de *hardware* específico e “possui aplicações úteis quando a aparência do recurso digital original é importante” (MÁRDERO ARELLANO, 2004, p. 21). Para aperfeiçoar a aplicação dessa estratégia, procede-se a uma descrição do ambiente tecnológico (LOPES, 2008).

Entretanto, a emulação é uma estratégia complexa e, conforme Casanovas (2008), faz uso de todo um aparato tecnológico obsoleto que é interpretado pelas tecnologias atuais a fim de continuar acessando os documentos digitais. Isso

Implica configurar un entorno simulado, en general de tecnología anterior, en un entorno de tecnología más actualizado permitiendo que los programas puedan seguir ejecutándose y los archivos puedan seguir siendo accedidos. El problema con esta estrategia es que son muchos los componentes que deben emularse (*hardware*, sistema operativo, aplicaciones y datos) y todos son elementos de un sistema interactuando entre sí; la falta o destrucción de uno de los componentes vuelve inservible al resto (CASANOVAS, 2008, p. 196).

A simulação de muitos componentes digitais torna o processo de emulação extremamente frágil, pois qualquer falha comprometerá o processo como um todo. Além disso, a emulação apresenta desvantagens como a própria obsolescência do emulador. Este aspecto negativo leva a necessidade de emular o emulador ou de reconstruí-lo em uma nova plataforma considerada atual toda vez que o emulador ficar obsoleto.

Considerando as complexidades dessa estratégia, Márdero Arellano (2008) relata que

A emulação deve ser usada nos casos em que os recursos digitais não podem ser migrados dada a sua complexidade, nem convertidos para formatos de *software* independentes. A implementação desta estratégia é complexa e difícil, devido principalmente, ao tipo de necessidades efetivas a serem preenchidas (MÁRDERO ARELLANO, 2008, p. 70).

Tanto as estratégias de emulação quanto as de migração poderão ser auxiliadas pelas estratégias de encapsulamento, podendo ser adotadas de forma complementar a estas. Sendo assim, o encapsulamento irá descrever o contexto dos objetos digitais e reunir todos os componentes digitais necessários para a correta representação do documento.

Esta estratégia permite a manutenção do formato original do objeto digital. Mas, para concretizar a preservação, é preciso que o objeto digital contenha um conjunto de instruções que permitam interpretar o documento, bem como o seu conteúdo (LOPES, 2008). Ou seja, é preciso saber usar os componentes digitais encapsulados a fim de interpretar corretamente o documento digital. Esta abordagem do encapsulamento visa à recuperação futura, guardando os objetos digitais para, então, realizar uma conversão ou emulação no futuro.

A reunião de componentes de um documento digital também poderá proporcionar o acesso imediato, como, por exemplo, com o PDF/A1, que incorpora as fontes utilizadas na escrita do documento. Segundo Thomaz (2006),

Com o PDF/A, as instituições arquivísticas poderão intercambiar seus conteúdos digitais com muita facilidade, pois o formato possui uma série de recursos que facilitam essa tarefa, tais como, suporte interno a metadados em XML, imagens supercomprimidas, imagens pesquisáveis pelo conteúdo, conexões de hipertexto etc. A expectativa é de que o formato PD-

F/A torne-se preferencial em projetos nos quais a permanência dos documentos seja fundamental, substituindo definitivamente os formatos de escritório proprietários (THOMAZ, 2006, p. 127).

O PDF/A é um padrão *International Organization for Standardization* (ISO), possuindo diversas variações, tais como A1 e A2, específicas para determinadas áreas e atividades. Em linhas gerais, o PDF/A1 representa uma alternativa para a preservação de documentos textuais dotados de forma fixa e conteúdo estável.

Partindo para o nível conceitual, cabe ressaltar que a pedra de Rosetta digital acaba por descaracterizar a natureza dos documentos digitais. A impressão em papel ou microfilme, técnica utilizada neste método, limita determinados recursos dos objetos digitais.

Esta é uma técnica questionável, pois descaracteriza a natureza dos documentos digitais, principalmente dos que apresentam conteúdos multimídia. Seu funcionamento pode ser satisfatório para documentos digitais que não sejam dinâmicos e/ou interativos. Porém, mesmo nestes casos, o documento deixará de ser digital, perdendo atributos como, por exemplo, possibilidade de cópia idêntica, busca interna, busca por palavras, manipulação de conteúdo e facilidade de distribuição (ROTHENBERG, 1999). Além disso, a perspectiva de recuperar um documento a partir do código fonte é muito reduzida e depende exclusivamente das tecnologias futuras.

A pedra de Rosetta digital deve ser implementada em situações em que todas as demais estratégias de preservação falharam. Pode-se dizer que ela é uma ferramenta de arqueologia digital e não uma estratégia para preservação de objetos digitais (HEMINGER & ROBERTSON, 2000). Por fim, ressalta-se que a reprodução em suporte analógico ainda é a estratégia de preservação mais

viável em um contexto onde complexos processos de preservação digital são inviáveis (CORRÊA, 2010).

Dentre as estratégias abordadas, a migração é encarada como a melhor prática em termos de estratégias de preservação digital, fato que se justifica tanto pelo seu baixo custo quanto por sua simplificação de procedimento. Esta estratégia “elimina a necessidade de conservar o *hardware* e/ou o *software*, pois o documento original é modificado para ser utilizado em uma nova plataforma diferente daquela em que foi criado” (GRÁCIO & FADEL, 2010, p. 76).

Entretanto, a migração implica mudanças na configuração que afeta o documento por inteiro. Após serem migrados, os documentos parecem ser os mesmos, mas não são. Sua forma física é profundamente alterada, com perda de alguns dados e acréscimo de outros (RONDINELLI, 2005). Além disso, Casanovas (2008) alerta para a necessidade de verificar os resultados das migrações, a fim de identificar a incompatibilidade provocada por uma sequência de migrações. Dessa forma, a migração:

Presenta el inconveniente de ser una actividad que requiere un gran esfuerzo, a lo que hay que agregar la tarea adicional de verificar que la nueva versión funcione correctamente y no se hayan producido alteraciones o incompatibilidad al acumularse migraciones. Todo el proceso descansa en un cumplimiento sistemático de la política de actualización fijada por la institución (CASANOVAS, 2008, p. 196).

Além de provocar alterações na cadeia de *bits*, a migração não pode ser aplicada a documentos de todas as naturezas, como objetos interativos. Entretanto, haverá casos em que as alterações proferidas sobre os objetos digitais não serão prejudiciais às propriedades significativas de preservação estabelecidas.

Ao considerar que o documento arquivístico não pode sofrer alterações, tanto as migrações quanto as respec-

tivas alterações na forma documental devem ser registradas em metadados, auxiliando na presunção de autenticidade do documento (CONARQ, 2014). Desse modo, pode-se garantir a sua autenticidade com o seguinte argumento: o documento migrado confere com o original.

Nos casos em que a migração seja a estratégia escolhida para determinado grupo de documentos arquivísticos digitais, deve-se considerar que “uma tarefa que diminui a sucessiva migração de formatos é a transferência de recursos digitais em formatos preserváveis no momento da submissão em um repositório” (MÁRDERO ARELLANO, 2008, p. 64). A preservação em repositórios possibilita um amplo controle das atividades e deve fazer parte dos procedimentos de preservação digital definidos nas políticas da instituição.

4.2 Repositórios digitais para documentos arquivísticos

Dentre os esforços realizados pela comunidade mundial para a preservação digital, pode-se destacar o modelo de referência OAIS, que é uma norma ISO que estabelece um sistema de preservação para garantir o acesso em longo prazo. Este modelo conceitual pode ser implementado por meio da adoção de repositórios digitais. A importância do uso de normas por parte dos repositórios é retratada por Márdero Arellano (2008)

Especificamente a preservação digital está sendo observada como o resultado do uso de normas nos processos de arquivamento dos documentos digitais. O uso de repositórios digitais que adotam essas normas está sendo apontado como uma melhoria na qualidade da produção da informação digital a ser preservada pelas organizações (p. 49).

Considerando que os estudos sobre preservação digital vêm resultando

em normas, as quais estabelecem metodologias específicas no que tange às atividades de preservação, entende-se que a conformidade com estas normas configura um procedimento mais próximo do ideal. Além disso, estar em conformidade com normas torna os documentos em custódia do repositório mais confiáveis. De tal forma, conforme Corrêa (2010, p. 57),

O responsável pela preservação de um patrimônio será respeitado se demonstrar ter controle e organização. A transparência é a melhor maneira de provar que o repositório é gerenciado de acordo com padrões e práticas consensuais e que o guardião está trabalhando para cumprir sua missão.

Além disso, devem ser adotados mecanismos que garantam a autenticidade dos documentos arquivísticos digitais. A confiabilidade é um ponto fundamental na preservação digital, de forma que “adicionar confiança aos repositórios digitais implica no estabelecimento da presunção de que um dado repositório digital é o que diz ser e que a informação armazenada lá está segura por longo prazo” (SAYÃO, 2010, p. 82). Logo, a confiança está intrinsecamente ligada à presunção de autenticidade através de mecanismos que registram alterações autorizadas ou não, adicionando confiança aos serviços de preservação digital.

O ponto mais pertinente a ser considerado no processo de preservação de documentos arquivísticos digitais é a sua própria qualidade arquivística, a qual pode ser atribuída à conformidade com determinados princípios. Santos (2015) prevê características fundamentais aos documentos de qualidade arquivística, os quais devem ser dotados de fixidez, organicidade, naturalidade, unicidade, autenticidade e imparcialidade.

Ressalta-se, ainda, a necessidade de sistemas informatizados, como, por exemplo, os repositórios digitais, que realizem a manutenção da fixidez, de

modo que os documentos mantenham sua forma e que seu conteúdo não seja alterado de maneira indevida. Da mesma forma, para manter a relação orgânica entre os documentos do acervo, é preciso estabelecer níveis de hierarquia conforme o plano de classificação organizacional. Além disso, não devem ser custodiados documentos acumulados de forma artificial, ou seja, que não são dotados de naturalidade. Observa-se, também, que os documentos devem ser únicos em seus respectivos contextos, dispondo de estruturas de metadados que registrem seu histórico.

Tais aspectos reunidos vêm a corroborar com a sua imparcialidade, registrando tudo o que é necessário, sem acréscimos e sem omissões, o que configura estes registros como documentos arquivísticos digitais. Nesse sentido, observa-se a importância dos metadados na presunção de autenticidade, pois eles criam um histórico de transformações, o qual é considerado um elemento fundamental para comprovar a fidedignidade documental. Ressalta-se que a informação digital possui peculiaridades e, por isso, precisa ser “alterada” de modo que sua preservação seja possível, como, por exemplo, no caso da migração. Dessa forma, o conceito de autenticidade poderá ser definido mediante a análise dos metadados, com o parecer final de que “confere com o original” (SANTOS & FLORES, 2015).

Além dos requisitos arquivísticos, ressalta-se que a infraestrutura também é considerada fundamental, pois apoia a confiabilidade e a sustentabilidade de repositórios digitais. Assim, as organizações e suas comunidades-alvo podem confiar no fato de que os documentos digitais serão preservados em longo prazo (THOMAZ, 2007). Conforme Thomaz (2007), organização, tecnologia e gerenciamento são bases fundamentais para atingir a confiança em repositórios digitais (Figura 3).



Figura 3 - Três pilares da certificação de repositórios digitais.
Fonte: Thomaz (2007, p. 85).

O pilar organizacional corresponde ao comprometimento e aos esforços da organização quanto a suas atividades de preservação digital. Já o pilar tecnológico corresponde à infraestrutura técnica – *hardware*, *software*, suportes, segurança de redes etc. O pilar gerencial, por sua vez, corresponde ao gerenciamento dos objetos digitais em conformidade com as entidades do modelo OAIS (THOMAZ, 2007). Através do desenvolvimento des-

ses três aspectos, é possível adicionar confiabilidade ao repositório digital.

Flores (2014) esquematiza a importância do uso de repositórios digitais confiáveis durante as três idades documentais: corrente, intermediária e permanente (Figura 4). O autor explicita, ainda, que o repositório digital confiável é tido como opcional nas duas primeiras idades documentais, tornando-se fundamental na terceira idade.

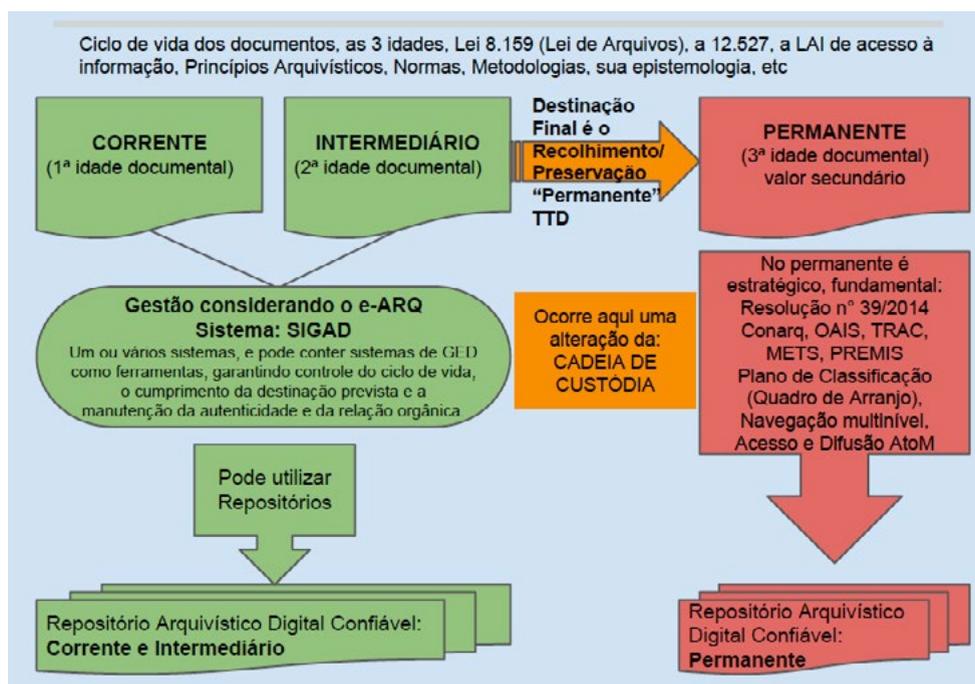


Figura 4 - Ciclo de vida dos documentos.
Fonte: Flores (2014).

O ciclo de vida dos documentos permite compreender a complexidade dos documentos arquivísticos digitais, pois evidencia a necessidade de usar dois sistemas distintos, um para a primeira e segunda idade e outro para a terceira idade documental. Da mesma forma, o Conarq (2014) também considera necessário, para a preservação dos documentos arquivísticos digitais, o uso de repositórios digitais confiáveis nas fases corrente, intermediária e permanente.

Além disso, no que diz respeito à idade permanente, percebe-se a existência da norma OAIS, das recomendações do *Trustworthy Repository Audit & Certification: Criteria and Checklist* (TRAC) e de padrões de metadados como *Metadata Encoding and Transmission Standard* (METS) e *Preservation Metadata: Implementation Strategies* (PREMIS). Outras recomendações, além do TRAC, podem ser adotadas na idade permanente: *Audit And Certification of Trustworthy Digital Repositories* (ACTDR), *Catalogue of Criteria for Trusted Digital Repositories da Network of Expertise in long-term STORage* (NESTOR) e *Digital Repository Audit Method Based on Risk Assessment* (DRAMBORA).

5 Considerações finais

A preservação de documentos arquivísticos digitais concentra-se na definição de políticas institucionais para posterior adoção de estratégias e repositórios. Logicamente, os objetivos devem ser definidos em um primeiro momento para, posteriormente, testar os métodos para sua comprovação e adoção em larga escala.

Com relação às estratégias de preservação digital, é fundamental que as propriedades significativas dos documentos digitais sejam identificadas inicialmente a fim de serem preservadas. Da mesma forma, é preciso manter os princípios arquivísticos como, por exemplo, fixidez e organicidade, implementando, quando necessário,

softwares e padrões de metadados que corroborem para a manutenção das qualidades arquivísticas. Sendo assim, as atividades de preservação digital devem ser orientadas com vistas à preservação destes princípios, adaptando os *softwares* conforme suas necessidades.

Posteriormente, escolhem-se as estratégias que melhor preservam os respectivos níveis escolhidos para cada um dos objetos digitais. Considerando que todas as estratégias apresentam vantagens e desvantagens, procede-se à adoção de mais de uma no acervo. Além disso, um mesmo documento pode ser preservado por meio de duas ou mais estratégias.

Entretanto, a ausência de uma ferramenta que verifique a autenticidade dos documentos irá ocasionar um estado de vulnerabilidade durante a execução de estratégias de preservação digital. Para serem considerados autênticos e confiáveis, os documentos arquivísticos digitais precisam estar inseridos em um sistema que proporcione estas qualidades. O sistema do qual se está falando é o repositório digital, que se torna confiável na medida em que contempla normas de preservação digital e requisitos arquivísticos.

Nesse sentido, é fundamental que um repositório digital esteja em conformidade com o modelo OAIS, sendo necessário fazer auditorias para verificar as responsabilidades institucionais quanto à preservação e ao acesso aos documentos. Qualquer instituição que deseje preservar documentos arquivísticos autênticos deverá verificar continuamente os estudos sobre preservação, pois muitos conceitos, normas e tecnologias estão sendo desenvolvidos.

Considerando que a grande maioria dos documentos arquivísticos digitais é de natureza estática, estratégias de migração tornam-se as mais usuais. Da mesma forma, o repositório digital proporcionará mais praticidade e garantia de autenticidade durante sua implementação.

Com o advento dos documentos digitais, colocou-se em questão a sua

presunção de autenticidade, fazendo dos repositórios digitais ambientes autênticos para a preservação de documentos arquivísticos. Desse modo, pode-se dizer que um acervo digital sem políticas, estratégias e repositórios não garante integridade nem autenticidade; logo, não é confiável. Por fim, observa-se a necessidade de estudos interdisciplinares que contemplem Arquivologia, Ciência da Informação, Diplomática e outras áreas que sejam necessárias na fundamentação das políticas de preservação digital.

Em linhas gerais, este artigo contribui para a iniciação de pesquisadores em preservação digital, fornecendo uma base teórica que perpassa as principais estratégias de preservação digital, os repositórios digitais e o modelo OAIS. Além disso, a ênfase depositada sobre o fator arquivístico tende a contribuir no preenchimento das lacunas teóricas da área.

Referências

1. CASANOVAS, I. **Gestión de documentos electrónicos**. Buenos Aires: Alfagrama, 2008.
2. CONSELHO NACIONAL DE ARQUIVOS – CONARQ (Brasil). Câmara Técnica de documentos eletrônicos. **Carta para a Preservação do Patrimônio Arquivístico Digital**. Rio de Janeiro: Arquivo Nacional, 2004. Disponível em: <<http://www.conarq.arquivonacional.gov.br/Media/publicacoes/cartapreservpatrimarqdigitalconarq2004.pdf>>. Acesso em: 10 ago. 2014.
3. _____. **e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos**. Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/earq/conarq_earqbrasil_model_requisitos_2009.pdf>. Acesso em: 05 ago. 2014.
4. _____. **Diretrizes para a presunção de autenticidade de documentos arquivísticos digitais**. Rio de Janeiro: Arquivo Nacional, 2012. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/diretrizes_presuncao_autenticidade_publicada.pdf>. Acesso em: 20 jun. 2014.
5. _____. **Diretrizes para a implementação de repositórios digitais confiáveis de documentos arquivísticos**. Rio de Janeiro: Arquivo Nacional, 2014. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/resol_conarq_39_repositorios.pdf>. Acesso em: 13 ago. 2014.
6. CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEM (CCSDS). **Reference Model for an Open Archival Information System (OAIS)**. Blue Book. Washington, Jan. 2002. Disponível em: <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>. Acesso em: 28 mai. 2014.
7. _____. Magenta Book. Washington, Jun. 2012. Disponível em: <<http://public.ccsds.org/publications/archive/650x0m2.pdf>>. Acesso em: 13 mai. 2014.
8. CORRÊA, A. M. G. **Preservação digital: autenticidade e integridade de documentos em bibliotecas digitais de teses e dissertações**. Dissertação (Mestrado em Ciência da Informação) 96 p, Universidade de São Paulo: São Paulo, 2010. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27151/tde-05112010-105831/pt-br.php>>. Acesso em: 03 jul. 2014.
9. FERREIRA, M. **Introdução à preservação digital: conceitos, estratégias e atuais consensos**, Portugal: Escola de Engenharia da Universidade do Minho, 2006. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 02 ago. 2014.
10. FLORES, D. Inovação em Arquivos: a formação e a pesquisa em arquivologia diante da complexidade e especificidade do documento arquivístico digital, da gênese à preservação. Portugal: Évora, 2014. In: Encontro Internacional de Arquivos (EIA). **Anais eletrônicos...** Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/earq/conarq_earqbrasil_model_requisitos_2009.pdf>. Acesso em: 05 ago. 2014.

- vel em: <http://eiarquivos2014.weebly.com/uploads/2/6/0/6/26068081/p1_3_df.pdf>. Acesso em: 05 nov. 2014.
11. GRÁCIO, J. C. A.; FADEL, B. **Estratégias de preservação digital**. In: VALENTIM, Marta. (org.) Gestão, mediação e uso da informação. São Paulo: UNESP: Cultura Acadêmica, 2010. 390 p. ISBN 978-85-7983-117-1. Available from SciELO Books. Disponível em: <<http://books.scielo.org>>. Acesso em: 01 dez. 2014.
12. HEMINGER, A. R.; ROBERTSON, S. B. The Digital Rosetta Stone: a model for maintaining long-term access to static digital documents. **Communications of AIS**. Vol. 3, Article 2, January 2000. Disponível em: <<http://delivery.acm.org/>>. Acesso em: 24 set. 2014.
13. INNARELLI, H. C. **Preservação de documentos digitais**: confiabilidade de mídias CD-ROM e CD-R. 174p. 2006. Dissertação (Mestrado) – Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2006. Disponível em: <<http://www.biblioteca-digital.unicamp.br/document/?code=-vtls000384479>>. Acesso em: 06 jun. 2014.
14. _____. Preservação digital e seus dez mandamentos. In: SANTOS, Vanderlei Batista (Org.). **Arquivística**: temas contemporâneos, classificação, preservação digital, gestão do conhecimento. Distrito Federal: SENAC, 2007. p. 21-75.
15. INTERPARES 2 PROJECT. **Diretrizes do Preservador**. A preservação de documentos arquivísticos digitais: diretrizes para organizações. TEAM Brasil. Tradução: Arquivo Nacional e Câmara dos Deputados. 2002 – 2007. Disponível em: <http://www.interpares.org/display_file.cfm?doc=ip2_preserver_guidelines_booklet--portuguese.pdf>. Acesso em: 09 ago. 2014.
16. LOPES, V. **Preservação Digital**. Portugal: Universidade do Minho, Guimarães, 2008. Disponível em: <http://www.vitorlopes.com/Trabalhos/Preservacao_Digital-Vitor_Lopes.pdf>. Acesso em: 28 ago. 2012.
17. MÁRDERO ARELLANO, M. Á. **Critérios para a preservação digital da informação científica**. 354 p. Tese (Doutorado em Ciência da Informação). Universidade Federal de Brasília, Departamento de Ciência da Informação, 2008. Disponível em: <http://btd.bce.unb.br/tesesimplificado/tde_busca/arquivo.php?codArquivo=4547>. Acesso em: 15 jun. 2014.
18. _____. Preservação de documentos digitais, **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/305/1452>>. Acesso em: 25 jul. 2014.
19. PAES, M. L. **Arquivo**: teoria e prática. 3. Ed. Rev. Ampl. Rio de Janeiro: FGV, 2004.
20. RONDINELLI, R. C. **Gerenciamento arquivístico de documentos eletrônicos**: uma abordagem teórica da diplomática arquivística contemporânea. 4. Ed. Rio de Janeiro: FGV, 2005.
21. _____. **O documento arquivístico ante a realidade digital**: uma revisão conceitual necessária. Rio de Janeiro: FGV, 2013.
22. ROTHENBERG, J. Avoiding technological quicksand: finding a viable technical foundation for digital preservation. **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, 1999. Disponível em: <<http://www.clir.org/pubs/reports/rothenberg/pub77.pdf>>. Acesso em: 24 set. 2014.
23. SANTOS, H. M. dos; FLORES, D. Políticas de preservação digital para documentos arquivísticos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.20, n.4, p.197-217, out./dez. 2015. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2542>>. Acesso em: 21 dez. 2015.

24. SANTOS, V. B. dos. **A Arquivística como disciplina científica**: princípios, objetivos e objetos. Salvador: 9Bravos, 2015.
25. SARAMAGO, M. de L. Metadados para preservação digital e aplicação do modelo OAIS. In: CONGRESSO NACIONAL DE BIBLIOTECARIOS, ARQUIVISTAS E DOCUMENTALISTAS, nº 08. 2004, Estoril. **Anais eletrônicos...** Disponível em: <<http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640>>. Acesso em: 04 jul. 2014.
26. SAYÃO, L. F. Repositórios Digitais Confiáveis para a Preservação de Periódicos Eletrônicos Científicos. **Ponto de Acesso**, Salvador, v.4, n.3, p. 68-94, dez. 2010. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/4709/3565>>. Acesso em: 08 ago. 2014.
27. SILVA, E. L. da; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4. Ed. Rev. atual. Florianópolis: UFSC, 2005. Disponível em: <https://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf>. Acesso em: 13 jun. 2014.
28. SOUSA, R. T. B. de. **A classificação como função matricial do que-fazer arquivístico**. In: SANTOS, Vanderlei Batista dos. Arquivística: temas contemporâneos, classificação, preservação digital, gestão do conhecimento. 2007 SENAC, p. 79-163.
29. THIBODEAU, K. **Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years**, presented at The State of Digital Preservation: An International Perspective, Washington D.C., 2002. Disponível em: <<http://www.clir.org/pubs/reports/pub107/thibodeau.html>>. Acesso em: 19 ago. 2014.
30. THOMAZ, K. P. Gestão e preservação de documentos eletrônicos de arquivo: revisão de literatura – parte 2. **Arquivística.net**, Rio de Janeiro, v.2, n.1, p.114-131, jan./jun. 2006. Disponível em: <www.brapci.ufpr.br/download.php?dd0=6733>. Acesso em: 07 set. 2014.
31. _____. Repositórios digitais confiáveis e certificação. **Arquivística.net**, Rio de Janeiro, v.3, n.1, p. 80-89, jan./jun. 2007. Disponível em: <http://www.brapci.inf.br/_repositorio/2010/05/pdf_fed0720dbb_0010726.pdf>. Acesso em: 07 set. 2014.