

Preservação de documentos digitais

Miguel Angel Arellano

Doutorando em ciência da informação.

E-mail: miguel@ibict.br

O conteúdo, a estrutura e o contexto da informação digital devem ser associados às funcionalidades do software que preserva suas conexões executáveis ou representações de suas relações, permitindo a sua reconstrução.

(David Bearman, 1994)

Resumo

O presente artigo tem como objetivo apresentar o resultado de uma pesquisa bibliográfica sobre a preservação digital. São analisadas as questões mais relevantes no estudo das práticas atuais relacionadas com a preservação dos documentos em formato digital. Dados foram coletados na literatura produzida por especialistas experientes na área, os quais proporcionam ampla lista de argumentos que mostram a emergência e importância das atividades de preservação relacionadas aos objetos digitais.

Palavras-chave

Preservação digital; Metadados; Repositórios digitais; Modelo de preservação OAIS; Biblioteca digital.

Preservation of digital documents

Abstract

The objective of this article is to present the results of a bibliographic research about digital preservation. The most relevant questions concerning the study of the current practices related to the preservation of documents on digital format are investigated. Data was collected on the literature produced by specialists, who put forth a wide list of arguments showing the emergency and importance of activities about preservation related to the digital objects.

Keywords

Digital Preservation; Metadata; Digital repositories; OAIS preservation model; Digital library.

INTRODUÇÃO

A natureza dos documentos digitais está permitindo ampla produção e disseminação de informação no mundo atual. É fato que na era da informação digital se está dando muita ênfase à geração e/ou aquisição de material digital, em vez de manter a preservação e o acesso a longo prazo aos acervos eletrônicos existentes. O suporte físico da informação, o papel e a superfície metálica magnetizada se desintegram ou podem se tornar irre recuperáveis. Existem, ademais, os efeitos da temperatura, umidade, nível de poluição do ar e das ameaças biológicas; os danos provocados pelo uso indevido e o uso regular, as catástrofes naturais e a obsolescência tecnológica. A aplicação de estratégias de preservação para documentos digitais é uma prioridade, pois sem elas não existiria nenhuma garantia de acesso, confiabilidade e integridade dos documentos a longo prazo.

É preciso chamar a atenção para a importância de informar o contexto do objeto digital a ser registrado (e preservado) para que, dessa maneira, futuros usuários possam entender o ambiente tecnológico no qual ele foi criado. A preservação dos documentos continua a ser determinada pela capacidade de o objeto informacional servir às utilizações que lhe são imputadas, às suas atribuições que garantem que ele continue a ser satisfatório às utilizações posteriores. Mas, no caso específico dos documentos em formato digital, a preservação dependerá principalmente da solução tecnológica adotada e dos custos que ela envolve.

CONTEXTO

A preservação é um dos grandes desafios do século XXI. Durante os últimos anos do século XX, apenas as bibliotecas, os arquivos e os centros e institutos de pesquisa e organismos governamentais criavam conteúdo digital relevante. Segundo Cunha (1999), nos últimos anos duas das funções básicas das bibliotecas estão sofrendo “perigo de extinção: a provisão de acesso à informação e a preservação do conhecimento para

futuras gerações”. Atualmente muitas coleções digitais importantes estão sendo construídas fora das bibliotecas por diferentes organizações, ou sendo publicadas diretamente na Internet. Com o aumento da produção de informação em formato digital, tem sido questionada cada vez mais a importância de se ter garantida a sua disponibilização e preservação por longos períodos de tempo. Essa preocupação envolve tanto os produtores dos dados quanto os órgãos detentores dessa informação. No início, as práticas relacionadas com a preservação digital estavam baseadas na idéia 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.

Em nível mundial, as iniciativas se multiplicam (NLNZ, 2003), e novas soluções são testadas por instituições detentoras de acervos de especial relevância para o desenvolvimento científico e tecnológico de seus países de origem. Muitos projetos e iniciativas têm conseguido estabelecer os benefícios de algumas metodologias e estratégias. A conclusão a que essas iniciativas chegam é que devem ser usados padrões e converterem-se os documentos nos formatos livres, para que eles sejam acessados após a obsolescência dos equipamentos e programas informáticos em que foram criados. Usar padrões abertos permite seu estudo e sua conversão para novos padrões.

Os objetos digitais* não podem ser deixados em formatos obsoletos para serem transferidos depois de longos períodos de negligência para repositórios digitais. O desafio é muito mais um problema social e institucional do que um problema técnico, porque, principalmente para a preservação digital, depende-se de instituições que passam por mudanças de direção, missão, administração e fontes de financiamento. Muitos materiais publicados digitalmente são produto de serviços de informação disponibilizados por organizações que adotam alguma infra-estrutura tecnológica. Essas instituições levam em consideração aspectos legais e culturais que afetam a oferta desses serviços orientados a atender as necessidades de determinados usuários. Mas muitas vezes essa preocupação não é a mesma que têm os produtores das novas tecnologias.

Para os detentores de acervos digitais, é cada vez mais imperiosa a necessidade de contar com mecanismos que garantam a preservação de seus documentos em formato digital. Especificamente essa preocupação parte das comunidades responsáveis pelas bibliotecas e pelos arquivos, para as quais o desenvolvimento de padrões e de mecanismo legais para lidar com arquivos eletrônicos precisa de estratégias metodológicas bem definidas.

Para Sant’Anna (2001), é responsabilidade dos arquivos adotar medidas preventivas e corretivas objetivando minimizar a ação do tempo sobre o suporte físico da informação, assegurando sua disponibilidade. A perspectiva arquivística da preservação parte da compreensão dos limites e significados dos documentos (autenticidade, capacidade probatória, integridade das informações, contexto de produção, manutenção etc.), dando ênfase às tarefas que as organizações e instituições arquivísticas que criam e são responsáveis pela guarda permanente desses documentos devem observar para lidar com objetos digitais autênticos. Os arquivos administrativos requerem o desenvolvimento de parâmetros de certificação da qualidade dos seus processos e serviços. O chamado “arquivamento” relacionado com o gerenciamento do armazenamento de registros passou a estar unido ao termo mais comum “arquivo”, que atualmente está mais identificado com formato eletrônico de documentos (*e-print*, *pré-print*, *pós-print* etc.).

Na área da ciência da informação, o uso da tecnologia digital que toma o lugar dos tradicionais meios de preservação, como a microfilmagem, trouxe consigo a preocupação com as normas para o uso das técnicas digitais e sua prontidão na tarefa da preservação a longo prazo (Chepesuik, 1997). Os especialistas da área que trabalham com informação em formatos digitais estão elaborando normas necessárias para armazenar e compartilhar de maneira adequada esses materiais, assim como buscam a formulação de políticas institucionais de preservação. Segundo Webb (2000), as bibliotecas são responsáveis por manter coleções para uso permanente, protegendo-as de ameaças, ou salvando-as e reparando-as para compensar seus impactos. Os sistemas de bibliotecas digitais atuais envolvem um grande esforço de gerenciamento de coleções digitais que vai além das tarefas tradicionais das bibliotecas (aquisição, seleção, classificação, arquivamento etc.), na construção da interoperabilidade de acervos digitais (arquiteturas, metadados, formatos padrão), que é possível por meio de sistemas relacionados e desenvolvidos para propósitos e comunidades específicas (Arms, 2000).

* Um **objeto digital** é aquele que foi criado em computador, podendo ser original ou uma versão depois de haver sido convertido (ou digitalizado).

Para alguns especialistas, as bibliotecas digitais são consideradas o caminho mais adequado para a preservação dos recursos de informação (Hildreth, 1996). As bibliotecas digitais são meios mais dinâmicos para preservação digital do que as bibliotecas tradicionais, no sentido da sua adaptação às freqüentes mudanças tecnológicas (Lesk, 1997). Para outros autores, os centros de preservação estão sendo considerados os lugares adequados para se testar e formular as metodologias e políticas a serem adotadas pelos provedores de informação científica. Os centros estariam localizados em instituições confiáveis e capazes de armazenar, migrar e dar acesso a coleções digitais (RLG, 2002). O treinamento de especialistas e técnicos é também contemplado nos custos das atividades desses centros, produzindo melhora na aplicação efetiva das estratégias (Chapman, 2003).

No quadro 1, está representado o lugar das estratégias de preservação no fluxo dos documentos digitais dentro da estrutura da biblioteca digital (Borbinha & Correia, 2001).

Na preservação de documentos digitais, assim com na dos documentos em papel, é necessária a adoção de ferramentas que protejam e garantam a sua manutenção. Essas ferramentas deverão servir para reparar e restaurar registros protegidos, prevenindo os danos e reduzindo os riscos dos efeitos naturais (preservação prospectiva), ou para restaurar os documentos já danificados (preservação retrospectiva). Para Margaret Hedstrom (1996), a preservação digital é "(...) o planejamento, alocação de recursos e aplicação de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável". A preservação digital compreende os mecanismos que permitem o armazenamento em repositórios de dados digitais que garantiriam a perenidade dos seus conteúdos.

As condições básicas à preservação digital seriam, então, a adoção desses métodos e tecnologias que integrariam a preservação física, lógica e intelectual dos objetos digitais. A preservação física está centrada nos conteúdos armazenados em mídia magnética (fitas cassete de áudio e de rolo, fitas VHS e DAT etc.) e discos óticos (CD-ROMs, WORM, discos óticos regraváveis). A preservação lógica procura na tecnologia formatos atualizados para inserção dos dados (correio eletrônico, material de áudio e audiovisual, material em rede etc.), novos *software* e *hardware* que mantenham vigentes seus *bits*, para conservar sua capacidade de leitura.

QUADRO 1

As estratégias de preservação no fluxo dos objetos digitais na biblioteca digital

| | | |
|-------------|---------------------------|--------------------------|
| Aquisição | Entrega pelo produtor | |
| | Captura pela biblioteca | |
| | Coletado pela biblioteca | |
| Verificação | Integridade física (meio) | |
| | Integridade do conteúdo | Integridade lógica |
| | | Autenticação |
| Registro | Metadados | Descrição bibliográfica |
| | | Instalação e manipulação |
| | | Acesso |
| | | Preservação |
| Preservação | Preservação física | Refrescamento do meio |
| | | Migração de suporte |
| | Preservação lógica | Conversão de formatos |
| | | Emulação |
| | Preservação intelectual | |
| Acesso | Condições de uso | Acesso local |
| | | Acesso remoto |

QUADRO 2

Relevância dos requisitos de preservação dos documentos impressos e digitais

| Requisitos | Documentos impressos | Documentos digitais |
|-------------------------|----------------------|---------------------|
| Preservação física | Relevante | Relevante |
| Preservação lógica | Pouco relevante | Relevante |
| Preservação intelectual | Não relevante | Relevante |

No caso da preservação intelectual, o foco são os mecanismos que garantem a integridade e autenticidade da informação nos documentos eletrônicos. Segundo Borbinha & Correia (2001), a preservação digital representa um novo agrupamento da perspectiva que se tinha dos requisitos associados com as atividades tradicionais nessa área. Como mostra o quadro 2, a preservação física continua relevante na mídia eletrônica, ainda que o seu armazenamento tenha mostrado a necessidade de constante atividade de migração para novos materiais (disquete, fita VHS, CD-ROM, DVD etc.).

No caso dos materiais impressos, a preservação lógica é pouco relevante, por estar garantida no formato específico em que foram publicados (periódico, revista, livro etc.). Na publicação digital, a preservação lógica está associada à necessidade de garantir a conversão dos formatos originais que tem se convertido em obsoletos ou de custosa manutenção. O quadro mostra também que a importância da preservação intelectual torna-se

maior no caso dos materiais digitais devido principalmente à capacidade de o objeto digital ser passível de modificação no seu desenho (leiaute), apresentação ou interação no formato de publicação. Com isso, a perda do conteúdo intelectual original pode ser declarada inaceitável pelo autor.

Esse último requisito envolve a preservação da propriedade intelectual Intellectual Property Rights (IPR), que tem um significado mais argüível do que na mídia tradicional. O IPR deve considerar não apenas o conteúdo, mas também qualquer ação relacionada ao *software* (cópias, encapsulamento de conteúdo, emulação de *software*, migração de conteúdo) que envolva atividades que podem infringir permissões específicas daqueles que mantêm os direitos. Para Bullock (1999), o IPR é uma das principais barreiras que interferem na preservação dos objetos digitais. Bullock também enumera os requisitos de preservação dos documentos digitais da seguinte forma:

- fixar os limites do objeto a ser preservado;
- preservar a presença física;
- preservar o conteúdo;
- preservar a apresentação;
- preservar a funcionalidade;
- preservar a autenticidade;
- localizar e rastrear o objeto digital;
- preservar a proveniência;
- preservar o contexto.

Esses novos requisitos especificam os elementos que deverão ser efetivamente mantidos, a série de *bits* que deverá ser recuperada; referem-se à capacidade de acesso aos conteúdos dinâmicos independentemente da sua apresentação, componentes de multimídia, hipertextualidade e interatividade. Observar essas condições significa identificar o objeto digital na sua origem e pelas suas dependências de *hardware* e *software*.

ESTRATÉGIAS

Pesquisadores em várias partes do mundo estão desenvolvendo modelos do que seria a infra-estrutura para a preservação a longo prazo de informação em formato digital. Muitas das iniciativas propostas pelos grandes centros de pesquisa continuam sendo utópicas, mas proporcionam uma base para a discussão sobre a implantação de soluções futuras ao problema (Granger, 2002). Para Beagrie & Greenstein (1998), algumas precauções devem ser tomadas para reduzir o perigo da perda dos materiais digitais:

- armazenar em ambiente estável e controlável;
- implementar ciclos de atualização (*refreshment*) para cópia em nova mídia;
- fazer cópias de preservação (assumindo licenças e permissões de *copyrights*);
- implementar procedimentos apropriados de manuseio;
- transferir para uma mídia de armazenamento padrão.

Segundo vários autores, o dado e a mídia que suportam a informação devem possuir um nível de funcionalidade representacional que permita a sua reprodução a qualquer momento que a instituição mantenedora precisar recuperar o dado. Cada mídia pode armazenar uma seqüência de *bits* de forma diferente, segundo as propriedades físicas da mídia. O *bit stream** precisa então ser interpretado, já que toda seqüência significativa de *bits* pode representar qualquer coisa. Os objetos digitais são salvos como coleções de *bits* representando documentos específicos, significativos apenas para o programa que os criou.

Os principais métodos recomendados para a preservação dos objetos digitais podem ser agrupados em dois tipos: os estruturais e os operacionais. Os estruturais tratam dos investimentos iniciais por parte das instituições que estão se preparando para implementar algum processo de preservação e que adotam ou adaptam um dos modelos de metadados existentes ou seu próprio esquema. As atividades operacionais são as medidas concretas aplicadas aos objetos digitais. No quadro 3, estão enumerados os métodos de preservação mais freqüentemente usados.

QUADRO 3
Métodos de Preservação Digital

| Métodos Estruturais | Métodos Operacionais |
|----------------------------------|---|
| Adoção de padrões | Conservação de <i>software</i> / <i>hardware</i> |
| Elaboração de normas | Migração de suporte |
| Metadados de preservação digital | Conversão de formatos |
| Montagem de infra-estrutura | Emulação |
| Formação de consórcios | Preservação do conteúdo |

As estratégias operacionais que englobariam os novos requisitos de preservação seriam a migração de suporte e o refrescamento do meio (preservação física), a conversão dos formatos, a emulação (preservação lógica) e a preservação do conteúdo (intelectual).

* **Bitstream:** a seqüência de zeros e de uns que compreende o dado.

OS METADADOS DE PRESERVAÇÃO

A criação e uso de metadados é uma parte importante em todas as estratégias operacionais de preservação digital, uma vez que elas estão baseadas na conservação de *software* e *hardware*, emulação ou migração, como um meio para garantir a autenticidade, registrar o gerenciamento de direitos e coleções de dados, e para a interação com recursos de busca (Rothenberg, 1996). Os metadados informam as partes importantes do objeto digital e indicam a sua localização. Os metadados de preservação são uma forma especializada de administrar metadados que podem ser usados como um meio de estocar a informação técnica que suporta a preservação dos objetos digitais. Os metadados para preservação visam a apoiar e facilitar a retenção a longo prazo da informação digital (OCLC/RLG, 2001).

Segundo Searle & Thompson (2003), a implementação de um esquema de metadados de preservação requer no mínimo a limitação do escopo dos metadados para preservação, a fim de maximizar o potencial para sua automação e garantir o controle da mudança nos metadados. Na parte operacional do desenvolvimento do esquema de metadados para preservação, algumas atividades são necessárias:

- a implementação do padrão de metadados;
- criação de repositórios de metadados de preservação integrados a outros repositórios de metadados usados pela instituição;
- definição do *script* para extrair os metadados de preservação, que produza um relatório em XML* do que os metadados identificaram como importante para a preservação, para depois serem transferidos para o repositório de metadados.

Segundo a National Library of New Zealand (2003), os metadados de preservação contêm informações sobre:

- políticas e técnicas de preservação adotadas;
- efeitos da estratégia adotada;
- gerenciamento de coleções;
- gestão de direitos autorais;
- autenticidade do recurso digital.

* **Extensible Markup Language**. XML é um formato padrão para publicação, armazenamento e transferência de documentos por via eletrônica, independente dos fabricantes de *software*, facilmente interpretável por múltiplas aplicações, em certa forma, auto-explicativo.

Várias instituições internacionais têm construído marcos conceituais para identificar elementos específicos de metadados de preservação. A National Library of New Zealand produziu um esquema de metadados para preservação orientado para ser um ponto de equilíbrio entre os princípios expressos no modelo de referência OAIS (*Open Archival Information System*) para preservação de metadados e a praticidade de implementação de um conjunto de metadados operacionais de preservação. Ele está dividido em quatro entidades: o objeto, o processo, o arquivo e a modificação de metadados. No artigo de Searle & Thompson (2003), o esquema da NLNZ é apresentado identificando-se as áreas nas quais ele pode ser desenvolvido: para a expansão dos dados de definição dos elementos do esquema; para criação de um repositório baseado nesses dados de definição; para o desenvolvimento de ferramentas de extração automática de metadados para alimentar o repositório. O esquema identifica os dados que a biblioteca coletará e manterá. O arquivo original ou *preservation master* deve ser preservado, gerenciado e disseminado através do tempo (NLNZ, 2003).

O MODELO DE PRESERVAÇÃO OAIS

O modelo para repositórios de metadados de preservação mais usado atualmente é o modelo de referência *Open Archival Information System* (OAIS), publicado pelo Consultive Committee for Space Data Systems (CCSDS). OAIS é uma iniciativa ISO (International Organization for Standardization), desde junho 2003, que define um alto nível de modelo de referência para arquivos que precisem de uma preservação de longo prazo. Originalmente era aplicado apenas para informações digitais obtidas de observações de ambientes espaciais e da Terra, mas, na verdade, ele é aplicável a qualquer tipo de arquivo digital e compatível com os sistemas baseados no protocolo da Open Archive Initiative (Day, 2001). Um arquivo nos termos do OAIS significa considerar os arquivos como organizações de pessoas e sistemas, que aceitaram a responsabilidade de preservar a informação e torná-la disponível e melhorar a comunicação e produtividade entre diferentes comunidades (CCSDS, 1998).

No modelo atuam quatro entidades: produtores, consumidores, administração e o arquivo propriamente dito. No esquema do OAIS, está incluído um modelo de informação para inserção dos metadados de preservação (figura 1), a seguir.

No modelo, objeto físico e objeto digital podem ser, coletivamente, referenciados como objetos de dados. Eles são interpretados pela combinação da base de conhecimento da comunidade alvo e a informação de

representação associada ao objeto de dados. O objeto de dados pode ser complementado com informação de representação, para ser entendido pela comunidade alvo. No modelo OAIS, existem quatro categorias de objetos de informação, entre os quais estão:

- informação de conteúdo: uma informação que é o objeto primário da preservação. Ela contém o objeto digital primário e informação representada necessária para transformar este objeto em informação com significado.
- informação de descrição de preservação: qualquer informação necessária para preservar adequadamente a informação de conteúdo com a qual está associada. Ela inclui:
 - informação de referência: e.g. identificadores;
 - informação de contexto: e.g. classificações por assunto;
 - informação de proveniência: e.g. *copyright*, histórico;
 - informação de integridade: documenta os mecanismos de autenticação.
- informação de pacote: informação que enlaça e relaciona os componentes de um pacote dentro de uma entidade identificável em uma mídia específica.
- informação descritiva: informação que permite localizar, analisar, recuperar ou ordenar a informação de pacote dentro de um OAIS.

No esquema conceitual do OAIS existem seis entidades funcionais: recepção, armazenamento, gerenciamento de dados, administração do sistema, planejamento de preservação e acesso (figura 2).

As entidades funcionais gerenciam o fluxo de informação entre as entidades que formam o ambiente OAIS e identificam os componentes funcionais dos arquivos relacionados com a preservação dos objetos digitais.

A estrutura conceitual para metadados de preservação do OAIS está sendo usada por várias instituições para identificar seus elementos de metadados específicos: Cornell University, as Bibliotecas Nacionais da Austrália e da Nova Zelândia, On-line Computer Library Center (OCLC), Michigan Institute of Technology (MIT) e outras. Por ser um modelo de referência, ele não é uma implementação específica, mas proporciona uma lista de condições

FIGURA 1
O modelo de referência *Open Archival Information System* (OAIS) ou Sistema Aberto para Arquivamento de Informação (SAAI)

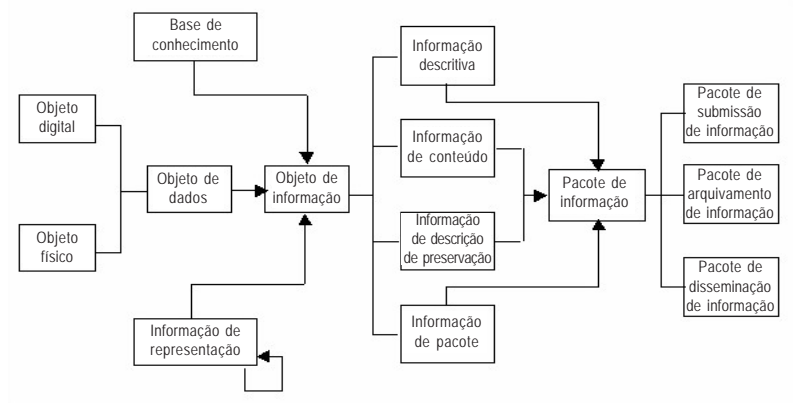
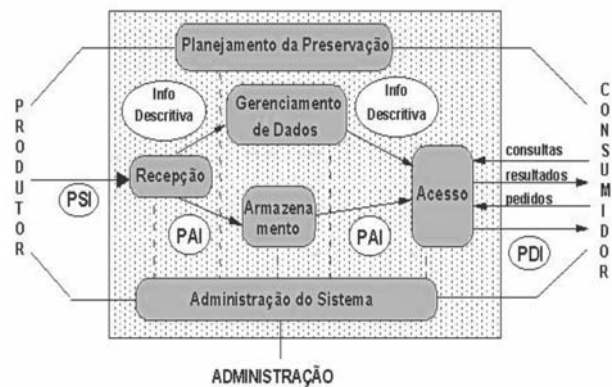


FIGURA 2
O Modelo Funcional OAIS



Fonte: LAVOIE, Brian. Meeting the challenges of digital preservation: the OAIS reference model. OCLC-Newsletter. n. 243. jan./fev. 2000, p. 26-30. (apud Thomaz & Soares, 2004)

do que deve ser considerado não estabelecimento de um projeto de preservação de qualquer tipo de documento, seja digital ou analógico.

MIGRAÇÃO E CONVERSÃO DE FORMATOS

A migração periódica da informação digital a partir de um ambiente de *hardware* ou de um *software* para outro é a estratégia operacional para a preservação digital mais freqüentemente usada pelas instituições detentoras de grandes acervos. Segundo a definição da *Task Force on the Archiving of Digital Information* (1996), migração é a transferência periódica de materiais digitais de uma configuração de *hardware/software* para outra ou, de uma geração de tecnologia computacional para a geração seguinte. O propósito da migração é preservar a integridade dos objetos digitais e assegurar a habilidade

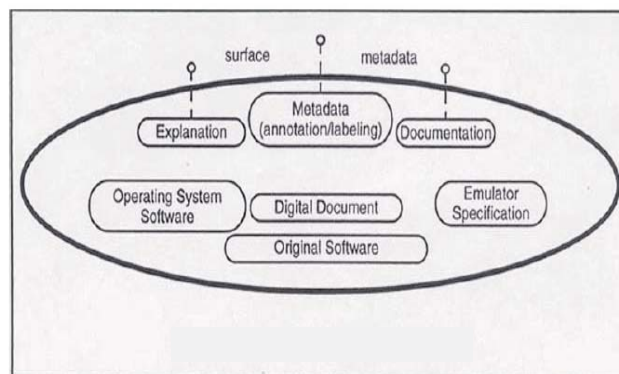
dos clientes para recuperar, expor e usá-los de outra maneira diante da constante mudança da tecnologia. A importância da migração é transferir para novos formatos enquanto for possível, preservando a integridade da informação. Um arquivo digital pode converter objetos digitais para um pequeno número de formatos “padrão”. Por exemplo, dados textuais podem ser estocados em um formato de *software* relativamente independente como ASCII, em formatos proprietários de grande difusão como PDF ou em formatos baseados em aplicações de SGML (*Standard Generalized Markup Language*) como XML (Coleman e Willis, 1997). Passado o tempo, os dados podem ser copiados, “re-criados” e atualizados como migrações periódicas necessárias em novos formatos para seu uso em novas gerações de *hardware* e *software*. Na área de documentos digitais, alguns modelos começam a surgir no sentido de converter materiais de formatos proprietários para a linguagem XML.

Os metadados têm um papel importante em qualquer estratégia de migração bem-sucedida. Esse tipo de estratégia dependerá dos metadados criados para registrar a história da migração de um objeto digital. Também, existe a necessidade de informação do contexto para ser registrada (e preservada) para que, dessa maneira, futuros usuários possam entender o ambiente tecnológico no qual um objeto digital foi criado.

EMULAÇÃO

As técnicas de emulação sugerem a preservação do dado no seu formato original, por meio de programas emuladores que poderiam imitar o comportamento de uma plataforma de *hardware* obsoleta e emular o sistema operacional relevante. O processo consiste na preparação de um sistema que funcione da mesma forma que outro de tipo diferente, para conseguir rodar programas. Essa estratégia está relacionada à preservação do dado original no seu formato original. No lugar de preservar o *software* e *hardware* hospedeiro, os engenheiros de sistemas poderiam construir programas emuladores (Rothenberg, 1995). Na prática, os dados podem ser encapsulados junto com a aplicação de *software* utilizada na sua criação, assim como uma descrição do ambiente de *software* e *hardware* requerido para seu funcionamento. Para facilitar seu uso no futuro, Rothenberg (1996) sugere que seja anexada uma anotação de metadados na superfície de cada encapsulação que poderia explicar como decodificar os dados obsoletos contidos e prover qualquer informação desejada sobre esses registros. A emulação precisa do desenvolvimento de técnicas de encapsulamento de documentos, seus metadados, *software* e especificações de emulador de forma a assegurar sua coesão e prevenir sua corrupção (figura 3).

FIGURA 3
Um documento digital encapsulado (Rothenberg, 1999)



Emulação é uma estratégia importante que possui aplicações úteis quando a aparência do recurso digital original é importante, mas onde não seja aconselhável investir em uma tecnologia da informação de alto custo. Hendley (1998) pede precaução no caso de se ater apenas nesta abordagem e comenta que os gerentes de coleções ficaram dependendo da habilidade técnica dos engenheiros de *software* para emular um ambiente específico e sustentá-lo, para rodar em futuros e desconhecidos computadores e recriar o comportamento de documentos digitais.

Um exemplo é o modelo *Digital Rosetta Stone* (DRS) desenvolvido por Steven Robertson da *United States Air Force* (Heminger e Robertson 1998). O modelo descreve três processos necessários para manter o acesso a longo prazo aos documentos digitais nos seus formatos originais:

- preservação do conhecimento;
- recuperação dos dados;
- reconstrução dos documentos.

O modelo desenvolvido garante o acesso aos repositórios de informação armazenados digitalmente. Os processos e a manutenção dos metadados sugeridos pelo DRS servirão para catalogar os diferentes aspectos das tecnologias digitais.

O processo de preservação do conhecimento

Ele dá suporte à descoberta dos dados e dos processos de reconstrução dos documentos. É a preservação do conhecimento das técnicas de armazenamento da mídia e dos formatos dos arquivos que serão mantidos nos arquivos de metaconhecimento. Metaconhecimento é o conhecimento ou alerta dos fatos, regras e do contexto no qual eles são usados e manipulados. O arquivo metaconhecimento *Meta Knowledge Archive* (MKA) é um dos fundamentos sobre o qual o DRS depende e preserva

o conhecimento em duas áreas-chave: as técnicas de armazenamento da mídia e os formatos dos arquivos.

Recuperação dos dados

É o processo de extração dos dados digitais da mídia obsoleta e sua migração para uma mídia que seja acessível aos sistemas de informação atuais. O conhecimento dos documentos digitais é maior que o custo da descoberta.

Reconstrução dos documentos

É o processo de interpretar documentos digitais dos arquivos originais usando informação de formato de arquivo reunida durante o processo de preservação do conhecimento. O arquivo de informação do formato descreve a formatação de informação usada para especificar as aplicações do *software*. É um *template* que pode ser usado para descrever a maneira como o dado é formatado e mostrado nos processadores de texto.

Um caso especial são os metadados da mídia. É o tipo mais fácil de dados coletado pelo DRS, isso porque as normas para armazenar mídia são rigidamente definidas antes que a mídia seja comercializada. Por exemplo, ISO 9660 é a norma que especifica como os dados são armazenados em um CD-ROM. Essa norma define estruturas do volume, arquivo e todos os atributos associados a um CD-ROM. Esse tipo de dado deve ser reunido para cada tipo de mídia e ser incluído no arquivo de metaconhecimento.

CONSERVAÇÃO DO SOFTWARE /HARDWARE

A perspectiva de preservação digital que propõe a conservação do *software* e do *hardware* (*technology preservation*) sugere que os dados digitais sejam preservados em uma mídia estável (e “refreshed” ou copiados para uma nova mídia caso seja necessário) e associados a cópias preservadas de uma aplicação de *software* original, um sistema operacional que será normalmente usado para ser lido sob uma plataforma de *software* relevante. Essa estratégia pode ter valor para casos particulares ou históricos de *software* e *hardware*, bem como ser útil para a comunidade dos museus (Swade, 1992). A longo prazo, entretanto, é cara e pouco prática. Segundo Hendley (1998) a perspectiva da tecnologia da preservação não pode ser apreciada como viável, mas apenas para o curto e meio termo, como uma medida de relativo desespero em casos em que valiosos recursos digitais não possam ser convertidos para formatos independentes de *software* e *hardware* e, assim, migrados para outros formatos.

OS CUSTOS

Cada estratégia técnica de preservação e de acesso implica diferentes custos e cronogramas. A preservação digital requer recursos disponíveis permanentemente que começam a se delinear no momento da criação do recurso.

Muitos materiais são criados fora dos centros de informação e das bibliotecas. Nesses casos, são os editores e outros criadores de conteúdos que devem adotar os padrões apropriados e as tecnologias que darão subsídios à preservação. Segundo Feeney (1999), “diferentemente da situação que se aplica aos livros, o arquivamento digital requer investimentos relativamente freqüentes para superar a obsolescência rápida produzida pelas mudanças tecnológicas”. Os provedores de material digital, no correr do tempo, precisam investir para criar documentação e metadados, gerando novas formas de material para manter o acesso. Esse investimento deve ser levado em consideração no momento de discutir os direitos de uso e reuso dos objetos digitais. Do mesmo modo, entre os aspectos que devem ser identificados em qualquer estratégia de preservação está a necessidade de contratação e de capacitação de pessoal. Se não houver nenhum plano administrativo que cubra esses itens, nenhuma manutenção de acervos a longo prazo será bem-sucedida.

EXPERIÊNCIAS

Desde 1995, várias iniciativas têm surgido internacionalmente. São projetos que envolvem o estudo e a prática de atividades de preservação digital em instituições preocupadas com a perenidade do crescente número de documentos digitais nos seus acervos. Desde o início, foram identificados os aspectos mais relevantes relacionados com essas atividades e propostas algumas ações. A maioria das soluções tomadas enfatizava o papel dos desenvolvedores na tomada de decisões que envolviam os primeiros estágios na criação desses repositórios, assim como o grau de confiança que deveria ser atingido pelos novos repositórios digitais por parte dos usuários. As pesquisas desenvolvidas até o momento indicam as atividades realizadas por meio de consórcios internacionais como o caminho mais viável para resolver as questões relacionadas com a preservação dos arquivos digitais. Na Inglaterra, a *Digital preservation coalition* foi estabelecida; na Austrália a National Library of Australia fundou o projeto PADI (*Preserving Access to Digital Information*), um portal especializado em recursos para preservação digital; o Council on Library and Information Resources (CLIR) e a Digital Library

Federation (DLF) realizaram vários relatórios para chamar a atenção sobre o problema e o que a pesquisa na área está fazendo.

Algumas organizações têm trabalhado no estabelecimento de padrões e melhores práticas. A OCLC e o Research Libraries Group (RLG)* desenvolveram documentos em que estabeleceram quais seriam as melhores práticas. Neles, estão os requisitos sugeridos para os modelos de metadados de preservação, a saber:

- abrangência: metadados estendidos a todos os aspectos do processo de preservação digital;
- estruturação: complemento aos componentes e processos funcionais de um sistema de arquivamento digital;
- ampla aplicabilidade: metadados aplicáveis a uma extensa faixa de tipos de objetos digitais, atividades de preservação digital e instituições.

Também foram determinados os atributos necessários para a construção de repositórios digitais confiáveis (*Trusted digital repository*):

- compilação com o *Reference Model for an Open Archival Information System (OAIS)*;
- responsabilidade administrativa;
- viabilidade organizacional;
- sustentabilidade financeira;
- sustentabilidade tecnológica e processual;
- sistema de segurança;
- responsabilidade processual.

Alguns experimentos têm sido elaborados em universidades. A Andrew W. Mellon Foundation financiou sete universidades para trabalhar com editores no planejamento de repositórios para o conteúdo de periódicos eletrônicos. Através do *PubMed Central*, a National Library of Medicine atua como um repositório de arquivos digitais para publicações da área médica e de outras informações da área da saúde.

O projeto NSDL (*National Science Digital Library*), centrado em ciências, tecnologia, engenharia, computação e matemática, e o Centro de Supercomputadores de San Diego na Califórnia (SDSC San Diego Supercomputer Center) possuem instalações imensas para comportar grandes acervos, assim como convertê-los e organizá-los

para serem acessados no futuro. Os trabalhos da National Archives and Records Administration (Nara) e o San Diego Supercomputer Center, assim como a Mellon Foundation, são dois exemplos relevantes das pesquisas em projetos sobre arquivos de periódicos eletrônicos.

No caso dos projetos da Nara, estão claramente separados os armazenamentos físicos da sua interpretação lógica, e existe a distinção entre o gerenciamento dos dados e o gerenciamento do conhecimento. O conceito de organizar a preservação digital segundo seus componentes dentro de um modelo de arquitetura da informação permite que as responsabilidades sejam distribuídas entre vários tipos de instituições. Existe um consenso atualmente de que a existência de um arcabouço é um passo importante, mesmo que não seja a única resposta ao problema da preservação digital. No momento são necessárias soluções em termos de escala, tipos de formatos e responsabilidades institucionais. Qualquer organização que esteja iniciando a construção de uma coleção digital deve seguir um roteiro e ter uma fonte financiadora. A NSF (National Science Foundation) e a Library of Congress trabalham unidas para desenvolver uma agenda de pesquisa sobre arquivamento e preservação digital a longo prazo.

A Library of Congress desenvolve uma estratégia nacional para preservar informação digital. Com um orçamento de 100 milhões de dólares do Congresso americano, a biblioteca estabeleceu um conselho nacional e trabalha com um número de agências privadas e governamentais para desenvolver o *Metadata Encoding and Transmission Standard (METS)*. A Library of Congress recebeu a encomenda de construir, unida a outras instituições, um programa nacional sobre preservação digital. O projeto *Making of America II (MOA2)* propôs um formato de codificação para metadados administrativos e estruturais para trabalhos textuais e baseados em imagens. METS é uma iniciativa que tenta construir um formato em XML para codificar metadados necessários tanto para a gestão de objetos de bibliotecas digitais em um repositório quanto para a troca desses objetos entre repositórios. O esquema METS oferece um mecanismo flexível para codificar metadados descritivos, administrativos e estruturais para um objeto de uma biblioteca digital e para exprimir as ligações complexas entre várias formas de metadados.

Em 2001, a Digital Library Federation e a University of Cornell, a University of Michigan e a Library of Congress começaram a discutir a importância das características dos documentos digitais para a sua interoperabilidade e preservação. Com base no trabalho de Daniel Greenstein

* RLG and OCLC Explore Digital Archiving <http://www.rlg.org/legacy/pr/pr2000-oclc.html>

Gerald George *Digital Reproduction Quality: Benchmark Recommendations* (2001), foi definido o “*preservation digital master*”, que seria uma cópia digital fidedigna de publicações impressas que observaria ou excederia um conjunto mínimo de características. O grupo recomenda também que a preservação do original digital tenha uma lista mínima de elementos de metadados (descritivos, estruturais e administrativos, disponíveis e documentados em XML). Esse conjunto de características foi primeiramente aplicado a manuscritos, obras raras e textos com ilustrações.

No caso da Inglaterra, o Joint Information Systems Committee (JISC) patrocinou inúmeros projetos de pesquisa e iniciativas com o intuito de conhecer os desafios da preservação a longo prazo. Projetos como o *CURL Exemplars in Digital Archives* (Cedar) (em colaboração com três universidades), que tinha como objetivo “*addressing strategic, methodological and practical issues and provide guidance for libraries in best practice for digital preservation*” e produziu cinco guias sobre aspectos específicos como gerenciamento de coleções digitais, propriedade intelectual, metadados, estratégias de preservação digital e protótipos de arquivamento digital. Outro esforço colaborativo foi o do projeto Cedar com o projeto Camileon (*Creative Archiving at Michigan & Leeds: Emulating the Old on the New*), este último fundado pelo JISC e a National Science Foundation dos Estados Unidos. O projeto Camileon testou a viabilidade da emulação como estratégia de preservação e desenvolveu ferramentas para dar suporte “*(...) in future rendering of digital materials*”. Todos os esforços realizados pelo JISC indicam a crescente necessidade de acelerar a implementação de estratégias de preservação, mas concentrando esforços para estabelecer uma estratégia principal “*(...) an infrastructure of collective and institutional services and repositories*”.

A European Commission fundou o projeto *Electronic Resource Preservation and Access Network* (Erpanet) em 2001, tendo como finalidade unir as diversas iniciativas na área de preservação. A principal contribuição do projeto foi mostrar a importância de participar de projetos cooperativos de preservação digital, nos quais são maximizados os esforços individuais, constroem-se parcerias, é alcançada uma melhora na administração dos riscos e se evita a redundância. A Erpanet vem monitorando as iniciativas sobre preservação digital que estão surgindo no mundo (<http://www.erpanet.org>). Até 2002 eram 63 projetos. As iniciativas existentes são limitadas na sua abrangência ou na sua cobertura geográfica.

O Modelo Lockss (*Lots of Copies Keep Stuff Safe*) para preservar através da replicação distribuída também

implica que, sem o repositório, a preservação é impossível. O projeto Lockss, da Stanford University, procura garantir a integridade das publicações eletrônicas pela manutenção de cópias em vários sites, checando periodicamente essas cópias para verificar a congruência informacional. Com a clonagem e o armazenamento distribuído, o Lockss espera minimizar antecipadamente o impacto de uma catástrofe em um site individual (Reich & Rosenthal, 2001).

A United Nations Educational, Scientific and Cultural Organization (Unesco) lançou, em 2002, uma nova iniciativa chamada “*Preserving Our Digital Heritage*”, unida ao *Memory of the World Programme*. Sua finalidade é construir uma linha de defesa contra a perda de informação valiosa que será implementada desde o ponto inicial da criação desses recursos. O grupo encarregado da preservação digital pronunciou uma estratégia de meio termo sobre o assunto. Atualmente o projeto está examinando as atividades na área de digitalização no mundo e já desenvolveu, em parceria com a International Federation of Library Associations (Ifla), uma base de dados *on-line* chamada “*Directory of Digitised Library Collections*”, com o objetivo de listar as principais coleções de bibliotecas e outras instituições com conteúdos relevantes para a herança cultural que já foram digitalizados. Nas últimas reuniões dos grupos interessados no diretório, foi enfatizada a importância de reconhecer a necessidade de que os diferentes esforços tenham um mínimo de consistência e de interoperabilidade.

A EXPERIÊNCIA BRASILEIRA

No Brasil, a Biblioteca Nacional está na fase inicial da definição dos metadados de preservação. Em contrapartida, a Câmara Técnica de Documentos Eletrônicos do Arquivo Nacional (CONARq) abriu para consulta pública no mês de abril de 2004 o anteprojeto da Carta para a Preservação do Patrimônio Arquivístico Digital (figura 4, a seguir). O objetivo do conteúdo da Carta é conscientizar as comunidades que possuem arquivos de documentos digitais a adotar práticas padronizadas de preservação digital. A carta coloca a necessidade de que sejam criadas políticas, estratégias e ações sobre o tema. Ela define o que é um documento arquivístico digital e propõe que todos os arquivos adotem algum sistema de gerenciamento arquivístico de documentos digitais. A proposta esteve baseada em estudos de projetos internacionais (InterPARES I e II*,

* International Research on Permanent Authentic Records in Electronic Systems.

o projeto *Pittsburgh Functional and Requirements for Evidence in Recording*, o modelo de requisitos para a gestão de arquivos eletrônicos – MoReq –, os requisitos usados pelo United States Department of Defense Records – DOD – e a metodologia DIRKS*).

O documento é introdutório no sentido de colocar na pauta de discussão a preservação digital. Baseado na Carta da Unesco, reproduz quase na íntegra as sugestões que aquela instituição pede para os estados membros adotarem com respeito às medidas para salvaguardar o patrimônio cultural de cada país. O texto fica mais no sentido de recomendações para “melhores práticas” e chama para um estudo mais aprofundado das limitações da tecnologia digital. Não foi lembrada a necessidade de uma especificação maior dos tipos de informação a serem arquivados a longo prazo (dados científicos, arquivos pessoais etc.). A preservação digital não é apenas realizar atividades arquivísticas ou de desenvolvimento de coleções. Ela inclui o gerenciamento da aplicação de estratégias de preservação apropriadas para cada tipo de acervo.

O documento não menciona os custos de qualquer estratégia de preservação e acesso aos documentos digitais ao longo do tempo. Por serem os arquivos administrativos o foco de atenção da Carta, a questão da “autenticidade” aparece no centro das preocupações. Ao colocar o problema no nível das prioridades nacionais que precisam de regulamentação, a Carta não aponta para a inclusão da problemática da preservação a longo prazo no âmbito dos programas oficiais de apoio ao uso de *software* livre, de inovação tecnológica e de desenvolvimento tecnológico.

Não existem dúvidas de que novos esforços aparecerão para entender as técnicas e os padrões que devem ser aplicados no arquivamento da informação de relevância e a manutenção do seu acesso no futuro. O Arquivo Nacional é o fórum adequado para a interação entre as organizações que criam, armazenam e disseminam informação digital. São essas instituições que podem definir os requisitos indispensáveis para chegar ao estabelecimento de estratégias estruturais e operacionais de preservação digital, dentro de uma política nacional de informação a ser considerada em qualquer investimento de redes de informação eletrônica, privado ou público.

* Designing & Implementing a Recordkeeping Systems.

FIGURA 4
Página do CONARq [<http://www.arquivonacional.gov.br/conarq/>]



CONCLUSÃO

Os documentos digitais são considerados, atualmente, registros oficiais e são gerenciados segundo leis e padrões que compreendem todo o ciclo de vida desses materiais. A preservação digital requer procedimentos específicos e técnicas apropriadas para cada tipo de formato e mídia. Pesquisas internacionais propõem o uso de técnicas de emulação, migração e preservação da tecnologia (Webb, 2000), assim como da preservação de metadados (Rothemberg, 1996) e a criação de repositórios que permitam o acesso e a recuperação dos dados.

Com a preservação digital, pretende-se garantir a inalterabilidade dos registros digitais. Todos os métodos mencionados procuram evitar o risco de criar barreiras para um uso pleno dos recursos no futuro. Para resolver esse problema, estão sendo elaborados e testados nos últimos anos esquemas conceituais que permitem identificar as partes integrantes do processo de preservação digital, assim como sua padronização em modelos de referência para iniciativas de repositórios digitais (CCSDS, 2002).

As tecnologias para a preservação digital estão sendo pesquisadas com o intuito de compreender as suas implicações dentro das políticas de criação e uso de repositórios de informação digital. Alguns estudos sobre a preservação digital têm estabelecido que a imediata implementação de políticas de preservação digital é a forma mais efetiva de garantir o armazenamento e uso de recursos de informação por longos períodos de tempo. A falta dessas políticas nos projetos de repositórios

digitais sugere a carência de conhecimentos técnicos sobre a importância das estratégias de preservação digital existentes. Essa lacuna informacional por parte dos responsáveis pelas políticas de implementação de informação digital precisa ser destacada. A segunda lacuna informacional que deve ser apontada é sobre a importância da existência de bibliotecas digitais e de centros especializados no estabelecimento e monitoração de estratégias de preservação digital bem-sucedidas.

O Brasil precisa de canais oficiais para a identificação, comunicação e avaliação na área de preservação digital de longa duração. O modelo de centros de preservação digital no Brasil poderá ser semelhante àqueles desenvolvidos pelas iniciativas observadas na comunidade internacional durante as últimas décadas: formulação de políticas que envolvam o uso de tecnologias e padrões testados e adaptados em projetos de coleções digitais. A organização desses centros não pode desconsiderar a sua dimensão internacional, já que, no momento em que se escolhe determinado modelo de preservação, ele deve ser integrado aos outros repositórios que adotem a mesma estratégia.

Artigo recebido em 29-07-2004 e aceito para publicação de 18 a 21/10/2004.

REFERÊNCIAS

- ARMS, William. *Digital libraries*. Massachusetts : Massachusetts Institute of Technology, 2000.
- BEAGRIE, N.; GREENSTEIN, D. *A strategic policy framework for creating and preserving digital collections*. London : Arts and Humanities Data Service. Disponível em: <<http://ahds.ac.uk/managing.htm>>. Acesso em: 28 jul. 2004.
- BEARMAN, David. *Electronic evidence: strategies for managing records in contemporary organizations*. Archives and Museum Informatics. [S.l. : s. n., 2000?].
- BORBINHA, José, CORREIA, RAMALHO, Ana Maria. Deposit of scientific and technical gray literature in Portugal: a case study. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 2001, Setúbal, Portugal, [S. l. : s. n.], 2001.
- BULLOCK, Alison. *Preservation of digital information; issues and current status*. Disponível em: <<http://www.collectionscanada.ca/9/1/p1-259-e.html>>. Acesso em: 28 jul. 2004.
- CHAPMAN, Stephen. Counting the costs of digital preservation: is repository storage affordable? *Journal of Digital Information*, v. 4, n. 2, 2003. Disponível em: <<http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Chapman/>>. Acesso em: 28 jul. 2004.
- CHEPESUIK, R. The future is here: America's libraries go digital. *American Libraries*, v. 2, n. 1, p. 47-49, 1997.
- COLEMAN, J.; WILLIS, D. *SGML as a framework for digital preservation and access*. Washington, D.C. : Commission on Preservation and Access, 1997. Disponível em: <<http://www.clir.org/pubs/cpanews/cpan1101.html>>. Acesso em: 28 jul. 2004.
- CUNHA, Murilo Bastos da. Desafios na construção de uma biblioteca digital. *Ciência da Informação*, Brasília, v. 28, n. 3, p. 257-268, set./dez. 1999. Disponível em: <<http://www.ibict.br/cienciainformacao/rst/viewarticle.php?id=323>>. Acesso em: 28 jul. 2004.
- DAY, Michael. *E-print services and long-term access to the record of scholarly and scientific research*. *Ariadne*, n. 28, June 2001. Disponível em: <<http://www.ariadne.ac.uk/issue28/metadata/>>. Acesso em: 28 jul. 2004.
- FEENEY, M. Towards a national strategy for archiving digital materials. *Alexandria*, v. 11, n. 2, p. 107-121, 1999.
- GRANGER, Stewart. Digital preservation and deep infrastructure. *D-Lib Magazine*, v. 8, n. 2, Feb. 2002. Disponível em: <<http://www.dlib.org/dlib/february02/granger/02granger.html>>. Acesso em: 28 jul. 2004.
- GREENSTEIN, Daniel; GEORGE, Gerald. Digital reproduction quality: benchmark recommendations. *RLG DigiNews*, v. 5, n. 4. Disponível em: <<http://www.rlg.org/preserv/diginews/diginews5-4.html#featured>>. Acesso em: 28 jul. 2004.
- HEDSTROM, Margaret. *Digital preservation: a time bomb for digital libraries*. Disponível em: <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>. Acesso em: 28 jul. 2004.
- HEMIGER, A. R.; ROBERTSON, S. B. Digital rosetta stone: a conceptual model for maintaining long-term access to digital documents. In: DELOS WORKSHOP PRESERVATION OF DIGITAL INFORMATION, 6., 1998. [S. l.] : Le Chesnay European Research Consortium for Informatics and Mathematics, 1998. p. 35-43. (ERCIM-98-W003). Disponível em: <<http://www.ercim.org/publication/ws-proceedings/DELOS6/>>. Acesso em: 28 jul. 2004.
- HENDLEY, T. *Comparison of methods & costs of digital preservation: British Library research and innovation*. London : British Library and Innovation Centre, 1998. (Report, n. 106). Disponível em: <<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html>>. Acesso em: 28 jul. 2004.
- HILDRETH, C. R. Preserving what we really want to access, the message, not the medium: challenges and opportunities in the digital age. In: HELAL, A. H.; WEISS, J. W. (Ed.). PROCEEDING OF THE INTERNATIONAL ESSEN SYMPOSIUM, 18., 1995, Essen. Festschrift in honour of Patricia Battin. Essen : Germany Universitätsbibliothek, 1996. p. 76-95.
- LESK, M. *Practical digital libraries: books, bytes and bucks*. San Francisco, CA : Morgan Kaufmann, 1997.
- NATIONAL LIBRARY OF NEW ZEALAND. *Initiatives*. 2003. Disponível em: <<http://www.natlib.govt.nz/en/whatsnew/4initiatives.html>>. Acesso em: 28 jul. 2004.
- _____. Metadata standards framework: preservation metadata. Disponível em: <<http://www.natlib.govt.nz/en/whatsnew/4initiatives.html#meta>>. Acesso em: 28 jul. 2004.
- OCLC/RLG. *Preservation metadata for digital objects: a review of the state of the art - a white paper*. Disponível em: <http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf>. Acesso em: 28 jul. 2004.
- REICH, L.; SAWYER, D. (Ed.). *Reference model for an open archival information system*. [S. l.] : Consultative Committee for Space Data Systems, 1998. (650.0-W-4.0). Disponível em: <http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html>. Acesso em: 28 jul. 2004.
- REICH, Vicky; ROSENTHAL, David S. H. *Looks: a permanent web publishing and access system*. [S. l. : s. n.], 2001.

Preservação de documentos digitais

- RESEARCH LIBRARY GROUP. *Trusted digital repositories: attributes and responsibilities and RLG-OCLC report*. Mountain View, Canada : OCLC, 2002. Disponível em: <<http://www.rlg.org/longterm/repositories.pdf>>. Acesso em: 28 jul. 2004.
- ROTHENBERG, J. Metadata to support data quality and longevity. In: PROCEEDINGS OF THE IEEE METADATA CONFERENCE, 1., 1996, Silver Spring, Md.[S. l. : NOAA Complex, 1996. Disponível em: <http://www.computer.org/conferences/meta96/rothenberg_Paper/ieee.data-quality.html>. Acesso em: 28 jul. 2004.
- SANT'ANNA, Marcelo Leone. Os desafios da preservação de documentos públicos digitais. *Revista IP*, v. 3, n. 2, dez. 2001. Disponível em: <<http://www.informaticapublica.mg.gov.br/revista0302/ip0302santanna.pdf>>. Acesso em: 28 jul. 2004.
- SEARLE, Sam.; THOMPSON, Dave. Preservation metadata: pragmatic first steps at the National Library of New Zealand. *DLib Magazine*, v. 9, n. 4, Apr. 2003. Disponível em: <<http://www.dlib.org/dlib/april03/thompson/04thompson.html>>. Acesso em: 28 jul. 2004.
- SWADE, D. The problems of software conservation. *History and Computing*, v. 4, n. 2, 1992. Disponível em: <http://www.cs.man.ac.uk/CCS/simulate/sim_home.htm>. Acesso em: 28 jul. 2004.
- TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION. *Preserving digital information: report of the task force on archiving of digital information*. Washington, D.C. : Commission on Preservation and Access, 1996. Disponível em: <<http://www.rlg.org/ArchTF/>>. Acesso em: 28 jul. 2004.
- THOMAZ, Kathia P. A preservação digital e o modelo de referência: open archival information system. *DataGramaZero - Revista de Ciência da Informação*, v. 5 n. 1, fev. 2004. Disponível em: <http://www.dgz.org.br/fev04/Art_01.htm>. Acesso em: 28 jul. 2004.
- WEBB, Collin. *The role of preservation and the library of the future*. [S. l.] : National Library of Australia, 2000. Disponível em: <<http://www.nla.gov.au/nla/staffpaper/cwebb9.html>>. Acesso em: 28 jul. 2004.