

A preservação digital e o modelo de referência Open Archival Information System (OAIS)

Digital preservation and the Open Archival Information System (OAIS)

por [Katia P. Thomaz](#) e [Antonio José Soares](#)

Resumo: O volume de informação em formato digital tem aumentado muito nas últimas décadas. Entretanto, os meios empregados para o armazenamento e o transporte desta informação são instáveis e as tecnologias necessárias para acessá-la são rapidamente superadas em sucessivas gerações. Diversas iniciativas surgem para tratar o problema da preservação digital e este artigo oferece uma visão geral deste cenário. Nesta primeira parte examina as questões, requisitos e principais estratégias adotadas. Numa segunda parte descreve o esquema conceitual para um sistema de arquivamento para preservação e manutenção de informação digital por longo prazo, proposto pelo modelo de referência *Open Archival Information System* (OAIS), como a alternativa de padronização mais promissora para a área.

Palavras-chave: Arquivologia; Gestão de Documentos de Arquivo; Gestão Eletrônica de Documentos de Arquivo; Documento Eletrônico de Arquivo; Preservação Digital; Requisitos de Preservação; Estratégias de Preservação.

Abstract: The information volume in digital forms has been increasing over the last decades. However, the means we use to carry and store it are unstable and the technology needed for access is quickly superseded by newer technologies, generation after generation. Many initiatives emerge to address the digital preservation problem and this paper provides an overview of the issues involved. This first part examines the main issues, requirements and proposed strategies. The second part describes the conceptual framework for an archival system to preserving and maintaining access to digital information over the long term, suggested by the Open Archival Information System (OAIS) reference model, as the most promising approach to the field.

Keywords: Archival Science; Records Management; Electronic Recordkeeping; Electronic Record; Digital Preservation; Preservation Requirements; Preservation Strategies.

“... apesar de a sua reprodutibilidade tornar a informação digital teoricamente invulnerável ao tempo, os suportes dos quais se serve estão longe de ser eternos.”

(LOPES, CARDOSO & MOREIRA, 2002).

1. Introdução

O volume de informação em formato digital tem aumentado muito nas últimas décadas. Lusenet (2002), em recente publicação, apresenta dados surpreendentes: estima-se a produção anual de informação, nos dias atuais, em torno de um a dois bilhões de Gigabytes, sendo 90% deste volume em formato digital, e ainda, grande parte deste último, existente exclusivamente em formato digital. Esta informação, entretanto, é transportada e armazenada em meios instáveis e as tecnologias necessárias para acessá-la são rapidamente superadas em sucessivas gerações.

É importante observar que, uma vez interrompido o suporte à tecnologia, o acesso à informação também é perdido. Esta é a problemática central da preservação digital como um todo, tratada especificamente na primeira parte deste artigo: manter o acesso a objetos digitais ao longo do tempo. O texto discorre sobre conceitos, problemas, requisitos e estratégias, construindo as bases para o entendimento do modelo de referência Open Archival Information System (OAIS) a ser abordado na segunda parte.

2. Definindo a preservação digital

Da revisão da literatura pode-se extrair duas definições básicas de preservação digital ou arquivamento digital:

* Planejamento, alocação de recursos e aplicação de métodos de preservação e tecnologias necessárias para que a informação digital de valor contínuo[1] permaneça acessível e utilizável por longo prazo, considerando-se neste caso longo prazo, o tempo suficiente para preocupar-se com os impactos de mudanças tecnológicas. A preservação digital aplica-se tanto a documentos "nato-digitais"[2] quanto a documentos convertidos do formato convencional para o formato digital. (HEDSTROM, 1997/1998)

* Capacidade de manter a integridade e a acessibilidade da informação digital por longo prazo. Esta preservação da integridade e acessibilidade não se limita, apenas, a proteger a informação digital contra o acesso não autorizado mas, também, contra o uso inadequado resultante da má interpretação ou má representação da informação por parte dos sistemas computacionais. Percebe-se, aqui, o aspecto da inseparabilidade entre as atividades de preservação e acesso do mundo digital. (TASK FORCE ON ARCHIVING OF DIGITAL INFORMATION, 1996)

3. O problema da preservação digital

A dificuldade fundamental da preservação digital advém da natureza dos próprios objetos que busca preservar. Diferentemente dos formatos tradicionais, os objetos digitais são acessíveis somente através de combinações específicas de componentes de *hardware*, *software*, mídia e pessoal técnico. Até bem pouco tempo atrás, a atenção de profissionais da informação concentrava-se, apenas, na longevidade do suporte físico onde a informação era armazenada. Esse posicionamento não é suficiente no mundo digital. Mesmo nas melhores condições de armazenamento, as mídias digitais podem ter sua vida interrompida pela falta ou inadequação de qualquer um dos demais componentes.

Segundo Lusenet (2001), para compreender o conjunto de problemas associados à preservação digital é preciso, principalmente e antes de mais nada, conscientizar-se de que: 1) as mídias são suportes transitórios que prestam sua função somente por um período limitado de tempo e que a transferência para novas mídias é absolutamente necessária; 2) o *software* e o *hardware* tornam-se obsoletos em questão de anos, ao invés de décadas, e que embora as versões sucessivas de programas possam ser compatíveis, os fabricantes de *software* normalmente não garantem a compatibilidade por um longo período; e 3) o *software* proprietário é problemático não somente porque é protegido e o código fonte não está disponível mas, também, porque normalmente está documentado de forma inadequada tornando a conversão de dados muito mais complexa.

Somam-se a esses problemas fundamentais da preservação digital, outras dificuldades de ordem técnica, política, econômica e social, dentre as quais destacamos:

* risco de agravamento do problema da exclusão social (p.ex., INTO..., 1997);

* dificuldade de definição de estratégias e custos associados às atividades de preservação digital (p.ex. CONWAY, 1996, TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION, 1996; HEDSTROM, 1997/1998; ROTHENBERG, 1999);

* agravamento do impacto ambiental com o fenômeno do “lixo eletrônico” (p.ex., ENTULHO..., 2000);

* falta de preparo de gestores de acervos e especialistas de preservação nas questões ligadas ao ambiente da tecnologia da informação (p.ex. ROUSSEAU & COUTURE, 1998);

* mercado altamente competitivo da tecnologia da informação levando a ciclos de renovação de tecnologia a cada 3 a 5 anos (p.ex. HEDSTROM, 1997/1998; BRAND, 1999);

* falta de linguagem própria nas diversas áreas do conhecimento para descrever os novos relacionamentos e tipos de documentos (p.ex. BELLOTTO, 1991; JARDIM, 1992);

* dificuldade para estabelecer as fronteiras dos documentos de natureza multimídia e hipertextual em

redes distribuídas de longa distância (p.ex., INTO..., 1997);

* no contexto digital a distinção entre documento original e cópia não é clara (p.ex. DURANTI, 2002);

* aumento da facilidade de uso e oferta de recursos da tecnologia da informação, acompanhado de aumento equivalente da complexidade e dependência tecnológica dos objetos digitais subjacentes (p.ex. INGWERSEN, 1995; TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION, 1996);

* aumento exponencial da capacidade de armazenamento, acompanhado de redução equivalente da expectativa de vida dos meios de armazenamento (p.ex., CONWAY, 1996.);

* abandono de fontes de informação digital quando as instituições deixam ou perdem o interesse pelos negócios (p.ex. TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION, 1996; LUSENET, 2002.);

* a tecnologia digital reformula antigos papéis e traz novos atores para o sistema sócio-econômico da informação (p.ex., CONWAY, 1996; TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION, 1996; HEDSTROM, 1997/1998);

* ausência de produtos de prateleira que considerem os requisitos de preservação por longo prazo (p.ex., BULLOCK, 1999);

* os atuais direitos de propriedade intelectual/copyright podem interferir na capacidade de preservar objetos digitais através de cópia sistemática (p.ex., BULLOCK, 1999);

* ênfase na geração e/ou aquisição de material digital numa era de redução de recursos, ao invés de manter a preservação e o acesso aos acervos eletrônicos existentes (p.ex., BULLOCK, 1999).

4. Requisitos da preservação digital

Medidas de preservação digital devem garantir que um objeto digital esteja acessível de forma utilizável ao longo do tempo. A partir dos problemas levantados no item anterior, percebe-se que manter a acessibilidade dos meios digitais é muito mais complexo, quando comparado aos meios não digitais, como o papel. Quando um relatório impresso é preservado em seu formato original, todos os seus aspectos são mantidos em sua presença física: seu formato, seu leiaute e seu conteúdo. Nesse caso, é praticamente impossível separar seus elementos individuais (p.ex., conteúdo sem leiaute) porque eles estão intrinsecamente ligados. Objetos digitais, ao contrário, podem, facilmente, ser decompostos em seus elementos individuais sendo necessário um esforço muito maior para preservá-los como um "todo". É possível, por exemplo, manter o conteúdo de um documento eletrônico mas perder o seu leiaute. Em outra situação, pode-se manter a presença física de um objeto digital (p.ex., o arquivo de dados gravado em um CD-ROM) mas deixar de preservar sua capacidade de leitura (p.ex., conservar o *hardware/software* adequado para apresentá-lo).

Uma instituição que pretende assumir a responsabilidade de preservar informação digital por longo prazo precisa, como primeiro passo, observar um conjunto mínimo de requisitos[3]. Neste sentido Bullock (2001), baseando-se nos conceitos do modelo de referência *Open Archival Information System* (OAIS) - que será objeto da segunda parte deste artigo - identificou nove itens, a saber:

* **Fixar os limites do objeto a ser preservado:** embora a natureza multimídia e hipertextual dos objetos digitais seja bastante vantajosa do ponto de vista da navegação, para fins de preservação é necessário definir, claramente, quais elementos serão efetivamente mantidos.

* **Preservar a presença física:** a presença física representa o(s) arquivo(s) físico(s), i.e., a camada

primitiva de suporte da informação a ser representada; refere-se, portanto, ao(s) arquivo(s) de computador, às séries de 0's e 1's que são a base para o significado de um objeto digital.

* **Preservar o conteúdo:** refere-se a manter a capacidade de acessar o conteúdo em seu nível mais baixo, como um arquivo texto em ASCII, independente do estabelecimento de variações de fontes e características de leiaute.

* **Preservar a apresentação:** o conteúdo é apresentado visualmente através da aplicação de fontes de diferentes formatos e tamanhos, uso de espaço em branco, colunas, margens, cabeçalhos, rodapés, paginação e assim por diante. Em alguns tipos de documentos digitais (p.ex., formatos padrão SGML e alguns formatos PDF) as especificações de apresentação ficam separadas do conteúdo.

* **Preservar a funcionalidade:** objetos digitais podem conter componentes multimídia (i.e., texto, gráficos, áudio e vídeo integrados), existir em formato hipertexto (i.e., podem desviar dinamicamente para outros pontos do próprio documento ou para outro documento), conter conteúdo dinâmico (i.e., gerado automaticamente a partir de bancos de dados) ou ter funções de navegação (i.e., barras de ferramentas, pesquisa a palavra-chave ou tabelas interativas de conteúdos).

* **Preservar a autenticidade:** é necessário confiar que o objeto acessado é exatamente aquele que se procura e que as possíveis transformações pelas quais passou, para manter sua acessibilidade, preservaram sua forma original.

* **Localizar e rastrear o objeto digital ao longo do tempo:** imediatamente após a sua criação, os objetos digitais tornam-se passíveis de serem alterados, copiados ou movimentados. Em qualquer referência ao objeto digital, é necessário localizá-lo na edição ou versão correta.

* **Preservar a proveniência:** identificar a origem de um objeto e detalhar seu histórico ajudam a confirmar sua autenticidade e integridade.

* **Preservar o contexto:** os objetos digitais são definidos por suas dependências de *hardware* e *software*, seus modos de distribuição e relacionamentos com outros objetos digitais.

5. Estratégias da preservação digital

Para enfrentar os novos desafios criados pelo mundo digital e, mesmo, aproveitar as oportunidades que oferece, algumas estratégias têm sido propostas. É preciso estudar cada uma delas, para compreender seus principais pressupostos, objetivos e resultados esperados. As estratégias podem ser agrupadas em dois conjuntos: estratégias estruturais e operacionais. As estratégias estruturais dizem respeito aos investimentos ou esforços iniciais por parte da instituição, preparando seu ambiente para o processo da preservação digital. As estratégias operacionais, por sua vez, representam as atividades ou medidas concretas de preservação digital.

5.1 Estratégias estruturais

5.1.1 Adoção de padrões

Esta estratégia recomenda o uso preferencial de padrões - de fato ou de direito - e formatos de arquivos de dados abertos, com amplo acesso e assistência técnica, para os quais exista uma crescente tendência de estabilidade e suporte por longo prazo. A intenção é simplificar a aplicação das outras estratégias de preservação e maximizar a sua efetividade. Esta estratégia pode estar relacionada à estratégia [5.1.3 Metadados para preservação digital].

Reunindo as idéias contidas no artigo de Bullock (1999) e no manual da NLA (2003), pode-se considerar que a estratégia de padrões consiste numa abordagem em quatro partes:

* definir um conjunto limitado de formatos para armazenar os dados;

- * usar padrões atuais para criar objetos digitais;
- * monitorar os padrões à medida que se modificam;
- * migrar para novos padrões uma vez estabelecidos.

Segundo Hedstrom (1997/1998), as bibliotecas e os arquivos estão, ainda, em fase inicial quanto ao teste de estratégias de normalização dos tipos de acervos através da conversão de documentos digitais da grande multiplicidade de formatos para uma quantidade menor e, conseqüentemente, mais gerenciável, de formatos padronizados. Ainda segundo a mesma autora, um repositório poderia, por exemplo, restringir o recebimento de material digital a apenas alguns formatos de processadores de texto mais comuns ou, mesmo, chegar a exigir que os documentos atendam a padrões como SGML (ISO 8879); os bancos de dados poderiam ser armazenados em um número reduzido de formatos comuns ou convertidos para um formato compatível com SQL (*Structured Query Language*), enquanto que os arquivos de imagem poderiam atender ao formato TIFF (*Tagged Image File Format*) com algoritmos de compressão padronizados.

5.1.2 Elaboração de manuais

Os arquivos e instituições de preservação de documentos na Europa, América do Norte e Austrália[4] têm liderado o desenvolvimento de melhores práticas e requisitos funcionais para abordar as questões da preservação, registrando suas descobertas em manuais ou guias para preservação digital ou gerenciamento de documentos eletrônicos (*Managing Electronic Records in an Electronic Work Environment*, Canadá, 1996; *Management, appraisal and preservation of electronic records*, Inglaterra, 1999; *Recomendações para a gestão de documentos de arquivo eletrônicos*, Portugal, 2000; *Designing and Implementing Recordkeeping Systems* (DIRKS), Austrália, 2001; *Design Criteria Standard for Electronic Records Management Software Applications* (DOD5015.2), EUA, 2002; *Les archives électroniques - Manuel pratique*, França, 2002). Estes manuais fornecem orientações gerais quanto ao tratamento de objetos digitais e o gerenciamento dos riscos envolvidos na sua preservação. A intenção é reduzir os riscos de perda de informação de valor contínuo, principalmente nas fases iniciais de seu ciclo de vida (corrente e intermediária), promovendo a sua movimentação através de sucessivas gerações tecnológicas. Segundo Bullock (1999), estes manuais têm em comum as seguintes recomendações:

- * “reconhecer a responsabilidade inicial do produtor na preservação de seus documentos;
- * identificar as responsabilidades da instituição arquivística;
- * adotar diretrizes adequadas para seleção [dos objetos a serem preservados] (que destacam a questão dos padrões quando existir possibilidade de escolha de formatos);
- * proteger os itens arquivados de alteração intencional e não intencional;
- * fornecer descrição de contexto incluindo histórico de criação, transferência e uso, e registros de auditoria;
- * descrever de forma completa os objetos digitais.”

5.1.3 Metadados para preservação digital

Existem, no mínimo, três fortes razões para a adoção de metadados – informação estruturada sobre dados – para descrever objetos digitais em detalhe: 1) facilitar a pesquisa e a identificação de suas fontes de informação; 2) gerenciar seu fluxo dentro de processos; e 3) representar suas estruturas para possibilitar o acesso. Os programas de preservação podem ter que optar entre aceitar, e possivelmente adaptar, um dos modelos de metadados atualmente em uso (p.ex., MARC, Dublin Core) ou especificar seu próprio esquema (como uma solução completa ou como uma solução mínima provisória até que surja um padrão).

Segundo o manual da NLA (2003), os metadados de preservação digital geralmente:

- * "identificam o material pelo qual o programa de preservação tem responsabilidade
- * informam o que é necessário para manter e proteger os dados
- * informam ao usuário, quando aplicável, o que é necessário para representar o objeto pretendido (ou seus elementos essenciais definidos), independente de mudanças nas tecnologias de armazenamento e acesso
- * registram o histórico e os efeitos dos fatos relativos ao objeto
- * documentam a identidade e a integridade dos objetos para garantir a autenticidade
- * permitem ao usuário e ao programa de preservação entender o contexto do objeto".

Existem dois pontos importantes a serem destacados na aplicação da estratégia de metadados. O primeiro diz respeito à definição e escolha do identificador único e permanente que deve ser associado ao objeto digital. Um identificador único, de ampla abrangência (de preferência universal), acompanha o objeto em todo o seu ciclo de vida, auxiliando na consolidação de sua autenticidade e proporcionando ao usuário a confiança de que está acessando a informação desejada (BULLOCK, 1999).

O segundo ponto diz respeito à forma de ligar os metadados ao conteúdo do objeto digital propriamente dito. Os metadados podem ser armazenados como parte integrante do objeto que descreve (p.ex., inseridos em um cabeçalho HTML) ou como parte de um arquivo de informação separado (p.ex., um registro MARC). Uma outra forma de ligar os metadados ao objeto digital é juntá-los em pacotes. O modelo de referência *Open Archival Information System* (OAIS), a ser descrito na segunda parte deste artigo, propõe um "pacote de informação" composto de "informação de conteúdo" e "informação de descrição de preservação". De forma similar, um grupo de trabalho da *Society of Motion Picture and Television Engineers* desenvolveu o *Universal Preservation Format* (UPF), um mecanismo de arquivo de dados que usa um container[5] para incorporar metadados dentro de objetos de mídias digitais. (BULLOCK, 1999)

5.1.4 Montagem de infra-estrutura para preservação digital

Uma instituição que decidiu assumir a responsabilidade de preservar objetos digitais por longo prazo só irá efetivamente concretizá-la através de uma infra-estrutura de *hardware*, *software* e pessoas, i.e. um *sistema de arquivamento digital* adequado a esta finalidade, tendo em vista seu acesso a futuras gerações. Alguns mecanismos são imprescindíveis para um ambiente operacional desta natureza, dentre os quais destaca-se:

- * sistema de cópia de segurança;
- * sistema de armazenamento redundante;
- * sistema de detecção e recuperação automática de falhas;
- * sistema de segurança de acesso físico e lógico;
- * sistema hierárquico de armazenamento.

Quanto ao aspecto humano, deve-se, obviamente, preparar um programa de formação e aperfeiçoamento para o pessoal que, em última instância irá operar e acessar o *sistema de arquivamento digital*. Este programa de capacitação deve considerar duas dimensões básicas: uma dimensão horizontal, dirigida às diferentes atividades ou funções desempenhadas, e uma dimensão vertical, com diferentes níveis de aprofundamento (básico, intermediário, avançado).

5.1.5 Formação de uma rede de relações

Um ambiente voltado para a preservação digital por longo prazo, para ser efetivo, deverá, certamente, organizar-se de diversas formas e, provavelmente, envolver corporações, federações, consórcios, i.e., uma rede distribuída de relações. Cada uma destas formações poderia, p.ex., especializar-se no arquivamento de um tipo de informação digital e organizar-se em fronteiras regionais ou nacionais. Tanto as cooperações informais (associações e alianças) quanto as parcerias formais entre contratantes e contratados poderiam, também, surgir, sendo as responsabilidades para o arquivamento distribuídas de acordo com interesses específicos como, p. ex., disciplina, tipo de informação, função arquivística. Outra especialização poderia advir de interesses na disseminação de informação dentro de fronteiras nacionais, levando ao projeto e implantação de instituições arquivísticas específicas. A utilização deste tipo de estratégia, provavelmente, exigirá das atuais instituições arquivísticas profundas adaptações estruturais. (TASK FORCE ON ARCHIVING OF DIGITAL INFORMATION, 1996)

5.2 Estratégias operacionais

5.2.1 Escolha do meio de armazenamento

Para uma boa preservação digital, o primeiro passo é a escolha adequada do meio de armazenamento da informação. Diversos fatores influenciam nessa escolha. As condições de acesso à informação irão, basicamente, determinar a disponibilidade (*online*, *near-line*, *off-line*) e a velocidade do dispositivo de leitura. Além disso, uma série de outras características físicas e lógicas devem ser avaliadas e ponderadas para se chegar a uma boa decisão. Dentre elas pode-se destacar: capacidade de leitura/gravação, capacidade de armazenamento, tamanho físico da mídia, nível de padronização, compatibilidade com o *hardware/software* já instalado, forma lógica de armazenamento, forma física de armazenamento, condições ambientais para armazenamento, durabilidade, robustez, tempo médio entre falhas, capacidade de detecção de falhas, capacidade de recuperação de falhas com interrupção/sem interrupção, presença de robótica, disponibilidade de suprimentos, nível de serviço de assistência técnica, facilidade de manuseio, preço por unidade de armazenamento (p.ex. Megabytes).

5.2.2 Migração

A migração é atualmente a estratégia mais utilizada pelas instituições que buscam preservar objetos digitais. Consiste de um conjunto de atividades para copiar, converter ou transferir, periodicamente, a informação digital existente em uma determinada geração de tecnologia para as gerações subsequentes. A migração significa copiar a informação digital de um suporte que está tornando-se obsoleto ou fisicamente deteriorado para um suporte mais novo (p.ex., um disco flexível para um CD-ROM) e/ou converter de um formato ultrapassado para outro mais atual (p.ex., Microsoft Word para ASCII) e/ou transferir documentos de uma plataforma de hardware/software em processo de descontinuidade para outra (p.ex., WINDOWS NT para LINUX). A migração preserva a presença física e o conteúdo do objeto digital. Entretanto, pode não preservar a apresentação, a funcionalidade e o contexto. Por exemplo, em uma migração sem um planejamento cuidadoso, pode ocorrer o desaparecimento de alguns elementos da apresentação. Além disso, a funcionalidade e o contexto envolvidos nas ligações entre entradas de banco de dados podem ser perdidos se estas ligações pararem de funcionar. Uma série de migrações sucessivas pode, eventualmente, resultar em perda inaceitável de dados. O foco desta estratégia é restringir a perda e manter o conteúdo de forma utilizável, prática.

As instituições arquivísticas têm um histórico de utilização bem sucedida de migração digital dentro de um contexto de informação relativamente simples e homogênea. As experiências com objetos complexos, entretanto, carecem de testes comprovados. (TASK FORCE ON ARCHIVING OF DIGITAL INFORMATION, 1996; ROTHENBERG, 1999; BULLOCK, 1999; CCSDS, 2002; NLA, 2003)

5.2.3 Emulação

A emulação refere-se à criação de novo *software* que imita o funcionamento do antigo *hardware* e/ou *software* para reproduzir seu comportamento. Dessa forma, não somente a presença física e o conteúdo são preservados mas os objetos digitais poderiam apresentar tanto as características originais (p.ex., *leiaute*) quanto a funcionalidade disponível no software anterior. Até o momento, a emulação tem sido usada para fornecer "compatibilidade retrospectiva" de jogos eletrônicos e para modelar futuros sistemas. Embora

existam emuladores para alguns sistemas obsoletos, a emulação para preservar objetos digitais por longo prazo ainda não foi suficientemente testada ou avaliada em termos de custo/preço final.

Mais recentemente, a emulação tem despertado a atenção como uma estratégia potencial para auxiliar a preservação, a partir do reconhecimento de que alguns materiais eletrônicos, altamente dependentes de determinado *hardware* e *software*, não se prestariam à migração. Apresenta, ainda, o apelo fascinante da conservação da "aparência e sensação" do ambiente original onde o documento era apresentado. (ROTHENBERG, 1999; BULLOCK, 1999; NLA, 2003)

5.2.4 Impressão em papel ou microfilme

Produzir uma cópia impressa de um arquivo digital é uma solução "*low tech*" que pode resultar em um produto bem padronizado com uma expectativa de vida de centenas de anos. Certamente, esta estratégia poderia fixar o objeto como um todo, preservar o conteúdo e, de certa forma, o leiaute. Entretanto, cada vez fica mais difícil transformar objetos digitais de natureza dinâmica e heterogênea em objetos fixos. Por exemplo, a impressão em papel para documentos hipertexto acarreta grande perda de funcionalidade e, para o caso de documentos multimídia, torna-se impraticável. Uma "estratégia híbrida"[6], com a produção de microfilme e cópias digitais, apesar de restrições por parte de alguns vanguardistas, está ganhando apoio como uma solução para reformatar originais em papel. A cópia digital melhora o acesso e a funcionalidade, enquanto a cópia em microfilme funciona como um substituto arquivístico. (BULLOCK, 1999; NLA, 2003)

5.2.5 Conservação da tecnologia

Outro método para garantir o acesso continuado aos objetos digitais seria simplesmente manter a tecnologia, que criou os objetos, disponível para uso. Embora isso preserve o conteúdo e permita que futuras gerações possam visualizar os objetos digitais em seu formato nativo com o leiaute e a funcionalidade originais, a criação de "museus" de hardware e software implicaria em requisitos de custo, espaço e suporte técnico impraticáveis. Na melhor das hipóteses, este método é uma medida transitória, enquanto não for possível a migração. (BULLOCK, 1999; NLA, 2003)

A seguir

A primeira parte deste artigo concentrou-se na problemática da preservação digital e principais estratégias propostas. A segunda parte irá descrever o modelo de referência *Open Archival Information System* (OAIS) como a proposta mais promissora para padronização de um sistema de arquivamento dedicado à preservação e acesso à informação digital por longo prazo.

PARTE II

1. Introdução

Ainda que se reconheça nas organizações alguns movimentos no sentido da preservação digital, materializados através da utilização de algumas estratégias específicas, não se pode afirmar ainda que as instituições, em geral, estejam preparadas para enfrentar os riscos associados à perda de objetos digitais que necessitem de preservação por longo prazo.

Uma alternativa recente, importante para a discussão desta problemática na ciência da informação, tem obtido boa repercussão internacional. Trata-se do modelo de referência *Open Archival Information System* (OAIS) ou Sistema Aberto para Arquivamento de Informação (SAAI)[7], como passamos a tratá-lo de agora em diante, do *Consultative Committee for Space Data Systems* (CCSDS). Este modelo é um esquema conceitual que disciplina e orienta um sistema para a preservação e manutenção do acesso à informação digital por longo prazo. O objetivo do modelo é ampliar a consciência e a compreensão dos conceitos relevantes para a preservação de objetos digitais, especialmente entre instituições não arquivísticas; definir terminologias e conceitos para descrever e comparar modelos de dados e arquiteturas de arquivos; ampliar o consenso sobre os elementos e os processos relacionados à preservação e acesso à informação digital; e criar um esquema para orientar a identificação e o desenvolvimento de padrões.

Embora o SAAI seja suficientemente genérico para abranger o arquivamento de objetos físicos e digitais, é no contexto destes últimos que extrai sua vocação. O modelo de referência tem sido bem recebido por diferentes comunidades interessadas na preservação da informação digital por longo prazo. Diversas iniciativas, de diferentes países, têm adotado o modelo conceitual, seja como um esquema básico para seus esforços de preservação digital, seja como uma ferramenta para a melhoria da comunicação e produtividade entre diferentes comunidades[8].

Esta segunda parte do artigo concentra-se nas duas principais fontes de informação sobre o modelo, quais sejam: o artigo de Brian Lavoie, *Meeting the challenges of digital preservation: The OAI reference model*, que fez a divulgação do primeiro esboço do modelo no ano de 2000, e a recomendação CCSDS 650.0-B-1 *Reference Model for an Open Archival Information System* (OAIS), que especifica o modelo em detalhes, com a versão final aprovada em 2002.

2. Antecedentes

A pesquisa no campo da preservação digital se avolumou nos anos 1990. A *International Organization for Standardization* (ISO)[9], sensível à nova necessidade, iniciou um processo de identificação de uma instituição de reconhecida competência para liderar o desenvolvimento de padrões e recomendações para o campo. Paralelamente, o *Consultative Committee for Space Data Systems* (CCSDS)[10] promovia negociações com o *Subcommittee 13 (space data and information transfer systems)* sob o *Technical Committee 20 (aircraft and space vehicles)* da ISO para que suas recomendações passassem pelo sistema normal de revisão e votação e, eventualmente, evoluíssem para padrões ISO. A ISO, percebendo então o potencial do CCSDS[11], convidou-o para coordenar o desenvolvimento de padrões para o armazenamento de informação digital por longo prazo. A partir daí, em 1995, foram firmadas parcerias com o *Federal Geographic Data Committee* (FGDC) e o *National Archives and Records Administration* (NARA) e, com o objetivo de tornar o processo mais aberto e participativo, foram convidados diversos representantes de instituições arquivísticas e usuários de diferentes países. Foram conduzidas oficinas internacionais de trabalho nos EUA e Europa (Inglaterra e França) tendo seus resultados sido divulgados na Internet. Decidiu-se que o ponto de partida deste processo seria o desenvolvimento de um modelo de referência para estabelecer conceitos e termos comuns, fornecer um esquema para esclarecer as entidades significativas e relacionamentos entre entidades em um ambiente de arquivos e servir como base para o desenvolvimento de padrões de apoio ao ambiente de arquivos. Os esforços do CCSDS resultaram na publicação da primeira versão do *Reference Model for an Open Archival Information System* (OAIS) - *RED BOOK* em maio de 1999, da segunda versão do *RED BOOK* em julho de 2001, da primeira versão do *BLUE BOOK* em janeiro de 2002, e aprovação da ISO 14721 *Space data and information transfer systems - Open archival information system - Reference model* em fevereiro 2003.

Observa-se, assim, que o modelo de referência SAAI surge de uma discussão aberta com a participação de diversas comunidades e com o referendo de duas importantes instituições de abrangência internacional e com larga experiência bem sucedida no desenvolvimento de padrões.

3. Sistema Aberto para Arquivamento de Informação (SAAI)

Para fins do modelo de referência, um SAAI é um tipo específico de arquivo que: 1) consiste de uma organização de pessoas e sistemas, que assumiu a responsabilidade de preservar informação e torná-la acessível a uma classe de usuários definida como Comunidade Alvo; e 2) atende a um conjunto de responsabilidades que o distingue do uso puro da palavra 'arquivo'. O adjetivo "aberto", em SAAI, refere-se ao fato do modelo e as futuras recomendações associadas serem desenvolvidos em fóruns abertos, não guardando nenhuma relação com o nível de acessibilidade da informação arquivada.

As responsabilidades que um arquivo tipo SAAI deve cumprir são as seguintes:

* negociar e aceitar informação adequada de produtores de informação;

- * manter o efetivo controle da informação para garantir a preservação por longo prazo;
- * determinar, por si mesmo ou em conjunto com outros parceiros, que comunidades devem tornar-se Comunidade Alvo e, portanto, devem ser capazes de entender a informação fornecida;
- * garantir que a informação seja compreensível para a Comunidade Alvo sem o auxílio dos produtores de informação;
- * seguir políticas e procedimentos documentados garantindo que a informação seja preservada contra todas as contingências cabíveis e possibilitando que a mesma seja disseminada como cópias autênticas do original ou rastreável até o original;
- * tornar a informação preservada disponível para a Comunidade Alvo.

O esquema SAAI detalha a arquitetura conceitual de um arquivo, no que diz respeito a seu ambiente, suas informações e suas funções para suportar estas responsabilidades.

4 O ambiente SAAI

A primeira preocupação é entender o contexto de um arquivo tipo SAAI, i.e, quais as suas interações com o ambiente no qual se insere. A FIG. 1 ilustra este ambiente.



FIGURA 1 - O AMBIENTE SAAI

Quatro entidades participam do ambiente SAAI: Produtores, Consumidores, Administração e o Arquivo propriamente dito. Os Produtores fornecem a informação a ser preservada. Os Consumidores usam a informação preservada. Uma classe especial de Consumidores é a Comunidade Alvo, i.e., o subconjunto de Consumidores que deve entender a informação preservada. A Administração é a entidade responsável pelo estabelecimento das políticas mais gerais do arquivo (a Administração não se envolve, p.ex., com a rotina do Arquivo; esta atividade é desempenhada por uma função dentro do próprio Arquivo).

5. O modelo de informação SAAI

O gerenciamento efetivo de todas as formas de preservação digital tende a ser facilitado pela criação, manutenção e evolução de metadados detalhados do arquivo. Estes metadados podem, p.ex., documentar os processos técnicos associados à preservação, especificar direitos de acesso e estabelecer a autenticidade do conteúdo digital. Podem registrar a cadeia de custódia de um objeto digital e identificá-lo de forma única, interna e externamente, em relação ao arquivo ao qual pertence. De fato, a criação e o desdobramento de metadados de preservação tende a ser um componente chave para a maioria das estratégias de preservação digital.

O esquema SAAI, seguindo esta tendência, inclui um modelo de informação para inserção dos metadados para preservação de informação digital por longo prazo, apresentado de forma simplificada na FIG.2 .

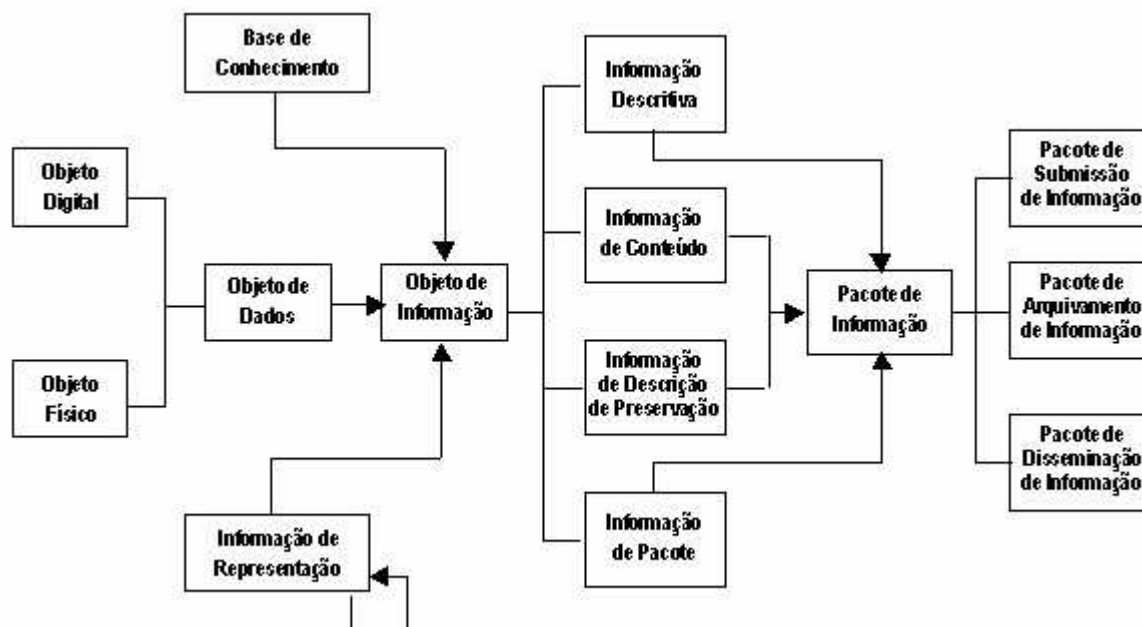


FIGURA 2 - O MODELO DE INFORMAÇÃO SAAI

No modelo, a informação, entendida como qualquer forma de conhecimento passível de intercâmbio por meio de dados, pode existir de duas formas: como um Objeto Físico (p.ex., um documento em papel, uma amostra de solo) ou como um Objeto Digital (p.ex., um arquivo PDF, um arquivo TIFF). Estes dois tipos - físico ou digital - podem ser, coletivamente, referenciados como Objetos de Dados.

A interpretação do Objeto de Dados como informação significativa pela Comunidade Alvo do arquivo é obtida através da combinação da Base de Conhecimento da Comunidade Alvo e a Informação de Representação associada ao Objeto de Dados. Cada indivíduo (ou classe de indivíduos, no caso de uma Comunidade Alvo) tem uma Base de Conhecimento, que é usada para entender e interpretar a informação. Por exemplo, uma Comunidade Alvo constituída de programadores Java deve ter uma Base de Conhecimento para entender a informação em formato de código fonte Java.

A Base de Conhecimento da Comunidade Alvo nem sempre é suficiente para entender integralmente a informação preservada. Nesse caso, o Objeto de Dados deve ser complementado com Informação de Representação, de forma que possa ser plenamente entendida pela Comunidade Alvo. Por exemplo, se a Comunidade Alvo abrange todos os tipos de programadores, e não somente programadores Java como no exemplo anterior, é necessária a informação referente à sintaxe e às convenções de programação Java para que essa classe de Consumidores entenda completamente o Objeto de Dados preservado (código fonte Java). É interessante observar que a Informação de Representação pode ainda necessitar de mais Informação de Representação para sua própria interpretação. Em princípio, essa recursividade continua até que sejam encontrados formatos físicos, totalmente compreensíveis pela Base de Conhecimento da Comunidade Alvo. Por exemplo, a Informação de Representação expressa em ASCII necessita de Informação de Representação adicional para ASCII, que poderia ser um documento físico explicando o padrão ASCII. Quando isso ocorre, o conjunto resultante de objetos de Informação de Representação é referenciado como uma Rede de Representação.

O Objeto de Dados, a Base de Conhecimento da Comunidade Alvo e a Informação de Representação, quando combinados, formam um Objeto de Informação que representa uma "informação significativa" para a Comunidade Alvo. O significado, obviamente, está relacionado com a definição da Comunidade Alvo atendida pelo arquivo.

Um Objeto de Informação pode ser de quatro tipos: Informação de Conteúdo, Informação de Descrição de Preservação, Informação de Pacote e Informação Descritiva. A Informação de Conteúdo é a informação principal, alvo da preservação associada à sua Informação de Representação. A Informação de Descrição de Preservação (IDP) contém a informação necessária para preservar adequadamente a Informação de

Conteúdo à qual está associada, podendo ser decomposta em quatro subcategorias a saber: Referência (descritores), Contexto (relacionamentos com o ambiente), Proveniência (proveniência e histórico) e Rigidez (informação para a comprovação de integridade e autenticidade)[6]. A Informação de Pacote reúne a Informação de Conteúdo e a Informação de Descrição de Preservação em um pacote identificável, enquanto a Informação Descritiva facilita o acesso à Informação de Pacote através de ferramentas de pesquisa e recuperação.

Dentro do esquema SAAI, são identificados três tipos de pacote de informação: Pacote de Submissão de Informação (PSI), pacote enviado do Produtor para o Arquivo; Pacote de Arquivamento de Informação (PAI), pacote de informação efetivamente armazenado dentro do Arquivo; e Pacote de Disseminação de Informação (PDI), pacote transferido do Arquivo para um Consumidor em resposta a uma solicitação.

6. O modelo funcional SAAI

O esquema SAAI inclui um modelo funcional, conforme FIG. 3. Em um primeiro nível de detalhamento, são identificadas seis entidades funcionais: Recepção, Armazenamento, Gerenciamento de Dados, Administração do Sistema, Planejamento de Preservação e Acesso.

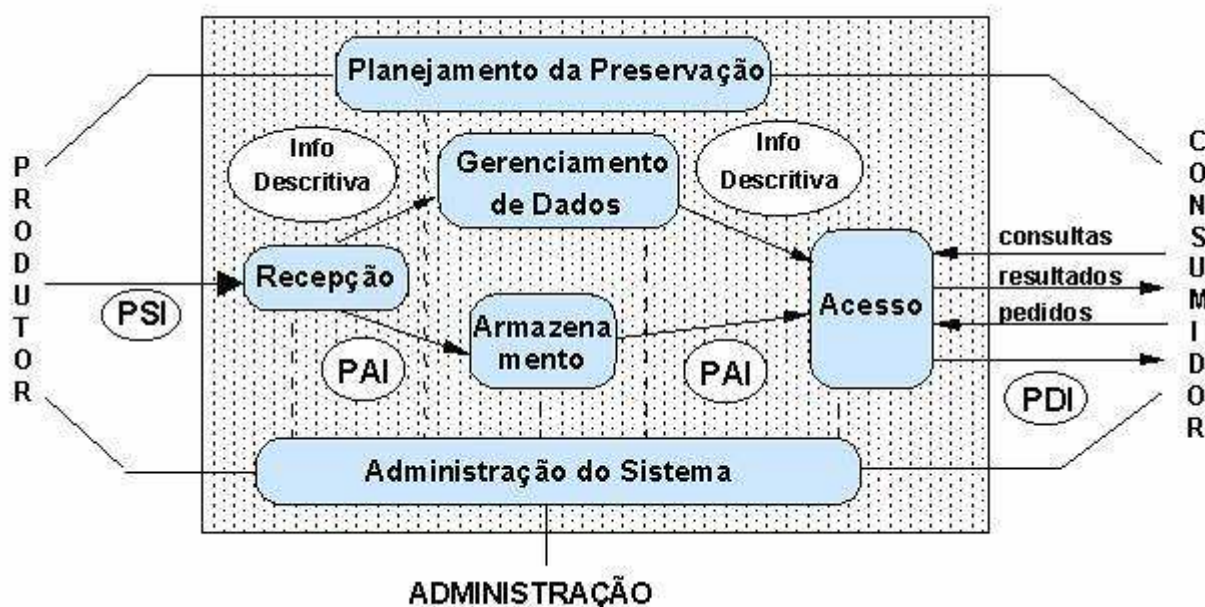


FIGURA 3 - O MODELO FUNCIONAL SAAI

A entidade **Recepção** é responsável pela aceitação dos Pacotes de Submissão de Informação (PSI's) dos Produtores (ou de componentes internos sob controle da Administração do Sistema) e preparação dos conteúdos para armazenamento e gerenciamento dentro do arquivo. Mais especificamente, a entidade Recepção recebe PSI's, verifica a qualidade dos PSI's, gera Pacotes de Arquivamento de Informação (PAI's), de acordo com a formatação de dados e padrões de documentação do arquivo, e gera Informação Descritiva dos PAI's (metadados para pesquisa e recuperação, ícones para navegação, etc). Finalmente, transfere os recém criados PAI's e Informações Descritivas associadas para a entidade Armazenamento e para a entidade Gerenciamento de Dados, respectivamente.

A entidade **Armazenamento** lida com o armazenamento, manutenção e recuperação de PAI's. Estas responsabilidades envolvem receber novos PAI's da entidade Recepção e posicioná-los na área de armazenamento permanente de acordo com diversos critérios (requisitos de suporte, taxas de utilização esperada, etc), gerenciar a hierarquia da área de armazenamento, renovar as mídias, executar rotinas de verificação de erro, oferecer capacidade de recuperação de falha e fornecer cópias de PAI's solicitados à entidade Acesso.

A entidade **Gerenciamento de Dados** mantém e acessa tanto a Informação Descritiva, que identifica e documenta os acervos do arquivo, quanto os dados administrativos usados para gerenciá-lo.

Especificamente, a entidade Gerenciamento de Dados administra a base de dados do arquivo (mantém os esquemas e definições de visões e integridade referencial), promove suas atualizações (carrega nova informação descritiva ou dados administrativos do arquivo) e consulta os dados da entidade para gerar relatórios.

A entidade **Administração do Sistema** gerencia a rotina operacional do arquivo como um todo. Suas funções incluem solicitar e negociar acordos de submissão com Produtores, auditar as submissões para garantir que estão atendendo aos padrões do arquivo e gerenciar a configuração do hardware e software do sistema. A entidade Administração do Sistema desempenha, também, funções mais técnicas para análise e melhoria do desempenho geral das operações do arquivo e migração/atualização de seus conteúdos. É responsável, ainda, por cumprir e manter os padrões e políticas do arquivo, fornecer suporte ao cliente e atender solicitações pendentes.

A entidade **Planejamento de Preservação** monitora o ambiente SAAI e fornece recomendações para garantir que a informação armazenada permaneça acessível por longo prazo à Comunidade Usuária Alvo, mesmo que o ambiente computacional original torne-se obsoleto. Suas funções incluem avaliar os conteúdos do arquivo e recomendar, periodicamente, migrações da informação arquivada, desenvolver recomendações para padrões e políticas do arquivo e monitorar as mudanças no ambiente tecnológico, nas demandas de serviço e na Base de Conhecimento da Comunidade Alvo. A entidade Planejamento de Preservação desenha, também, modelos de pacotes de informação e faz adaptações de PAI's e PSI's para submissões específicas. É responsável, ainda, pelo desenvolvimento de planos detalhados de migração, protótipos de software, e planos de teste para permitir a implementação das metas de migração da entidade Administração do Sistema.

A entidade **Acesso** apóia os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no SAAI e permite que os Consumidores solicitem e recebam produtos de informação. Suas funções incluem comunicar com os Consumidores para receber solicitações, aplicar controles para limitar o acesso (principalmente à informação protegida), coordenar a execução de solicitações para que se completem com sucesso, gerar respostas (Pacotes de Disseminação de Informação, resultados, relatórios) e entregar as respostas aos Consumidores.

As seis entidades funcionais SAAI gerenciam o fluxo de informação entre os **Produtores** e o **Arquivo** e entre o **Arquivo** e os **Consumidores**. Vistas juntas, identificam os processos chave típicos da maioria dos arquivos dedicados à preservação de informação digital. Um arquivo digital deverá, provavelmente, conter componentes funcionais similares àqueles descritos acima, embora cada implementação específica tenha suas peculiaridades.

7. Conclusões

Como vimos na primeira parte deste artigo, o crescente volume de informação em formato digital não permite mais que as instituições mantenham um posicionamento conservador ou indiferente quanto à adoção de estratégias para a preservação digital. O cenário de relativa estabilidade alcançado com a preservação de informação na “era papel”, está definitivamente ameaçado pela propagação da nova informação digital através de *hardware*, *software* e mídia, com ciclos de vida cada vez menores.

Para uma instituição que se dispõe a preservar informação digital por longo prazo, duas constatações surgem deste cenário:

* Necessidade de, tomando emprestada as palavras de Lusenet (2002), transformar o antigo procedimento de *preservar* o objeto para o acesso para um novo procedimento de *preservar* o acesso ao objeto.

* O acompanhamento da evolução tecnológica como fator crítico de sucesso para a preservação digital. Uma vez que a tecnologia evolui, a intervalos cada vez menores, e juntamente com ela, toda a infra-estrutura de suporte da informação a preservar, é preciso perceber e agir imediatamente a cada alteração de um componente do sistema de arquivamento digital. Não se trata de escolher essa ou aquela corrente ou linha tecnológica. O que realmente importa é a adoção de uma filosofia de monitoramento tecnológico continuado, tornando a mudança aliada da preservação.

Há, ainda, outra questão a considerar nas decisões sobre preservação digital: o *trade-off* entre o desejado, do ponto de vista da funcionalidade, dependência tecnológica e custo, e o possível, com relação às tecnologias e metodologias correntes.

Todas as questões, ligadas à conservação do conteúdo digital por longo prazo, como sua estrutura, semântica, referências, contexto, proveniência, autenticidade e integridade de objetos digitais suscitaram, nos últimos anos, a discussão da sistemática mais adequada para preservação digital. Inicialmente, algumas iniciativas tiveram uma abordagem localizada, sendo direcionadas, em sua grande maioria, para comunidades específicas.

O modelo de referência SAAI tenta preencher esta lacuna, na medida em que integra os diversos esforços em torno de um sistema conceitual padrão que servirá de base ou linguagem comum para futuras iniciativas quanto ao arquivamento de informação digital. É bom lembrar que o modelo teve ampla participação de diversas comunidades e países, além do referendo de duas conceituadas instituições de âmbito internacional com larga experiência bem sucedida no desenvolvimento de padrões.

O desenvolvimento de outros padrões, em apoio ao modelo de referência SAAI, poderá servir para promover a interoperabilidade entre bibliotecas, arquivos e outras instituições que mantêm informação digital por longo prazo. Com isso, instituições que no passado não viam oportunidades para o desenvolvimento da preservação digital, ou não tinham meios para explorá-las, passarão a contar com recursos que facilitam a promoção de esforços cooperativos.

A adoção do esquema SAAI poderá, ainda, produzir benefícios econômicos, uma vez que a padronização de entidades e processos comuns favorece a redução de custos, através do compartilhamento de componentes de sistemas. Além disso, a padronização promove o desenvolvimento do mercado fornecedor, transformando produtos e serviços altamente personalizados em versões padronizadas menos onerosas.

Ainda não está claro se a iniciativa SAAI será selecionada como a abordagem padrão para a preservação por longo prazo. Mas, provavelmente, qualquer outra solução que venha a ser adotada seguirá um caminho similar, com ênfase na ampla participação em fóruns abertos. O modelo SAAI está, no mínimo, construindo bases importantes para uma solução coordenada e, sem dúvida, perfeitamente alinhada aos desafios da preservação digital.

Notas

[1]] Hedstrom opta pela expressão "valor contínuo" ao invés de "valor permanente" para evitar - segundo O'TOOLE, James M. *On the Idea of Permanence*. Washington, D.C: Commission on Preservation and Access, 1993 - o absolutismo e o idealismo que o termo permanente envolve.

[2] Documentos "nato-digitais" são documentos produzidos originalmente em formato digital, em contraposição aos documentos produzidos originalmente em formato tradicional e que, por razões diversas, são convertidos posteriormente para o formato digital através do processo de digitalização.

[3] Requisitos são objetivos e restrições, estabelecidos pelas partes envolvidas, que definem as propriedades de um sistema ou área de assunto. Os requisitos podem ser funcionais e não funcionais. Os requisitos funcionais descrevem as funções ou processos a serem desempenhados e os requisitos não-funcionais descrevem as qualidades globais, a partir de dimensões básicas sugeridas pelo Controle de Qualidade Total:

qualidade intrínseca do produto, custo/preço do produto, entrega/atendimento do produto, moral das pessoas, segurança das pessoas e impacto ambiental.

[4] No caso específico da Austrália, uma de suas iniciativas, o projeto *Strategic Partnerships with Industry Research & Training* (SPIRT), conduzido pela Universidade de Monash no período 1998/1999, desdobrou-se na norma australiana AS 4390 Records Management, posteriormente gerando a ISO 15489 Records Management cuja primeira edição foi liberada em 15/09/2001. Outra contribuição mais recente foi o manual “*Guidelines for the Digital Heritage*”, preparado pela *National Library of Austrália*, editado em março de 2003 e disseminado através do sítio da Unesco <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>.

[5] Um arquivo contendo objetos vinculados incorporados (Microsoft PRESS, 1998).

[6] Uma estratégia híbrida envolve a utilização de suportes convencionais (papel, microfilme, etc) e digitais (disquete, CD-ROM, fita magnética, etc).

[7] Encontra-se em fase de elaboração projeto para tradução e adaptação do documento CCSDS 650.0-B-1, cuja responsável é autora deste artigo, na CE 08:001.06 Sistema Espaciais de Dados e Informações da Associação Brasileira de Normas Técnicas (ABNT), previsto para liberação para revisão em março/2004.

[8] Para maiores informações, o Research Libraries Group (RLG) estabeleceu uma página na Internet para acompanhar as iniciativas e implementações do modelo OAIS em <http://www.rlg.org/longterm/oais.html>.

[9] A *International Organization for Standardization* (ISO) é uma rede de instituições de padronização nacionais de 148 países - um membro por país -, com uma Secretaria Central em Genebra na Suíça. Foi estabelecida, oficialmente, em 1947 com o objetivo de facilitar a coordenação e a unificação internacional dos padrões da indústria.

[10] O *Consultative Committee for Space Data Systems* (CCSDS) foi estabelecido em 1982 para servir como um fórum para agências espaciais interessadas no desenvolvimento cooperativo de padrões para tratamento de dados em apoio à pesquisa espacial. O Comitê se reúne periodicamente para tratar problemas de sistemas de dados comuns a todos os participantes e formular soluções técnicas adequadas a esses problemas. Uma vez que a participação no CCSDS é totalmente voluntária, os resultados destas ações são denominados Recomendações que não ficam ligadas a nenhuma Agência específica. Até o momento da produção deste artigo, duas instituições brasileiras participavam do CCSDS: o Instituto Nacional de Pesquisas Espaciais (INPE), como agência membro e o Centro Técnico Aeroespacial (CTA), como agência observadora.

[11] A *National Aeronautics and Space Administration* (NASA), principal agência membro do CCSDS, que vem acumulando objetos digitais desde 1966 como resultado de suas experiências espaciais, tem passado por diversas mudanças tecnológicas. Os volumes de informações nessa área de conhecimento são expressivos com um intenso intercâmbio entre países de culturas bastante diversificadas.

[12] Para fins de simplificação, estas subcategorias não foram retratadas no modelo de informação apresentado na FIG. 2.

Referencias Bibliográficas

BELLOTO, Heloísa L. *Arquivos permanentes*; tratamento documental. São Paulo:, T.A. Queiroz, 1991. 198p.

BRAND, Stewart. Scaping the digital dark. *Library Journal*, v. 124, n. 2, p. 46-8, feb. 1999.

BULLOCK, Alison. *Preservation of digital information*; issues and current status. April 22, 1999. Last updated on February 27, 2001. Disponível em: <<http://www.nlc-bnc.ca/publications/1/p1-259-e.html>>. Acesso em: 25 nov. 2003.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). *Reference model for an Open Archival Information System (OAIS)*; recommendation. Washington, 2002. 139p. Disponível em: <http://www.ccsds.org/documents/650x0b1.pdf> Acesso em: 14 nov. 2002.

CONWAY, Paul. *Preservation in the digital world*. March 1996. Disponível em: <<http://www.clir.org/pubs/reports/conway2/>>. Acesso em: 23 nov. 2003.

DURANTI, Luciana. Merging the old with the new; the management of electronic records in Canada in theory and practice. In: SEMINÁRIO INTERNACIONAL DOS ARQUIVOS DE TRADIÇÃO IBÉRICA, 2., 2002, Rio de Janeiro. *Anais eletrônicos...* Rio de Janeiro: Arquivo Nacional, 2002. Disponível em: <<http://www.arquivonacional.gov.br>>. Acesso em: 16 nov. 2003.

ENTULHO digital. *Veja*, v. 33, n. 23, p. 105, jun. 2000.

HEDSTROM, Margaret. Digital preservation; a time bomb for digital libraries. *Computer and the Humanities*, v.31, n.3, 1997/1998, p.189-202. Disponível em: <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>. Acesso em: 25 nov. 2003.

INGWERSEN, Peter. Information and information science. *Encyclopedia of Library and Information Science*, v. 56, n. 19, p. 137-174, 1995.

INTO the future. Direção de Terry Sanders. Santa Monica, CA: American Film Foundation, 1997. 33 min. color. legendado. (Fita vídeo VHS).

JARDIM, José Maria. A arquivologia e as novas tecnologias da informação. *Estudos Históricos*, Rio de Janeiro, v. 5, n. 10, p. 251-60, 1992.

LOPES, Pedro F., CARDOSO, Gustavo, MOREIRA, Maria V. Preservação de publicações electrónicas na internet; os arquivos imperfeitos. *Cadernos BAD*, V.2, p. 33-53, 2002.

LAVOIE, Brian. *Meeting the challenges of digital preservation*; the OAIS reference model. OCLC-Newsletter. n. 243. jan/feb. 2000, p. 26-30.

LUSENET, Yola. Digital heritage for the future. *Cadernos BAD*, v. 2, p. 15-27, 2002.

MICROSOFT PRESS. *Dicionário de informática*. Rio de Janeiro: Campus, 1998. 804p.

NATIONAL LIBRARY OF AUSTRALIA (NLA). *Guidelines for the preservation of digital heritage*. Paris: UNESCO, 2003. 177p. Disponível em: <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>. Acesso em: 25 nov. 2003.

ROTHENBERG, Jeff. *Ensuring the longevity of digital information*. Revision February, 22 1999. Disponível em: <<http://www.clir.org/pubs/archives/ensuring.pdf>>. Acesso em: 26 nov. 2003.

ROUSSEAU, Jean-Ives, COUTURE, Carol. *Os fundamentos da disciplina arquivística*. Lisboa: Publicações Dom Quixote, 1998. 356p.

TASK FORCE ON THE ARCHIVING OF DIGITAL INFORMATION. *Preserving digital information*; report of the Task Force on Archiving of Digital Information; commissioned by the Commission on Preservation - CPA and Access and the Research Libraries Group - RLG. Washington, D.C.: Commission on Preservation and Access, 1996. Disponível em: <<http://www.rlg.org/ArchTF>>. Acesso em: 7 nov. 2003.

THIBODEAU, Kenneth. Building the archives of the future. *D-Lib Magazine*, v.7, n.2, p. 1-13, feb. 2001.

YAKEL, Elizabeth. Digital preservation. In: WILLIAMS, Martha E. (Ed.) *Annual Review of Information Science and Technology*. v. 35. p. 337-378. Medford, N.J: Information Today Inc., 2001.

Sobre os autores / About the Authors:

Katia P. Thomaz

katia.thomaz@uol.com.br

Bacharel em Ciência da Computação pela UFMG

Especialização em Engenharia da Qualidade pela PUCMG

Certified Document Imaging Architect - CDIA™ pela CompTIA

Certified Project Management Professional - PMP™ pelo PMI

Membro da CE 21:102.10 Metadados, da Associação Brasileira de Normas Técnicas (ABNT)

Membro da CE 08:001.06 Sistemas Espaciais de Dados e Informações, da Associação Brasileira de Normas Técnicas (ABNT), responsável pela tradução e adaptação da recomendação CCSDS 650.0-B-1 Reference Model for an Open Archival Information System (OAIS) para norma brasileira

Aluna do Programa de Pós-Graduação em Ciência da Informação - nível Doutorado, da Escola de Ciência da Informação da Universidade Federal de Minas Gerais

Antonio José Soares

ajsoares@uai.com.br

Bacharel em Engenharia Elétrica pela UFMG

Especialização em Análise de Serviços para Computadores pela IBM do Brasil Ltda

Certified Project Management Professional - PMP™ pelo PMI

Diretor de Gestão de Projetos da Sucesu/MG gestão 2002/2003

Sócio-Diretor da Tech-in Gestão, Projetos e Tecnologia

Professor de Gestão de Projetos em nível de especialização do IETEC – Instituto de Educação Tecnológica, da PUC/IEC - Instituto de Educação Continuada, da UFF - Universidade Federal Fluminense e da FDC – Fundação Dom Cabral

Presidente do Capítulo Minas Gerais do Project Management Institute - PMI™, gestão 2002/2003