

Os metadados como recurso à preservação digital do patrimônio documental

Denise Frigo - denisefrigo@yahoo.com.br- UFSM

Daniel Flores - dflores@gmail.com- UFSM

Resumo

Essa pesquisa propõe analisar a importância dos metadados como recurso para preservação do patrimônio documental. Explanando sobre a preservação digital no contexto do patrimônio documental e apresentando os aspectos teóricos que caracterizam os metadados. O estudo para definir-se a importância dos metadados, mostra-se de fundamental valor, além disso, destaca-se a escassa produção literária existente. Os procedimentos metodológicos foram baseados na consulta a obras, artigos e materiais afins sobre o tema de estudo, através da Internet, em bases de dados das áreas científicas estudadas na pesquisa. Contudo, os projetos analisados neste trabalho e os autores têm defendido a abordagem de metadados para a preservação do patrimônio documental por considerar mais adequada à nova realidade do alto volume de informação e complexidade do ambiente eletrônico.

Palavras-Chave: preservação digital, patrimônio documental, metadados

Introdução

Como parte da construção cultural da realidade, a informação reflete as visões do mundo de uma sociedade em um determinado tempo. No século XX as novas tecnologias de informação colocaram em evidência o papel dos meios e técnicas de gestão da informação digital, pois é fato que na era digital se está dando muita ênfase à geração ou aquisição de material digital, em vez de manter a preservação e o acesso em longo prazo aos acervos eletrônicos existentes.

Atualmente há um despertar da nossa sociedade pela busca de soluções e medidas simples para salvaguardar adequadamente os nossos bens culturais. A era da informação valorizou ainda mais os dados vitais e estratégicos que precisam ser preservados, divulgados e acessados rapidamente para uso presente e futuro. Então, nos deparamos com danos ou perdas irreparáveis dos acervos, somente a partir daí percebemos a importância da manutenção desses para a continuidade da memória do patrimônio histórico e cultural da nação. (SICHMANN, 2003, p.6)

Sendo assim a natureza dos documentos eletrônicos está permitindo ampla produção e disseminação de informações no mundo atual. Entretanto, ocorre a ênfase à

geração e/ou aquisição de material eletrônico, em vez de manter a preservação e o acesso a longo prazo dos acervos existentes. Com isso, o propósito desta investigação é analisar a importância dos metadados como recurso para preservação do patrimônio documental.

Arellano (2008, p. 20) relata que “a preservação digital é uma preocupação dos profissionais da informação, devido principalmente ao uso extensivo que tem sido feito das tecnologias digitais e do registro de informações em soluções cada vez mais sofisticadas e complexas”.

Por isso, a aplicação de estratégias de preservação digital do patrimônio documental é uma prioridade, pois sem elas não existiria nenhuma garantia de acesso, confiabilidade e integridade dos documentos em longo prazo.

Para os detentores de acervos, 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 os quais o desenvolvimento de padrões legais para lidar com arquivos eletrônicos precisa de estratégias metodológicas bem definidas. (ARELLANO, 2004, p.16)

Portanto, nesse contexto de registros e na falta de padronização que possa ajudar no tratamento desses dados. A pesquisa a ser relatada pretende contribuir para as investigações das transformações que os sistemas de informação vêm enfrentando.

Preservação digital: uma abordagem conceitual

A informação digital tem um lado frágil e fugaz, necessitando de gerenciamento cuidadoso desde o momento da sua criação, e merece uma abordagem política e estratégica, por parte da unidade de informação, de modo a garantir a preservação.

Os documentos digitais trouxeram uma série de vantagens na produção, transmissão, armazenamento e acesso que, por sua vez, acarretaram alguns problemas. A facilidade de criar e transmitir documentos traz como consequência a informalidade na linguagem, nos procedimentos administrativos, bem como o esvaziamento das posições hierárquicas.

A facilidade de acesso pode acarretar intervenções não autorizadas que podem resultar na adulteração ou perda dos documentos. A rápida obsolescência tecnológica (software, hardware e formatos) e a degradação das mídias digitais dificultam a preservação de longo prazo dos documentos e sua acessibilidade contínua. Os problemas em questão tornam necessária a

adoção de medidas preventivas para minimizá-los. (ARQUIVO NACIONAL, 2006, p.16)

Para BOERES; ARELLANO (2005, p. 2) a preservação digital é um processo distribuído que envolve o planejamento, alocação de recursos e aplicação de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável. Neste trabalho considera-se esta definição a mais completa, uma vez que privilegia a questão do “valor” da informação digital a ser preservada.

Muitas vezes os documentos eletrônicos necessitam de uma constante atualização de suporte de formato, além de estratégias para possibilitar a recuperação das informações, que passam pela preservação da plataforma de hardware e software em que foram criados, pela migração ou emulação. Estas são algumas iniciativas que vêm sendo tomadas, mas que não são ainda respostas definitivas para o problema da preservação de longo prazo. Não há soluções únicas e todas elas exigem investimento financeiro e contínuo em infra-estrutura tecnológica, pesquisa científica e capacitação de recursos humanos.

“a preservação digital vem então se tornando um imperativo pela agilidade de expansão, e um desafio, uma vez que o processo de preservar a informação por longo período e mantê-la acessível exige um grande e contínuo esforço. Esforço esse, justificado pela fragilidade da mídia, pela volatilidade tecnológica, pela necessidade de provas documentais para se interpretar o passado e por perdas já sofridas pela humanidade.”(LIMA, 2007, p.19)

Os principais agentes causadores de danos sobre a fragilidade do suporte digital são causados pela temperatura, umidade relativa do ar, nível de poluição do ar e das ameaças biológicas, manipulação das mídias, tempo de uso e qualidade da mídia, falhas no processo de fabricação, as catástrofes naturais e a obsolescência tecnológica. Vale ressaltar que grande parte dos princípios centrais da preservação, desenvolvidos e adotados no universo analógico, podem ser levados.

O tema da preservação dos documentos digitais está cada mais presente entre os criadores e gerenciadores de coleções digitais. Para aqueles que começam a enfrentar a possibilidade de não poder recuperar informações produzidas há 5 ou 10 anos, é vital contar com estratégias que garantam a persistência desses dados. (Arellano 2008, p. 23)

Outro ponto relativo á preservação digital, que tem preocupado a muitos, é a autenticidade dos dados, pois os usuários precisam ter certeza de que informação que estão utilizando é original e não foi alterada nem por outros usuários, nem por alguma atualização de dados, não prevista.

Em um contexto geral, sobre preservação digital Ferreira (2006, p. 32), explica que consiste na capacidade de garantir que a informação digital esteja acessível e com autenticidade no futuro, utilizando-se plataformas tecnológicas diferentes das utilizadas no momento de sua criação.

Hedstrom (1996 apud SANT'ANNA, 2008, p.07) considera que preservação é assegurar proteção à informação de valor permanente para acesso pelas gerações presentes e futuras [...] planejamento, alocação de recursos e aplicações de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável.

Sendo assim, a preservação digital consiste em estabelecer meios para a manutenção do acesso a objetos informacionais em longo prazo. Esses meios precisam ser definidos de forma que obstáculos conhecidos sejam superados ou, pelo menos, amenizados.

No contexto digital, as atividades de preservação podem ser divididas em dois grupos: os cuidados físicos, incluindo procedimentos aos suportes físicos (obsolescência tecnológica de hardware) garantindo que existam equipamentos apropriados para leitura das mídias (suportes), pelo menos até uma transposição das informações para um suporte com tecnologia recente. Em segundo, haveria os cuidados lógicos que envolvem a forma como a informação é representada nos suportes (bits 0 e 1). Caberá aos softwares ler essas seqüências de bits, sendo necessários que estes sejam compatíveis mesmo com o passar do tempo.

Bodê (2007, p.08) complementa que “os cuidados físicos e lógicos precisam ser levados em consideração numa política de preservação, no universo digital. O descuido com um deles implicará na inviabilidade de acesso aos documentos eletrônicos para a manutenção destas”.

Para isto, a preservação de documentos eletrônicos se concentra na obtenção da estabilidade do suporte da informação, com foco na manutenção do acesso, que pode implicar na mudança de suporte e formatos, bem como na atualização do ambiente tecnológico. Sendo que, a fragilidade do suporte digital e a obsolescência tecnológica de hardware, software e formato exigem intervenções periódicas através de estratégias de preservação.

Contudo, CASTRO; CASTRO; GASPARIN (2007, p. 122) revelam que “para a preservação da integridade é necessário que os documentos e seus respectivos metadados permaneçam seguros, completos e autênticos.

Portanto, Rotthenberg (1996, apud ARELLANO, 2008, p.19) revela que 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 interação com recursos de busca.

Metadados

O termo metadados antecede a Web tendo, aparentemente, sido cunhado por Jack Myers nos anos 60 para descrever arquivos eletrônicos (MILSTEAD & FELDMAN, apud BAPTISTA, 1999, p.184), mas começou a aparecer mais freqüentemente na literatura sobre Sistemas Gerenciadores de Banco de Dados (SGBD) nos anos 80, para descrever as características das informações armazenadas nos bancos de dados (VELLUCCI, 1998, apud CAMPOS, 2007 p. 191).

Surgiu, em função das necessidades das organizações de conhecer melhor os dados que elas mantêm e conhecer com mais detalhes os dados de outras organizações através de intranets e extranets. A catalogação dos dados propiciará a maior utilização deles por usuários com múltiplos interesses. Sem uma documentação eficiente dos dados é dificultada aos usuários a localização de dados necessários para suas aplicações.

Organizações que não documentam seus dados, freqüentemente, com o decorrer do tempo, ficam sujeitas à superposição de esforços de coleta e manutenção de seus dados, vulneráveis a problemas de inconsistências e, principalmente, pagarão um alto custo pelo não uso ou uso impróprio dessa informação. Além disso, Boeres e Arellano (2005, p. 7) revelam “na preservação digital as informações devem ser preservadas de modo a evitar que sejam corrompidas, ao criar uma estrutura que guarde o conteúdo e a estrutura da informação, para isto podemos usar os metadados”.

Todas as formas de preservação digital, exceto as mais simples, podem se beneficiar pela criação, manutenção e evolução de Metadados detalhados para apoio aos processos de preservação. Por exemplo, metadados podem documentar o processo técnico associado com a preservação, especificar informações de direitos autorais e estabelecer o conteúdo digital. Eles podem registrar a cadeia de custódia de um objeto digital e identificá-lo individualmente tanto interna como externamente em relação ao arquivo em que reside. Em resumo, a criação e instalação de metadados para preservação parece ser um componente chave para as estratégias de preservação. (OCLC/RGL, 2001, p. 2 apud BODÊ, 2008, p. 63).

Contudo, metadado é um termo que remete às tradicionais atividades de representação da informação (catalogação, descrição arquivística e outros), aquelas com

as quais os profissionais da informação já estavam às voltas antes da chegada das atuais tecnologias.

Pode se considerar metadado como a informação que descreve e explica qualquer dado que, de modo geral, possa vir a aparecer em meio eletrônico, sabendo que os metadados não são utilizados apenas em aplicações digitais.

Assim, a utilização de metadados vai além de descrever e catalogar objetos de informação. Enquanto profissionais de museus, arquivos, e de biblioteca podem estar muito familiarizados com o termo em associação com descrição ou catalogação, metadados também podem indicar o contexto de criação, gerenciamento, processamento, preservação e uso dos recursos que estão sendo descritos. (FERREIRA, 2006, p.60)

Quanto ao seu tipo, os metadados recebem a seguinte distinção: descritivos, administrativos e estruturais. Os administrativos contêm os metadados técnicos e de preservação. Os metadados de preservação, um pouco mais abrangentes, mantêm informações acerca da origem do arquivo, sobre o suporte que o gerou e também a respeito das ações efetuadas no arquivo dentro do repositório digital. Já os metadados técnicos, ainda mais específicos, por exemplo, um arquivo sonoro, engloba informações como: o formato do arquivo, a taxa de bits, a taxa de amostragem, equipamentos e softwares utilizados.

Os metadados de preservação informam sobre a origem do material, os detalhes técnicos dos registros como qual foi a versão de software usado, como foi construído o registro, etc. esse método é uma forma especializada de administrar metadados os quais podem ser usados como um meio de estocar a informação técnica que apóia a preservação dos objetos digitais e visam apoiar e facilitar a retenção, a longo prazo, da informação digital. (BOERES, ARELLANO, 2005, p. 12)

E Innarelli (2007, p. 69) complementa que a finalidade dos metadados tecnológicos de preservação é a garantia das informações funcionais sobre o documento eletrônico, permitindo assim o entendimento da tecnologia utilizada na sua produção e manutenção, assim como todo o seu histórico de acesso, migrações, rejuvenescimento, mídias, sistemas etc.

Além disso, o principal objetivo dos metadados é registrar e organizar de uma forma estruturada os dados de uma determinada comunidade ou organização, visando à padronização e fácil recuperação de informações representadas e organizadas sob sua estrutura, além de permitir melhoras na busca e na avaliação das informações, bem como facilitar a pesquisa e manutenção desses dados.

De acordo com Hodgson (1998 apud FERREIRA, 2006, p. 61) as funções desempenhadas pelos metadados, as quais podem ser operacionalizadas por pessoas ou por agentes automatizados são:

- Fornecer um resumo do conteúdo do recurso de informação(isto é, do que trata o recurso), permitindo aos usuários determinar se este é exatamente o que eles desejam;
- Dar condições aos usuários de procurar, recuperar e usar recursos de informação;
- Quando formatos múltiplos são providos, ajudar aos usuários na escolha de recurso de informação apropriado ao seu uso;
- Prevenir usuários quanto às limitações no acesso aos dados, no que diz respeito a softwares necessários, padrões e formatos adotados;
- Instruir em como interpretar os dados (formato, codificação, criptografia);
- Fornecer informações que afetam o uso do recurso (por exemplo, condições legais de uso, idade mínima do usuário, cobrança de taxas, usos permitidos e não permitidos para o recurso etc.);
- Fornecer a história ou proveniência do recurso de informação, tal como sua fonte original e suas transformações subsequentes (filtragem, atualização, descarte etc.);
- Fornecer especificações para o gerenciamento dos recursos de informação (data da última modificação data de criação, identidade do administrador etc.);
- Fornecer especificações sobre relações entre recursos (por exemplo, o relacionamento entre um conjunto de artigos e um jornal, entre um trabalho original e sua versão subsequente, entre componentes de um trabalho multimídia etc.);
- Listar os componentes lógicos de recursos de informação complexos e oferecer meios para acessar esses componentes (por exemplo, tabelas, lista de componentes de um software, etc.).

A utilização de metadados em uma organização permite melhorias na busca e avaliação de informações, bem como facilidades na pesquisa e manutenção dos dados. Os esquemas de metadados possibilitam a definição dos elementos de metadados a serem usados por uma comunidade para descrever suas informações em suas bases.

Em geral, essas abordagens baseadas em metadados para localizar informação permitem ao usuário realizar uma busca, por um objeto armazenado em um repositório, utilizando campos de consulta típicos, como título, autor, ou ano de publicação. No entanto, metadados não apenas descrevem o conteúdo de outros dados, mas também indicam muitos outros aspectos, por exemplo, os aspectos relacionados ao gerenciamento, preservação, contexto, e histórico de uso de recursos de informação.

Os metadados podem, por exemplo, documentar procedimentos associados à preservação, especificar direitos de acesso e estabelecer a autenticidade do conteúdo digital. Podem unir os diversos componentes constituintes de um objeto complexo e, ainda, as diversas versões do mesmo objeto. Em resumo, a criação e o desdobramento de metadados de preservação tende a ser um componente chave para a maioria das estratégias de preservação digital.

Contudo, os desenvolvimentos de padrões de metadados também possibilitam que a produção científica possa ser disseminada através de documentos eletrônicos, até mesmo representando a disseminação do acesso a acervos.

Considerações finais

Boa parte das expressões humanas são apenas digitais, tanto oficiais como pessoais. Essas são questões de mais alta relevância para a humanidade e a responsabilidade por essa preservação estará na esfera de atuação dos arquivos, que custódia e garante a transmissão e preservação da diversidade para as futuras gerações. Sendo, preservação digital a capacidade de garantir que a informação digital permaneça acessível e autêntica com o passar do tempo, identifica-se assim a necessidade de utilizá-la. Além disso, sem a aplicação de estratégias de preservação não existiria garantia de confiabilidade e integridade desses documentos em longo prazo.

A preservação digital é um constante controle de riscos, no qual o reconhecimento de que perdas são inevitáveis desenvolve a relevância dos modos de evitá-las. Não existe consenso quanto às melhores estratégias e metodologias a serem empregadas num plano de preservação digital, pois nenhuma medida é capaz de atender todos os requisitos de todos os documentos eletrônicos. No entanto, quanto mais a gestão de documentos eletrônicos estiver conforme padrões, maior probabilidade haverá da sustentabilidade do patrimônio digital. Embora o conhecimento e a experiência que existem hoje não sejam capazes de garantir a durabilidade dos documentos eletrônicos, a adoção de planos de preservação digital, mesmo que básicos, se faz necessária.

Apesar da crescente relevância que os meios e técnicas digitais de gerenciamento da informação digital vêm assumindo no mundo, existe pouca avaliação dos fundamentos teóricos e metodológicos das estratégias de preservação digital. Porém a utilização de metadados e padrões de metadados tornam possível promover a

integração, interpretação, localização e reutilização dos dados, assegurando sua existência através do tempo.

Salientando-se que os metadados não se configuram em solução final para o problema da preservação, mas são relevantes para a redução dos riscos e o aumento das chances de sobrevivência da informação digital, assim como dão uma grande contribuição para a gestão, armazenamento, recuperação e uso da informação.

Portanto, o uso de metadados permite uma organização eficiente dos dados armazenados, facilitando o acesso a esses dados e permitindo a construção de mecanismos de busca interativos. As pesquisas podem ser realizadas baseando-se nas características semânticas da informação procurada, ou seja, pode-se descrever uma informação para que o mecanismo de busca verifique sua existência na base de dados.

Referências

ARELLANO, Miguel Ángel Márdero; Andrade, Ricardo Sodré. **Preservação digital e os profissionais da informação**. DataGramZero – Revista de Ciência da Informação – v.7 n.5 out/06 ARTIGO 05. 2006.

_____. **Preservação de Documentos Digitais**. Artigo. Ci. Inf., Brasília, v. 33, n.2, p. 15-27, maio/ago. 2004. Disponível em:< <http://revista.ibict.br/index.php/ciinf/article/viewFile/305/271>>. Acesso em 5 fev.2009.

_____. **Critérios para a preservação digital da informação científica**. 2008. 354f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2008. Disponível em:<http://bdtd.bce.unb.br/tesesimplificado/tde_busca/arquivo.php?codArquivo=4547> Acesso em: 10 mai. 2009.

ARQUIVO NACIONAL (Brasil). Conselho Nacional de Arquivos. Câmara Técnica de Documentos Eletrônicos. **Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos: e-ARQ Brasil**. Rio de Janeiro: Arquivo Nacional, 2006.

BAPTISTA, Dulce Maria. **O impacto dos metadados na representação descritiva**. Florianópolis: Revista ACB. 2007. Disponível em:< <http://www.acbsc.org.br/revista/index.php/racb/article/viewFile/529/663>> Acesso em: 06dez.2008

BODÊ, Ernesto Carlos. **Preservação de documentos digitais: o papel dos formatos.** 2008.153f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2008.

CAMPOS, Luiz Fernando B. **Metadados digitais: revisão bibliográfica da evolução e tendências por meio de categorias funcionais.** Bibli: R. Eletr. Bibliotecon. Ci. Inf. Florianópolis, 2007. Disponível em:< http://www.lfbcampos.com/Campos_Metadados_n23_1_semestre_2007.pdf> Acesso em: 16 jan.2009.

CASTRO, Ástrea M.; CASTRO, Andresa M. , GASPARIAN, Danuza M. C. **Arquivos físicos e digitais.** Brasília: Ed. Thesaurus. 2007.

FERREIRA, Miguel. **Introdução à preservação digital – Conceitos, estratégias e actuais consensos.** Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. Disponível em:<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf> > Acesso em: 10 mai. 2009.

INNARELLI, H.C. **Preservação digital e seus dez mandamentos.** SANTOS, V. B. (Org.) ; _____(Org.) ; SOUSA, R. T. B. (Org.) . Arquivística: temas contemporâneos. 1. ed. Brasília: Editora Senac-DF, 2007. cap.01, p. 21-75.

LIMA, Clarissa Costa e. **Preservação digital: a experiência da pesquisa Guignard.** Dissertação apresentada ao Programa de Pós- Graduação em Artes da Escola de Belas Artes da Universidade Federal de Minas Gerais. Belo Horizonte: Escola de Belas Artes, 2007.

LUCCA, Giana; CHARÃO< Andrea S.; STEIN, Benhur O. **Metadados para um sistema de gestão eletrônica de documentos eletrônicos.** Arquivística. net: Rio de Janeiro. 2006. Disponível em: <<http://www.arquivistica.net/ojs/include/getdoc.php?id=188&article=54&mode=pdf>> Acesso em: 7 nov.2008.

SANT’ANNA, Marcelo Leone. **Os desafios da preservação de documentos públicos digitais.** Artigo. 2008. Disponível em:< www.ip.pbh.gov.br/ANO3_N2_PDF/ip0302santanna.pdf>. Acesso em: 12 mar.2009.

SICHMANN, M. **O reconhecimento da importância de preservação de acervos na região.** Campinas, 2003. Disponível em:<[HTTP://bibmemoria.cmu.unicamp.br/sarao/Revista15/sarao](http://bibmemoria.cmu.unicamp.br/sarao/Revista15/sarao)>>. Acesso em 27 nov.2008