

NORMAS ARQUIVÍSTICAS E PADRÕES DE DESCRIÇÃO DE METADADOS APLICADOS À PRESERVAÇÃO DO PATRIMÔNIO ARQUIVÍSTICO DIGITAL

André Ricardo de Andrade Vasconcellos Luz

Graduando em Arquivologia / 6º período - UNIRIO / andrericardoluz@iq.com.br

Resumo

Esta comunicação pretende apresentar um panorama da importância das normas Arquivísticas e dos Metadados para a Preservação do Patrimônio Arquivístico Digital. Assim como a publicação da ‘Carta para a Preservação do Patrimônio Arquivístico Digital’ e a Resolução nº 20, de 16 de julho de 2004, este trabalho busca consolidar as inquietações sociais frente à preservação digital e fomentar as propostas da Arquivística para equacionar e colaborar a esta preservação. Aborda ainda os padrões nacionais e internacionais de Metadados, atuais e em pesquisa, bem como o conceito de interoperabilidade.

Palavras-chave: Arquivística. Arquivo. Metadado. ISO. Preservação Digital. XML. Interoperabilidade.

1. INTRODUÇÃO

Fruto da inquietação da maioria dos profissionais da informação, bem como da sociedade em geral, o Conselho Nacional de Arquivos (CONARQ), através da Câmara Técnica de Documentos Eletrônicos, publicou neste ano de 2004 a ‘Carta para a Preservação do Patrimônio Arquivístico Digital’¹ e a Resolução nº 20, de 16 de julho de 2004, que “dispõe[m] sobre a inserção dos documentos digitais em programas de gestão arquivística de documentos dos órgãos e entidades integrantes do Sistema Nacional de Arquivos” (CONARQ, 2004), fomentando ações quanto a políticas, normalizações, incentivo e disseminação do conhecimento na área da informação, frente ao patrimônio arquivístico digital.

Duas considerações destacam-se como precípuas e justificam plenamente a investida:

“Considerando que a informação arquivística, produzida, recebida, utilizada e conservada em sistemas informatizados, vem constituindo um novo tipo de legado: o patrimônio arquivístico digital;

Considerando que este patrimônio arquivístico digital se encontra em perigo de desaparecimento e de falta de confiabilidade, e que sua preservação em benefício das gerações atuais e futuras é uma preocupação urgente no mundo inteiro;[...]”(CONARQ, 2004).

Por sua vez, o próprio documento do CONARQ apresenta como referências carta e diretrizes da UNESCO ⁱⁱ, contendo preocupações semelhantes e igualmente fundamentais à sociedade contemporânea:

“Considerando que, a Carta de Preservação do Patrimônio Digital da UNESCO alerta sobre a possibilidade de desaparecimento do legado digital e sugere que os estados membros adotem um conjunto de medidas para salvaguardar esse patrimônio, o CONARQ está propondo um documento que atenda às **especificidades do documento arquivístico digital, fonte de prova e de informação, que necessita de proteção especial**” (*grifo nosso*, 2003).

A Carta do CONARQ propõe ainda:

“Definir **estruturas padronizadas de metadados** e determinar a sua utilização nos sistemas eletrônicos de gestão arquivística, com o propósito de gerir a preservação e a acessibilidade dos documentos digitais” (*grifo nosso*, 2004).

Tais inquietações despontam como força motriz deste trabalho, explicitado a seguir.

2. OBJETIVOS

O objetivo deste estudo é apresentar e divulgar algumas normas aplicadas no âmbito da díade Arquivística-Ciência da Informação e também padrões de descrição de metadados nacionais e internacionais que vêm sendo adotados com intuito de aplicar com segurança e eficiência os conceitos de interoperabilidade e preservação digital, obedecendo aos preceitos arquivísticos. Pretende com isso favorecer a cultura da preservação do patrimônio arquivístico digital. O assunto carece de um aprofundamento teórico, em parte coberto por este trabalho.

3. METODOLOGIA

O método deu-se por meio do levantamento, leitura e análise de fontes como: normas arquivísticas (ISO/IEC 17799, ISO 15489-1 e ISO 15489-2, ISO 15836/2003), *sites* da Internet referentes a padrões de metadados e bibliotecas digitais (já consolidados e algumas pesquisas em desenvolvimento) e literatura arquivística.

Importante salientar o foco arquivístico do resultado, ou seja, o que estes padrões têm a contribuir à Arquivística e por fim esta à sociedade, no âmbito da preservação não somente digital, mas principalmente da memória social.

4. RESULTADOS E DISCUSSÃO

Esta comunicação é em grande parte baseada no artigo “Os Arquivos e os sistemas de gestão da qualidade”, enviado em co-autoria com Julio César Cardoso para publicação pela Revista Arquivo & Administração da AAB. Porém, aqui o foco está no uso do conhecimento e aplicação das normas e padrões visando a preservação digital.

Para que um projeto de preservação funcione a contento deve-se considerar o planejamento, a implantação, auditoria e manutenção de um sistema de gerenciamento de informações (incluindo o sistema de informações arquivísticas). O objetivo é tornar estas informações disponíveis a qualquer tempo, de modo

inteligível e seguro, assim que solicitadas para em seguida disseminá-las às pessoas ou sistemas autorizados. Apesar de não estar garantindo, estas ações estarão ao menos colaborando de modo relevante para a preservação da memória social. Neste sentido, abordei as seguintes normas:

- **ISO/IEC 17799 – Segurança da Informação**

A atividade arquivística deve sempre considerar, entre outros aspectos, o grau de sigilo da documentação em questão, qualquer que seja o suporte. No meio digital, os níveis de acesso determinam este grau de sigilo.

A Informação “é um ativo que (...) tem valor para a organização e conseqüentemente necessita ser adequadamente protegida” [ABNT(h), 2001, p.2]. Logo, a segurança da informação reconhece-se como a preservação da confidencialidade (garantia de acesso somente por autorizados), integridade (exatidão e completeza da informação e métodos de processamento) e disponibilidade (acesso sempre que necessário) (Ibid., p. 3). Os esforços não deverão estar concentrados apenas em abordagens técnicas, mas também nos aspectos culturais e políticos envolvidos. É preciso comprometimento do mais alto nível da instituição visando identificação e controle dos riscos e ameaças através de uma avaliação sistemática dos requisitos de segurança.

Importante considerar os riscos de acesso não autorizados, não apenas físicos, mas também lógicos. É fundamental estabelecer um perímetro de segurança física sempre que necessário, bem como políticas e procedimentos de segurança dos sistemas (*Firewall*, antivírus, rotina de *backup*, níveis de acesso, *logs* de operação e falhas etc).

Cabe ressaltar que o mais interessante e funcional para o profissional da informação será atuar interdisciplinarmente interagindo e dialogando com outros profissionais especialistas, como por exemplo, profissionais de informática (analistas de sistemas, programadores, etc.), bem como com outros envolvidos nas atividades.

- **ISO 15.489-1/ISO 15.489-2 - Gestão Arquivística – Gestão de Documentos**

Na norma, entende-se Gestão de Documentos como:

“... campo da gestão responsável pelo eficiente e sistemático controle de criação, recepção, manutenção, uso e disposição (arranjo) dos documentos, incluindo os processos de captura e manutenção das evidências das informações em torno das atividades de negócio na forma de documentos de arquivo” [ISO(a), 2001, p.9, tradução do autor].

Como dito na própria norma, esta deve ser utilizada em consonância com as normas ISO 9000 e ISO 14001. Promovem orientações (gerais e técnicas) para o planejamento e implementação de políticas, procedimentos, sistemas e processos para a gestão arquivística. Oriunda da AS 4390 – Australian Standards/Records Management, possui em sua segunda parte [Guidelines – (Technical Report)], metodologia para implantação nas organizações. Trata-se de um referencial bastante completo. Aborda em linhas gerais: as políticas de gerenciamento arquivístico; a estratégia, criação e implementação do sistema de arquivo; os processos de arquivamento e seus meios de controle; monitoramento e auditoria do sistema e finalmente, treinamento da equipe de trabalho e dos usuários.

A norma fornece diretrizes para um projeto de gestão arquivística com maior segurança, precisão e controle das atividades e principalmente dos resultados, contribuindo em muito para a satisfação dos clientes, a melhoria contínua e a abordagem factual para a tomada de decisão, requisitos fulcrais para as instituições segundo a ISO 9000:2000. Proporcionando com isso uma atuação arquivística (podendo e devendo ser aí incluída a preservação) já na criação dos documentos. Esta atuação terá efeitos mais pontuais ainda no caso dos documentos digitais.

- **ISO 15.836/2003 – Dublin Core Metadata element set (DCMES)**

Para uma adequada compreensão do contexto exposto a seguir é fundamental a apresentação do conceito de “Metadado” bem como os benefícios de sua aplicação. *Lato Sensu* significa “dado sobre dado”. Numa abordagem mais completa, entende-se como sendo o

“[...]conjunto de dados estruturados que identificam os dados de um determinado documento e que podem fornecer informação sobre o modo de descrição, administração, requisitos legais de utilização, funcionalidade técnica, uso e preservação” (DCMI, 2004).

Elaborado pelo DCMI - *Dublin Core Metadata Initiative*, organização não governamental dedicada à organização de arquivos eletrônicos, este padrão (DCMES) para cruzamento de descrição de recursos de informação figura como um conjunto mínimo de informações (metadados) sobre objetos (arquivos) digitais. Um padrão de catalogação composto por um conjunto de quinze descritores, divididos em três áreas: conteúdo, propriedade intelectual e instancialização (realização ou ocorrência física). Pretende co-existir com outros padrões de metadados com semântica diferenciada e promover a interoperabilidade entre sistemas. Sugere ser usado como ponto de partida para padrões de descrição mais complexos e personalizados a cada necessidade. Os elementos descritivos são:

- **conteúdo** → *Título (Title), Assunto (Subject), Descrição (Description), Fonte (Source), Língua (Language), Relação (Relation), Cobertura (Coverage)*;
- **propriedade intelectual** → *Autor (Creator), Editor (Publisher), Contribuidores (Contributor), Direitos (Rights)*;
- **instancialização** → *Data (date), Tipo (Type), Formato (Format) e Identificador (Identifier)*.

Para facilitar a interoperabilidade já citada entre sistemas, apresenta também os “Qualificadores do Dublin Core”, objetivando tornar mais específicas a criação e redação dos conteúdos dos elementos. Propõe ainda a adoção de vocabulário controlado para facilitar este monitoramento.

Conforme demonstrado no curso “Formação de Diretores de Projetos ETD-Net”, de autoria da professora Ana Pavani, outros padrões de conjuntos de metadados foram elaborados para usos específicos; portanto, cada necessidade poderá gerar padrões customizados. Por exemplo, para caracterizar as ETDs (*Electronic Theses and Dissertations*) adequadamente, seus conjuntos de metadados necessitam de um número muito maior de elementos descritivos. Ilustraremos a seguir o padrão internacional (ETD-ms) e nacional (MTD-Br), padrões adotados nos catálogos coletivos internacional (NDLTD) e nacional brasileiro (BDTD). Isto não impede que as instituições construam seus próprios conjuntos de metadados, facilitando programas de preservação digital posteriores, porém sempre em consonância com os supracitados.

- **ETD-ms - *Interoperability Metadata Standard for Theses and Dissertations***

Alguns elementos do ETD-ms são comuns ao *Dublin Core Metadata Element Set*. Os elementos são: Título, Título Alternativo, Autor, Assunto, Descrição, Descrição/Resumo, Descrição/Nota, Descrição/Versão, Editor, Contribuidor, Contribuidor/Papel, Data, Tipo, Formato, Identificador, Língua, Cobertura, Direitos, Tese/Grau/Nome, Tese/Grau/Nível, Tese/Grau/Disciplina, Tese/Grau/Outorgante. Tanto o ETD-ms quanto o MTD-Br atendem ao DCMES.

- **MTD-Br – Padrão Brasileiro de Metadados de Teses e Dissertações**

Concebido com os seguintes objetivos: identificação para localização das ETDs, coletar informações para indicadores nacionais referentes às pós-graduações e intercâmbio com outras bases do país de ensino e pesquisa, como a plataforma Lattes, o ScieLO, o INPI, entre outras. Compõe-se de 17 elementos, com subcampos, podendo ter ainda subcampos de segundo nível (sub-subcampo, seguindo o padrão XML – níveis hierárquicos consecutivos de subcampos), sendo alguns campos obrigatórios, outros facultativos e outros apenas desejáveis (logo, também facultativos). São eles: Controle, Biblioteca Digital, Biblioteca Depositária, Título, Idioma, Grau, Titulação, Resumo, Cobertura, Assunto, Local Defesa, Data Defesa, Autor, Contribuidor, Programa, Agência Fomento, Direitos.

- **WBE (IMS Project e IEEE LTSC)**

Nos dias atuais, cada vez mais projetos de educação à distância vêm sendo desenvolvidos tanto para uso acadêmico quanto para uso institucional, em ambas as esferas: pública ou privada. A sigla WBE significa *Web Base Education*, ou seja, o uso da Tecnologia da Informação na Educação. Tal uso fez surgir o termo Learning Object (LO), objeto digital voltado para a educação e, por conseguinte, com metadados específicos: os chamados *Learning Object Metadata* (LOM). Dois grupos trabalham no desenvolvimento de um conjunto de metadados voltados para a educação: o IMS – *Instructional Management System Project* (composto de 43 elementos – 15 do DCMES e mais 28 adicionais) e o IEEE LTSC – *Institute of Electrical and Electronics Engineers Learning Technology Standards Committee* (composto por quase 80 elementos). A preservação deste material será imprescindível para referenciar uma época da sociedade. A educação determina e

reflete grande parte da cultura de um povo. O fato de parte desta cultura estar em ambiente eletrônico, em documentos digitais, gera a necessidade de uma atenção específica a suas nuances e detalhes técnicos ou conceituais.

- **XML – Extensible Markup Language**

Segundo PAVANI (2004), significa uma “linguagem utilizada para transferência de metadados entre sistemas de bibliotecas digitais (...) usada em outros sistemas que transferem metadados e/ou informações, como por exemplo, os sistemas de comércio eletrônico”. Trata-se de uma linguagem contemporânea de aplicação da maior importância para registro, troca, armazenamento e recuperação de informações, devendo ser, por este motivo, objeto de atenção dos profissionais da informação em geral, logo, também dos arquivistas, cada vez mais levados a apresentar soluções de gerenciamento para documentos e informações em suporte ótico-magnético. Para exemplificar, apresentamos no quadro abaixo um trecho dos metadados em XML do artigo citado anteriormente, caso fosse publicado em formato PDF, no padrão DCMES:

```
<!DOCTYPE rdf:RDF SYSTEM "http://dublincore.org/2000/12/01-dcmes-xml-dtd.dtd">
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://dublincore.org/">
<dc.title>Os Arquivos e os sistemas de Gestão da Qualidade</dc.title>
<dc.author>Julio Cesar Cardoso e André Ricardo Luz</dc.author>
<dc.subject>gestão da qualidade, arquivologia, informação, metadados, ISO</dc.subject>
<dc.description> Reflexões sobre a inserção da Arquivística no campo estudos da gestão da
qualidade. A partir de consulta à literatura sobre sistemas de gestão da qualidade, procura-se
demonstrar a importância de estudos e pesquisas voltadas para o campo da informação e dos
arquivos, aplicados em organizações que atuam buscando padrões de qualidade. São comentadas
algumas normas e outros documentos referentes à gestão da qualidade e informação. Apresenta
parte do universo das normas arquivísticas nacionais e internacionais, incluindo padrões de descrição
de metadados.</dc.description>
<dc.publisher>AAB</dc.publisher>
```

Figura 1: Metadados na linguagem XML, padrão DCMES.

O conhecimento dos padrões de transferência de informações é um componente a mais a ser considerado no âmbito da preservação digital. Não obstante o

conhecimento técnico, a própria existência destes os fazem objeto de interesse do profissional da informação, em nosso caso o Arquivista/Conservador.

- **PDF/A**

Há ainda um esforço internacional liderado pela AIIM – *Enterprise Content Management Association* no sentido de elaborar um conjunto de metadados específico para o padrão PDF (*Portable Document Format*). O objetivo é a manutenção arquivística e a preservação dos documentos por extensos períodos de tempo, por eles nomeado de *e-preservation*. A idéia é aprová-lo junto a ISO – *International Organization for Standardization* - como padrão internacional. A versão mais recente oferece um rico conjunto de metadados conhecido como *Extensible Metadata Platform* (XMP), Plataforma de metadados Extensível, baseado nas especificações XML e RDF (*Resource Description Framework*). A última reunião deste grupo até a data de fechamento desta comunicação foi em 14 de julho de 2004, onde foi definida a agenda dos próximos encontros, já que a expectativa é obter a aprovação em 2005.

5. CONCLUSÕES

São muitos os aspectos a serem considerados para que se diga estar promovendo a preservação do patrimônio arquivístico digital, seja de uma pessoa, instituição e até mesmo uma nação.

No que tange a atuação do Estado, pode-se dizer que os primeiros e principais passos já foram dados: a própria elaboração da 'Carta para a Preservação do Patrimônio Arquivístico Digital', através de consulta pública e a Resolução nº 20, de 16 de julho de 2004, sinalizam mobilidade e preocupação com a preservação digital. Não só o Estado, como alguns setores da sociedade e os profissionais da informação, entre eles os arquivistas, já atentaram para a necessidade do fomento da discussão visto a urgência e as particularidades que cercam o assunto. Por exemplo, muito do legado arquitetônico projetado atualmente encontra-se em arquivos CAD (*Computer-Aided Design*), estarão estes arquivos acessíveis quando necessário?

Neste sentido, algumas proposições são cabíveis:

- Para preservação digital, ações puramente técnicas ou ainda apenas preventivas podem não ser suficientes. Aspectos políticos também influenciam e são influenciados pelos tecnológicos;
- Durabilidade (resistência mecânica) e permanência (estabilidade química), até então pilares do conceito de preservação, não satisfazem quanto se trata de preservação digital. Há de se considerar também: obsolescência de software, hardware, formatos, mídias e linguagens (de programação); de forma que o paradigma de imutabilidade para a preservação seja rompido;
- A preservação do conteúdo de um documento digital é estratificada, ou seja, deve-se preservar a cadeia binária (dados) e também a linguagem para decodificá-los (linguagem de programação, softwares e sistema operacional envolvidos).
- A atuação arquivística faz-se da maior importância já que pressupõe controle e avaliação, fundamentais para a preservação desde a origem dos documentos, neste aspecto as normas proporcionam um arcabouço técnico de grande importância;
- Os padrões de metadados fornecem subsídio para as questões de descrição, acessibilidade e interoperabilidade, propiciando o acesso futuro, objetivo da preservação.

Para atingir com segurança e eficiência a preservação da memória social, cada vez mais residente em meio digital, é fundamental que os arquivistas se aprofundem nas pesquisas em busca do levantamento de novas questões e de algumas soluções, verticalizando seus estudos, mas não se desfazendo da atuação interdisciplinar, tão necessária nos dias atuais.

6. NOTAS E REFERÊNCIAS

_____. **NBR 12676**: Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: 1992.

_____. **NBR ISO 9000**: Sistemas de gestão da qualidade – fundamentos e vocabulário. Rio de Janeiro, 2000.

_____. **NBR ISO 9001**: Sistemas de gestão da qualidade – requisitos. Rio de Janeiro, 2000.

_____. **NBR ISO 9004**: Sistemas de gestão da qualidade – diretrizes para melhoria de desempenho. Rio de Janeiro, 2000.

_____. **NBR ISO/IEC 17799**: Tecnologia da informação – código de prática para a gestão da segurança da informação. Rio de Janeiro, 2001.

CARDOSO, Julio Cesar; LUZ, André Ricardo. **Os Arquivos e os sistemas de gestão da qualidade**. 2004. 17 p. No prelo.

CONARQ. **Carta para a Preservação do Patrimônio Arquivístico Digital**. Rio de Janeiro, 2004. Disponível em: http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/preservacao/cartapreservacao.a_sp. Acesso em: 23 jul. 2004.

DICIONÁRIO DE TERMINOLOGIA ARQUIVÍSTICA. São Paulo: Associação dos Arquivistas Brasileiros / Núcleo Regional de São Paulo: Secretaria de Estado da Cultura, 1990.

DCMI (Dublin Core Metadata Initiative). **Dublin Core Metadata Element Set (DCMES)**, version 1.1: Reference Description. [S.l., 2003]. Disponível em: <<http://dublincore.org/documents/dcmes/>>. Acesso em 8 ago. 2004.

EMERY, Gail Repsher. e-documents needs e-preservation. *Washington Technology*. [S.l., v.17, n. 23, out/2003]. Disponível em: <<http://washingtontechnology.com/news/17-23/federal/20235-1.html>>. Acesso em: 20 ago. 2003.

ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION) – **ISO 15489-1:2001** ; Information and documentation – Records Management. Part 1: General. [S.l., 2001]. Disponível em : <<http://www.iso.org>>. Acesso em: 25 abr. 2004.

_____. – **ISO 15489-2:2001** ; Information and documentation – Records Management. Part 2: Guidelines. [S.l., 2001]. Disponível em : <<http://www.iso.org>>. Acesso em: 25 abr 2004.

LeFurgy, William G. **PDF/A: Developing a File Format for Long-Term Preservation**. RLG DigiNews. [S.l., v. 7, n. 6, 2003]. Disponível em : <<http://rgl.org/index.php>>. Acesso em: 20 ago. 2004.

NDLTD (Networked Digital Library of Theses and Dissertations). **ETD-ms – an Interoperability Metadata Standard for Electronic Theses and Dissertations**. [S.l., 2004?]. Disponível em: <<http://www.ndltd.org/index.en.htm>>. Acesso 22 ago. 2004.

PAVANI, Ana. **Formação de Diretores em Projetos ETD-Net**. [S.l., 2004]. Disponível em: <<http://ibict.br>>. Acesso em: 10 ago. 2004.

SAYÃO, Luis Fernando. **Preservação Digital: uma brevíssima introdução**. Rio de Janeiro, 2003. Disponível em: http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/preservacao/preservacao.asp Acesso em: 03 dez. 2003

SILVA, Armando Malheiro da et al. **Arquivística**: teoria e prática de uma ciência da informação. Porto: Edições Afrontamento, 1998. V.1. (Biblioteca das ciências do homem, série plural 2).

THOMAZ, Kátia P; SANTOS, Vilma Moreira. Metadados para o gerenciamento eletrônico de documentos de caráter arquivístico – GED/A: estudo comparativo de modelos e formulação de uma proposta preliminar. **Revista DataGramZero**. [S.l., v. 4, n. 4, artigo 4, ago/03]. Disponível em: http://www.dgz.br/ago03/Ind_art.htm. Acesso em : 22 ago. 2004.

UNESCO. **Carta para la Preservación del Patrimonio Digital**. [Paris?], 2003. Disponível em:

http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/preservacao/cartapreservacao.asp . Acesso em: 23 jul. 2004.

W3 Consortium. **Extensible Markup Language (XML)**. [S.l., 2004]. Disponível em: <<http://www.w3.org/XML/>>. Acesso em: 23 ago. 2004.

ⁱ Disponível em http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/preservacao/cartapreservacao.asp

ⁱⁱ Disponível em http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/preservacao/cartapreservacao.asp