



Processos e métodos de preservação digital do Património Arquivístico

Ana M^a Rodrigues

19 de Setembro de 2013

Objectivos da apresentação

Relativamente ao **Património Arquivístico Digital**, apresentar os principais :

- ◆ conceitos;
- ◆ necessidades / problemas
- ◆ Processos, estratégias e métodos de preservação

Património Arquivístico Digital / Património Digital

Inf. produzida, recebida, utilizada e conservada em **sistemas informatizados das organizações públicas e privadas**; distingue-se de qualquer outra pelo facto de ser produzida com o propósito primário de constituir **prova** de uma actividade organizacional

Estrutura dos OD:

física;
lógica;
conceptual.



Patrimónios constituídos por **ODs**, contendo **Inf. digital**

Recursos de conhecimento ou expressão humana de tipo cultural, científico, administrativo ou legal (Carta da Unesco sobre PD)

Valor do Património Arquivístico Digital

- ◆ Recurso social extremamente **valioso** por gerar **riqueza** e **conhecimento**, e poder contribuir, em última instância, para o **progresso científico, cultural**;
- ◆ **Património arquivístico digital** é importante pelo valor **económico** que lhe está associado e que as novas tecnologias (sobretudo a *world wide web*) têm vindo a potenciar, tornando mais fácil a sua **produção, disseminação** e **fruição** , possibilitando a satisfação das necessidades crescentes das organizações e dos clientes (traduz-se basicamente em **valor continuado** possibilitado pela existência de **arquivos digitais**).

Todavia

a preservação deste património, dada a vulnerabilidade da **informação digital**, comporta **riscos**, mas também **desafios**

Conceito e Objectivo da Preservação Digital do Património Arquivístico

PD consiste na capacidade de garantir que a informação digital permanece acessível, interpretável e autêntica, no longo prazo, na presença de uma plataforma tecnológica diferente da que foi inicialmente utilizada no momento da sua criação

O **objectivo** é assegurar:

- ◆ o **acesso continuado** a informação digital autêntica, avaliada como património arquivístico;
- ◆ acções de **gestão faseada** e de **rotinas de preservação** levadas a cabo ao longo da totalidade do ciclo de vida dos objectos digitais (processos de: ingestão, descrição da identidade, documentação da informação, armazenamento, legibilidade do suporte, formatos - manutenção ou conversão - e verificação da autenticidade). Em suma, aceder ao conteúdo dos ODs ou seja à sua informação tal como foi depositada na instituição de memória, mediante acordo de custódia.

Será mesmo possível ?



Estes processos pressupõem a existência de **Repositórios Digitais**

Objectivo da Preservação Digital do Património Arquivístico

Este objectivo é **paradoxal**: em ambiente digital **preservar** implica manter a **autenticidade** e a **integridade** dos **ODs** originais ao longo do tempo, **no entanto**, não é possível preservar ODs, no longo prazo, sem alterá-los nas suas componentes **física** e/ou **lógica**

A aplicação de diferentes **acções, processos, métodos** e **estratégias** de PD implica, nalgumas situações, a alteração de algumas **propriedades significativas** dos ODs

Este **paradoxo** poderá desencadear algumas questões sobre o que é importante preservar em contexto digital, mas tb a compreensão e reformulação de conceitos

Preservar um OD pressupõe preservar **toda a sua complexidade** ou apenas a **informação** que ele contém ? (a qual pode ser codificada de várias formas, dependendo do formato)

Em suma,

O objectivo último da PD deve ser o de preservar **OD autênticos** para que o resultado do processo de PD se traduza em **ODs idênticos**, em todos os aspectos essenciais ao **OD original**

Necessidades específicas de preservação do Património Arquivístico Digital

Garantir **características essenciais** dos documentos de arquivo, definidas nos programas de gestão de documentos de arquivo de acordo com **NP 4438-1: 2005** tais como:

- ◆ suporte às **necessidades do negócio** com que se relaciona e ser usado para fins de atribuição de responsabilidade;
- ◆ prova do **conteúdo e da metainformação** associada ao documento de arquivo para documentar uma transacção, acção (incluindo elementos da estrutura; do contexto organizacional; da gestão);
- ◆ **autenticidade** (um documento de arquivo autêntico é aquele do qual se pode provar ser aquilo que pretende ser; ter sido produzido ou enviado por alegado produtor ou remetente; ter sido produzido ou enviado no alegado momento de produção ou envio;

Necessidades específicas de preservação do Património Arquivístico Digital

- ◆ **Fidedignidade** (documento cujo conteúdo é digno de crédito enquanto representação completa e fiel das transacções, actividades ou factos que atesta, podendo dele depender subsequentes transacções ou actividades;
- ◆ **Integridade** (refere-se ao facto dos documentos permanecerem completos e inalterados e deverem ser protegidos contra alterações não autorizadas);
- ◆ **Utilização** (refere-se a doc. utilizável que pode ser localizado, recuperado, apresentado e interpretado; as ligações entre os documentos de arquivo que representam uma sequência de actividades devem ser mantidas).

Necessidades específicas de preservação do Património Arquivístico Digital

Necessidades presentes e futuras das organizações preservando:

- ◆ informação relativa **decisões** ou **acções passadas e presentes** (**memória organizacional**);
- ◆ **elementos de prova** de acções passadas e presentes para satisfazer obrigações (**documentos probatórios**);
- ◆ **contexto do documento arquivado** para permitir a futuros utilizadores avaliar a sua autenticidade e fidedignidade, mesmo nos casos em que o sistema de arquivo tenha sido descontinuado ou sofrido modificações significativas.

Necessidades presentes e futuras de utilizadores internos e externos referentes a:

- ◆ **obrigações** e a **interesses legítimos** de utilizadores (parceiros de negócio, clientes ou outras pessoas sobre as quais recaem decisões ou actos da organização como autoridades administrativas, judiciais, órgão de coordenação arquivística, para efeitos de auditoria e fiscalização;
- ◆ avaliação de **benefícios legais, financeiros** ou **outros** que se prendam com a conservação dos documentos de arquivo para fins de investigação ou de interesse da sociedade no seu todo.

Obrigações de conservação de documentos de arquivo

Decorre da análise de exigências legais, tais como:

identificação do **ambiente regulador** que afecta as actividades de uma organização, bem como os **requisitos específicos para as documentar**;

fornecimento de **prova adequada da sua conformidade com o ambiente regulador**, na forma como documenta as suas actividades;

obrigação de **prestar contas** sobre as actividades desenvolvidas;

obrigação de determinar **destino final e prazos de conservação de documentos de arquivo** para garantir necessidades presentes e futuras das organizações .



Legislação aplicável – **Dec. Lei nº 447/88, de 10 de Dez. e Portarias de Regulamentação Arquivística**

Documentos de conservação permanente

De acordo com a autoridade arquivística competente (DGLAB), os **documentos de conservação permanente** que constituem o **Património Arquivístico Digital** são:

os que fornecem **elementos de prova** ou de informação sobre políticas e acções da organização;

os que constituem **prova** ou informação da interacção entre a **organização** e os seus **clientes**;

documentam **direitos** e **obrigações** de **pessoas** singulares ou colectivas;

contribuem para construir a **memória da organização** para fins científicos, culturais ou históricos;

contêm **prova** e **informação** sobre **actividades** de **interesse interno** e **externo**

Principal processo de preservação digital

Normalização dos formatos

- Consiste na utilização de formatos estáveis largamente adoptados e preferencialmente não proprietários; tem sido a proposta mais defendida no âmbito da PD;
- É entendida como um **processo transversal** que facilitará aplicação de todas as estratégias de PD e não como uma acção isolada;
- A **redução do nº de formatos** reduz os problemas de preservação da informação com conseqüente diminuição de custos e complexidade do processo de PD;

Crítérios de escolha

- Formatos conhecidos pela comunidade de interesse e baseados em normas internacionais abertas (para evitar eventuais complicações ex: com direitos de autor);
- Os formatos abertos garantem mais viabilidade da preservação futura e facilitam a interoperabilidade entre os sistemas;
- Riqueza do formato normalizado é importante considerar para que as características dos vários formatos possam ser devidamente incorporadas.

Principais métodos ou estratégias de preservação digital

Preservação da tecnologia - preservar o ambiente tecnológico original utilizado para aceder aos documentos digitais nele concebidos; pressupõe conservar e manter operacionais as plataformas que, entretanto se tornaram obsoletas, de forma a garantir que os documentos são acedidos no seu ambiente tecnológico original;

Emulação – permite o acesso aos documentos digitais no seu ambiente tecnológico original sem necessidade de preservar ou manter em funcionamento o hardware /software originais; o ambiente tecnológico original é mantido através da criação de um novo software (emulador) que imita o funcionamento do hardware e /ou software originais;

Migração - consiste na transferência periódica de material digital de uma dada configuração de hardware e software para outra ou de uma geração de tecnologia obsoleta para outra subsequente; mantém os ODs compatíveis com as tecnologias actuais, embora não preserve a sua vertente física e lógica;

Refrescamento – consiste na transferência sistemática de Inf. Digital registada e armazenada num dado suporte digital (CD-ROM, DVD) para outro do mesmo tipo, de forma a combater a deterioração física dos suportes; esta acção permite que a informação se mantenha acessível.

Estratégia alternativa: Repositórios Digitais Confiáveis (*Trusted Digital Repositories*)

Tanto a integridade de um OD como a sua autenticidade são difíceis de garantir em ambiente digital, por isso urge **repensar** e **reformular** estes **conceitos** em **sintonia** com o novo **contexto** ;

A autenticidade passa a ter como objectivo a **capacidade de demonstrar que o OD está conforme o original** em vez de demonstrar que o OD preservado é o original;

Tudo se resume a uma **questão de confiança** :

- ◆ na instituição que custodia os ODs e recursos digitais;
- ◆ nos métodos e estratégias utilizados durante a custódia e no ODs recuperados;

A solução proposta passa pela:

- ◆ criação e **certificação** de **Repositórios Digitais Confiáveis** (de acordo com norma ISO: 16363 *Audit and certification of trustworthy digital repositories*;
- ◆ descrição pormenorizada das propriedades significativas - conjunto de propriedades que se consideram essenciais para a interpretação do OD - para provar a sua autenticidade e integridade relativamente ao original e metainformação das alterações que foram levadas a cabo para assegurar a sua preservação.

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital

- **Base de Dados Relacionais** - a preservação só é possível para as **tabelas** e respectivos dados e **relações**; tudo o que respeita a funcionalidades do sistema de gestão da base de dados não é preservável;
- **Documentos áudio, vídeo e multimédia** - a preservação destes documentos passa pela identificação das suas **propriedades significativas**, ou seja as características essenciais que determinam a sua **forma**, **conteúdo**, **estrutura**, **qualidade**, **processamento** e **comportamento** e que um OD deve manter ao longo do tempo, independentemente das acções e estratégias de PD realizadas;

A definição do conjunto das propriedades significativas depende da natureza da organização responsável pela preservação, as características da colecção e as exigências da sua comunidade de interesse.

- **Sítios Web** - preservar sítios web é igualmente complexo, já que pressupõe:
 - ◆ uma **categoria** de sítio web;
 - ◆ um **tipo** de sítio web.

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital

Riscos e questões específicas dos sítios web

- ◆ Natureza heterogénea dos sítios web;
- ◆ Conectividade;
- ◆ Abolição da noção de tempo e de espaço ou fronteira;
- ◆ Frequência da mudança;
- ◆ Quantidade e diversidade de sítios web que precisam de preservação;
- ◆ Continuidade;
- ◆ Integridade dos sítios web;
- ◆ Propriedade;
- ◆ Base de Dados e sítios web profundos;
- ◆ Dados digitais e multimédia.

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital

Recursos web

Considerando que os sítios web são uma **estrutura tecnológica** constituída por diversos objectos digitais, armazenados em servidores, sistemas de gestão de conteúdos ou noutro tipo de sistema, apresentam, contudo, uma **característica comum**: o serem **distribuídos e acedidos** por um *browser*.

A forma de **disseminação** e **acesso** confere-lhes o **estatuto** de **Recursos web**.

Em termos de **Património Arquivístico Digital**, são considerados **recursos web** o(a)s:

- ◆ documentos de arquivo;
- ◆ publicações oficiais;
- ◆ artefactos;

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital

Critérios para escolha dos aspectos e elementos essenciais dos recursos web

Preservar a informação (todo o conteúdo significativo do recurso, incluindo texto, imagens, áudio) que a experiência torna disponível

e

Preservar uma experiência (o acesso a um conteúdo web, incluindo todos os seus aspectos e comportamentos inerentes, pressupondo para tal, a preservação de funcionalidades do sistema)

Aspectos essenciais e elementos dos recursos web a capturar e a preservar

Aspectos essenciais dos recursos web:

- ◆ o valor probatório de certos recursos web (documentos de arquivo);
- ◆ a reorientação da finalidade e reutilização do recurso web.

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital

Elementos essenciais dos recursos web:

- ◆ o conteúdo;
- ◆ a apresentação (visualização da página, incluindo dispositivos de navegação, imagens e layout da página);
- ◆ o comportamento (ou seja as características mais dinâmicas e interactivas que provocam mudanças, como acontece em sítios de blogs ou mesmo administrativos, com conteúdos sujeitos a mudança).

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital: documentos electrónicos assinados

Validade e força probatória dos Documentos electrónicos assinados pode ficar comprometida pela:

Falta de segurança jurídica a qual só pode ser comprovada **na transmissão / recepção de dados de documentos electrónicos no imediato** (em momento posterior à sua produção/recepção, a fidedignidade original garantida pela aposição de uma AD ao documento não pode ser considerada suficiente nem indispensável para provar a sua autenticidade (regime jurídico da AE e dos DE, definido no Dec-Lei nº 290D, de 2 de Agosto de 1999));

Obsolescência tecnológica (os equipamentos, tecnologias, formatos sucedem-se, sendo descontinuados com frequência, podendo inviabilizar a consulta a longo prazo);

Impossibilidade de demonstração de inalterabilidade consiste na incapacidade de demonstrar que o documento se manteve inalterado desde a sua produção até à data em que é necessária a sua consulta; depende da possibilidade de demonstração de que o certificado com base no qual assinatura foi aposta se encontrava válido à data em causa;

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital: documentos electrónicos assinados

Falta de segurança nos algoritmos criptográficos utilizados na tecnologia de criptografia assimétrica (consiste na pouca confiança na ligação única entre a chave pública e a chave privada baseada na confiança do algoritmo de encriptação utilizado (a evolução constante da capacidade de processamento informático pode conduzir à destruição periódica da confiança nos algoritmos utilizados, exigindo assim a respectiva substituição por outros mais seguros e menos vulneráveis).

Aspectos do problema

- ◆ A preservação, no médio e longo prazo, de documentos assinados com AD, sujeitos a migrações periódicas, não garante a inalterabilidade dos documentos;
- ◆ As migrações para outros formatos alteram o código binário do ficheiro da mensagem e os seus valores hash – da versão codificada – inviabilizando o mecanismo de verificação da AD;
- ◆ Quando aplicados processos de preservação, a AQ deixa de validar o documento.

Problemas específicos de preservação de certo tipo de Património Arquivístico Digital: documentos electrónicos assinados

A lei portuguesa (Dec-Regulamentar nº 25 /2004, de 15 de Julho) no que respeita à conservação a longo prazo de documentos electrónicos assinados opta pela **re-assinatura**, baseada em boas práticas como:

- a recolha e preservação do documento original (indissociavelmente associado à assinatura que lhe foi aposta) sem qualquer alteração;
- a recolha e preservação dos certificados utilizados para gerar a assinatura original (recolha dos certificados da entidade certificadora e todos os demais na **cadeia de certificação**);
- a validação cronológica do estado dos certificados à data de submissão ao arquivo.
- a re-assinatura electrónica pode ficar sob a responsabilidade de uma **terceira parte**, encarregue da tarefa de preservação;
- a **garantia da autenticidade dos documentos electrónicos assinados é uma das tarefas que mais dificuldades suscita**, pois é essencial manter evidências, em juízo, de que a assinatura foi validamente aposta aos documentos, mesmo que entretanto os certificados usados tenham expirado e os algoritmos sejam considerados inseguros. Juridicamente é possível garantir a manutenção da força probatória.

MAIS INFORMAÇÃO

Contactos:

ana.rodrigues@dglab.gov.pt

Obrigada pela vossa atenção !