



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

Preservação Digital

Um problema multidimensional

Miguel Ferreira

mferreira@dsi.uminho.pt

<http://www.dsi.uminho.pt/~ferreira>



24 de Novembro de 2008
Lisboa



Conteúdo

- ▶ **Contextualização**
- ▶ **Exemplos de obsolescência** tecnológica
- ▶ **Conceitos/definições**
- ▶ **Estratégias** de preservação
 - Preservação de **tecnologia**
 - Preservação do objecto **conceptual**
- ▶ **Boas práticas**
- ▶ **Notas finais**

Contextualização

Domínio físico ou analógico
“das coisas tangíveis”

Contextualização

Domínio físico ou analógico “das coisas tangíveis”



Contextualização

Domínio físico ou analógico
“das coisas tangíveis”



Digitalização



Contextualização

Domínio físico ou analógico
“das coisas tangíveis”



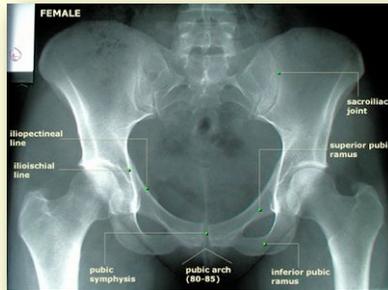
Digitalização

Domínio digital



Contextualização

Domínio físico ou analógico “das coisas tangíveis”



Domínio digital



Digitalização



Contextualização

- ▶ Vantagens das representações digitais
 - Fáceis de **partilhar, disseminar, transmitir**
 - **Facilidade** de **criar e manipular**
 - Ocupam pouco **espaço físico**
 - **Evitam** a **manipulação** do objecto original
 - **Fáceis de copiar** sem perder qualidade

- ▶ Desvantagens
 - Necessitam de um **contexto tecnológico** para que possam ser interpretadas
 - Rápida **obsolescência** da tecnologia
 - Os objectos acabam **prisioneiros** dos seus próprios **suportes** e/ou **formatos**



Contextos de preservação digital

- ▶ Domínio doméstico
 - **Avaria** no disco rígido...
 - Novo computador não tem **drive de disquetes...**
 - Serviço de **e-mail online** que desaparece...
 - O **vídeo de casamento** está em VHS e já não temos leitor de vídeo...
 - **Baby-blogs...**
 - Documentos em formatos obsoletos, ...
- ▶ Porta aviões USS Carrier
 - 1/2 milhão de diagramas (eléctricos, propulsão, etc...)
 - Após **actualização do software** de CAD

“Wow, this doesn’t look exactly like the drawing did before”, Brad Cumming





Preservação digital

Conjunto de
actividades ou **processos**
responsáveis por garantir o **acesso**
continuado a **longo-prazo** aos **objectos digitais** e
restante **património** cultural existente em
formatos digitais sem que seja
comprometida a sua
autenticidade



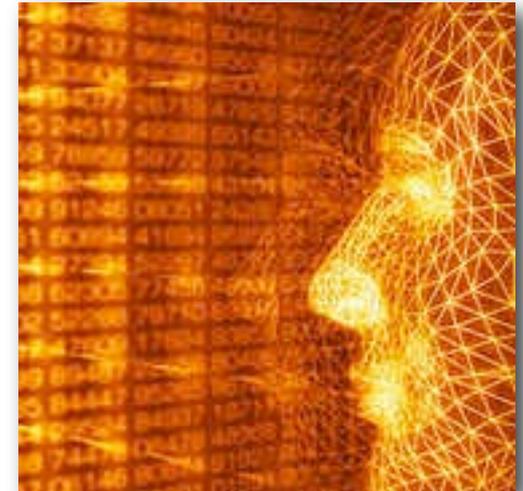
Objecto, representação ou recurso digital

Tudo aquilo que possa ser **representado através de uma sequência de dígitos binários**

- Nascido digital
- Digitalizado

▶ **Exemplos:**

- bases de dados
- folhas de cálculo
- documentos de texto
- páginas Web
- correio electrónico
- vídeo, audio, imagens
- modelos 3D
- software
- etc...



Analógico vs Digital



UNIVERSIDADE DO MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

TER24AGO
EDIÇÃO PORTO

27 de Agosto de 2004
ANO XV • Nº 1251
€0,80 (IVA incluída)
Diretor JOSÉ MANUEL FERRADARES
Secretário-Geral NUNO FERRADARES
e MANUEL CARVALHO
www.publico.pt

PÚBLICO

SINÓNIMOS E ANTÓNIMOS
Dicionário
COM MAIS 10 VOLUMES
HOJE PRECISAMOS VOLUME POR 0,3 EURO

TRÊS MILHÕES POR DISTRIBUIR ÀS VÍTIMAS DOS INCÊNDIOS DE 2003

Um ano depois dos combates de salitre, o Estado tem em mãos os 3 milhões de euros necessários para distribuir 3 milhões de euros às vítimas dos incêndios de 2003. A distribuição será feita em 12 parcelas de 250 mil euros cada, durante o ano de 2004. O governo da Associação Nacional de Municípios, Francisco Ruas.

Verdes das campanhas de solidariedade do ano passado podem ir para casos deste ano

Se um utilizador não estiver, "isto faz sentido que o dinheiro vá para os beneficiários. É preciso avaliar se a distribuição é mais eficaz com o dinheiro". O que se sabe de 2003 é que se trata de um ano de incêndios de casas.

CULTURA
Polícia à procura do DNA dos ladrões dos quadros de Manet

A polícia portuguesa está a analisar as molduras de "O Cristo" e "Madame", as últimas obras de Manet roubadas no mundo, à procura de impressões digitais. Já se fez a identificação do crime. A ministra da Cultura, Irina Cabaco, critica a falta de investigação crítica antes da compra.

Assessor de Rui Rio ganha mais que Jorge Sampaio

Finanças com débitos no reembolso de 40 milhões de IVA ao BCP

A Inspeção-Geral de Finanças nacionais denunciou a falta de transparência do reembolso de IVA ao BCP. O BCP não quis comentar o caso, mas considerou que, entre outros, se trata de diferentes interpretações da lei sobre a distribuição de valores.

FUTEBOL
Gilberto Madaladinite chamada de Derki à seleção nacional

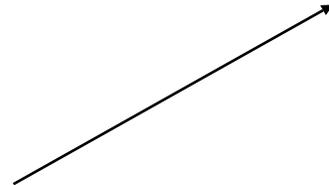
Britânica Kelly Holmes acaba com reinado de Maria Moutão nos 800 metros

Francis Obikwela recebeu a medalha de prata e disse cansado para os 200 metros

Edivaldo Monteiro sem ilusões para as meias-finais dos 400 metros barreiras

Como Angola defrontou os gigantes norte-americanos do basquetebol

Central Oklahoma's Premiere Shopper's Guide Since 1977



Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

Exclusividade

publicopt

Edição em PDF

2 de Maio de 2006 - 22h50

Edição Impressa Última Hora 10 Últimas Dossiers Guia do Lazer Cinecartaz Blogues Coleções

PLUS

Utilizador: [input type="text"]
[input type="password"] [input type="button" value="OK"]
Registo

HOME

CANAIS

- Edição Impressa
- Última Hora
- Última Hora UL
- Carga e Transportes
- Última Hora **15:30**
- Bolsa
- Desporto
- Consultórios
- Podium
- Guia do Lazer
- Cinecartaz

Protesto inflamado
Um apicultor sul-coreano salta para cima de uma bandeira japonesa e de uma imagem do primeiro-ministro Junichiro Koizumi, em Seul, Ahn

Dez detidos e 19 armas apreendidas no Bairro da Torre
Dez detidos e 19 armas de fogo apreendidas é o balanço provisório da operação policial hoje realizada no Bairro da Torre, em Camarate, segundo o Comando Metropolitano da PSP de Lisboa.

- › **Moradores do Bairro da Torre queixam-se da actuação da polícia**
- › **Operação policial no Bairro da Torre resultou em várias detenções**

China e Rússia opõem-se a aprovação de sanção vinculativa contra o Irão

Conselho de Segurança continua dividido sobre crise nuclear iraniana

- O meu analista avisou-me

ESCREVA AO PROVEDOR DOS LEITORES

Mail Site Blog

Find: paper Find Next Find Previous Highlight all Match case Phrase not found

http://www.publico.clix.pt/sites/blogues/

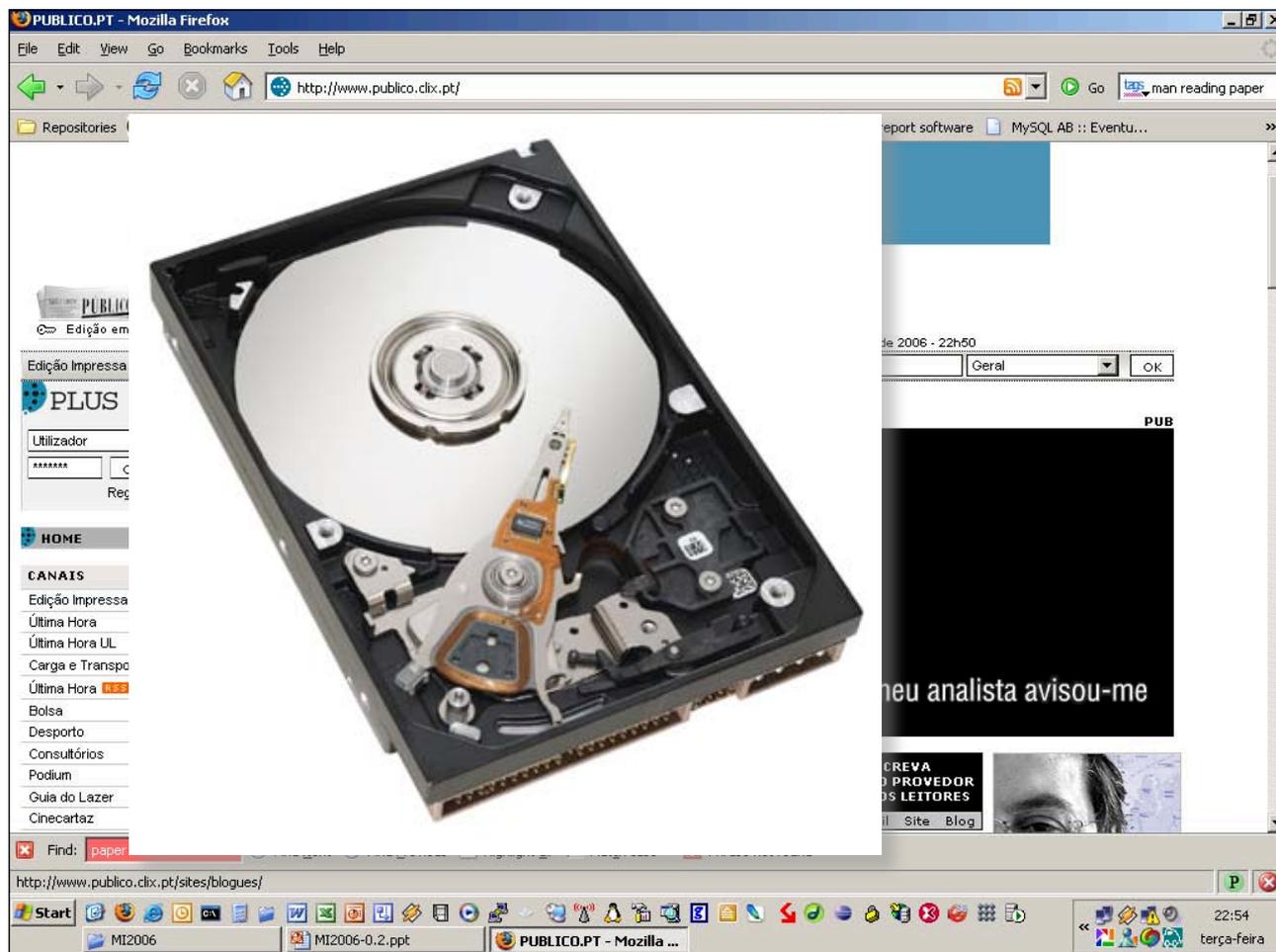
Start MI2006 MI2006-0.2.ppt PUBLICO.PT - Mozilla ... 22:54 terça-feira

Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos



Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

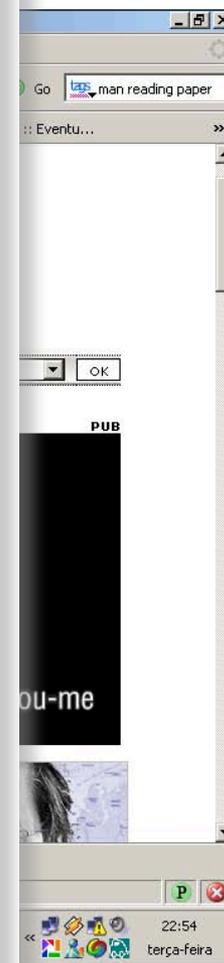
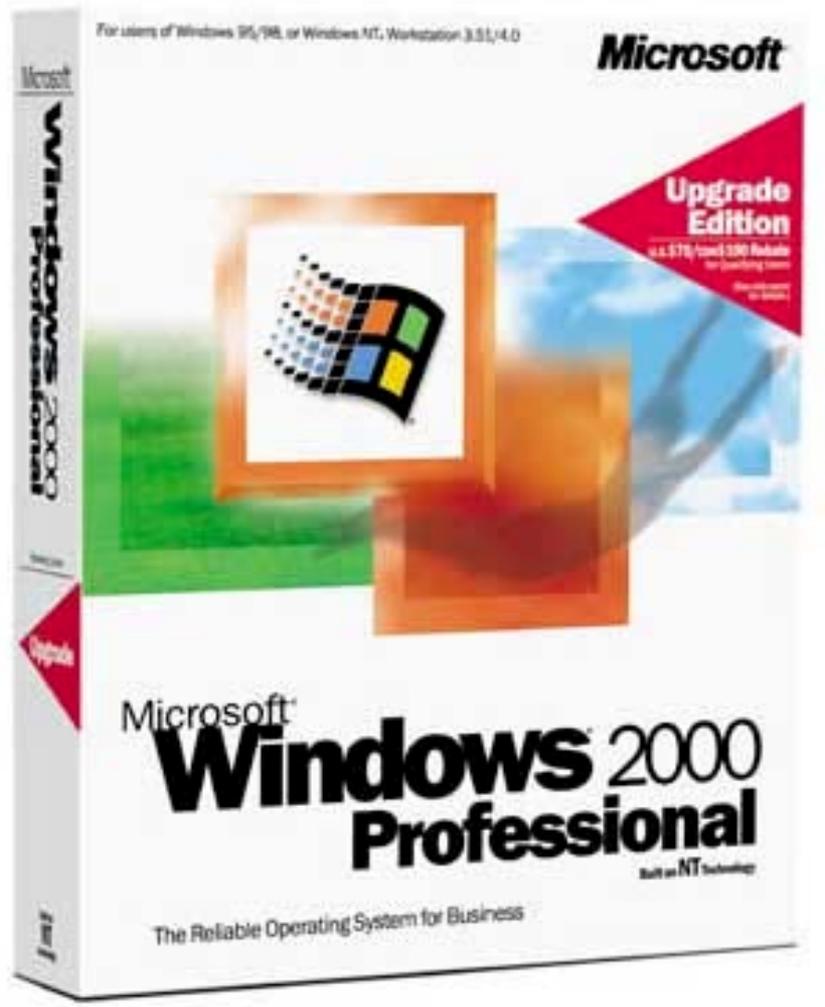


Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos



Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos



Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos



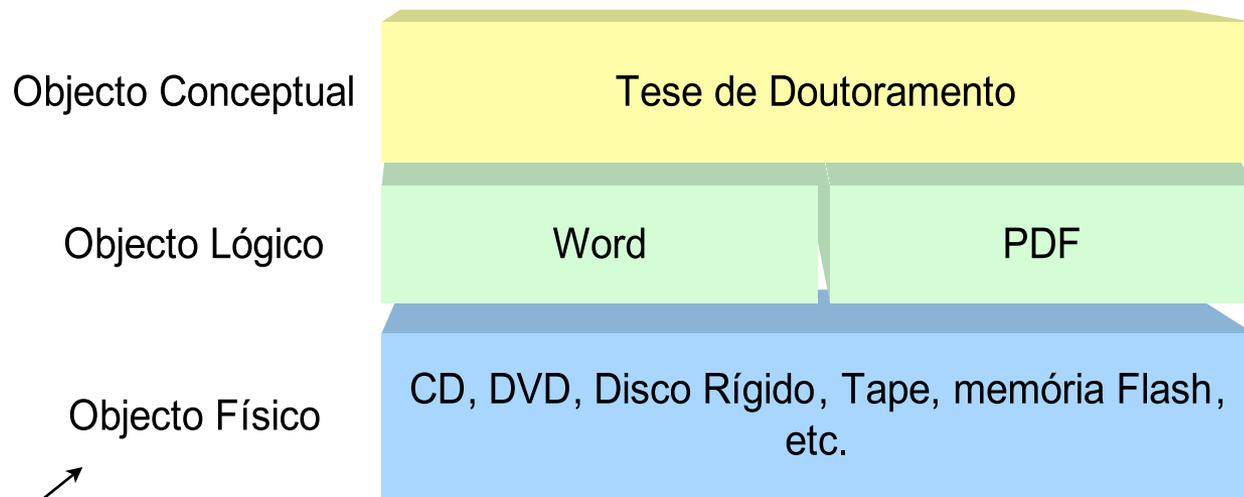
Analógico vs Digital



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

A anatomia de um objecto digital



Não confundir
com o objecto
físico que foi
digitalizado

Se um dos níveis se tornar obsoleto o objecto conceptual deixa de estar acessível



Estratégias de preservação



Preservação de tecnologia

- ▶ **Preservar** a informação nos seus **formatos** lógicos e/ou físicos **originais**
- ▶ Utilização da **tecnologia originalmente associada** a estes **objectos** para garantir o acesso aos mesmos
- ▶ Foco no objecto **físico/lógico**



Preservação de informação

- ▶ Preservar as **características essenciais** dos objectos de forma **independente** do hardware e do software
- ▶ Foco no objecto **conceptual**



Refrescamento

Transferência

de **informação** de um **suporte físico** de armazenamento **para outro mais estável** antes que o primeiro se deteriore ou se torne **irremediavelmente obsoleto**



Não é tanto uma estratégia, é mais uma **precondição** necessária à preservação!



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

Estratégias de preservação

Preservação de tecnologia



Preservação de informação





UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

Estratégias de preservação

Preservação de tecnologia





Preservação de tecnologia

▶ **Museus** de hardware

- **Preservação** do **contexto** de criação/interpretação **original**
- Dificuldade em **manter o hardware operacional** a longo-prazo (oneroso)
- **Restrições** ao nível do **acesso** e **reutilização** da informação

▶ **Emuladores**

- **Software** capaz de **reproduzir** o **comportamento** de uma **plataforma** de hardware e/ou software
- A **construção de emuladores é um processo complexo** e muito dependente da publicação de informação por parte do fabricante de hardware ou software
- **Não garante a reutilização da informação**
- Necessidade de saber **operar tecnologias** há muito **desaparecidas**
- **Obsolescência do próprio emulador**

Emulação [exemplo]



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

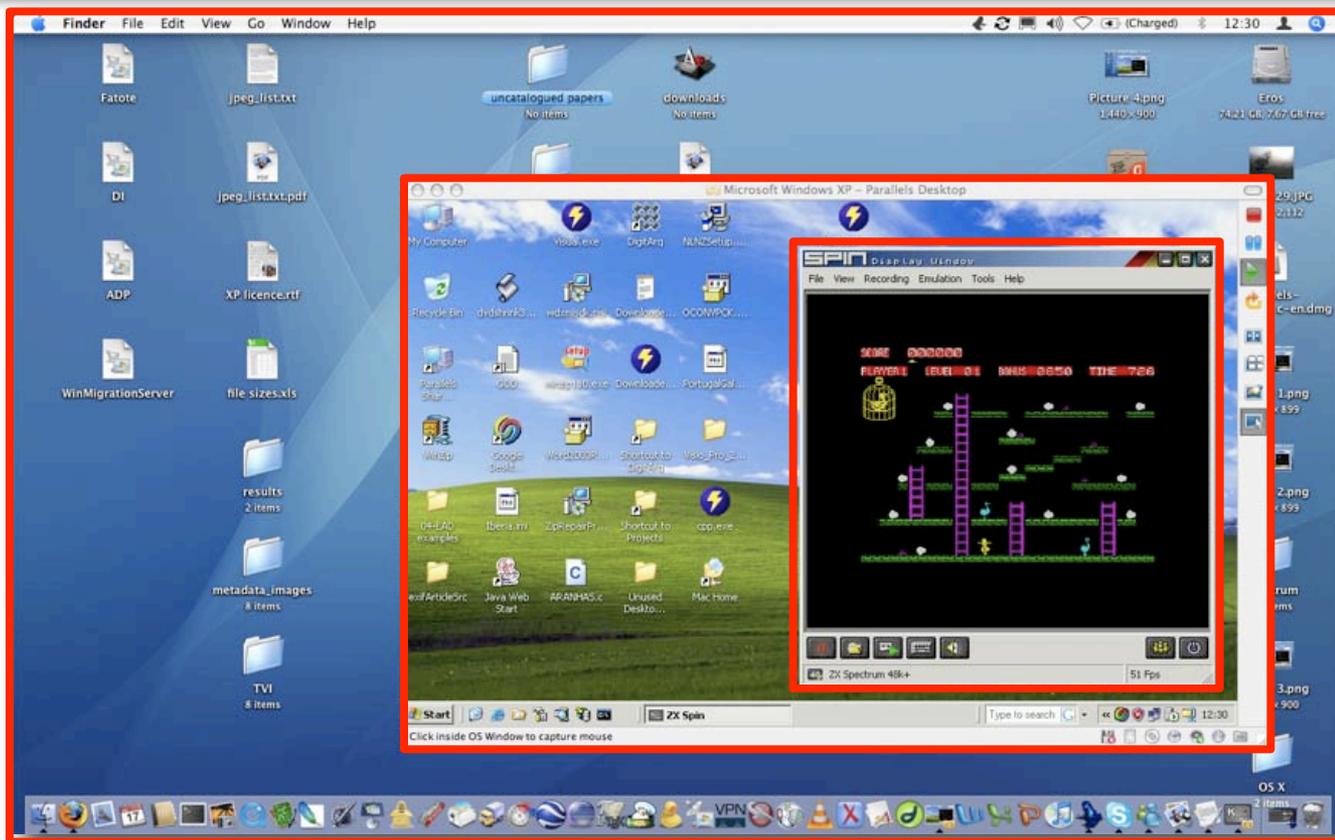




UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

Emulação [exemplo]





UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direção-Geral de Arquivos

Estratégias de preservação

Preservação de
tecnologia



Preservação de
informação





UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

Estratégias de preservação

Preservação de informação





Migração de formatos

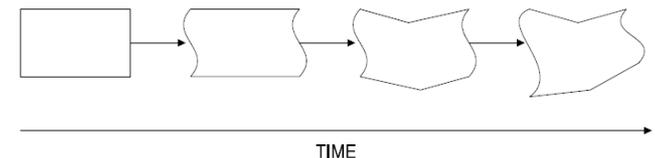
Transferência periódica de **material** digital de uma **configuração** de hardware/software para outra mais actual

▶ Vantagens

- **Disseminação** dos materiais em **formatos actualizados** que os consumidores são capazes de reconhecer
- **Sem necessidade** de preservar o **software original**
- A estratégia **mais aplicada** na maioria dos domínios
 - e.g. doméstico, organizacional, etc.

▶ Desvantagens

- Possível **perda de informação** durante a conversão
- Necessidade de **diligência contínua**
- **Dispendiosa** a longo-prazo





Outras estratégias

- ▶ **Encapsulamento**
- ▶ **Normalização** para formatos de preservação
- ▶ Migração **a-pedido**
- ▶ **Serviços** de preservação
 - Identificação de formatos
 - Migração distribuída
 - Planeamento de preservação
 - Notificação de obsolescência



Boas práticas de preservação

- ▶ Utilizar um sistema de gestão de informação (e.g. **repositório digital**)
- ▶ Definir um conjunto de **políticas de preservação** de nível:
 - Físico
 - Lógico
 - Conceptual
 - Social
 - Económico
 - Organizacional



Políticas de preservação

- ▶ **Nível físico**
 - Escolha de **tecnologia de armazenamento** adequada
 - multi-vendedor,
 - multi-plataforma,
 - resistente a situações adversas (e.g. temperatura, humidade)
 - **Política de refrescamento** de suportes
 - **Plano de contingência**
 - recuperação de dados em situação de desastre
- ▶ **Nível lógico**
 - **Replicação local** (backups)
 - **Replicação remota**



Políticas de preservação

- ▶ **Nível conceptual/semântico**
 - **Seleccção de formatos de preservação**
 - **Normalização** de formatos durante a incorporação
 - **Migração** de formatos obsoletos para formatos mais actuais

- ▶ **Nível social**
 - **Regras de acesso aos dados**
 - Todos os **utilizadores** deverão estar **autenticados**
 - Toda a sua **actividade é registada**
 - O sistema deve **evitar** a ocorrência de **erros humanos**
 - Definição de esquemas de **metainformação**
 - **descritiva, técnica** e de **preservação** que garantam o **acesso** e a **autenticidade** dos objectos
 - Direitos de autor e propriedade intelectual



- ▶ **Nível económico**
 - Plano de negócios para **sustentabilidade económica** do repositório a médio e longo-prazo
 - **Custos** de preservação deverão fazer parte do **orçamento anual** da instituição preservadora

- ▶ **Nível organizacional**
 - Política de **avaliação e selecção**
 - Definição de **limites de responsabilidade**
 - Definição de um **plano de sucessão**



Características de um formato de preservação

- ▶ **Disseminação do formato**
- ▶ Elevado nível de suporte
- ▶ **Standard** (gerido por um organismo normativo)
- ▶ **Especificação aberta**
- ▶ Ausência de compressão ou compressão sem perdas
- ▶ Isento de royalties
- ▶ **Retrocompatibilidade**
- ▶ Sem possibilidade extensões não normalizadas
- ▶ Idade (e.g. TIFF)
- ▶ Facilidade de descodificação
- ▶ **Suporte multiplataforma**
- ▶ **Software** de leitura/escrita em **código-aberto**



Notas finais

- ▶ **Não existe uma solução única** e adequada a todos os tipos de material
- ▶ A decisão sobre que estratégia de preservação adoptar **depende de vários factores**:
 - Simplicidade de implementação a curto e a longo-prazo
 - Custo
 - Restrições legais
 - Satisfação da comunidade de interesse
 - Tipo de material
- ▶ **Emulação** versus **Migração**
 - **Emulação** para preservação de objectos **dinâmicos/interactivos** (e.g. software, animações flash)
 - **Migração** para a preservação dos restantes **objectos digitais** (e.g. documentos texto, imagens, bases de dados)



Notas finais

▶ Autenticidade

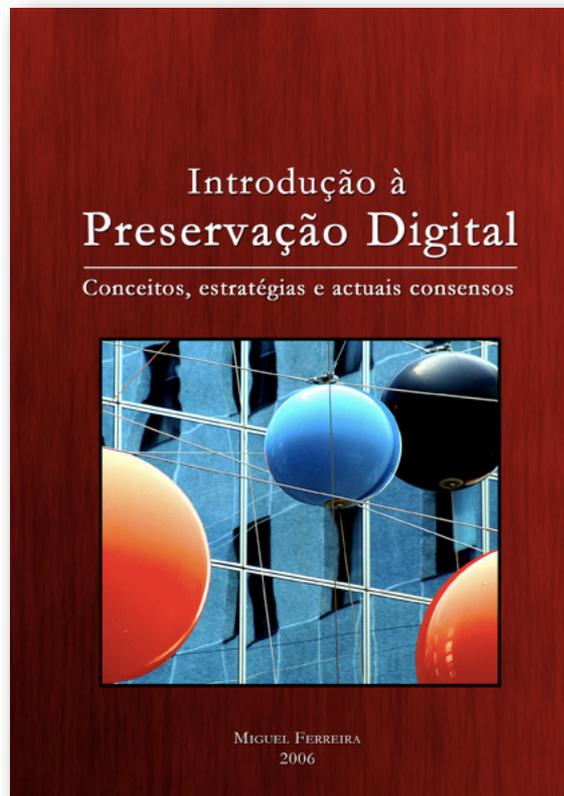
- Capacidade de **demonstrar** que um objecto **é aquilo que se propõe ser**
- Capacidade de **identificar** e apresentar os **elementos diplomáticos** que permitem aferir se um objecto é autêntico
- Metainformação de preservação - **PREMIS**



UNIVERSIDADE
DO
MINHO

Seminário de Preservação Digital
Direcção-Geral de Arquivos

... para mais informação...



Disponível em

<http://hdl.handle.net/1822/5820>