

28 de Março de 2008  
Braga



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# Preservação Digital

---

## das normas às boas práticas

Miguel Ferreira  
[mferreira@dsi.uminho.pt](mailto:mferreira@dsi.uminho.pt)  
<http://www.dsi.uminho.pt/~ferreira>

José Carlos Ramalho  
[jcr@di.uminho.pt](mailto:jcr@di.uminho.pt)  
<http://www.di.uminho.pt/~jcr>





# Conteúdo

- ▶ **Contextualização**
- ▶ **Exemplos** de obsolescência tecnológica
- ▶ **Conceitos**/definições
- ▶ **Estratégias** de preservação
  - Preservação de tecnologia
  - Preservação do objecto conceptual
- ▶ **Projecto RODA**
- ▶ Normas e boas práticas

# Contextualização

Domínio físico ou analógico  
*“das coisas tangíveis”*

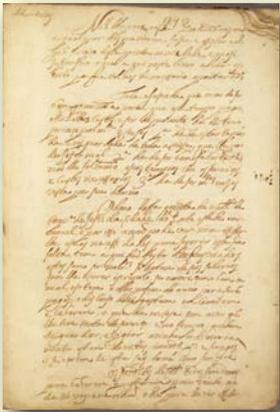
# Contextualização

## Domínio físico ou analógico “das coisas tangíveis”



# Contextualização

Domínio físico ou analógico  
“das coisas tangíveis”

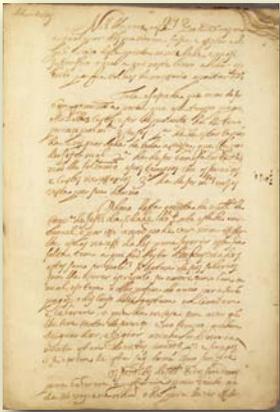


Digitalização



# Contextualização

Domínio físico ou analógico  
*“das coisas tangíveis”*



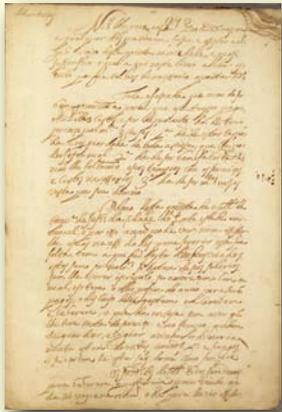
Digitalização

Domínio digital



# Contextualização

## Domínio físico ou analógico “das coisas tangíveis”



## Domínio digital





# 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
- ▶ Rápida **obsolescência** da tecnologia
  - Depende de um **contexto tecnológico** para que possa ser interpretado
  - 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 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**



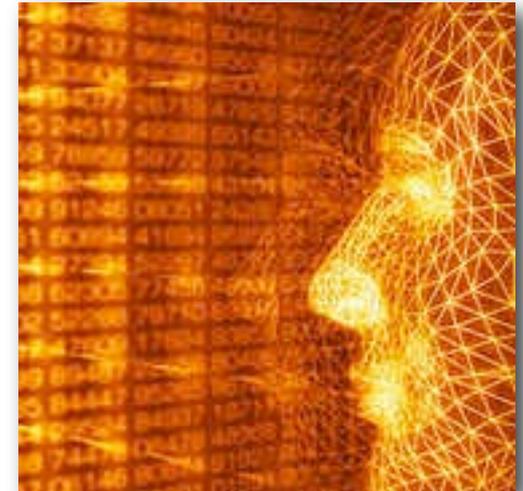
# Objecto, representação ou recurso digital

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

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

**TER24AGO**  
EDIÇÃO PORTO  
24 de Agosto de 2004  
Ano XV • Nº 1251  
450 000 VOTA ANUALMENTE  
Diretor JOSÉ MANUEL FERREIROS  
Secretaria-geral NUNO FERRAZES  
e MANUEL CARVALHO  
www.publico.pt

**SINÓNIMOS E ANTÓNIMOS**  
Dicionário  
COM MAIS 10 VOLUMES  
QUE PRIMEIRO VOLUME FOR O A TERCEIRO

**PÚBLICO**  
www.publico.pt

**TRÊS MILHÕES POR DISTRIBUIR  
ÀS VÍTIMAS DOS INCÊNDIOS DE 2003**

Um ano depois das campanhas de solidariedade para com as vítimas dos incêndios florestais de 2003, ainda restam por ser pagos 3,2 milhões de euros dos 3 milhões de euros prometidos. O presidente da Associação Nacional de Municípios, Fernando Ruas,

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

Em uma entrevista aos autarcas, "vão ter sentido que o dinheiro cobria a terra e não os mortos". É preciso trabalhar para uma situação mais discreta com o cidadão. "O que sobra de 2003 reparte-se para as vítimas do Incêndio de 2004", afirma.

**LE CARRONNIER DE LA JUSTICE  
RESCOURRONS LA JUSTICE**

**A jornalista  
de Pinson Penh**  
P1813

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

A polícia portuguesa está a analisar as moléculas de "O Crime" 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 inspecciona criticamente a falta de investigação, mas

**Assessor de Rui  
Rio ganha mais que  
Jorge Sampaio**

**PROPOSIÇÕES**  
**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 pedido pelo Banco BCP e pagou em 2003 pela administração fiscal, no valor aproximado de 40 milhões de euros, segundo o ICFE. O BCP não quis comentar o caso, mas consideramos que, neste contexto, se trata de diferentes interpretações da lei sobre a distribuição de responsabilidades. p184

**FUTEBOL**  
**Gilberto Madal  
admitido chamado de Derki  
à seleção nacional**  
p184

**BRITÂNICA KELLY  
Holmes acaba  
com reinado  
de Maria Montoi  
nos 800 metros**

**FRANCIS OBIKWELA  
recebeu a medalha  
de prata e diz-se  
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**

**Reportagem  
de uma  
partida no  
Jornal. Lado  
americano e  
Brasil. Foto  
P. A. M.**

**A festa de vitória  
no Estádio Olímpico**



# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

**TER24AGO**  
EDIÇÃO PORTO  
24 de Agosto de 2004  
ANO XV • Nº 1251  
€0,80 (IVA incluída)  
Director JOSÉ MANUEL FERREIROS  
Secretaria-geral NUNO FERRAZ  
e MANUEL CARVALHO  
www.publico.pt

## PÚBLICO

SINÓNIMOS E ANTÓNIMOS  
Dicionário  
COM MAIS 10 VOLUMES  
HOJE PRECISAMOS VOLTAR PÁG. 33

### 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 de 2003 destinados para os 125 milhões de euros dos 100 milhões de vítimas. O presidente da Associação Nacional de Municípios, Fernando Ruas,

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

Se um utilizador não autoriza, "isto faz sentido que o dinheiro destinado seja usado para ajudar a quem precisa". O que se sabe de 2003 é que se trata de uma vítima dos incêndios de 2003.

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

A polícia portuguesa está a analisar as moléculas de "O Cristo" e "Madalena", as últimas obras de Manet roubadas no mundo, à procura de impressões digitais. Já se fez a recolha de impressões digitais. A ministra da Cultura, Irina Cabaco, critica a falta de investigação crítica antes da aquisição.

**Assessor de Rui Rio ganha mais que Jorge Sampaio**

**Fimancas com débitos no reembolso de 40 milhões de IVA ao BCP**

A Inspeção-Geral de Finanças multou o Banco Comercial Português por um reembolso de IVA pedido pelo 3.º Milionário BCP por valor de 40 milhões de euros, apesar de o BCP não ter apresentado o caso, mas considerou que, entre outros, se trata de diferentes interpretações da lei sobre a distribuição de responsabilidades.

**FUTEBOL**  
**Gilberto Modest admite chamada de Derki à selecção nacional**

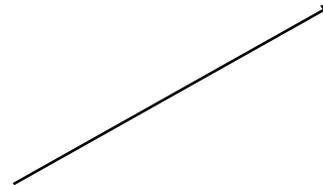
**Britânica Kelly Holmes acaba com reinado de Maria Mutola nos 800 metros**

**Francis Obikwela recebeu a medalha de prata e diz-se 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**

A fonte da notícia, no Estado Olímpico.



Central  
Oklahoma's  
Premiere  
Shopper's Guide  
Since 1977

# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

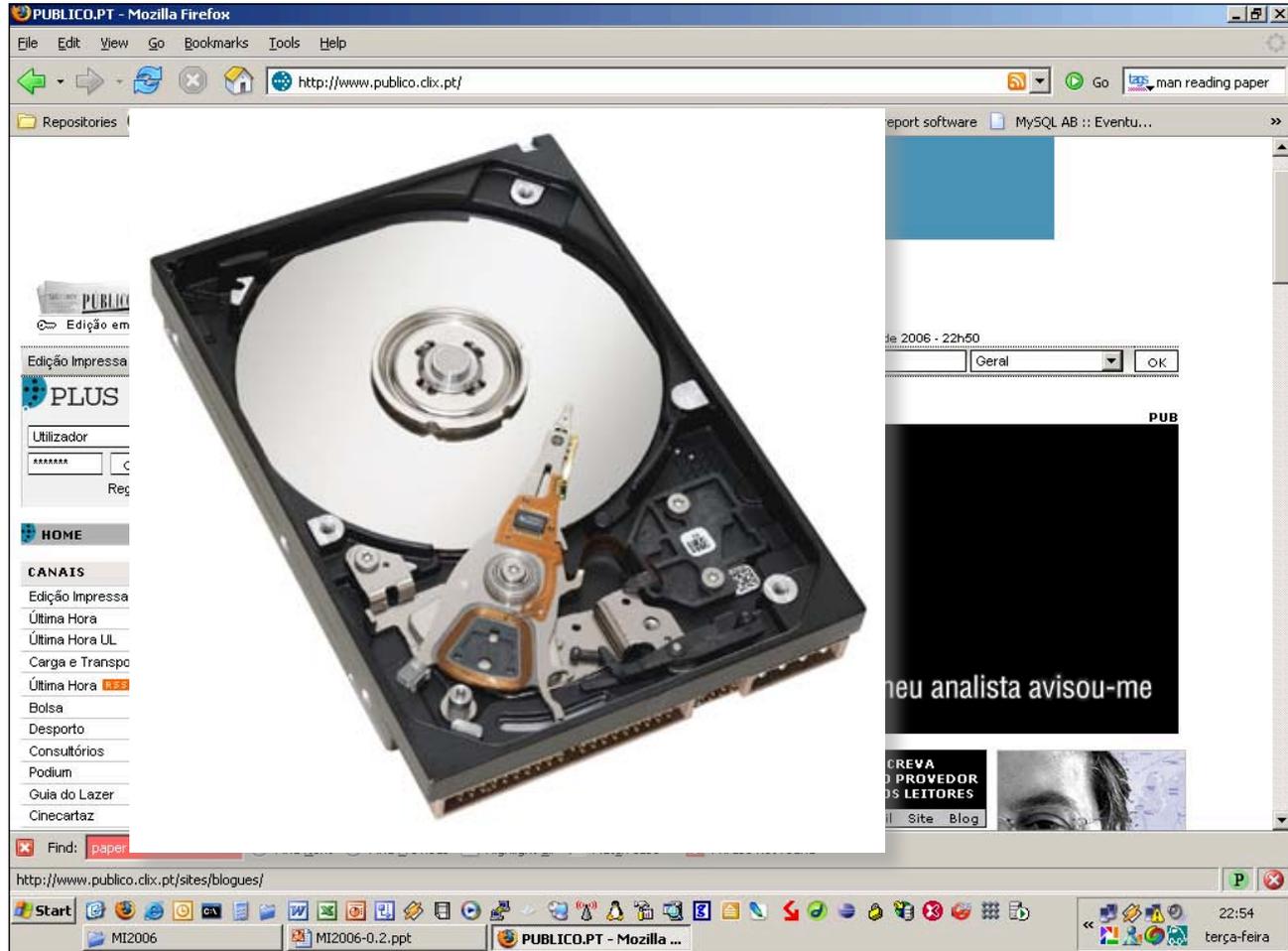


# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

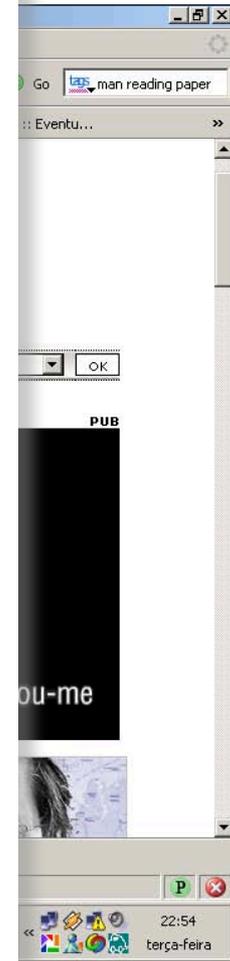


# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

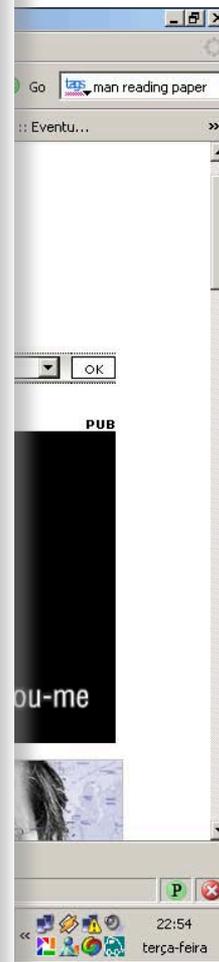
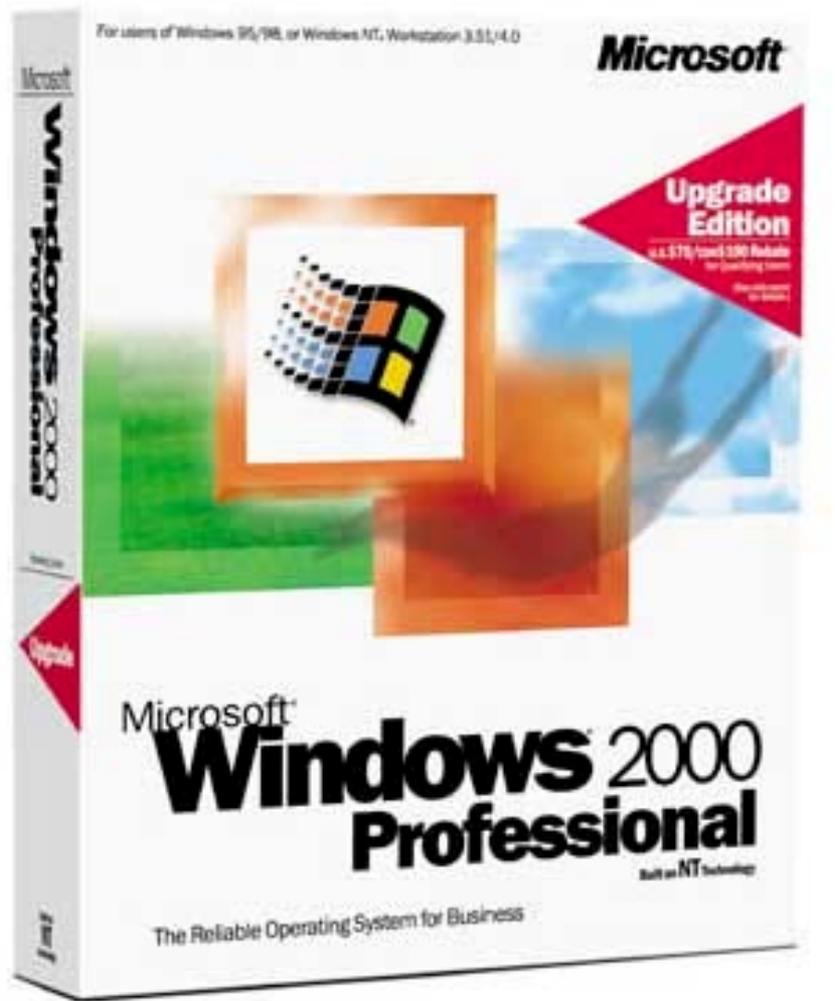
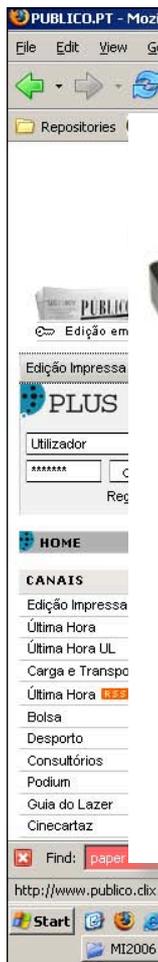


# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's





UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# Analógico vs Digital



# Analógico vs Digital



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's



# Analógico vs Digital

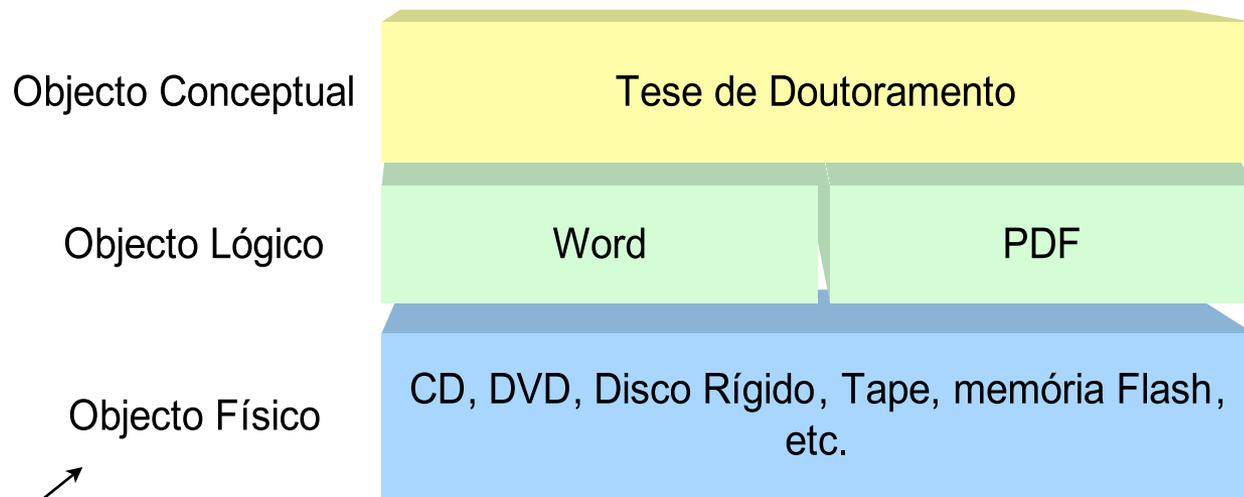


UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's



# 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 da 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 para a preservação!



UNIVERSIDADE  
DO  
MINHO

# Estratégias de preservação

## Preservação de tecnologia



## Preservação da informação



Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# 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**



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# Emulação [exemplo]

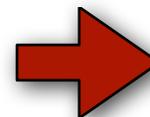
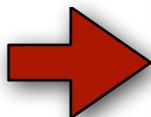
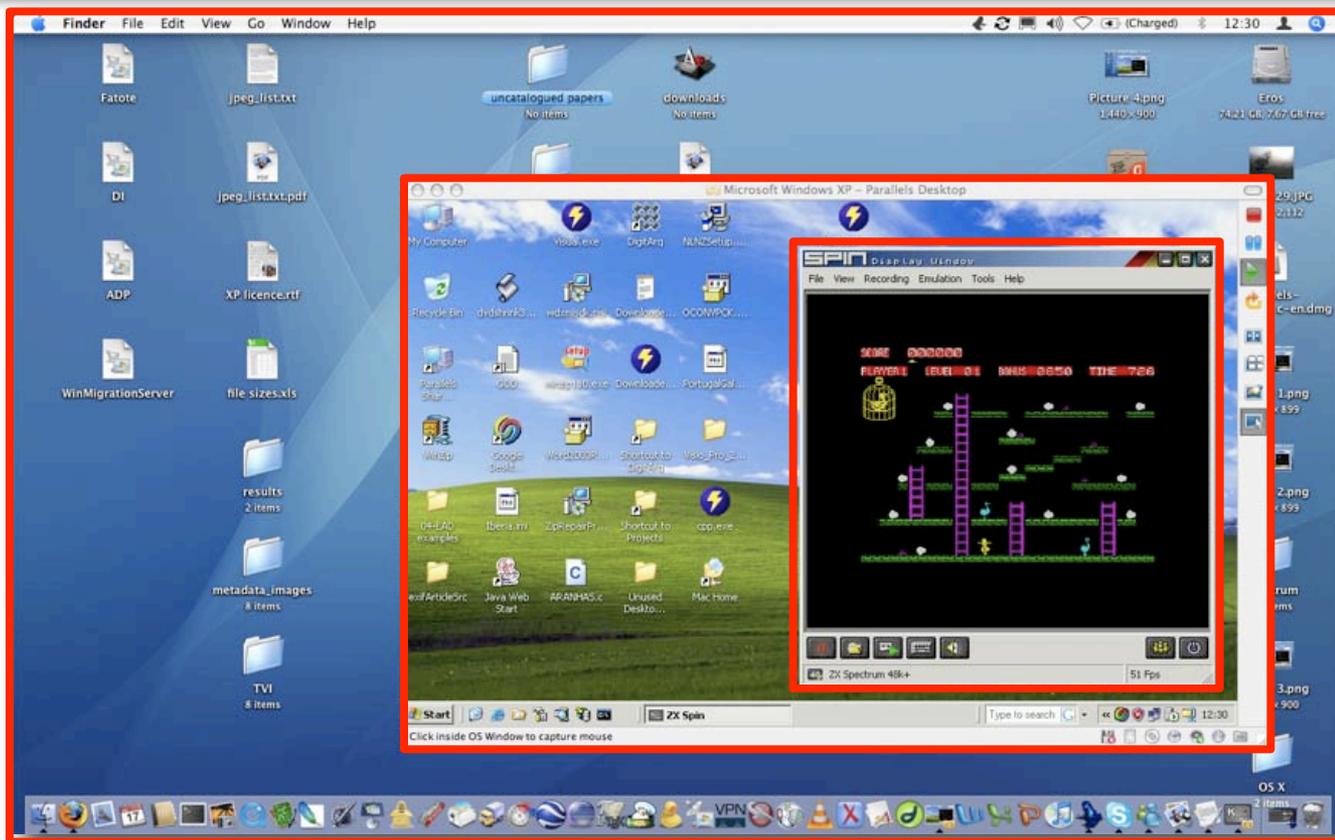




UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# Emulação [exemplo]





UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# Estratégias de preservação

Preservação de  
tecnologia



Preservação da  
informação





UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

16

# Estratégias de preservação

## Preservação da informação





# Migração de formatos

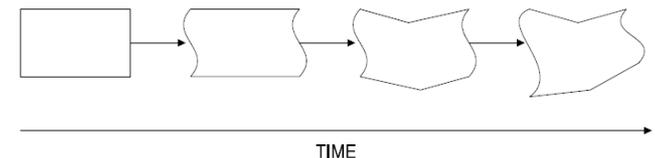
**Transferência** periódica de **material** digital de uma **configuração** de hardware/software para outra mais actual

## ▶ 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 continua**
- **Dispendiosa** a longo-prazo





# Outras estratégias

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



# Características de um formato de preservação

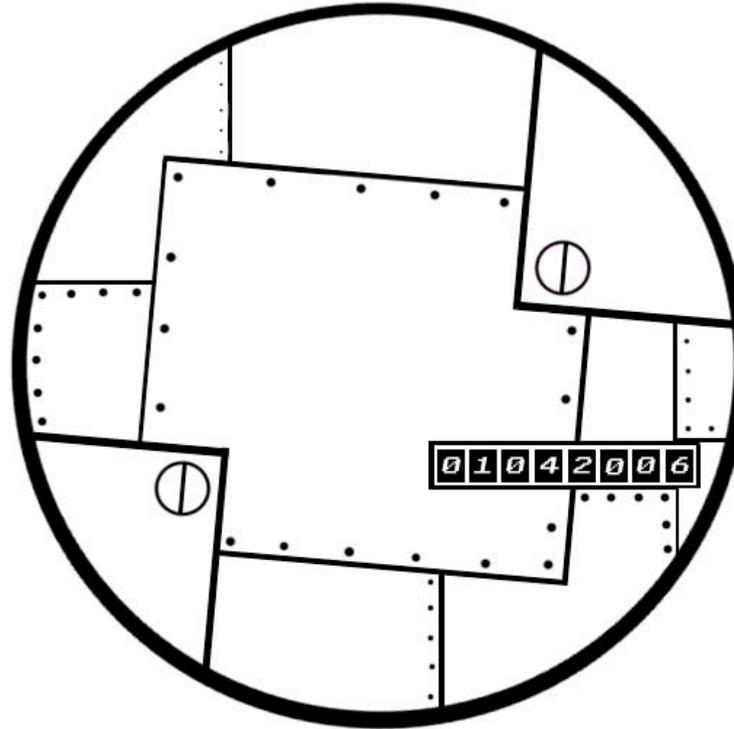
- ▶ **Ubiquidade**
- ▶ Elevado nível de suporte
- ▶ **Standard**
- ▶ **Especificação aberta**
- ▶ Compressão
- ▶ Isento de royalties
- ▶ **Retrocompatível**
- ▶ **Estável em termos de versões**
- ▶ Suporta extensões
- ▶ Idade (e.g. TIFF)
- ▶ Fácil de decodificar
- ▶ **Multiplataforma**
- ▶ **Software** de leitura/escrita em **código-aberto**



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

# RODA



## Repositório de Objectos Digitais Autênticos





# O que é?

- ▶ Projecto de **I&D** iniciado em 2006
  - Final do desenvolvimento previsto para fins de **2008**
  - Parceria entre a DGARQ e a UMINHO
- ▶ **Repositório** digital
- ▶ **Guardar, preservar, dar acesso** a objectos digitais
  - Imagens,
  - Documentos de texto,
  - Bases de dados relacionais
- ▶ Preservação por **longos períodos** de tempo
  - Na ordem das dezenas de anos
- ▶ Assegurar o **valor probatório** da informação custodiada

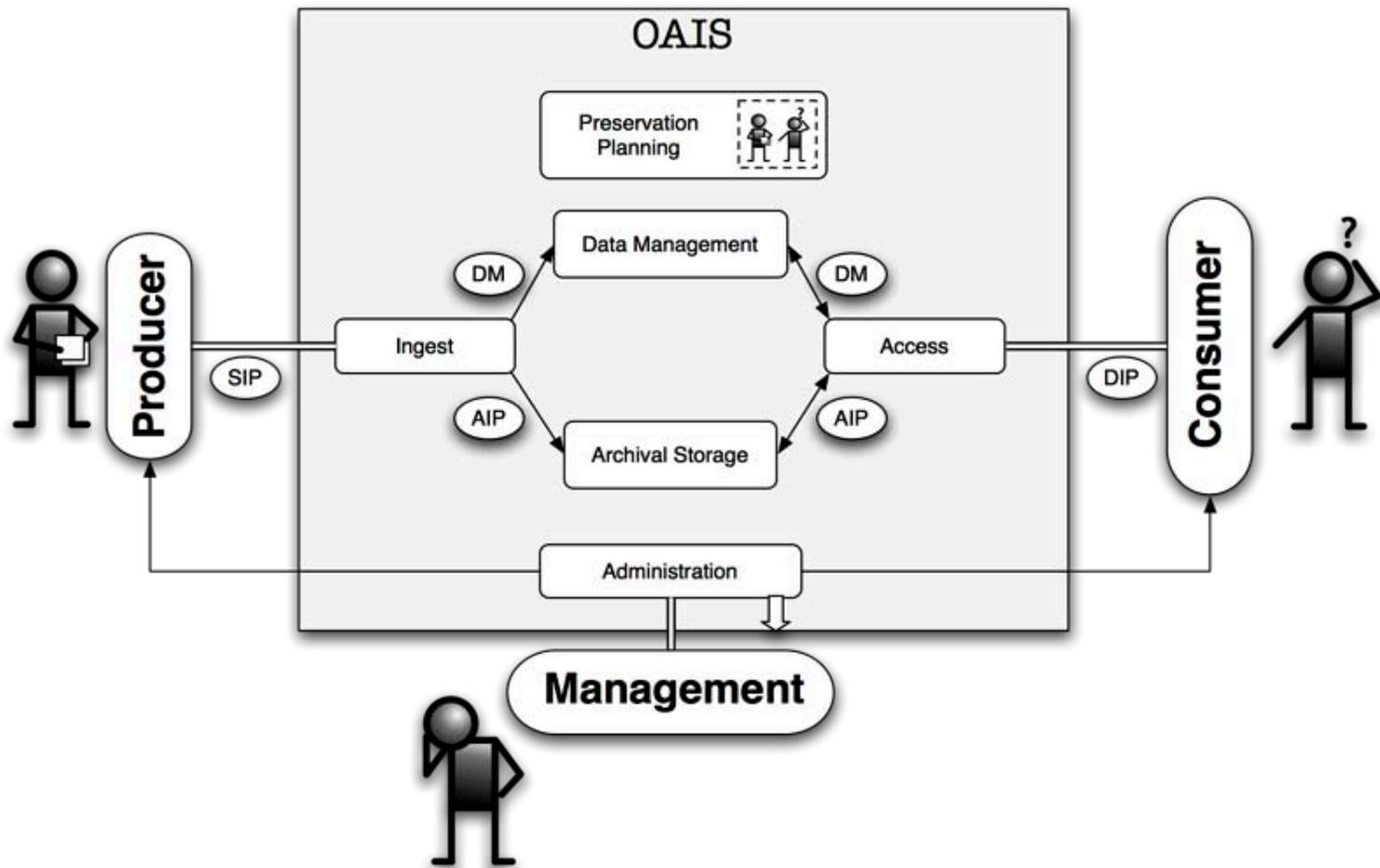
# A quem se destina?

- ▶ Organismos **públicos**
  - clientes privilegiados
- ▶ Organismos **privados**
  - estabelecimento de contratos de incorporação
- ▶ **Arquivos pessoais** - Não nesta fase!
- ▶ Todos os organismos que queiram **preservar objectos digitais produzidos** no decorrer da sua **actividade**
  - Valor **patrimonial**
  - Valor **probatório**
  - Incorporação **obrigatória**





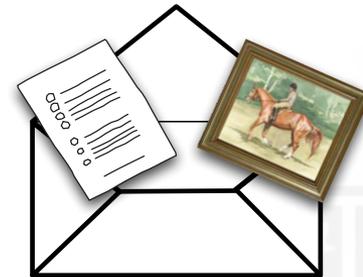
# Arquitectura geral





# Ingestão

- ▶ **Contrato pré-ingestão**
  - Arquivo e o produtor acordam
    - a **informação** a incorporar
    - os **formatos** de ingestão
    - a estrutura do **SIP** (informação de depósito)
- ▶ **Produtor prepara SIPs**
  - Metainformação **descritiva**
  - **Representações** digitais
    - nos formatos adequados
  - Envelope XML (**METS**)
- ▶ O RODA disponibiliza uma **ferramenta para construir SIPs**





# Ingestão

## ▶ **Validação automática do SIP**

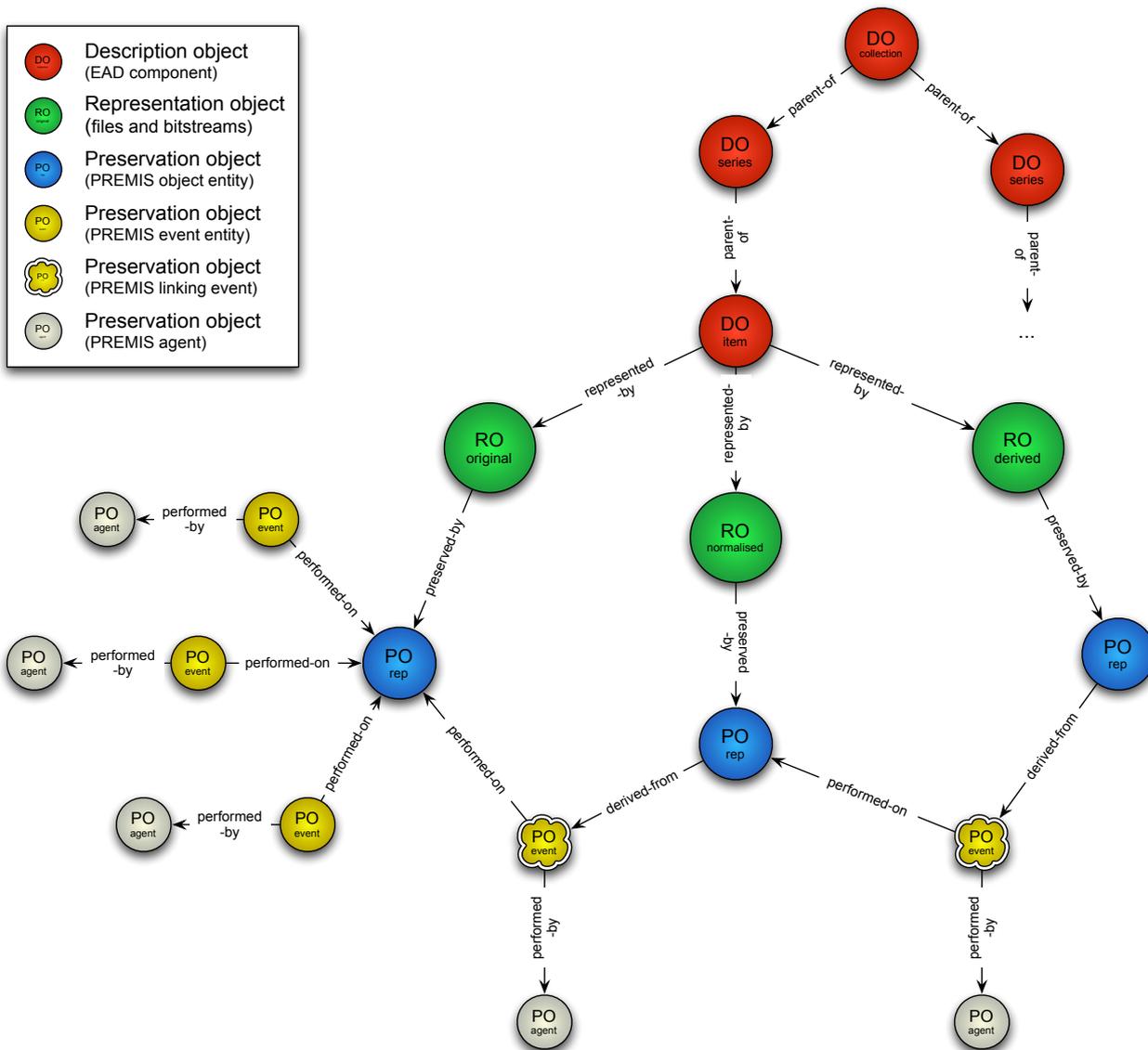
- verificação de **vírus...**
- verificação de **integridade** do **SIP...**
- verificação de **integridade** dos seus **constituintes...**
- geração de **MI técnica...**
- geração de **MI preservação** (PREMIS)...
- normalização de **formatos...**

## ▶ **Validação manual da SIP**

- sujeitos a escrutínio arquivístico
- aceitação/rejeição



# Modelo de dados atómico



# Autenticidade

- ▶ Garantir que os **objectos** semânticos **não foram adulterados** e que se mantêm de acordo com o original
  - História custodial/**Proveniência**
  - Verificação sistemática da **integridade**
  - **Documentação** de todos os processos de preservação
  - Análise das **propriedades significativas** após uma migração
  - Estabelecer um **contexto de confiança** em torno dos objectos





# Acesso

- ▶ **Pesquisa** básica
- ▶ Pesquisa **avançada**
- ▶ **Navegação** nos fundos documentais
- ▶ **Visualização** das representações
- ▶ **Descarregamento** (*download*)



# Screenshot

REPOSITÓRIO DE OBJECTOS DIGITAIS AUTÊNTICOS

# RODA

Bem-vindo visitante

Nome de utilizador

Palavra-passe

[Registar](#) · [FAQ](#) · [Ajuda](#)

[SOBRE O RODA](#)  
[CATÁLOGO](#)

*catálogo / navegar fundos*

## NAVEGAR FUNDOS

3 fundos

- [PSS-Francisco](#)
- [PSS-MFERREIRA](#)
  - [001](#)
  - [002](#)
  - [003](#)
  - [004](#)
  - [0001](#)
- [RODA](#)
  - [APRESENTAÇÕES](#)
  - [RELATÓRIOS](#)

**Descrição**

Identificação	
Referência	PSS-MFERREIRA
Título	Fundo pessoal de Miguel Ferreira
Nível de descrição	fundo
Data inicial	1979-06-09
Data final	2007-01-01
Código do país	PT
Código do repositório	DGARQ
Produtor	Miguel Ferreira

Contexto	
História administrativa e biográfica	Miguel Ferreira graduated as a Systems and Informatics Engineer and has worked as a consultant at the OPorto's Regional Archive (Arquivo Distrital do Porto) and as a researcher at the University of Minho. Since 2003, has been publishing in the fields of digital archives, libraries and preservation. Currently, Ferreira is developing work as a Ph.D. student and coordinating several research projects at the OPorto's Regional Archive and the

[FECHAR](#)



# Processos de preservação

- ▶ **Nível físico**
  - Tecnologia de **armazenamento de elevada disponibilidade** baseada em NAS
  - **Refrescamento** de suportes (hot-swap)
  - **Plano de contingência** (recuperação perante desastres)
- ▶ **Nível lógico**
  - **Replicação local** (backups)
  - **Replicação remota**
- ▶ **Nível conceptual/semântico**
  - **Normalização** de formatos durante a ingestão
  - **Migração** de formatos obsoletos



# Processos de preservação

- ▶ **Nível social**
  - Todos os **utilizadores** têm de estar **autenticados**
  - Toda a sua **actividade é registada**
  - O sistema procura **evitar** a ocorrência de **erros humanos**
- ▶ **Nível económico**
  - Modelo de **sustentabilidade económica** para o repositório
  - Custos de preservação deverão fazer parte do **orçamento anual** da DGARQ
- ▶ **Nível organizacional**
  - Definição de um **plano de sucessão**

# Normas suportadas



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

- ▶ **ISADg/EAD**
  - Encoded Archival Description
  - Metainformação **descritiva**
- ▶ **METS**
  - Metadata Encoding and Transmission Standard
  - Metainformação **estrutural**
- ▶ **OAIS**
  - Open Archival Information System
  - **Componentes funcionais** do sistema e identificação dos **actores**

# Normas suportadas



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

- ▶ **PREMIS** Data Dictionary
  - Metainformação de **preservação**
- ▶ **ANSI/NISO Z39.87**
  - **Metainformação técnica** sobre imagens digitais
- ▶ **TRAC**
  - Trustworthy Repositories Audit & Certification: Criteria and Checklist
  - Auditoria e **auto-avaliação**
- ▶ **DRAMBORA**
  - Digital Repository Audit Method Based on Risk Assessment
  - **Certificação** de repositórios digitais



# Formatos de preservação

- ▶ Documentos de texto
  - PDF/A
- ▶ Bases de dados relacionais
  - DBML/XML
- ▶ Imagens digitais
  - TIFF não comprimido



### Navegação

- Notícias
- Sobre o RODA
- Disseminação
- Ferramentas e normas
- Projectos associados
- Eventos
- Fóruns
- Votações
- Equipa & contactos

## Society of Archivists 2007 Conference

Submetido por [lfaria](#) em Quarta, 18/10/2006 - 10:57 :: [Arquivística](#)

Start: 28/08/2007 - 08:00

End: 31/08/2007 - 17:00

Timezone: Etc/GMT

The Society of Archivists is hosting a major international conference in 2007 - Differing Directions: Challenging Communities

The Conference will be held at Queen's University Belfast, Northern Ireland on 28 - 31 August 2007. The Society of Archivists (UK and Ireland) covers archivists, records and information management, and archive conservators - and the programme will include three parallel sessions relating to all these professional areas.

For more information see the [Call for Papers](#)

[calendar](#)

## Participação do RODA no 4<sup>a</sup> Congresso Nacional de Administração Pública

Submetido por [rcaastro](#) em Terça, 17/10/2006 - 16:38 :: [Arquivística](#)

Start: 02/11/2006 - 07:00

Timezone: Etc/GMT+1

A Administração Pública portuguesa está a viver dias de mudança profunda. Mudança

A Administração Pública portuguesa está a viver dias de mudança profunda. Mudança

Timezone: Etc/GMT+1

Start: 02/11/2006 - 07:00

Submetido por [rcaastro](#) em Terça, 17/10/2006 - 16:38 :: [Arquivística](#)

## Administração Pública

## Participação do RODA no 4<sup>a</sup> Congresso Nacional de

### Idiomas

[Português](#)

[English](#)

### Próximos eventos

[Digital rights and asset management: access and digital preservation forum](#) (evento)  
(10 horas)

[mais](#)

### Novos tópicos no fórum

[Questionario sobre PREMIS e EAD no DSpace](#)

[Requisitos do RODA](#)

[Criação de uma identidade para o grupo](#)

[Colecções fotográficas: como estruturá-las?](#)

[Abertura do forum...](#)

[mais](#)



UNIVERSIDADE  
DO  
MINHO

Workshop sobre representação  
digital de objectos com baixo relevo usando  
PTM's

... para mais informação...

Introdução à  
**Preservação Digital**

Conceitos, estratégias e actuais consensos



MIGUEL FERREIRA  
2006

**Disponível em**

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