



# Past, present, future



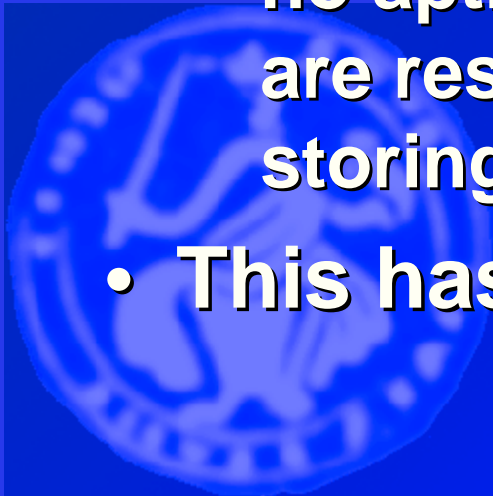
***Digital Curation Seminar***

***19 October 2001***

***David Ryan***

# The past

- Logical issues avoided or not understood in paper world
  - ‘today, business and professional people with no training and usually no aptitude for managing records are responsible for creating and storing them’
- This has always been true!



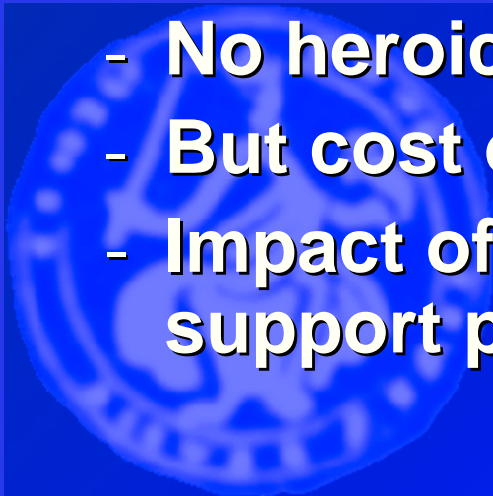
# Speed of Entropy

- Paper lasts
  - Over 400 years old issues of
    - Language
    - Script
    - Physical condition
- Computer files don't
  - Over 5 years old issues of
    - Technical obsolescence
    - Physical deterioration



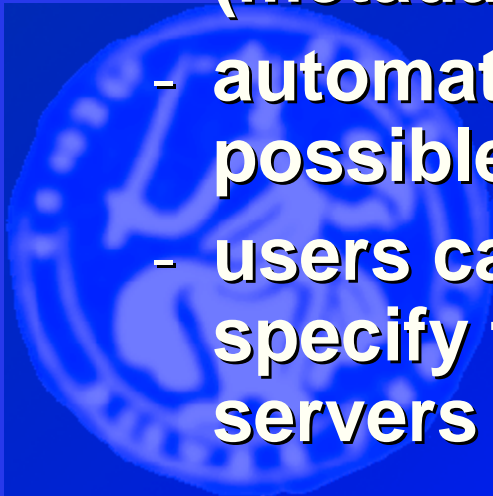
# The CEA

- **Collaboration of scientists (users), IT data centre staff (technical) and archivists (context)**
  - complex data storage requirements
  - Long term retention (30 years plus)
  - No heroic measures (Dept. of Health)
  - But cost of re-creation
  - Impact of not having data available to support products



# The CEA

- **Tessella Project - Electronic Archive**
  - user initiated copying of files from network file servers to the CEA for secure long term storage
  - association of additional information (metadata) with each archived file
  - automated collection of metadata where possible from existing databases
  - users can interrogate the CEA index and specify files to be retrieved to network file servers



# People who have to, do

- Responsibility and authority
  - Impacts 'The bottom line'
  - Or Legal 'duty of care'
- Societal assumptions of who will provide services
  - Acts at corporate level
  - Provides framework for credibility of records



# The present

- E-preservation at the PRO mandate to store and make available comprehensive 'born digital' public records
  - Support statutory duty to have records for business and historical reference
  - Deliver Value For Money solutions
- Core part of PRO e-business strategy
- Timetable must run in parallel with 2004 e-government project



# Key Functions

Incorporate  
them

**Stakeholders'  
Ideas and plans**

**Architecture**

Devise and Deliver  
Technical Strategy

**Project  
Delivery**

Vision  
and  
Policy

**PR / Marketing**

Deliver Projects  
On time all the time

**Live Service**

Tell The  
World

Manage Live Service  
No downtime



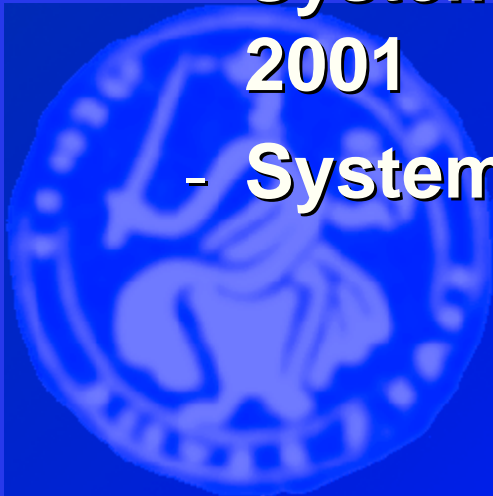


# Goals for 2001-02

- Present strategy to Management Board in Nov 2001
- Organise service delivery
- Liaise with other PRO departments e.g. RMD, ICT, E-Access
  - Establish process interfaces and dependencies
- Complete systems requirements and OJEC Expression of Interest letter by 31 March 2002 (BP11)

# Goals for 2001-02

- **Release Digital Preservation Assessment Database (PRONOM)**
  - User Requirements drafted
  - System purchased by December 2001
  - System operational by March 2002



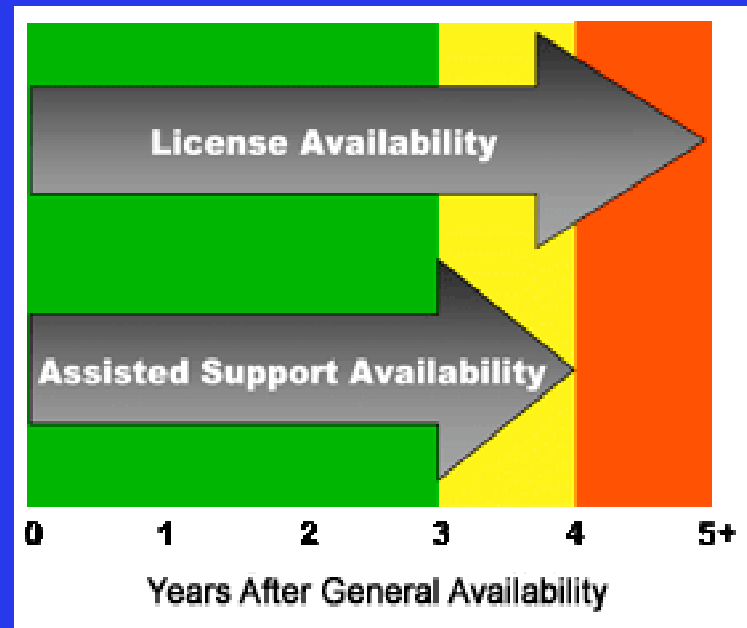
# External Goals for 2001-02

- **Contribute to cross sector initiatives in e-preservation**
  - **Government advisory bodies**
  - **Digital preservation coalition membership**
  - **Set up working group with cognate bodies on collecting software and hardware products to support electronic records preservation**



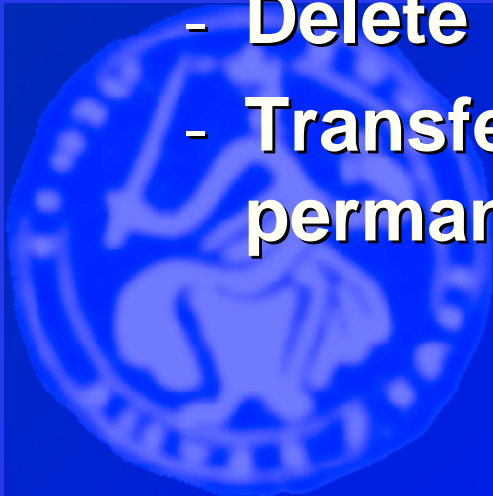
# Appraisal

- Content of records
- Product support defined in Lifecycle Guidelines
- Cost (volume and complexity)



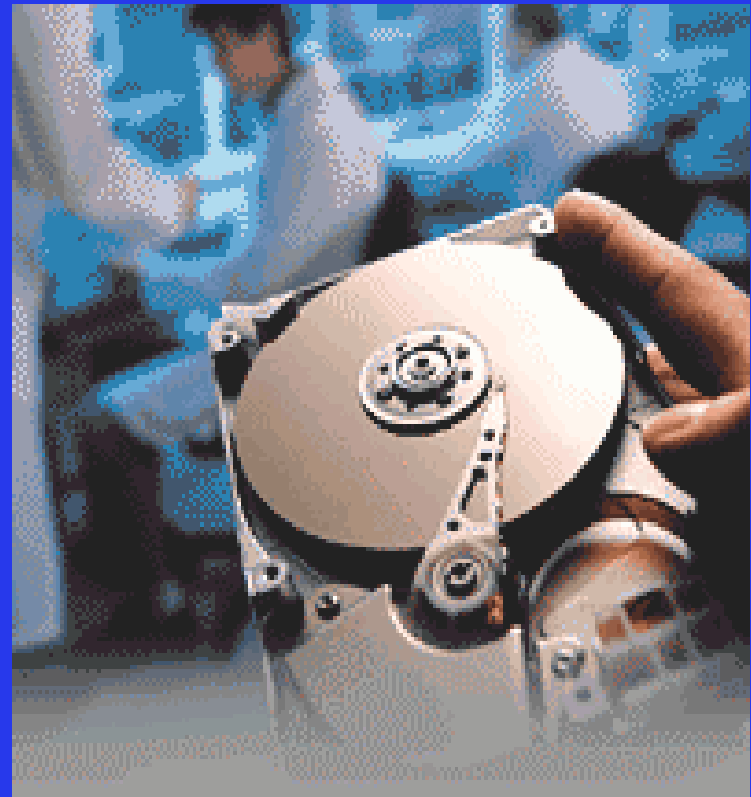
# Transfer

- Create and store in ERMS or other work application repository
- Actively appraise
  - Maintain locally
  - Delete
  - Transfer selected records to PRO for permanent storage



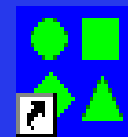
# Storage - how

- UK government data storage – what's the % selected for PRO
  - 3%?
  - 20 TB per year by 2010?
- Data centre
- System for receipt
- Release
  - Fluid rules
  - Through PRO Online

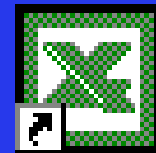


# Storage - what

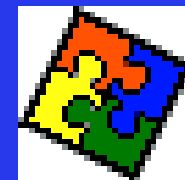
- Users will use records not just read them
- Record as it was when it was declared
  - Or as near as we can get it
- Work on standards for future records
  - Accept early files may be in proprietary format



Visio.Lnk



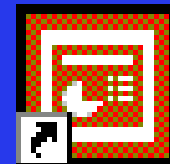
Microsoft Excel.Lnk



Microsoft Word.Lnk



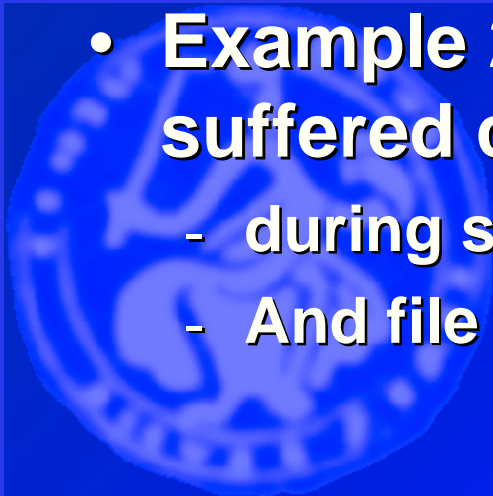
Microsoft Access.Lnk



Microsoft PowerPoint.Lnk

# The future

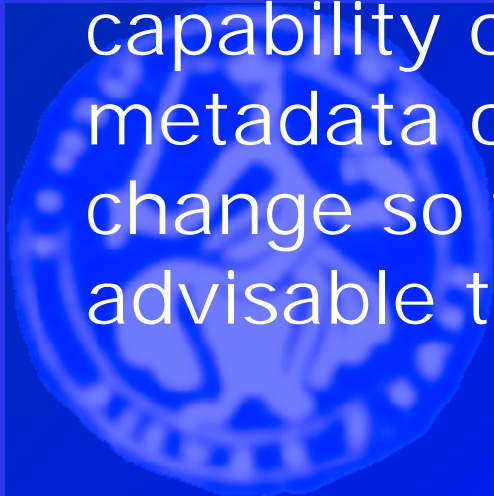
- **Lots of methods and lots of models**
  - because lots of records
- **Example 1: Reversion testing for programs' functionality**
- **Example 2: Ensuring record has not suffered content variance**
  - during storage media migration
  - And file format migration





# The future

- Emerging products
  - metaPackager™ allows you to **use your common practices to use rich metadata to catalog and archive files** that do not have the capability or facility to embed metadata or where the standards change so frequently that it is not advisable to do.



# The future

- Diplomatic rules for the e-world



# The future

- Need to keep up with Computer industry



# The future

- **Listening to our stakeholders**
  - **Users of the service**
    - ask the diner, not the cook
  - **Budget holders**
    - Develop an information democracy and avoid the data liability zone
  - **Vendors, service providers, etc**
    - Maximise relationships



# Any questions?

## Thank you

