Digital Preservation Research at the PRO

David Ryan
Head of Archive Services
DPC Forum on Digital Preservation R&D
5 June 2002
The Public Record Office

- The UK National Archives
  - Select
  - Preserve
  - Make available
- Kew, Surrey
- 450 staff
- 2005 E-business plan
Digital preservation raison d’etre

- 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  - Support statutory duty to have records for business and historical reference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
  - Advise other UK archives on e-preservation
- Core part of PRO e-business strategy
- Timetable must run in parallel with 2004 e-government
Why undertake research?

- To maintain its evidential quality (the fundamental nature of a record) it must be
  - unaltered from the original
  - what it purports to be
  - Its representation is transparent (the rules are stated, and if possible, reversible)
Respect the constraints

- Documented decisions to abide with greater or lesser fidelity to constraints of the original need to be made.
- The specific degree to which all transformations respect formal properties of the original is the responsibility of those managing records.
- And they must report on their methods.
Computer lab raison d’être

- Knowledge of digital representation and modelling methods
- Implications of specific approaches to re-presentation when original technical environment is no longer available
R&D focus of PRO

- Proprietary format migration
- Emulation and simulation
- Open format export
- Open format migration
- Open format export and migration
- Museum route (myth slaying)
- Product reviews
How much?

- UK government data storage – what’s the volume?
- % selected for PRO
  - 3%?
  - 2 TB per year by 2010?
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 Access.lnk
Microsoft Excel.lnk  Microsoft Word.lnk
Microsoft PowerPoint.lnk
Our Initial Scope

- Public Enquiries
- Agenda
- XML
- PRONOM
- No 10 web site
- Open Source

E Preservation Team
Current Status

- On-track for background research
  - Even if sources partial and fragmented
- Computer lab equipped to base level
  - Not expensive!
- First utilities/programs purchased or acquired
- Started analysis of what is coming
- Trials and pilots plan approved
Issues

• Recruitment
• Software/hardware acquisition
  - For older systems
• Need to have precise info on what and how much
• We need to anticipate more rapid technological change
Next steps

- Establish what other organisations have done
  - Don’t repeat work unless essential as expensive
  - Build up relationships with other programs and with DPC
  - Share experiences where appropriate

- Build up technical e-library of reference material
Some thoughts

• Over riding need for archives community to establish credibility with stakeholders in e-preservation
  - IT industry
  - Media
  - Public
  - Finance!

• Quick wins and key tasks
  - Pick out the obvious and the urgent