

# The Paradigm Project: Practical Lessons

Janette Martin,  
Digital Archivist

**PARADIGM** Project:

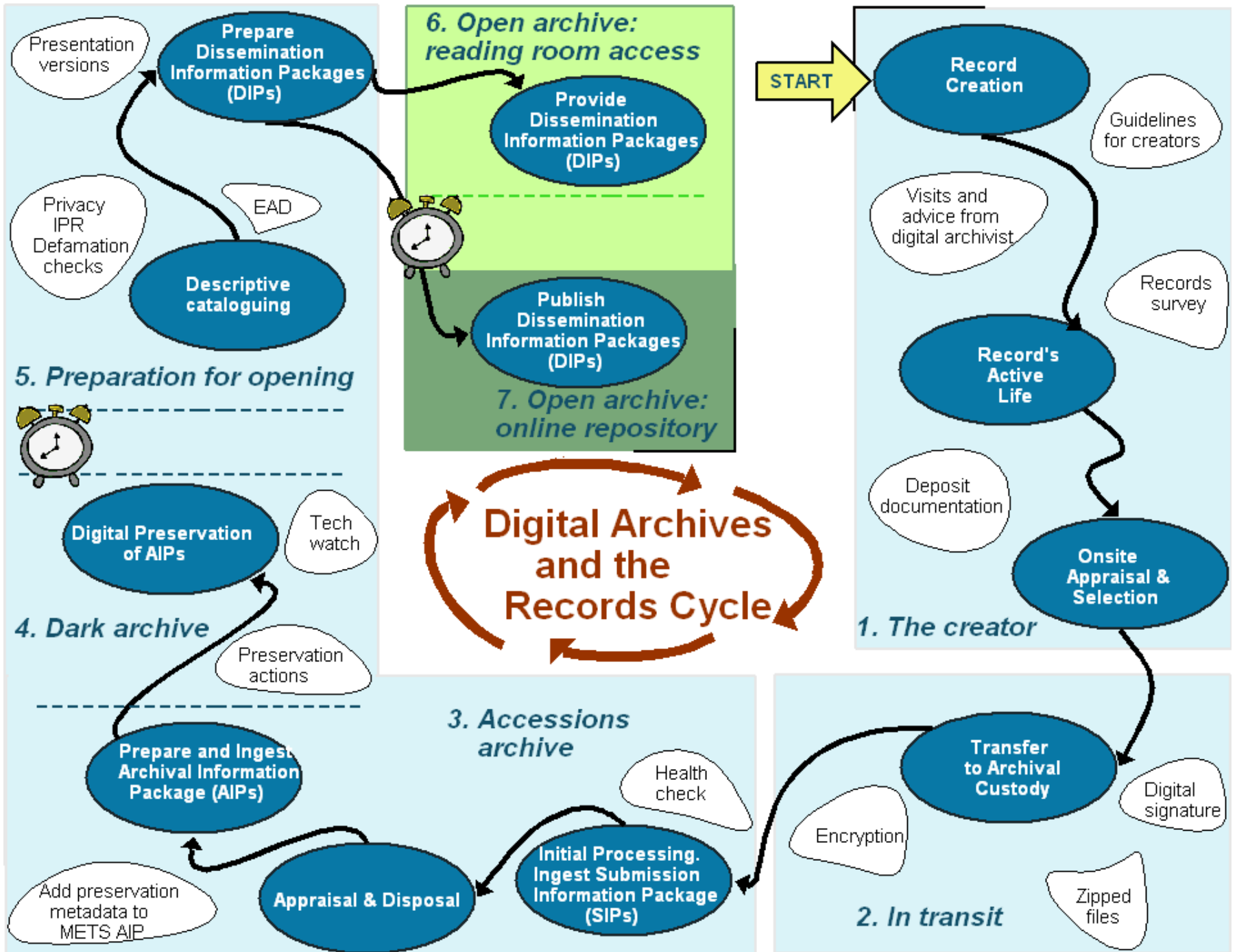
Challenges of the e-environment  
for HE records managers &  
archivists

Kings College, London, 20<sup>th</sup> July 2006



# PARADIGM

- Oxford University Library Services & John Rylands University Library, Manchester
- Exploring from 'personal' and 'collecting' perspectives
- Exemplar strategies for preserving digital private papers – using politicians as our example
- Technical issues and tools (Fedora & DSpace repositories; standards - OAIS & METS)
- Cultural issues (creator attitudes, legal issues, etc.)
- Results – an online workbook for curators



# Creation of Digital Record

- Records are created, collected or received through daily transactions
- What softwares and hardwares are used
- File and folder naming conventions
- Variety of record keeping behaviours
- Influence of the digital archivist at the creation stage (guidelines and visits)

# Active Life of Digital Record

- Select records with long term historical significance only (records kept purely for business or legal reasons should be retained by the creator)
- 5 years and under – likely no preservation required
- Good records management eg naming conventions, record file locations, regular back ups and delete what is no longer required

# Onsite Appraisal and Selection

- Paradigm records survey for personal papers - see work book
- Record IPR, DPA and privacy issues
- Job description of creator/support staff
- Technical environment (CD writer, USB ports, are office PCs networked etc)
- Extent, date(s) and types of file series

# Transfer: Acquisition Documentation

- Transfer protocol establishes procedures to protect original order, authenticity and keep records secure during transit
- Deposit agreement (DPA, IPRs, permission to make multiple copies for digital preservation)
- Transfer forms, basic record of extent, type, date(s) and content of acquisition signed by creator and collecting institution

# Transfer : Technical Procedures

- Digital signature: checksums used to guarantee integrity and authenticity of files
- Copied folders compressed into zip or tar files to retain directory structure (original order)
- Portable media protected during transit by encryption (USB mass storage device) or finger print recognition (portable hard drive)

## SCHEDULE OF TRANSFERRED MATERIAL



<b>Media ref. no.</b> <i>[Ref. no of CD-R or USB stick, e.g. CD-R-1]</i>	USB 1
<b>MD5 checksum(s)</b> <i>[record values of MD5 checksum(s)]</i>	e072dad61e152c96889d3e117038e5741mydocs.zip
<b>Extent</b> <i>[In bytes]</i>	I 32.7
<b>Technical description</b> <i>[description of file formats, passwords]</i>	Word files and Excel spreadsheets created by Joe Bloggs using Microsoft Office 2002.  No pass words required
<b>Content Description</b> <i>[Covering dates, subjects, record types, etc.]</i>	Folders include: Press releases Constituency party information General Election 2005 General Committee reports Annual reports



# OAIS Model and METS

- METS acts as a wrapper for digital object and metadata (descriptive, structural, administrative and preservation) to form Information Package
- SIP: Submission Information Package
- AIP: Archival Information Package
- DIP: Dissemination Information Package
- For a simple overview of OAIS see
  - <http://www.paradigm.ac.uk/workbook/oais/index.html>

# Ingest: Accessions Archive (SIP)

- The SIP is the package sent to an OAIIS by the Producer (or, in our case, accessioned onsite by a digital archivist)
- The SIP comprise 3 units :
  - The digital object(s) to be preserved
  - The Metadata required at this stage of cycle
  - The Packaging information required at this stage of cycle
- Quarantine and authenticity checks

# Appraisal & secure disposal

- Appraisal take places at the SIP stage while the records are held in the accessions archive
- Dealing with duplicate files
- How do we appraise an email directory?
- Delete software files which are not historical records eg clip art
- Ensuring deleted records are entirely removed

# Preparation of AIP

- PREMIS = core preservation metadata set supported by a data dictionary and schemas
- At the AIP stage the OAIS model specifies that Content Data Objects (digital objects) are combined with Preservation Description Information (PDI)
- PDI is divided into 4 types - Reference, Provenance, Context and Fixity information

# Ingest of AIP into Dark Archive

- A METS file is used to wrap the digital object with its metadata to form the Archival Information Package
- See METS files and Fedora's directory ingest service section in Paradigm workbook.
- The completed AIP is ingested into the standalone environment of the 'Dark Archive'



# Digital Preservation: Overview

- Migration – potential loss of look and feel
- Normalisation – potential loss of look and feel
- Emulation – costly and IPR issues
- Likely that certain digital objects will suit particular strategies
- Strategies will change over time

# Closed Storage

- DPA and Copyright and Sensitivity issues mean that a large proportion of collections will be closed for several decades
- A 'dark archive' is one that operates in a secure stand alone environment
- Sections opened as copyright, DPA, defamation and privacy issues expire or permissions are granted



# Descriptive Cataloguing- EAD

- Practical issues arise, the need for two screens (digital object and EAD)
- Use surrogates to prevent accidental contamination
- Advantages of 'cut' and 'paste' and sort capabilities of digital records
- Problem of determining creation date

# Barriers to Access

- Legal (Copyright and other IPRs; DPA, Privacy issues and Defamation)
- Technical restraints
- Institutional requirements and policies
- Access conditions imposed by the creator

# Dissemination Information Packages

- DIPs will be created when a user requests digital object(s) – presentation version
- A DIP comprises object + metadata (likely to be only descriptive metadata)
- Authenticity of the DIP – how will the user know that the record is what it purports to be?
- Philosophical questions: concept of original is difficult in a digital world

# Open Archive

- 2 stages of 'open' in a reading room or online
- Reading room: access to material under legal protection (DPA, Copyright etc) in controlled environment
- Remote access: once legal restrictions have expired material can be made available via an online repository

# Conclusions

- Impact of digital records on cataloguing and arrangement will be interesting
- Need to balance legal requirements with desire to provide access
- Collecting institutions must act now to ensure that the 'personal' is not missing from the historical record
- Exciting time, lots of challenges ahead

Contact me at: [Janette.martin@manchester.ac.uk](mailto:Janette.martin@manchester.ac.uk)

Or further information from:

- Susan Thomas  
Project Manager, Paradigm  
Oxford University Library Services  
Osney One Building, Osney Mead  
OXFORD OX2 0EW  
Web : <http://www.paradigm.ac.uk>  
Email : [susan.thomas@ouls.ox.ac.uk](mailto:susan.thomas@ouls.ox.ac.uk)
- Tel: 01865 283821
- For regular project announcements join  
[paradigm-announce@jiscmail.ac.uk](mailto:paradigm-announce@jiscmail.ac.uk)