

**Susan Thomas,
Project Manager**

PARADIGM

Making links with potential donors

**DCC & LUCAS joint workshop on pre- and post-ingest
activity for digital archive collections**

30 November 2006



The University of Manchester



Summary of today's talk

- Motives for making links with depositors
- Possible approaches to collection development
- Paradigm – piloting approaches and changes in practice to support them
- What next in this area?

Waiting for the Ghost Train?

“Failure to pursue a more active agenda will leave us patiently waiting at the railway station for the goods train of life to deliver the unreliable electronic leavings of our society’s movers and shakers. Years of passive and patient pacing of the platform will come to an end, I fear, when the whistle blows and the train pulls in to the station and we finally come to the realization that it is full of ghosts and that ghosts do not satisfy our researchers’ need for solid, reliable, and authentic evidence of the past.”

Adrian Cunningham, 'Waiting for the Ghost Train: Strategies for Managing Electronic Personal Records Before it is too Late' (1999)



2006, by Doeth

<http://www.paradigm.ac.uk/>

Why a Ghost Train?

- Collecting archives traditionally reactive. We accept deposits offered once an individual has retired or passed away – potentially a long time after creation
- Short-term survival of historically significant digital material uncertain
 - Media degradation
 - Corruption
 - Accidental loss
- Medium- to long-term accessibility hampered by technological change
 - Media
 - Hardware
 - Software
- Individuals have limited support from Information and Information Technology professionals to manage these risks
- Usage of third party storage solutions growing, so likelihood of capturing entire archive without active engagement reduces

Comparing analogue with digital

<u>Intellectual manifestation</u>	Draft speech (Object A)	Recording of speech (Object B)	Draft speech (Object C)	Personal website (Object D)	<u>Risks</u>
<u>Physical manifestation</u>	Paper and ink	Audiotape	3" Amsoft disk	CD	<ul style="list-style-type: none"> ● Security ● Degradation ● Disaster
<u>Hardware stack</u>	X	✓	✓	✓	<ul style="list-style-type: none"> ● Obsolescence of 1+ components in the stack
<u>Software stack</u>	X	X	✓	✓	<ul style="list-style-type: none"> ● Obsolescence of 1+ components in the stack
<u>Representation</u>	X	X	1 locoscript file	1 css file 5 html files 6 jpg files 1 pdf file 1 javascript file	<ul style="list-style-type: none"> ● Relationships between component files broken

Solving the Ghost Train Problem

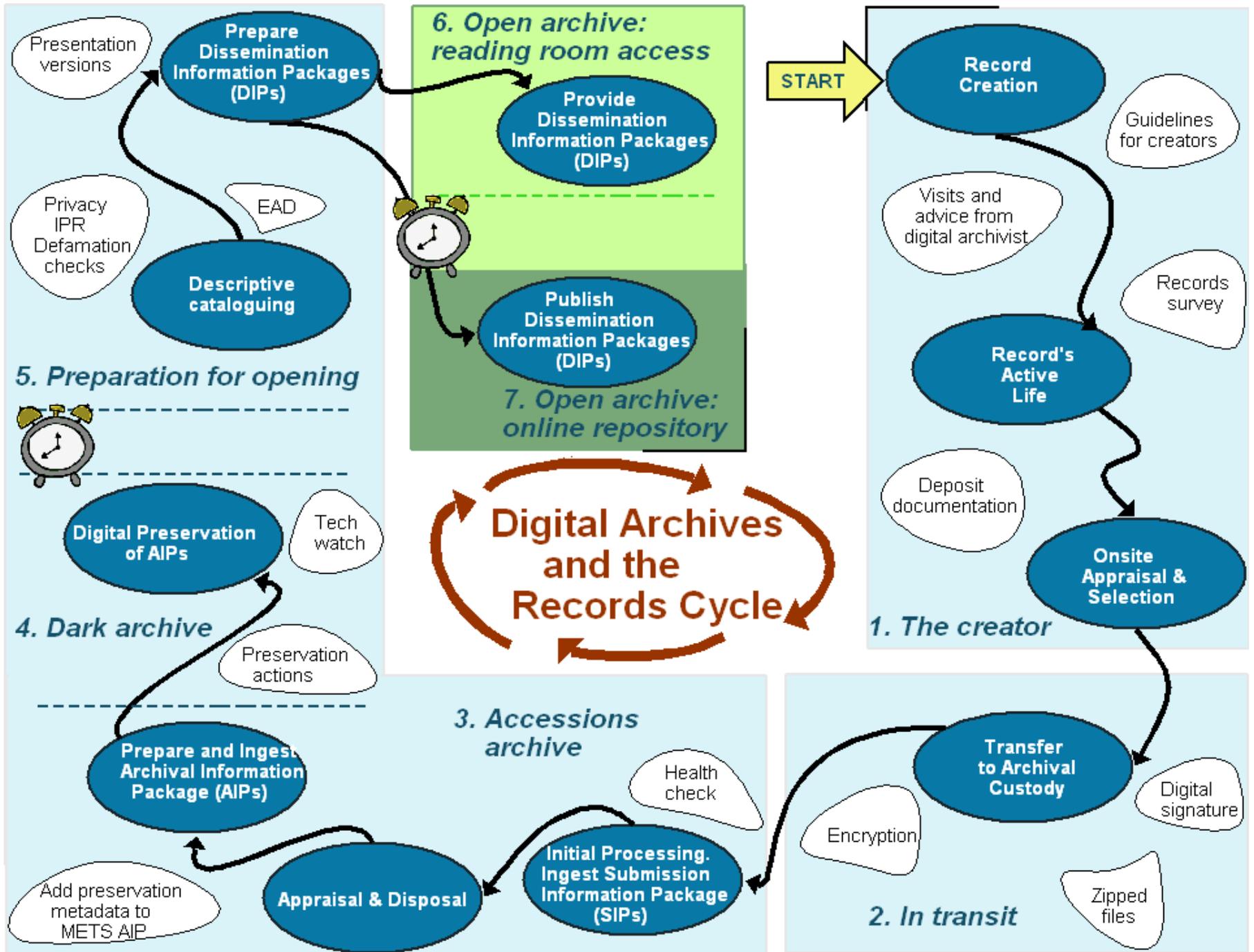
- By intervening earlier in the lifecycle of personal archives
- Increased support for 'records continuum' approaches to records and archives which embrace the entire life of a record
- Typical in organisations with a formal relationship with an archival unit where connection between records management and archival management is clearer
 - National archives
 - Local authorities
 - Other organisational archives
- Long-suggested in a personal archive context – 'pre-custodial intervention' (Adrian Cunningham, 1994)
- Now piloted in the Paradigm project

PARADIGM

- Funded for 2 years by the JISC, ends Feb. 2007
- Collaboration between Oxford University Library Services (lead) & John Rylands University Library, Manchester
- 1.5 fte archival, 1 fte developer plus input from Oxford Digital Library and Special Collections departments
- Explores digital preservation from 'personal' and 'collecting' perspectives in the context of a 'hybrid archive'
- To gain hands-on experience of:
 - an early-intervention approach to developing hybrid archive collections
 - soft issues - by working with politicians and their materials (selection and acquisition, creator attitudes, legal issues, etc.)
 - relevant technical issues and tools

Outcomes

- Have started to harmonise long-standing archival standards and workflows with digital repository standards and workflows
- Prototype preservation repository developed
 - Key standards: OAIS, Fedora, METS and PREMIS
 - Focus on acquisition, ingest and preservation rather than access
- Developed expertise in management of hybrid archives at partner institutions and provided a platform for future activity
- Have developed and shared strategies for personal digital archives
 - That are based on experiences with politicians and their digital archives
 - Through Paradigm's Online Workbook
<http://www.paradigm.ac.uk/workbook>
 - Through 'roadshow' events like this talk



Overview of Possible Approaches to Collection Development

- Regular snapshot accessions
 - Archivist visits regularly to make snapshot accessions of digital materials and to provide advice on caring for archives
- Post-custodial approach
 - Creator retains all material, but archivist provides guidance in caring for it
- Combining snapshot accessions and post-custodial approach
 - Combines above approaches, allowing the creator to transfer some material sooner than other material, and to receive advice in caring for it
- Traditional approach
 - The repository is offered digital archives, or purchases older archival collections
- Transfer via retired media
 - The archivist asks the donor to hand on retired hardware and media to the repository
- Self-archiving
 - Donors submit material electronically to the repository
- See 'Approaches to Collection Development' section in Paradigm Workbook
<http://www.paradigm.ac.uk/workbook/collection-development/index.html>

Proactive Selection 1

- Invited a range of politicians to participate in piloting early-intervention
 - Three political parties
 - MPs, MEP, Peers
 - Local, national and international portfolios
- Move to macro-appraisal – what functions do we want to document rather than what records do we want to keep
- UK politics
 - important organisations (e.g. House of Commons, House of Lords, political parties, think tanks)
 - key functions (e.g. policy making, law-making, constituency work, treasury)
 - record creators in each function (e.g. politicians, politicians secretaries, think tank employees, members of committees)
 - key political events (e.g. elections, PM's Question Time, debates, votes)

Proactive Selection 2

- Relies on a detailed knowledge of area to be documented
- More difficult to determine 'historical significance' early-on
 - Simpler for politics than other areas?
- Selection may confer significance – researchers can only work with material that survives
- Makes creators self-conscious?
- Requires better co-ordination of collecting across organisations
- Results in a more considered collection of archives
- An up-front investment which may yield no dividend?
- More thoughts on selection
<http://www.paradigm.ac.uk/workbook/appraisal/index.html>

Surveying

- Developed a records survey to identify:
 - Functions and roles
 - Technical environment
 - Working practices
 - Rights and responsibilities
 - Record series of historical interest and if and when they could be accessioned
- Allows collection of metadata that will be needed to:
 - Transfer the archive, manage its preservation, administer it and describe its context

or

 - Provide guidance to the creator to manage the archive in situ
- See <http://www.paradigm.ac.uk/workbook/record-creators/index.html> and <http://www.paradigm.ac.uk/workbook/introduction/structure.html>

Guidance for Creators 1

- Flexible and general rather than prescriptive
- Allow users to pick and choose the advice they follow
- Hardware/software neutral
- Designed to facilitate not burden
- Enable creators to make informed decisions about using services, hardware, software and formats
- Enable creators to make informed decisions about archiving
 - Hidden material
 - Data mining

Guidance for Creators 2

- Paradigm is producing generic guidance for creators which covers:
 - Backup
 - Caring for hardware and media
 - Administering your system
 - Selecting file formats
 - Filing and naming
 - Passwords and encryption
 - Keeping up-to-date
 - Managing emails
 - Handling legacy material
 - Ask digital curators for advice
 - Rights
- Would be useful to supplement this with domain-specific guides
- Best-supplemented by guidance tailored to an individual's needs
- Also producing advice for archivists supporting creators on the production and communication of such guidance

Acquisition and Transit

- Agree which archives should be transferred and when, or how often.
- Digital archiving allows exact copies to be taken - creator can retain the material.
- Must learn how to extract typical personal digital archives from popular desktop software and web services.
- Developed transfer procedure and toolkit for secure and authentic transfer
 - Ideal process – use biometric protected USB-powered external hard-disk with forensic software
 - Captures material as structured by creator
 - Records checksums for each item acquired, which can be used to validate the continuing authenticity of items in the accession
 - Ideal process not always possible. Depends on the hardware and software in place



Terms of Agreement 1

Aims

- Produce a legally sound agreement, yet understandable agreement
- *Optimum scenario* - single template for agreements, which secures licences required by the archive and offers protections to the donor. Simplifies the administration of the archive's holdings.
- *Reality* - necessary to have some flexibility

Agreement is with the donor of the archive; need to remember that the archivist has a duty of care to the third-parties represented in the archive too

Some questions

- What licence does the archive need to undertake necessary work relating to the preservation and access of the archive?
- Can the donor grant the archive that licence?
- What protections does the donor require?
- Will one agreement cover all accessions?

Terms of Agreement 2

Permissions required by the archive

- Explicit permission to undertake preservation actions from simple backup to format migration
- Permission for third-parties to work on/store the material?
- Permission to act on behalf of the donor re. publishing requests
- Permission to provide access in a particular environment?

Protections required by the donor

- Explicitly document terms of agreement in relation to closure periods
- State that closure periods will be applied to material containing sensitive personal data or commercial sensitivity, etc.
- Template at:
<http://www.paradigm.ac.uk/workbook/accessioning/documentation/index.html>
- Thoughts on legal issues
<http://www.paradigm.ac.uk/workbook/legal-issues/index.html>

Early-intervention pilot : Lessons

- Digital increasingly used as 'master', but poorly managed
- Poor understanding of archiving for historical purposes
- Privacy and security concerns – own and third party – increased by recent date of material. Reluctance to deposit some material now, or at all
- Repository must manage material with legal protections for longer
- Finding time for history in the present
- Impact on resources
- Authority to act
- Variety: individual concerns; technical set-up; organisational set-up; IT literacy or support
- Frequency and scope of accessions; dealing with duplication
- What about the paper, audio, video, photographs, etc.?
- Opportunity to acquire valuable contextual information
- Contemporary formats are easier to access and normalise

Early-intervention: Conclusions

- A worthwhile approach
 - Individuals have lost material!
 - Can obtain excellent context
- But relies on
 - Headhunting individuals
 - Good will and trust of individuals
 - Sustaining relationships over long periods of time
 - May produce **different** collections
 - May not work so well in instances where archives are to be purchased
- Digital archaeology inescapable
- Need to repeat with other groups
- What happens if the creator wishes to sell, rather than donate?

What next in this Area?

- Sustain relationships over longer period of time
- Work with other kinds of creator and their archives
- Pilot other approaches to collection development
- Integrate digital archives into existing policies for archives

Questions?

- Ask me now
- Or later:

Susan Thomas (Project Manager, Paradigm & Cairo)
Oxford University Library Services
Osney One Building, Osney Mead
OXFORD, OX2 0EW

Web : <http://www.paradigm.ac.uk>
<http://cairo.paradigm.ac.uk>

Email : susan.thomas@ouls.ox.ac.uk

- Tel: 01865 283821