planets

Are you ready?: Assessment of readiness of organisations for digital preservation

Tessella: Pauline Sinclair, James Duckworth, Lewis Jardine, Ann Keen, Robert Sharpe

British Library: Clive Billenness, Adam Farquhar, Jane Humphreys

Introduction

1. Overview of the survey purpose & responses
2. Analysis of digital preservation problem
3. How are people trying to solve it?
4. Summary

1. Survey Purpose

- Establish state of readiness of organisations to use digital preservation solutions:
- Targeted at national libraries & archives in Europe
  - 96 institutions specifically targeted
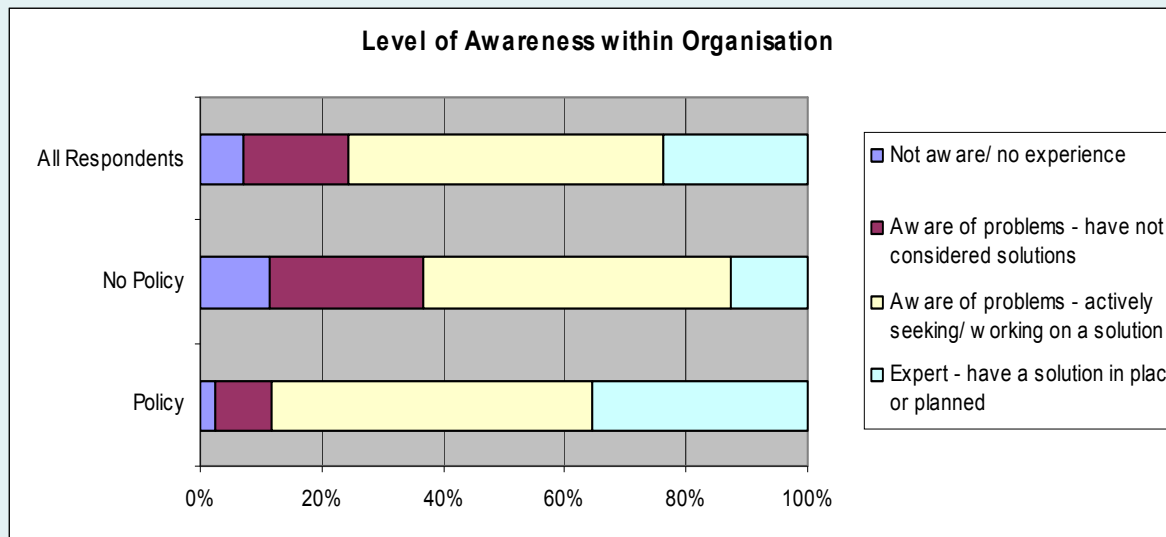- BUT also general invite
  - 2000 individuals
  - Publicity

1. Survey Response

- 206 responses
- 70% European
- Diverse community
  - ⅓ libraries
  - ¼ archives
  - Also government departments / public sector, commercial organisations, suppliers & vendors
- Range of roles
  - 15% digital preservation
  - 16% preservation in general
  - 22% curation (archivist, librarian, records management)
  - 16% IT
  - Also directors, managers, researchers, & data producers
- All responses can be categorised by sector/region as needed

planets

2. Awareness, policies & budgets

❑ Good start on getting digital preservation message out:
- 93% respondents aware of challenges of digital preservation
- Probably informed community responded

❑ Even in this group, more work required:
- 17% organisations have not considered any solutions
- 52% organisations do not have a digital preservation policy
- Policy important: 3 times more likely to have a budget if have a policy

Level of Awareness within Organisation

All Respondents	
No Policy	
Policy	

0%   20%   40%   60%   80%   100%

□ Not aware/ no experience

■ Aware of problems - have not considered solutions

□ Aware of problems - actively seeking/ working on a solution

□ Expert - have a solution in place or planned

planets

2. Digital information to preserve: sources

Source System	% Respondents
File System	77%
Document Scanning Programme	58%
Internet	55%
Electronic Document Management System	55%
E-mail System	54%
Media Digitisation Programme	54%
Media Store	34%
ERMS Database	32%
Computer Aided Design (CAD)	29%
Lab Systems	18%
Other	16%

2. Digital information to preserve: types

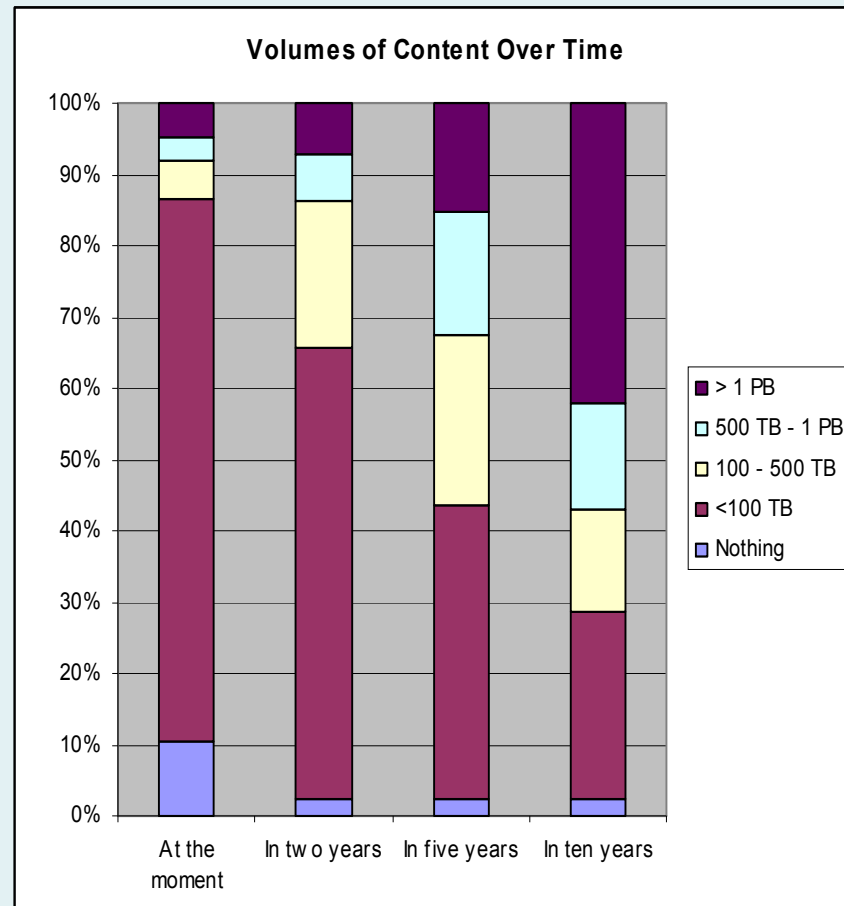| Digital Information Type	2009	2019
Documents	81%	99%
Images	80%	95%
Databases	49%	85%
Audio	49%	80%
Websites	47%	79%
Video	51%	76%
E-mails	39%	66%
GIS	23%	52%
Scientific Data	22%	51%
Software	21%	49%
eBooks	21%	49%
eJournals	25%	49%
ISO or disc images	20%	40%

planets

2. Digital information to preserve: formats
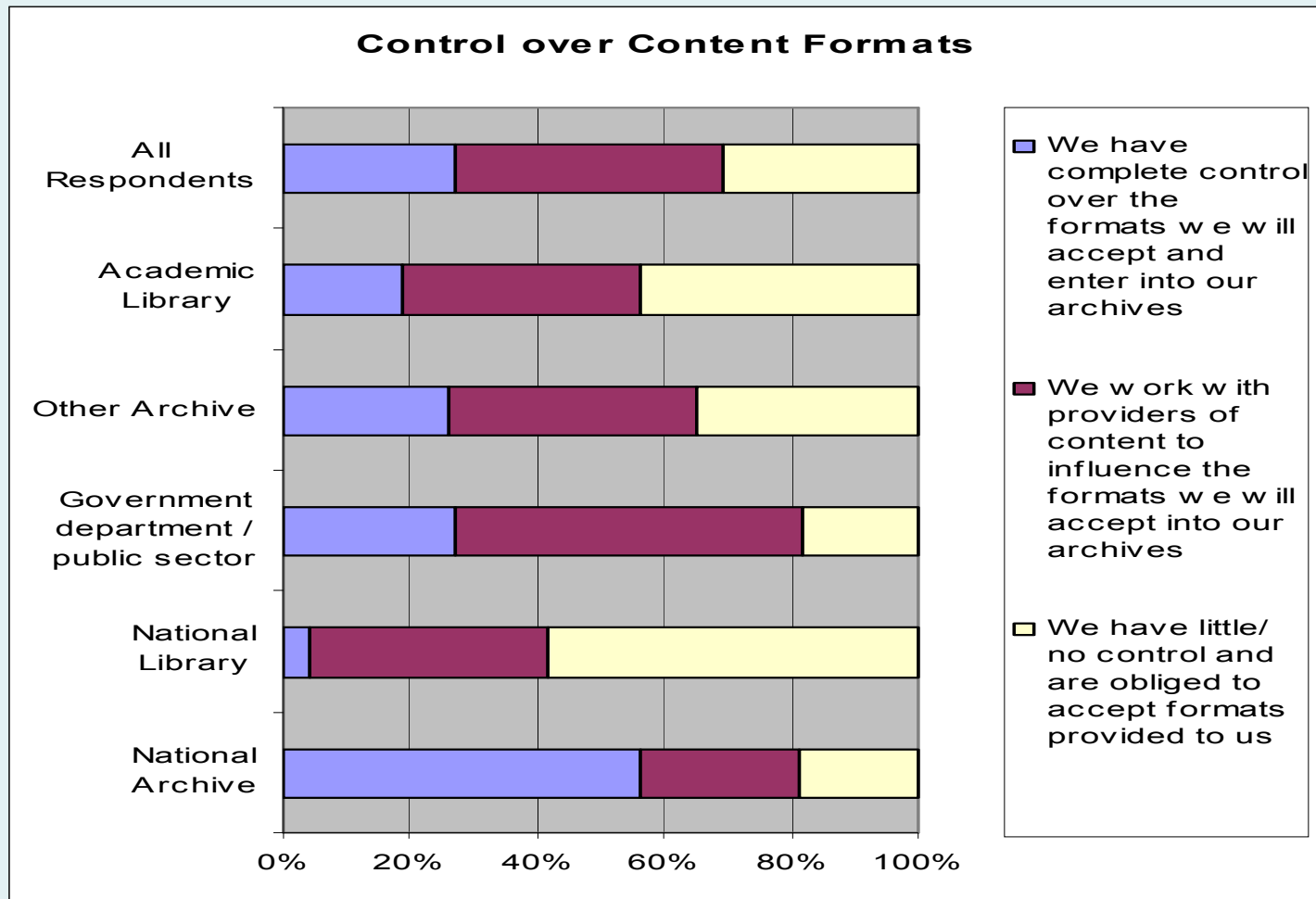
2. Data Quantities to Preserve

- Now:
  >80% org.s hold <100TB

- In 10 years:
  70% org.s hold >100TB
  42% org.s hold > 1PB

- Scalability of solution important for
  - High content volumes
  - High ingest rates

Volumes of Content Over Time

Legend:
- > 1 PB
- 500 TB - 1 PB
- 100 - 500 TB
- <100 TB
- Nothing

X-axis: At the moment | In two years | In five years | In ten years

Y-axis: 0% to 100%

3. Solutions

❑ People are trying to solve problem:
- 85% have a solution or are working on one
- 77% make a (further) investment inside 2 years

❑ Generally expecting "plug-and-play" components:
- Could be in-house, open source, off-the-shelf
- But still expect components

3. What functionality is important?

Capability	Mean Rating
Maintains authenticity, reliability and integrity of records	3.8
Checks records have not been damaged	3.5
Plans the preservation of content to deal with technical obsolescence	3.4
Complies with established data or digital information management standards	3.4
Ensures records are accessible for up to 50 years	3.4
Performs migrations to deal with technical obsolescence	3.4
Is able to store many different types of content	3.3
Handles a wide variety of file formats	3.3
Ensures records are accessible for more than 50 years	3.3
Adheres to metadata standards	3.2
Retrieves content by description	3.0
Characterises records by extracting technical metadata	3.0
Integrates with content delivery systems	2.7
Retrieves content using full text	2.7
Supports emulation to deal with technical obsolescence	2.5
Integrates with content producing and holding systems	2.5
Checks for duplicate items	2.3

3. Standards

❑ Are important – but not clear which ones…

Metadata Standards in Use
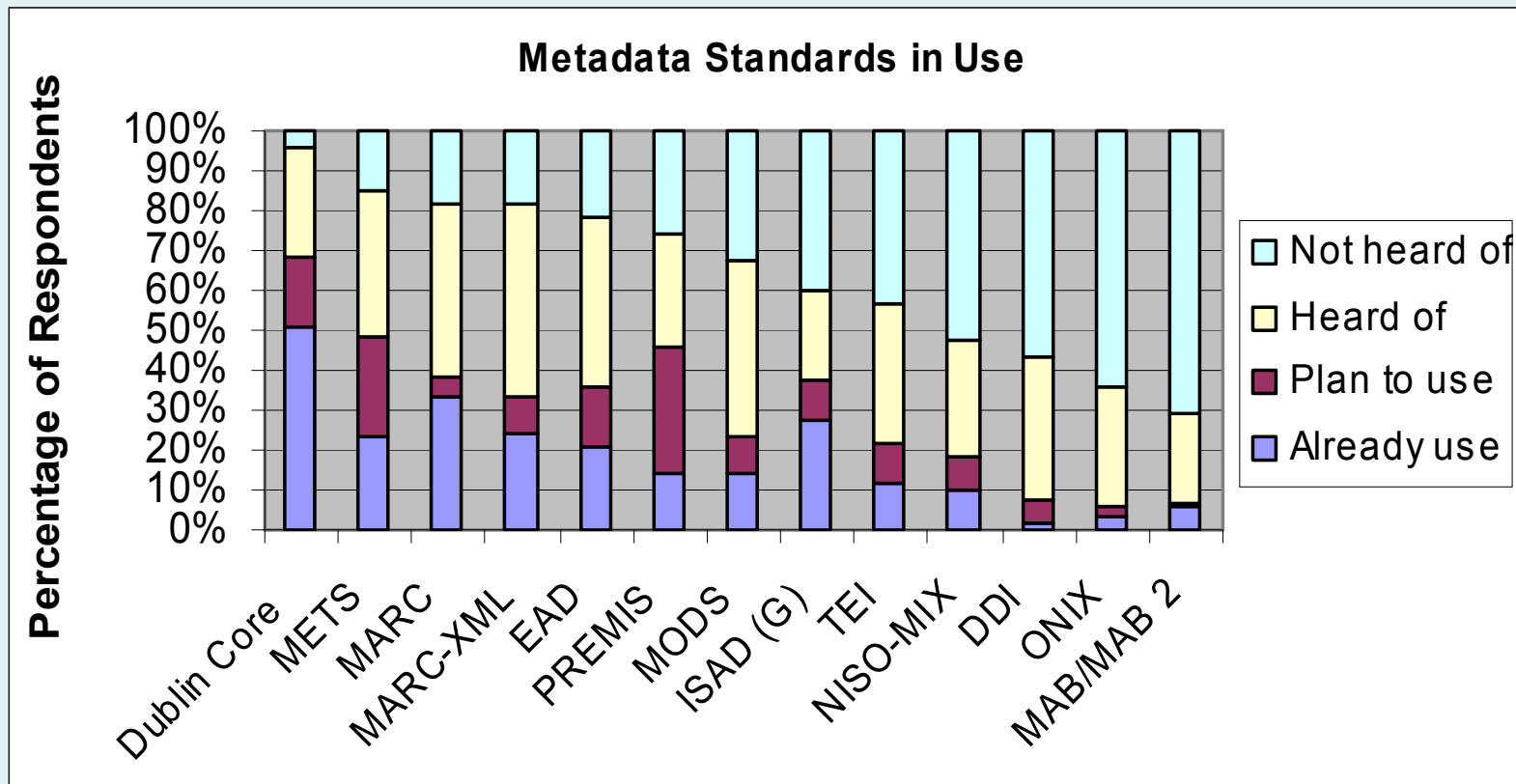
4. Summary

- Excellent start on getting digital preservation message out
- More work needed on policies & budgets
- Wide range types of digital info. from variety of sources
- Significant quantities data to preserve => scalability
- Component-based solutions required
- Best practice not yet clear (e.g., standards)
- Early adopters are busy & planning to do more
- Are you ready?