The European Repositories Landscape –
The view from 20,000 feet

Ian Dolphin
Head of eStrategy
The University of Hull – United Kingdom

i.dolphin@hull.ac.uk / i.dolphin@jisc.ac.uk
Europe Overall

- Evidence produced by the DRIVER Project
  - Digital Repository Infrastructure Vision for European Research

- 3 Key Reports
  - The European Repository Landscape
  - Investigative Study of Standards ...
  - DRIVER's Guide to European Repositories

- http://www.driver-community.eu/
DRIVER Evidence 1: Penetration, Type

- **Respondents** –
  - 70% > 1 repository, 23% > 1+, 7% Outsourced
  - Majority full text, eg: articles, theses, working papers, books
  - Minority metadata of those materials without full-text
  - Small minority contain non textual – video, images, data sets

- **Content Type**
  - 53% Journal articles
  - 18% Book/s/chapters
  - 30% "Grey literature"

  - Articles – published and post-print version dominant type
95% of respondents access type – Open Access

DRIVER Evidence 2: Discipline Division

- 30% Humanities & Social Science
- 25% Natural Sciences
- 20% Engineering
- 13% Life Sciences
- 12% Other
DRIVER: European "Desirable Services"

- Top 3 – selected by >33%
  - General search engines, gateways and portals
  - Disciplinary and thematic search engines, gateways and portals
  - Citation index services

- Next 3...
  - Preservation services
  - Advisory services – Open Access advocacy
  - Usage statistic services
However, nota bene ... 

- DRIVER defines a Digital repository as
  - Containing research results
  - Institutional and/or thematic
  - OAI-PMH Compliant

- There are other definitions ...
  - Even where the target was a research repository, DRIVER quantitative analysis responses: 5% "learning materials +"
  - UK examples, EPrints for LO at Southampton, York multimedia etc

- Growth of open educational resources in Europe
- Growth of media specific repositories in Europe
Preservation activity @ European Level

- Planets, Caspar, DP Europe; Pan European projects
  - Mainly Higher Ed and National Libraries
- DPE – Advocacy role, risk management toolkits
  - http://www.digitalpreservationeurope.eu/
- CASPAR – Impact of preservation on activity lifecycle
  - http://www.casparpreserves.eu/
- Planets – Technical infrastructure enablers of preservation
  - http://www.planets-project.eu/
  - SOA/Web Services infrastructure preservation services
  - Assumption that DP has to be a distributed activity
UK – Mapping business functions...

Thanks to Paul Walk, UKOLN

Figure 3: Business Functions Supporting User-facing Search

- RSP supports
- ZETOC provides journal titles
- RoMEO identifies for harvesting
- OpenDOAR reports status
- Heterogeneous Metadata Store persists
- DEPOT subset retrieved by
- Repository informs
- Harvester is harvested by
UK – Negotiating the joins…

Thanks to Paul Walk, UKOLN
Directions: Integration with *flow

- Institutional level; integration with *flow a priority
  - Virtual Research Environment
  - Virtual Learning Environment/Learning Management System
  - Administrative purposes (eg. UK Freedom of Information Act)

- Two perspectives of “repository”
  - An application with defined functionality (cf DRIVER)
  - A bunch of services which form part of an infrastructure

- **Not** an either/or choice – flexibility is key

- Cambridge–Sakai/DSpace, Hull–uPortal/Sakai/Fedora

- Is the application or environment receptive?
REPOSITORY SERVER

LDAP / IdMS

Image Magick
MIME Magic
Data Fountains
JHOVE
BuildObject
Base64Encode

Fedora
Fedora web svcs

BPEL

MVC

University SAN

HTTP

FTP

Return MIME type(s). Is it a new object?
It's a new object so get next persistent identifier (PID)
We are not alone ...

- European activity is not “stand alone”
- DSpace, Fedora, Eprints usage internationally
- Engagement with Open Source communities
- JISC Involvement:
  - SWORD lightweight deposit API DSpace, Fedora, EPrints
  - Projects using ORE
- Shared problem spaces – Data
  - “Dealing with Data”

http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/publications.html