The Rise and Rise of Digital Repositories

Institutional Web Management Workshop 2006
14th June 2006, University of Bath

Julie Allinson and Mahendra Mahey

Digital Repositories Support Officers

UKOLN, University of Bath

UKOLN is supported by: JISC MLA MUSEUMS LIBRARIES ARCHIVES COUNCIL

UNIVERSITY OF BATH

www.ukoln.ac.uk

www.bath.ac.uk

UKOLN   A centre of expertise in digital information management

MM

Format of the Workshop

Time	Content	Duration
16.15	Presentation and overview of Digital Repositories	25 mins
16.40	Group Activity	30 mins
17.10	Same Role Activity	10 mins
17.20	Feedback, Discussion and Conclusions	25 mins
17.45	End	

Content of Workshop

- Digital Repositories Presentation
  - Definition, types and examples
  - Repositories, the Information Environment and interoperability
  - Issues (General, Legal, Policies, Technical)
  - Work of JISC Digital Repositories Programme

- Activities
  - Based around hypothetical scenario of institution implementing a repository
  - Group discussion where participants adopt chosen 'roles', e.g. web manager
  - Same roles meet up and discuss their findings and present to group
  - Discussion and conclusion

What is a digital repository?

- Vaguely medical (suppos….)
- A store of digital information where information can be put in, taken out and searched for
- However, definition needs more clarity to avoid confusion with other technologies
- A digital repository is differentiated from other digital collections by the following characteristics:
  - content is deposited in a repository, whether by the content creator, owner or third party
  - the repository architecture manages content as well as metadata
  - the repository offers a minimum set of basic services e.g. put, get, search, access control
  - the repository must be sustainable and trusted, well-supported and well-managed

Heery, Rachel and Anderson, Sheila. Digital Repositories Review, UKOLN and AHDS, 2005 (Final version)
<www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf>

Repositories

More than just a store…

Institutional Repository more than just software/hardware…

"a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution. ….. An institutional repository is not simply a fixed set of software and hardware."

Clifford Lynch. 'Institutional repositories : essential infrastructure for scholarship in the digital age'. *ARL Bimonthly Report*, February 2003
http://www.arl.org/newsltr/226/ir.html

www.ukoln.ac.uk

Repositories in the UK

- 130 Higher Education Institutions in the UK (Source: HEFCE)
- 50 have Institutional and/or Department repositories
  - OpenDOAR (www.opendoar.org) lists 56 UK repositories
  - ROAR (archives.eprints.org) lists 72, 44 of which are classed as Institutional or Departmental
- JISC (www.jisc.ac.uk) has been funding directly in this area since 2002
  - FAIR Programme (2002-2005)
  - Exchange for Learning (X4L) Programme (2002-2005)
  - Digital Repositories Programme (2005-2008)
  - JISC Capital Programme Repositories and Preservation strand (2006-)

The many faces of repositories

- Content
  - Eprints, learning materials, corporate records, research data, multimedia objects
- Coverage
  - Departmental, Institutional, Regional, National, Personal (web site / blog)
- User Group / Domain / Community
  - Learners, teachers, administrators
- Access Policies
  - Open, restricted
- Distribution
  - Centralised, federated
- Function / Purpose
  - Open access, subject access, publication, sharing and re-use, preservation

www.ukoln.ac.uk
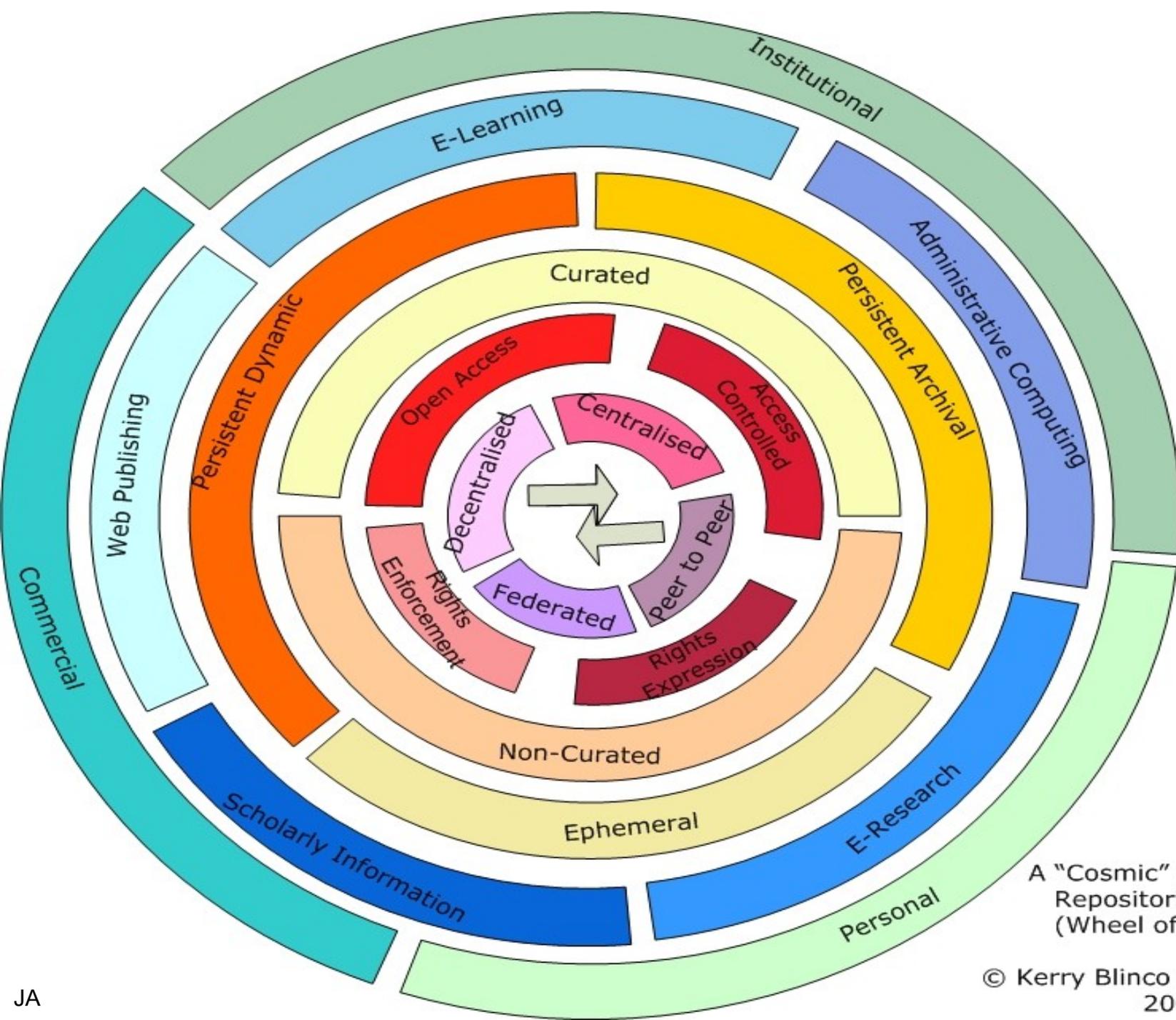
A "Cosmic" View of the Repositories Space (Wheel of Fortune)

© Kerry Blinco & Neil McLean, 2004

Jorum

- Content: Learning objects
- Coverage: National
- Function: re-use and sharing
- Community / Domain: Teachers (primarily) / E-Learning
- Centralised
- Access controlled
- Curated
- Intrallect software

http://www.jorum.ac.uk

Edinburgh Research Archive (ERA)

- Content: Eprints
- Coverage: Institutional
- Function: Access (primarily)
- Community / Domain: Researchers / Scholarly communication and E-Research
- Centralised
- Open access, controlled deposit
- Curated
- DSpace software

http://www.era.lib.ed.ac.uk/

UKOLN A centre of expertise in digital information management

eBANK UK

- Content: X-ray Crystallography Data
- Coverage: Institutional (Southampton) and National
- Function: Access and reuse
- Community / Domain: Researchers in sciences
- Centralised
- Open access, controlled deposit, repository used as part of laboratory workflow
- Curated
- Eprints software (modified for Crystallography data)

http://ecrystals.chem.soton.ac.uk/

Flickr

- Content: Images
- Coverage: International; community-based; personal
- Function: sharing and re-use
- Community / Domain: Anyone / Public
- Centralised
- Part open access, part access controlled

    http://www.flickr.com/

- Flickr API

    – Allows innovative re-use of Flickr data, e.g.
    http://metaatem.net/words/

Repositories and the Information Environment

JA

Information Environment

www.ukoln.ac.uk

UKOLN | A centre of expertise in digital information management

JA
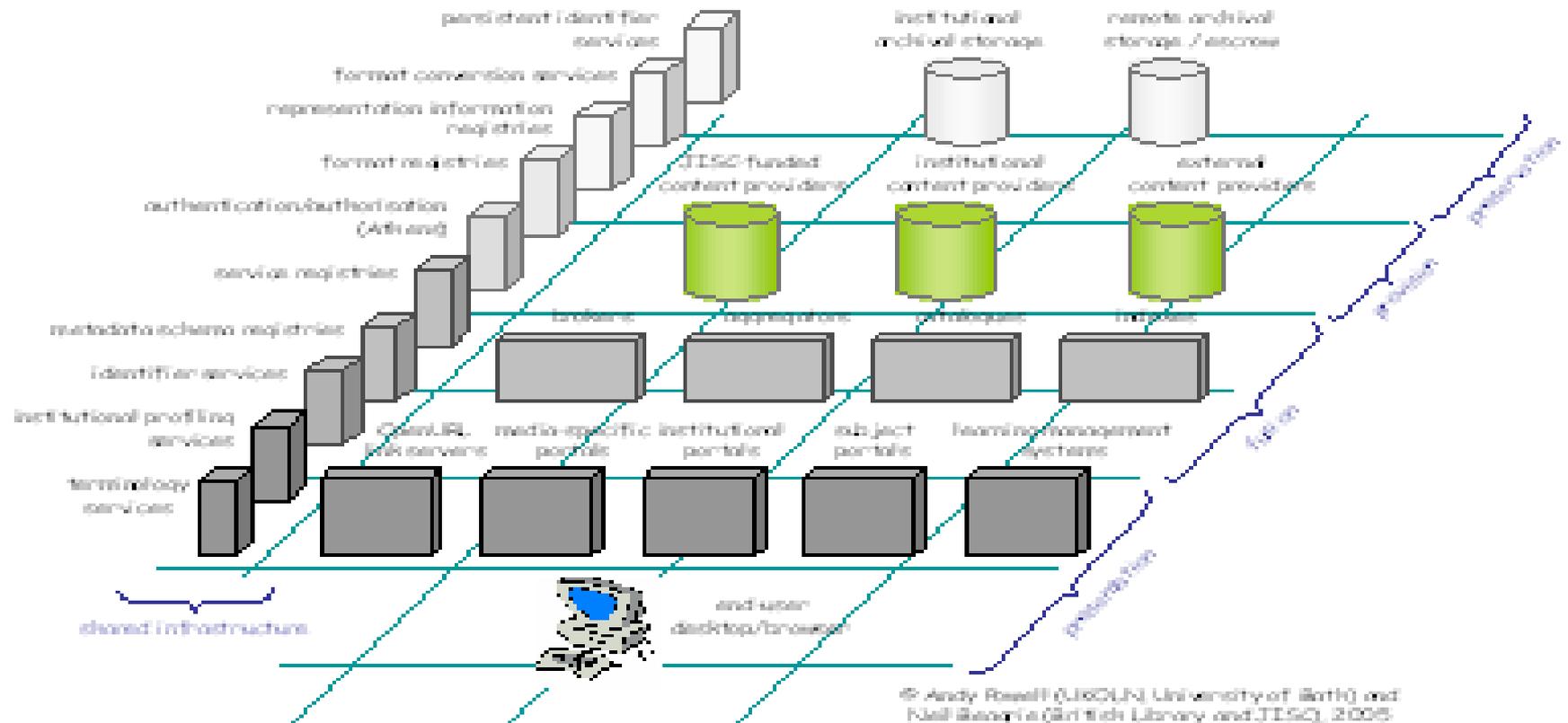
Issues

General Issues

- Scholarly publishing model
  - Peer review
  - Impact Factor
  - Citations
- Sharing vs Competition
- RAE
- Personal resource management
- Rewards / Motivators
- Trust
- Open access and culture change
- Preservation
- Funding for research and implementation

Legal

- Intellectual Property Rights (IPR)
  - Copyright
  - Other: database rights, moral rights, performers' rights, trademarks, patents
- Copyright
  - Who owns? Author, Institution
  - Publisher agreements - Sherpa Romeo (http://www.sherpa.ac.uk/romeo.php)
  - Moral rights are also important – paternity and integrity rights
- Institution as publisher
  - Securing the right to publish, store, preserve
  - Plagiarism/copyright infringement, other infringements
  - Defamation, inaccurate information, confidentiality, libel
  - Freedom of Information and Data Protection
- Risk Assessment
- Licences / policies
  - deposit AND end-user

www.ukoln.ac.uk

UKOLN   A centre of expertise in digital information management

Policy

- Having well developed policies are important if an institution wants to implement a digital repository
- Policies can cover; legal issues, who deposits, who adds metadata, long term preservation, etc
- One current concern is whether it is good to mandate deposit into the repository or not and whether this is a good thing
- It's one thing having a policy, it's another thing complying to it

Carrot or Stick?
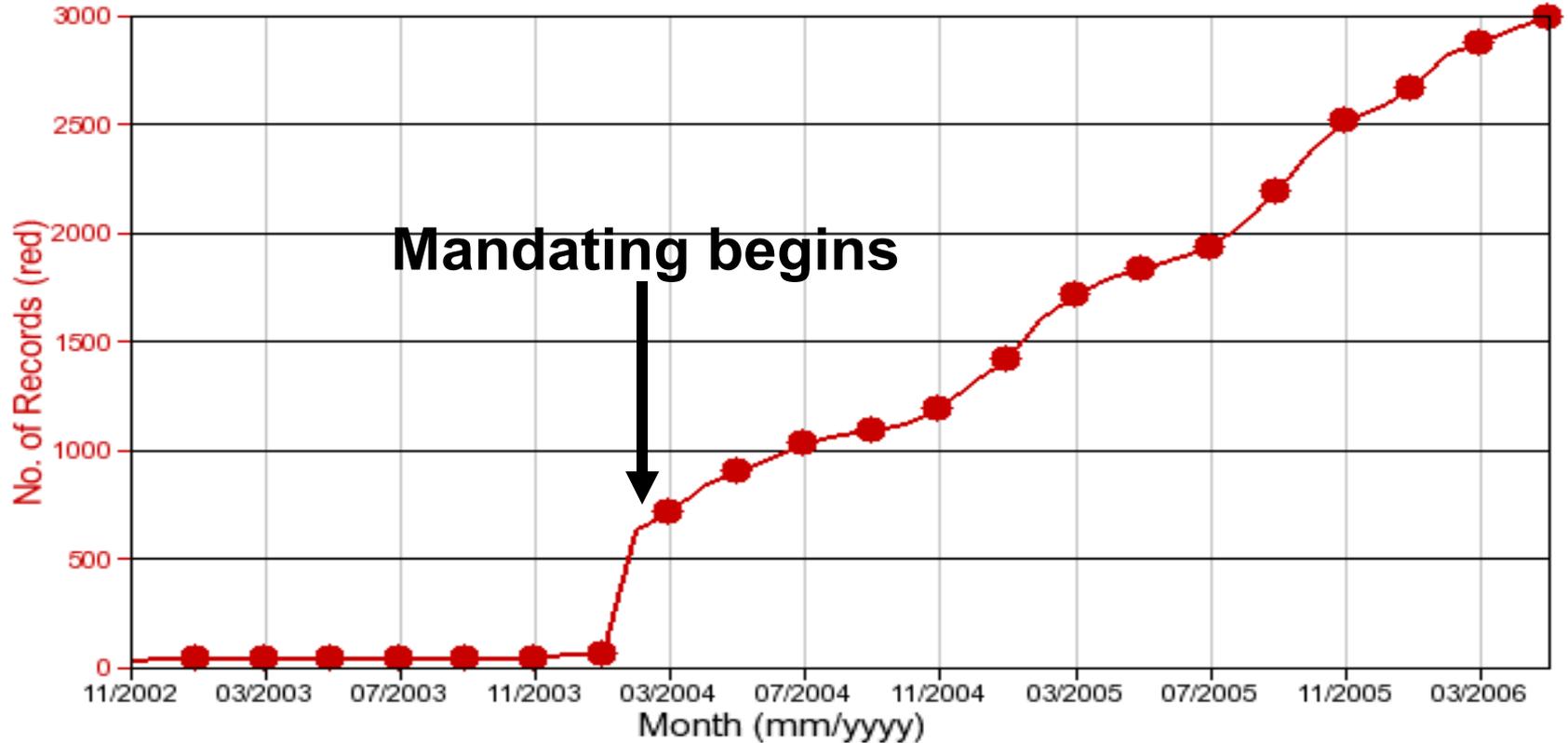

University of Queensland ePrint Archive
Generated by http://archives.eprints.org/

Mandating begins

Technical

- Most repositories built on open source (Linux, Apache, MySql and PHP/PERL)
- Requires specialist skill set, e.g. LAMP
- Requires relatively high processing power
- Most repositories can expose metadata for harvesting if not the actual data
- In a service-oriented architecture a repository may be built up of a number of service components

JISC
Digital Repositories Programme

UKOLN   A centre of expertise in digital information management

JA

Digital Repositories Programme Jigsaw
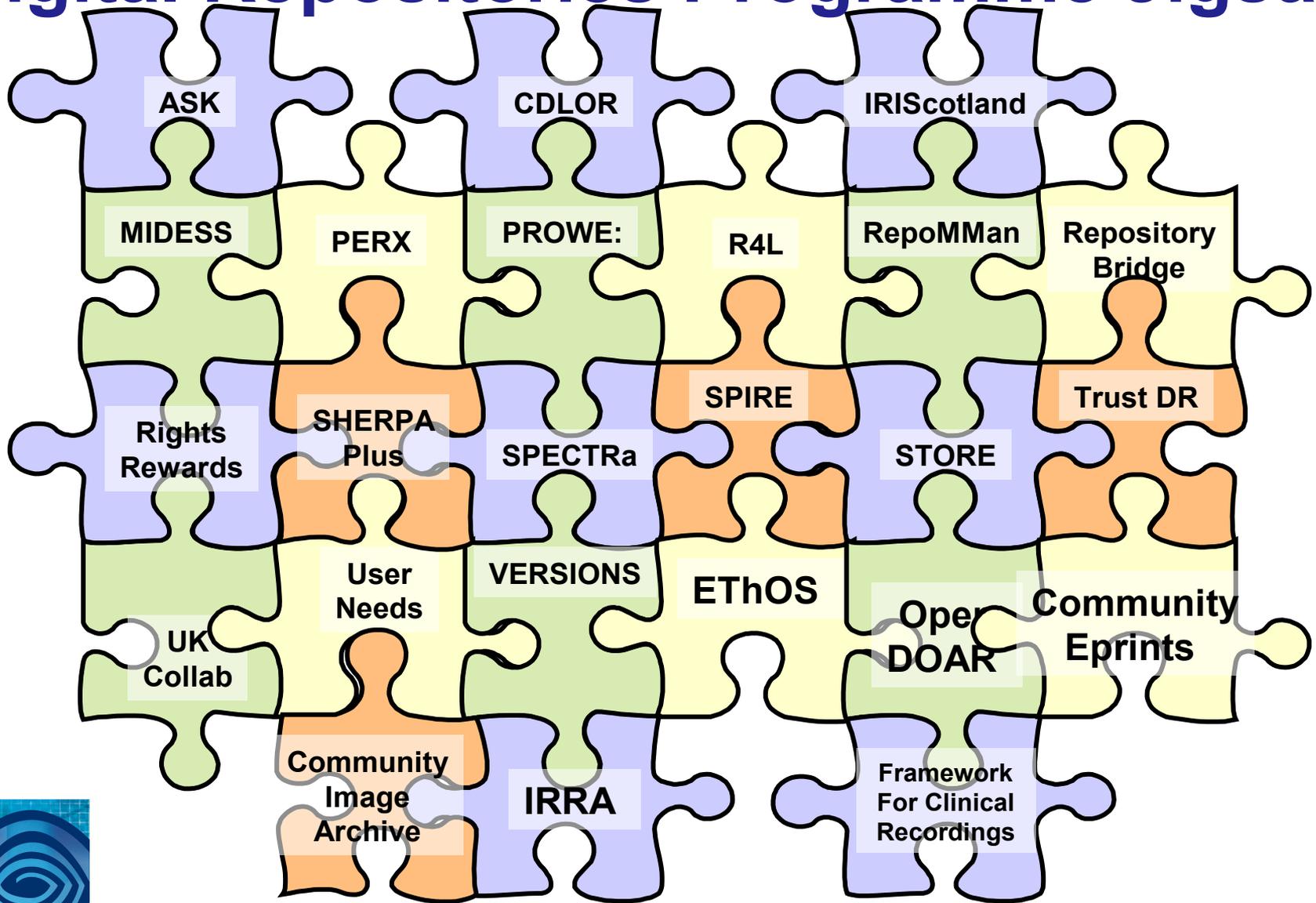
ASK
CDLOR
IRIScotland
MIDESS
PERX
PROWE:
R4L
RepoMMan
Repository Bridge
Rights Rewards
SHERPA Plus
SPECTRa
SPIRE
STORE
Trust DR
User Needs
VERSIONS
EThOS
Open DOAR
Community Eprints
UK Collab
Community Image Archive
IRRA
Framework For Clinical Recordings

Summary

- Repositories come in many shapes and sizes
- They are more than simply hardware and software
- Many are built on open standards and open source which allow them to interoperate, e.g enabling users to cross search across repositories
- Many issues involved in their implementation
- A lot of funding focusing on research and implementing repositories from funding bodies
- The number of repositories being set up around the world growing rapidly

Activity – Group Work

Overall Scenario – the context

The **Combe Down University**, near Bath, is a medium-sized Higher Education Institution with around **15,000 students**. It teaches across a wide-range range of subject disciplines and has 5* Departments for Chemistry, English and Psychology, each with a **strong International research presence**.

The University is committed to a creating an Institutional Repository, but has **yet to establish the scope and policies of this repository**.

To facilitate the successful implementation of a repository, the University has created a new post of **Repository Manager** and has also seconded **technical support** for development. The University have **selected** a popular repository software package, **DPrints**.

Overall Scenario – the context (2)

The University wants its repository to be a success, both in terms of its being populated with materials by academics and also used, internally and externally by student and academic researchers. As a first step towards establishing its repository, the University want to assess the needs and requirements of the different stakeholders and to identify barriers and issues, which might include: the purpose of the repository, the types of materials accepted, its target users, access policies, other policies, legal considerations, usability, design, metadata creation, workflow and technical considerations.

UKOLN   A centre of expertise in digital information management

Overall Scenario - Task (3)

The following people have been invited to join a focus group to discuss their own requirements:

- **Student / Researcher**
- **Lecturer**
- **Web / IT Manager**
- **Librarian / Information Manager**
- **Departmental Manager**

Previously they were all given an advocacy presentation about what repositories are, examples and related issues to their Implementation.

Through this group it is hoped that issues and barriers can be identified and the group can come up with some possible solutions and agreements as to the way forward.

Activity Instructions

Step 1 – Getting into Groups, choosing a role and character

- Get into groups of five
- Each group will have:
  - Students and researchers
  - Lecturers
  - Librarians and information managers
  - Department and institution managers
  - Web and IT managers
- Each of you will need to decide the role you want to be
- Once you have done this, pick up the relevant handouts for your 'role'. You will notice that for each role you can choose one of two characters

Activity Instructions

Step 2 – Ground rules for discussion

- Each of you read the information about the character you have chosen for your role
- Decide who is going to go first
- Tell people who you are and what your needs may be. Try not to read directly from the sheet. You may 'adlib' and add other features to your character if you so wish and try to stay in role. Part of the exercise is to see how effectively you can communicate your requirements and express your views and opinions
- Other members of the group may ask more questions for clarification
- Have a brief discussion which focuses on the barriers, issues and possible solutions t o the points identified for success (no more than 5-6 minutes – someone keep an eye on the time – remember there are five of you!)
- The person who initiated the discussion notes down the barriers, issues and solutions on their sheet
- Move on to the next person until everyone has had a chance to present
- Remember all of you need to keep the discussion very focused as you only have 5-6 minutes. There whole activity takes 30 minutes.

1. Each person presents their needs
2. The group discuss Barriers, Issues and Solutions (Max 5-6 minutes), the initiator writes these down
3. Move onto the next person

1 - Student/Researcher
2 – Lecturer
3 – Librarian
4 – IT Manager
5 – Head of Department
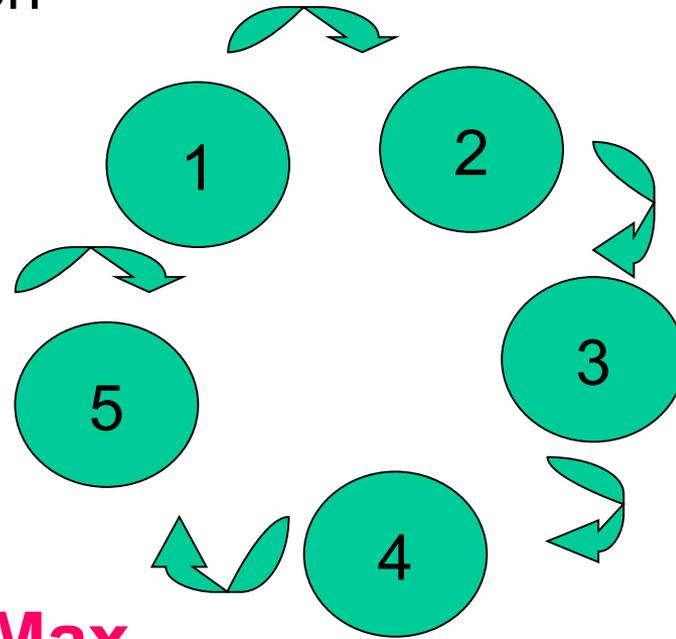
Time limit 30 minutes Max
Make sure everyone has a go!

Activity – a reminder

Activity 2 – same roles together

Activity 2 – same roles together

- Get together with person(s) of same role

- Tell each other the character that you chose, was it the same one?

- Discuss what you have found and make a note of any differences

- Create a combined list of *barriers issues* and *solutions*
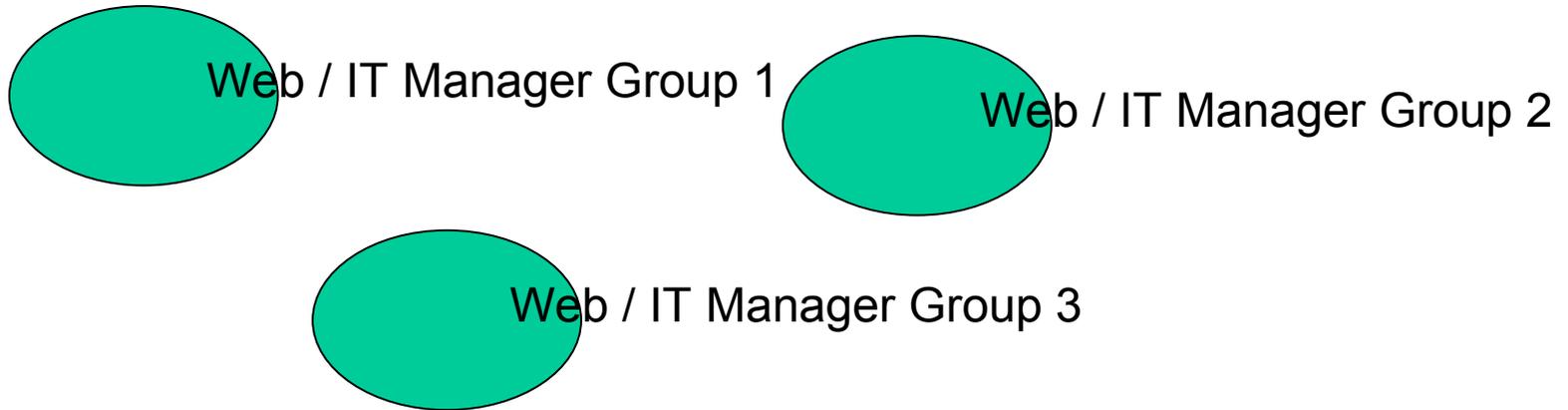
- Report back to whole group

MM

Activity 2 – same roles together

Reminder – 10 minutes MAX

Web / IT Manager Group 1

Web / IT Manager Group 2

Web / IT Manager Group 3

1. Compare your list of Barriers, Issues and Solutions
2. Note any differences and possible reasons why, e.g. did you have the same characters?
3. Combine the list
4. Report back to whole group when called out

Activity 3 - Reporting Back
5 minutes per role

- Say whether your character was supportive or had reservations about repositories

- Present a compiled list of Barriers, Issues and Solutions

- If there were different points raised, what were the reasons for this

Student / Researcher

Barriers / Issues

- **Lack of trust**
- **Quality control and management**
- **Peer review**
- **Preservation**
- **Replacing print with electronic**
- **Access control - who controls?**

Solutions

- **Training**
- **Departmental 'management'**

Lecturer

Barriers / Issues

- Control of IPR and copyright
- Access control
- Logging/usage statistics
- Quality issues
- Differences of opinion
- Repository as 'official' barrier to informal processes

Solutions

- Secure delivery
- Lockable content formats for teaching materials
- Policies on rights
- Quality workflow (metadata, peer review)
- Communication and education

Librarian / Information Manager

Barriers / Issues

- **Data quality**
- **Advocacy**
- **Copyright**
- **Policy issues need tackling before library will buy into the repository**
- **Resourcing, e.g. cataloguing, training, documentation**

Solutions

- **Extra staff and resources**

Institutional / Departmental Manager

Barriers / Issues

- **Management responsibility**
- **Metadata creation**
- **How do you join up with other universities?**
- **Legacy systems**
- **Access management (LDAP)**
- **Product choice?**
- **Multiple repositories?**
- **Flawed procurement process**
- **Content types**
- **Institutional barriers**
- **IPR**

Solutions

- **Resource**
- **Senior management buy-in;**

Web / IT Manager

Barriers / Issues

- **IT strategy**
- **Risk of new technology, scalability etc.**
- **LDAP and integration with other institutional services**
- **Metadata management**
- **Library involvement**

Solutions

- **IT director buy-in**
- **Library buy-in**

Conclusions

- Why communicate?
  - Gathering requirements
  - Advocating
  - Populating repository
  - Ensuring quality of metadata/service
- Communication methods
  - Are beyond verbal
  - Other methods include:
    - Scenarios and Use Cases; surveys; focus groups; interviews

www.ukoln.ac.uk

UKOLN A centre of expertise in digital information management

MM

Conclusions (2)

- Achieving buy-in and uptake from stakeholders is crucial to a successful repository.
- To do this, we need
  - EFFECTIVE COMMUNICATION
  - Advocacy
  - Rewards and motivators
  - No (few?) barriers in the deposit process
  - A high-quality and trusted service
- Can't solve the issues with technical solutions alone,
- or with policy and mandate alone

www.ukoln.ac.uk

UKOLN   A centre of expertise in digital information management

JA