

Thinking about the ecology of repositories

La ecología de los repositorios institucionales, Antigua Universidad Laboral de Gijón

R. John Robertson,
Repositories Research Officer
JISCCETIS
13th December 2007,
robert.robertson@strath.ac.uk

Overview

- Introductions
- Repositories domain- opportunities and challenges
- Why ecology
- Basic concepts of the metaphor
- An ecological mindset for repositories

Introductions

- R. John Robertson
 - ◆ JISC CETIS
 - ◆ Part of Repositories Research Team. We support JISC's development programmes in digital repositories and preservation
 - ★ Through training events, synthesis, and targeted research
 - ★ Team involvement in national and international developments (CRIG; OAI-ORE)
- Acknowledgements - ongoing input into and development of ecology work from:
 - ◆ Mahendra Mahey, UKOLN, RRT
 - ◆ Phil Barker, JISC CETIS

Repositories domain, opportunities and challenges (1 - standards)

- **The repositories domain is well served by technical specifications, standards, protocols, and architectural models** (WSDL, http, SRU/W, OAI-PMH, IMS CP, DC, IEEE LOM, MARC, JISC IE, CORDRA)
- They are generally stable, well defined, and understood
- There are various initiatives to move the community towards service-oriented approaches for software development (SOA, and soa) and selection.
- There are high level programming tools that allow the assembly of custom combinations of services and content (DELOS dlms)
- **Objects and metadata can be transported around such systems with relative ease** (- or at least the problems are understood)

Repositories domain, opportunities and challenges (2 - implementation)

- Implementers of repositories can have difficulty in planning and managing their particular service in relation to the rest of the information environment, when trying to:
 - ◆ Establish service connections – the required awareness of how an external service has actually used a specification and standard
 - ◆ Express complex dependencies – the need to communicate why a low-profile service is vital or why the same architecture can work in one institution and fail completely in another.
 - ◆ Identify opportunities – pinpoint what is missing and needed in a service network and take advantage of the gap
 - ◆ Manage ongoing development – understand and communicate how the technical and non-technical dimensions interact
- “Planning and articulating these interactions requires a way of thinking that can capture and address the *untidy complexity of specific interactions* found in the real world”

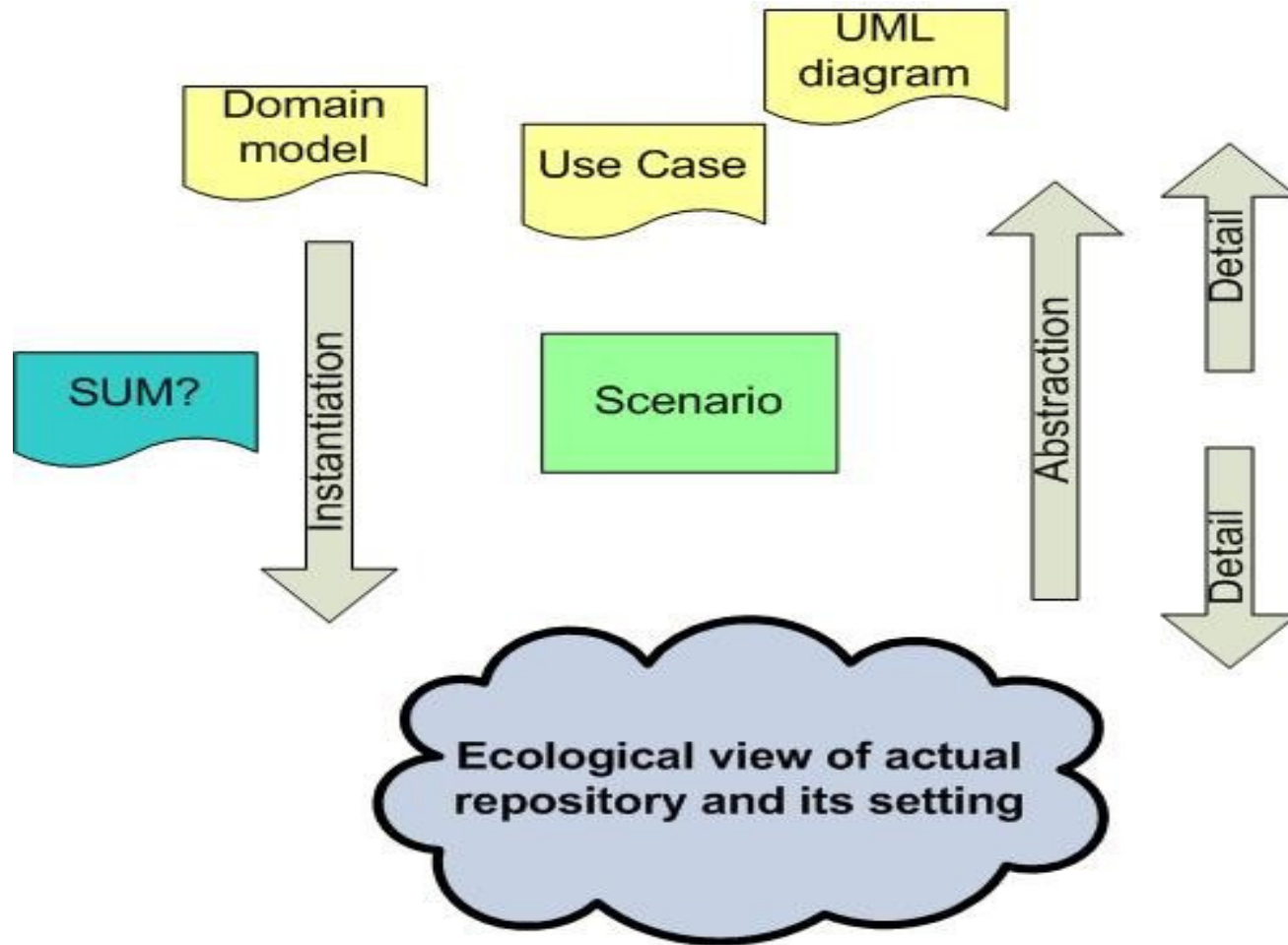
Repositories domain, opportunities and challenges (3 - building service networks)

- **There is a need to express and understand the impact of local decisions and culture on interoperability with particular communities, populations, or ecosystems**
 - ◆ Content issues include: different access restrictions: Intellectual Property Rights, non-digital objects; formats.
 - ◆ Metadata issues include: standards, element selection/ application profile, vocabulary choices, assumed knowledge (regional resources/ language codes)
 - ◆ Local attitudes: commitment to Open Access, concern about plagiarism, ability to find materials, preservation state of original, willingness to expose full text.
- These factors present a barriers to participation but a clear articulation of such interoperability boundaries is necessary to address them (e.g. by adjusting metadata at creation, export, or by use of third party service).

Repositories domain, opportunities and challenges (4 -describing specific practice)

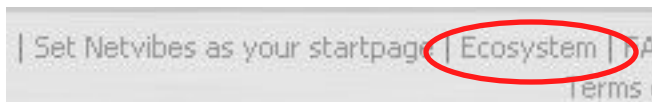
- We have suggested that we need a way to capture and articulate local practice in its relationship with external services
- Established methods strive for:
 - ◆ unitary granularity
 - ◆ abstract representations of technical interactions
 - ◆ Implicitly static representations of practice
- Established methods struggle with
 - ◆ presenting general conditions (e.g. university policy, or the impact of funding bodies on networks)
 - ◆ asking why
- **Something else is required...**
 - ◆ **To articulate key issues across technology, practice and setting.**
 - ★ One option is an ecologically influenced approach

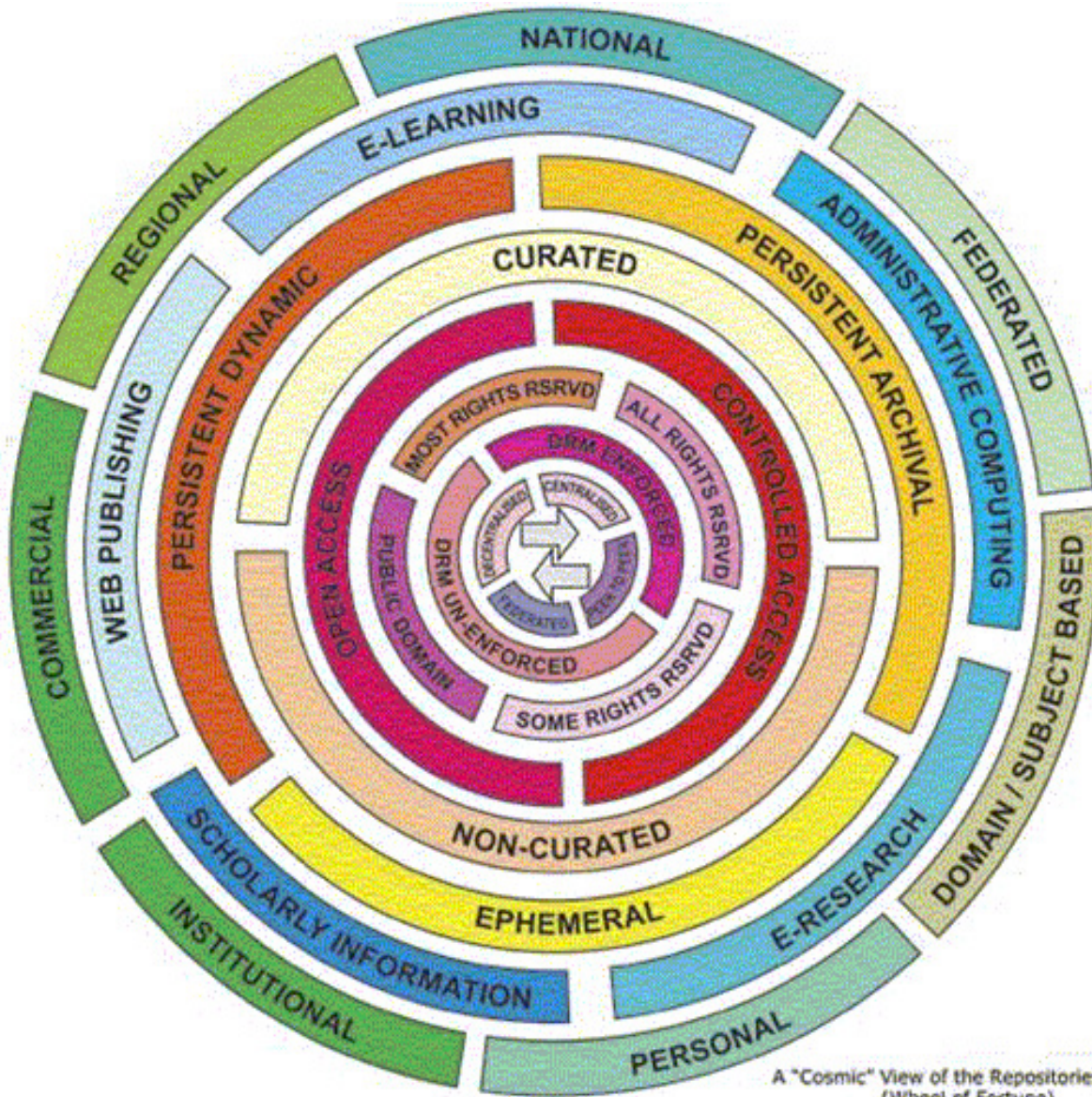
Repositories domain, opportunities and challenges (5- the relation of specific tools)



Context of our work

- **Ecology presents a metaphor that has been adopted by other domains and is – to a degree – intuitive**
- Some previous use of ecology in information science:
 - ◆ Nardi & O'Day introduced some concepts from ecology in *Information Ecologies: using technology with heart*
 - ◆ Davenport addressed similar questions in the context of ICT in a business setting in *Information Ecology: Mastering the Information and Knowledge Environment*
 - ◆ They present relevant issues and concepts from ecology, but present little in terms of methodology
 - ◆ Blinco and Maclean's 'Cosmic view' on repository ecologies offers a tool to identify or type repositories.
- Some natural usage of the term ecology or ecosystem
 - ◆ Blog posts (recent posts using metaphor by: Dave Cormier, George Siemens, Kia Pata, Alan Dix)
 - ◆ Netvibes 'ecosystem'



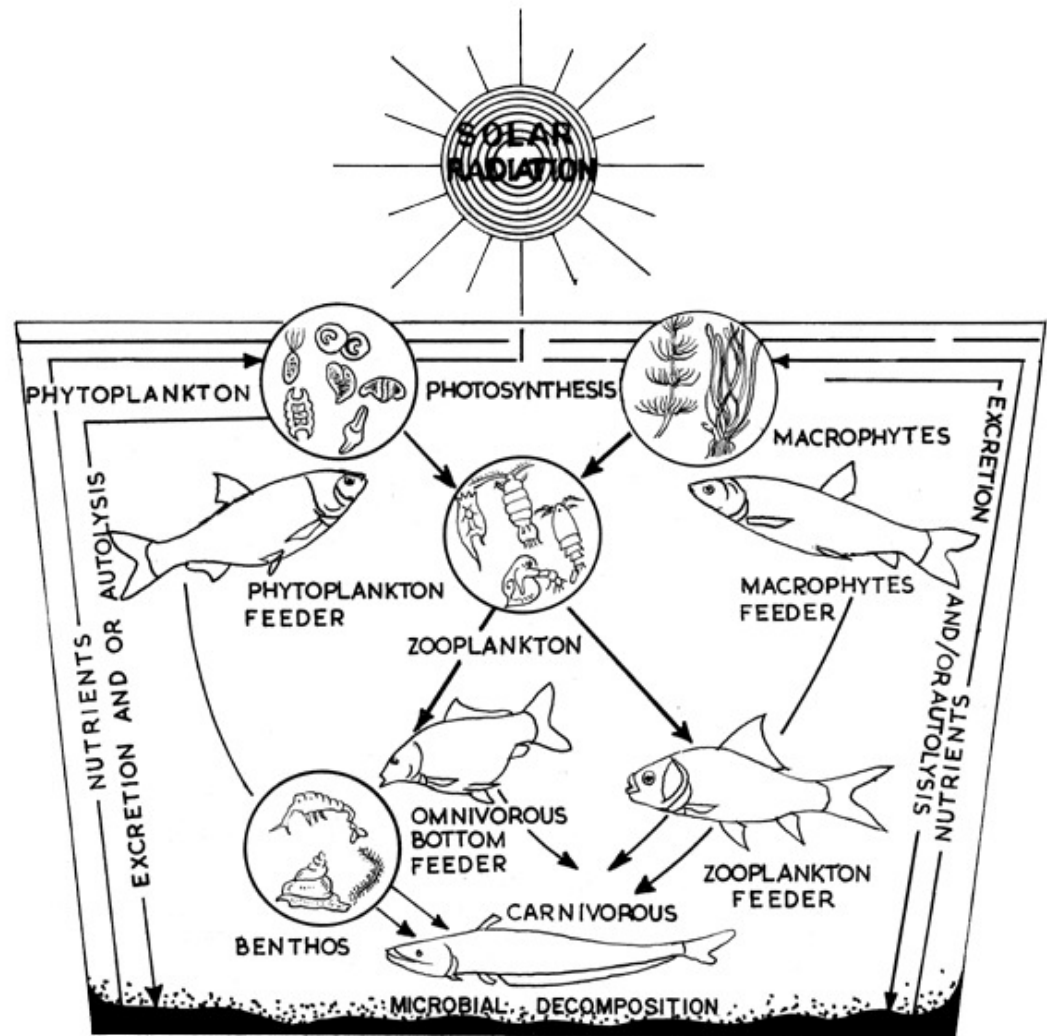


A "Cosmic" View of the Repositories Space
(Wheel of Fortune)

© Kerry Blinco & Neil McLean, 2004

Why Ecology?

- Ecology may be defined as:
“the branch of biology dealing with the relations and interactions between organisms and their environment, including other organisms”
(Dictionary.com)
- A classic example is a pond.



Pond ecology © FAO

Why Ecology? -statement of purpose

- What are we trying to do?
 - ◆ Articulate the human alongside the technical
 - ◆ Articulate the complexity and messiness
 - ◆ Support better communication and management
 - ◆ i.e. ask 'what' questions to support understanding the 'why' questions
- What are we not trying to do?
 - ◆ This is not a replacement for architectures; whenever you start to design an system or implementation you still need to start here
 - ◆ This is not a replacement for service usage models, service expressions, or related soa work.
 - ◆ This should not be seen as an attempt to model everything but rather as the creation of a view with a purpose.
- To ask if it is useful to have:
 - ◆ [a branch of information science dealing with the relations and interactions between network components and their environment, including other network components] which is metaphorically informed by ecology.

An aside: defining a repository

- “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. While operational responsibility for these services may reasonably be situated in different organizational units at different universities, **an effective institutional repository of necessity represents a collaboration among librarians, information technologists, archives and records managers, faculty, and university administrators and policymakers**. At any given point in time, an institutional repository will be supported by a set of information technologies, but a key part of the services that comprise an institutional repository is the management of technological changes, and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services. An institutional repository is not simply a fixed set of software and hardware.”

(Clifford Lynch ARL: A Bimonthly Report, no. 226 (February 2003) “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age” <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml> ;

Basic concepts from ecology that are relevant to digital libraries, repositories, and services

- A dynamic system
- A view at particular scale
- Entities and species
- Interactions
- Environmental factors

A dynamic system

- A system
 - ◆ The actions of one entity in the system can effect other entities and possibly the whole system
 - ◆ Changes in the environment of the system can affect entities and their interactions
- Constantly changing
 - ◆ There is an implicit understanding that the environment is constantly changing (and any representation only produces a snapshot)
- Character
 - ◆ Successful changes need to work with the ecology

A dynamic system (2)



Brewbrooks (2007)
<http://creativecommons.org/licenses/by-nc-sa/2.0/>
<http://flickr.com/photos/brewbrooks/397238796/>

A view at a particular scale

- There is a clear need to articulate what level of interaction you are trying to describe or model. Are you describing a microbe, a herd of giraffes, a river valley, or a city?
- There is overlap and interaction between different scales but, to be detailed enough, an analysis should focus on one at a time.
- For the repository domain:
 - Are interactions between repository networks, repositories, objects, or metadata?
 - Is it about the policies of a region, country, institution, class, academic?

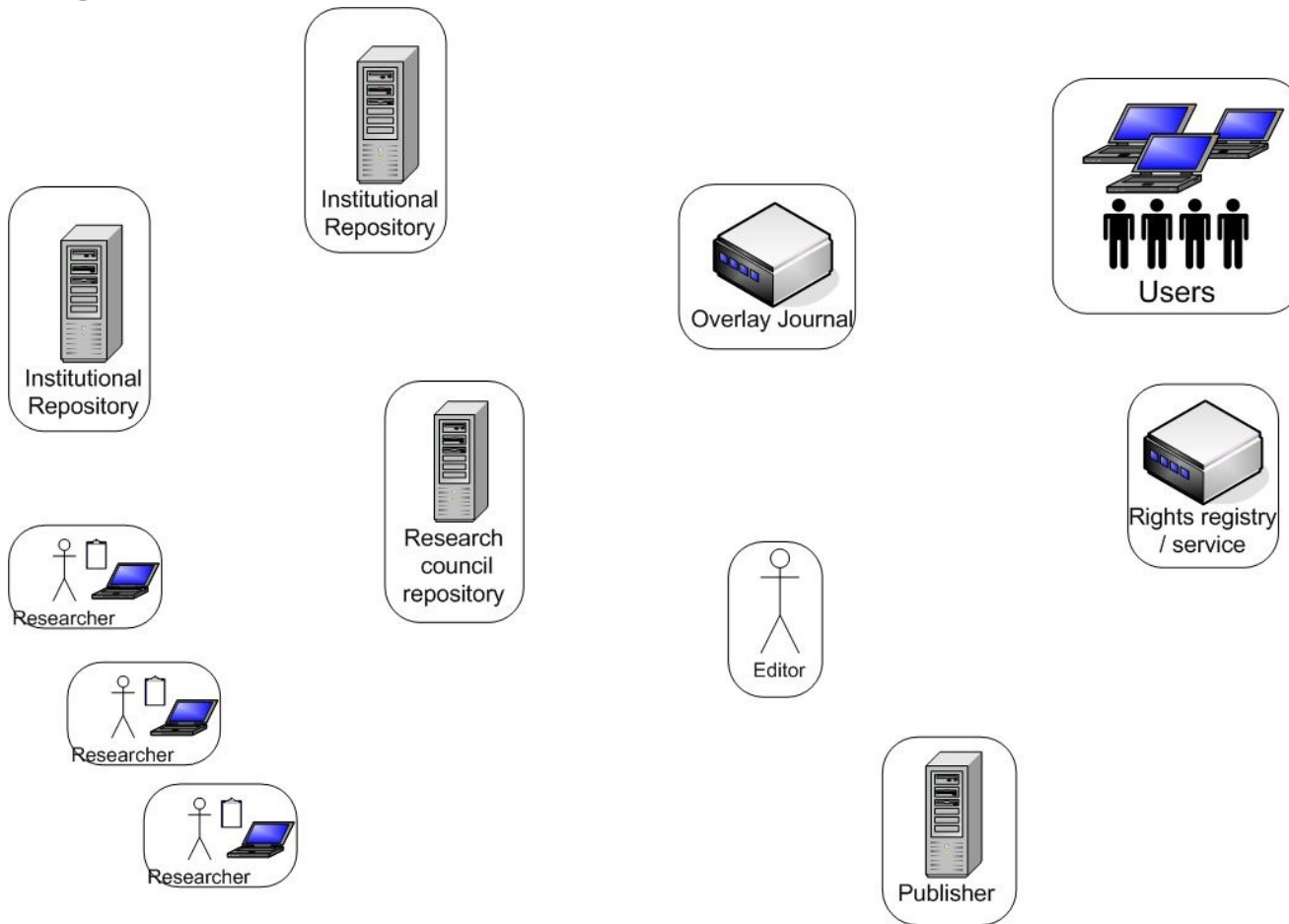
A view at a particular scale



Entities (and species)

- “An entity is a tangible thing that exists within a repository ecosystem.”
 - ◆ William, the repository administrator
 - ◆ E-lis, a subject repository
 - ◆ An article by Alan (<http://eprints.cdlr.strath.ac.uk/1062/>)
 - ◆ Oaister, an aggregated search service
- “A species within an ecosystem is a collective name for a particular type of entity.”
 - ◆ It allows what is known about the behaviour of the species to help understand the actions of a particular entity.
 - ◆ Common types of entities are: users; repositories; services; objects; metadata records.

Example: entities in an overlay journal



Interactions

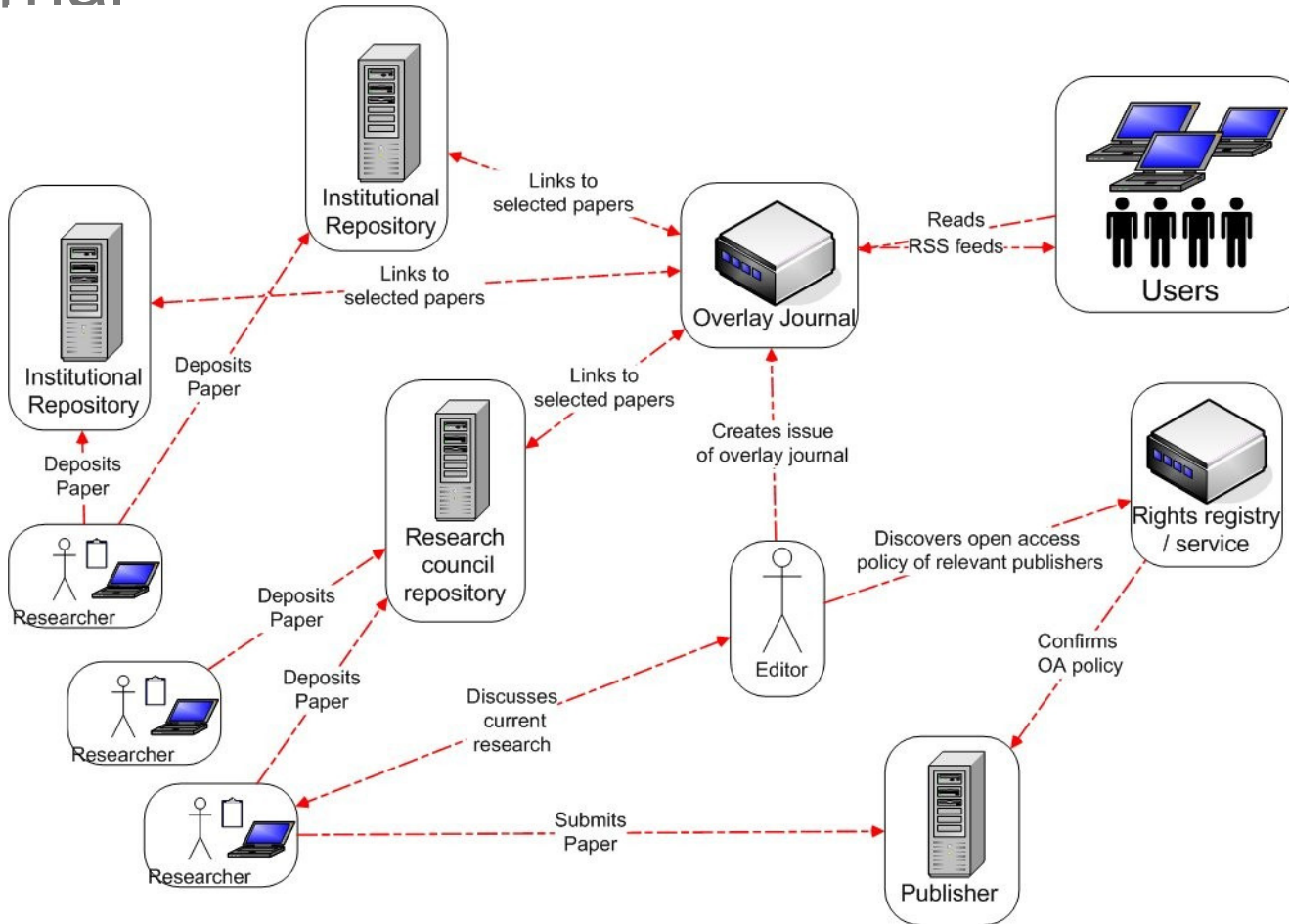
- “An interaction is a connection, relationship, or link between two or more entities or species in a population, community, or ecosystem”
- This is a broad definition, it is not just related to a function or technical request and can be between any type of entity; for example:
 - ◆ harvest records using oai-pmh
 - ◆ contact repository administrator
 - ◆ locate article
 - ◆ share information
- It is useful to include the nature of the interaction [submits, approves, informs, talks about] and the object (pdf , thesis, image, knowledge, rights)

Interactions (2)



Symbiosis by "MrUllmi" <http://flickr.com/photos/mrullmi/245555310/>
http://creativecommons.org/licenses/by-nc-sa/2.0/deed.en_GB

Example: interactions in an overlay journal



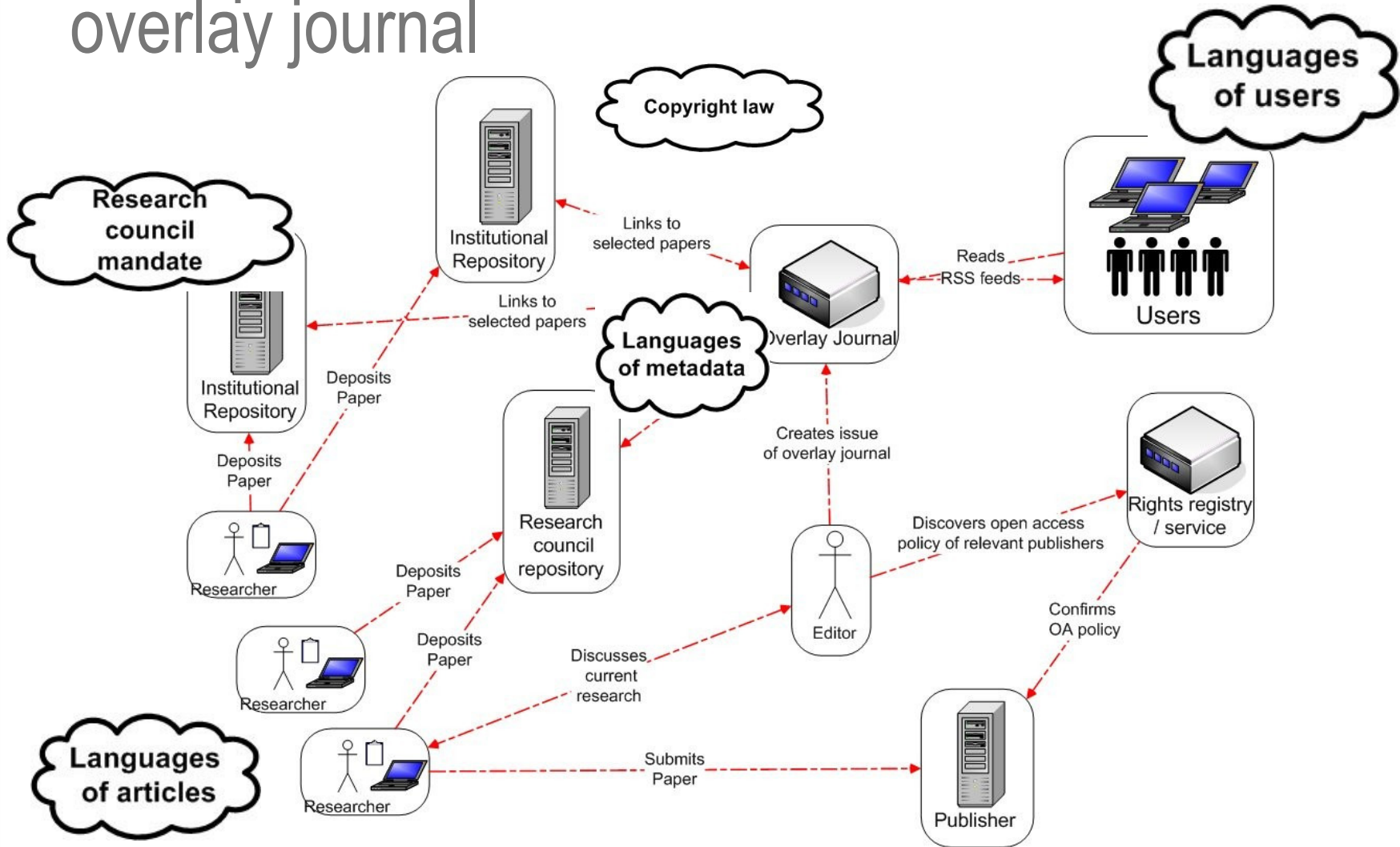
Environmental factors

- “An environmental factor is something that influences an entity, community, or ecosystem but is more general [i.e. wide ranging] than an interaction between constituent entities. ”
- It can be
 - ◆ Something that affects the whole environment (a change in copyright law, a university-level review of research output)
 - ◆ Something affecting particular entities or species (the availability of a funding stream for Open Access publishing costs)
 - ◆ A direct effect from an entity not involved in the particular environment under consideration (competition for funding or political manoeuvring)
 - ◆ Something that would be an interaction if considered on another scale (the release of a new specification, or an new trend in the web environment)

Environmental factors



Example: environmental factors in an overlay journal

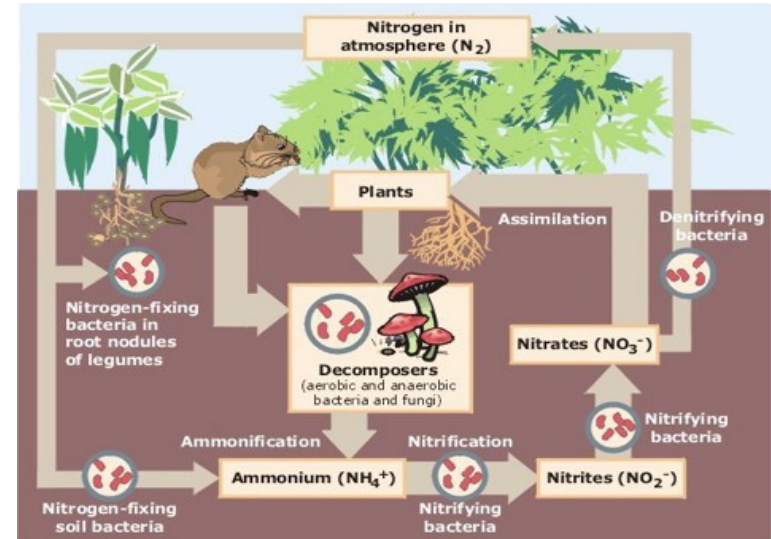


An ecological mindset

- Modelling questions
 - ◆ What sort of thing (repository or service) is this? – identifying species and entities
 - ◆ What does it relate to (other repositories or services)? – specify interactions
 - ◆ What does it depend on? – environmental factors and key interactions
- Analysing questions
 - ◆ How adaptable is it?
 - ◆ What helps it to thrive?
 - ◆ Are there keystone entities/ species
 - ◆ Is there sufficient diversity?

An ecological mindset (2)

- Two key methods:
 - ◆ Resource tracking
 - ★ Public domain image, source US Environmental Protection Agency
 - ◆ Habitat mapping

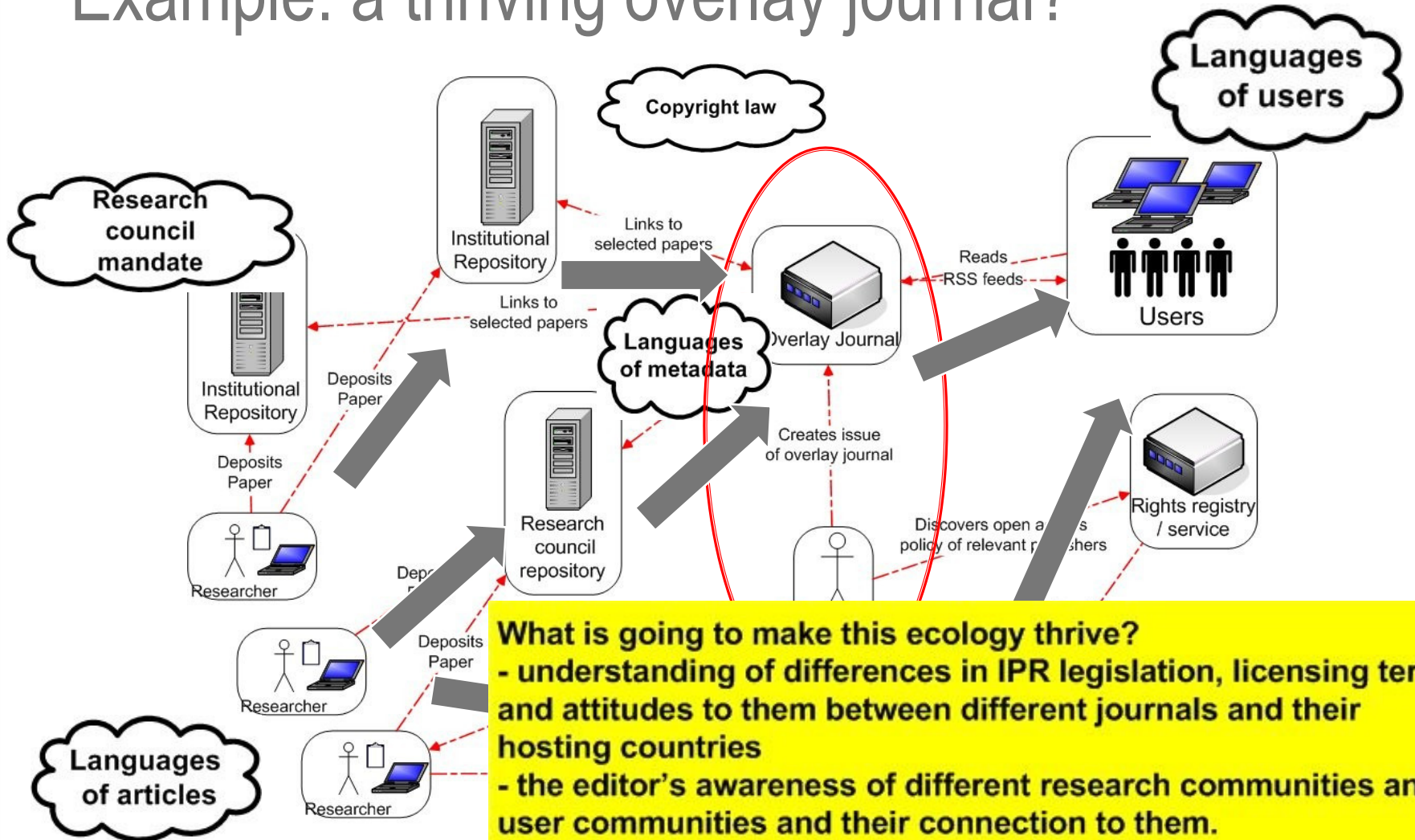


The term *ecological niche*] “is often used loosely to describe the sort of place in which an organism lives as in the sentence: 'Woodlands are the niche of woodpeckers'. Strictly, however, where an organism lives is its *habitat*. A niche is not a place but an idea: a summary of the organism's tolerances and requirements”
Begon *et al* “Ecology from Individuals to Ecosystems” (2006) p

An ecological mindset (3)



Example: a thriving overlay journal?



What is going to make this ecology thrive?

- understanding of differences in IPR legislation, licensing terms, and attitudes to them between different journals and their hosting countries
- the editor's awareness of different research communities and user communities and their connection to them.
- understanding the possible effects of web2.0 and social software on how users want to access and use information (adding RSS/ Atom feeds, rights issues with using flickr/ scribd)

Future work

- Review of draft report
- Exploration of connections with
 - ◆ Digital Business Ecosystems
 - ★ Digital Business Ecosystems EC projects
 - ◆ Aspect oriented programming
- Further development and publication of case studies

- Questions ?
 - ◆ robert.robertson@strath.ac.uk

Further information

- Draft report
 - ◆ An ecological approach to repository and service interactions
<http://www.ukoln.ac.uk/repositories/digirep/index/Ecology>
- Related reading
 - ◆ **ECDL/ JISCCETIS Conference**
 - ◆ Bonnie A. Nardi, and Vicki L O'Day, First Monday Vol 4 No 5 May 3, 1999. Information ecologies: using technology with heart. Chapter Four: Information ecologies.
http://www.firstmonday.org/issues/issue4_5/nardi_chapter4.html
 - ◆ Thomas H. Davenport, Information ecology, OUP, 1997
 - ◆ R. Heery and A. Powell, Digital Repositories Roadmap: looking forward
<http://www.ukoln.ac.uk/repositories/publications/roadmap-200604/>
 - ◆ Rachel Heery and Sheila Anderson, Digital Repositories Review, UKOLN and AHDS, 2005 (Final version)
http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf
 - ◆ R. J. Robertson and J. Barton, "Optimising Metadata Workflows in a Distributed Information Environment", Digital Repositories: Interoperability and Common Services, The Foundation for Research and Technology, 11-13 May 2005 , Hellas (FORTH), Heraklion, Crete, 2005,
<http://delos-wp5.ukoln.ac.uk/forums/dig-rep-workshop/robertsonbarton.pdf>
 - ◆ Robertson, R.J. et al. EThOS, the new UK e-theses service, national and institutional repository interaction. JISC Conference, March 2007.
<http://www.jisc.ac.uk/media/documents/events/2007/03/ethos.ppt>
 - ◆ Kerry Blinco and Neil McLean, 'A 'Cosmic' View of the Repositories Space (Wheel of Fortune)', 2004, <http://www.rubric.edu.au/extrafiles/wheel/main.swf>