An ecological approach to repository and service interactions

Version 1.2

R. John Robertson (JISCCETIS, University of Strathclyde)
Mahendra Mahey (UKOLN, University of Bath)
Julie Allinson (University of York)
March 2008

1 Statement of intent and acknowledgements

This work began in response to a perceived need to express something of how and why repositories and services interact.[1] As a community we have well understood technical models and architectures that provide conceptual mechanisms to promote interoperability. Articulating the details and challenges of actual interactions that occur, however, is not so widely supported understood and knowledge about them is not often shared. This is, in part at least, because we tend to share in the abstract through architectures and use cases and in these we focus on the technical. Articulating interactions or connections well requires an engagement with and presentation of specific local details. Beginning to consider why particular interactions succeed or fail over time requires us to factor in more than the technical.

We think that the systems that the community is constructing benefit from being described and understood well and that such understanding allows the possibility of better future development. Within the context of the Information Environment, a greater understanding of the details of particular interactions may enable richer services, more efficient interoperability, or may simply support new connections.

Ecology is the study of systems that are complex, dynamic, and full of interacting entities and processes. Although the nature of these interactions and processes may be highly detailed, a higher level view of them is accessible and intuitive. We think that ecology and the ecosystems it studies may offer a useful analogy to inform the task of understanding and articulating the interactions between users, repositories, and services and the information environments in which they take place. This report outlines some concepts from ecology that may be useful and suggests some definitions for a common conversation about the use of this metaphor. It concludes by examining how this approach relates to other initiatives currently ongoing in the JISC Information Environment.

We hope that this report suggests a new way to conceptualise and analyse interactions. It should as a minimum provoke and support some useful discussions about repository networks, communities, and what we need to be able to express about them.

As well as all those cited, we would like to acknowledge the contributions made by colleagues at CETIS and UKOLN as well as other who commented on the draft report and those who participated in the workshops at ECDL2007 and JISC CETIS2007.[2]

John, Mahendra, and Julie
rrt@ukoln.ac.uk

[2] http://www.ukoln.ac.uk/repositories/digirep/index/ECDL2007Workshop:Towards_an_European_repository_ecology ; http://wiki.cetis.ac.uk/Learning_Resources_in_the_Ecology_of_Repositories

2 Thinking about repository interactions

The JISC Information Environment is steadily becoming populated with repositories and services that make available or enhance a wide range of digital assets which "support researchers, learners, teachers and administrators in their work and study". (http://www.jisc.ac.uk/whatwedo/themes/information_environment.aspx).

As these different types of repositories and services develop there is a growing need to consider the interactions between repositories and between repositories and other systems. 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 as well as supporting the consideration of abstract types of interaction.

In the *Digital Repositories Review*, Rachel Heery and Sheila Anderson commented that, as the use of repositories matures,

> A framework needs to be established for repositories that would encompass:
> - relation between repositories
> - data flow between repositories
> - workflow issues
>
> This would begin to address fundamental questions, such as how institutional repositories relate to thematic, subject repositories? Within institutions, how do repositories relate across the 'service domains' of research, learning, administration? A meeting point is required at various levels, both as regards service provision and technical infrastructure. (Heery and Anderson, 2005)

The *Repositories Roadmap* regarded the development of such a framework as a necessary milestone in moving the organisational viewpoint forward to achieve their 2010 vision for the sector (Heery and Powell, 2006).

This report is a first step in the development of this framework - it introduces the idea of using ecological concepts to describe repository interactions; illustrating its application with specific reference to examples from the UK FE/HE sector, and outlines how this approach fits with other related work. Such a discussion must also make reference to the distributed services that use information made available by repositories and the part they play in a repository ecosystem.

It is anticipated that the report will be of most relevance to the communities of repository and service developers, JISC, and other funding bodies as they plan future developments and as they articulate the current state of the Information Environment.

The term 'repository' is used by the report in a similar sense as Heery and Powell use it, that is (following Lynch, 2003) to view a repository as a commitment by a group (typically an institution) to support its members and users by providing set of services. As much of the discussion about repository interactions takes place in the context of distributed services, repositories are sometimes described in terms of their function as a data provider to these distributed services. The term 'service' refers to a specific function offered to a user community; most often in this document a 'service' exists independently of one or more data providers and interacts with them.

3 Why technical architectures need help

In its aim "to allow discovery, access and use of resources for research", the JISC Information Environment (IE) provides an overview of its expected technical architecture.[3] Such technical architectures are required because they:

1. allow developers and implementers to gain an overview of the potential components in complex distributed systems.
2. specify the component types of information environments and protocols for their technical interoperability.
3. consequently support planning for the deployment of such large-scale systems
4. provide an abstraction of the participants in an information environment that may assist the identification of missing types of participant (data or service provider) within that environment.

Architectures and architects are needed to plan the way forward and enable the development and construction of great edifices. This is as true for the deployment of software and the creation of local approaches for its use, as it is for building a skyscraper. Architectures are, however, not, in themselves, sufficient for the longer-term development and management of interactions in an information environment – whether large or small. Other approaches are also needed to help those developments mature, take root, and flourish.

Some suggested reasons why the ongoing growth of an information environment requires that the technical architecture be supplemented with other 'models'.

1. architectures are implicitly static and only address the structures that have been built into them. In particular, technical architectures formalise the technologies they include and the inclusion of a new technology may require the entire architecture to be updated.[4]
2. actual implemented practice rarely matches the architectural specification exactly.
3. information environment architectures tend to focus on types of technical deployment and interaction. Non-technical interactions or constraints are not often represented.
4. architectures deal with abstractions of data or service providers (n, n+1 components) and types of interactions– they do not represent the peculiarities of actual instances of services and their interactions.
5. architectures aren't designed to express influences or similar technical or cultural factors, that affect an information environment without being part of it.
6. architectures may not be the best way to engage managers and users without a technical background and may not address the questions they want to ask.

---

[3] http://www.jisc.ac.uk/whatwedo/themes/information_environment.aspx

[4] It must be acknowledged that many technological architects, so to speak, are quite aware of the indicated problems, and often hold and use much more flexible mental models of their architectures and edifices. Indeed their architectures may never actually be finished until the systems they are creating are deployed and in particular use. Even then, their view of the may continue to be dynamic entities. This does not negate the point that many technical architectures are perceived as static – especially by those from outside the particular local context or by higher-level managers.

Managers, developers, and implementers need to be able to do some of the things that an architectural approach can't articulate. They need to understand, articulate, and reflect on the specific details and conditions of their project and how it will actually relate to other repositories and services. Questions like 'How will staff in the education faculty want this information presented?' or 'If I want to use JHOVE how do the outputted fields map onto our settings?' or 'Which members of the university's senior staff do I need to persuade?' inevitably arise.

This report suggests that ideas and concepts from ecology can complement the architectural view of an information environment and help developers, implementers, and managers articulate and analyse their setting and connections.

4 Why a service oriented approach isn't the whole solution

There are of course other models that have recognised that the traditional architectural approach is not entirely satisfactory. One major effort to address some of these issues has been the development of Service-Oriented Architecture and service oriented approaches. A service oriented approach to software design and implementation recasts the traditional architecture in terms of "networks of loosely-coupled communicating services"(http://www.e-framework.org/Resources/Glossary/tabid/642/Default.aspx). Such an approach allows for the creation of software architectures that are conceptually constructed from interchangeable components or groups of components that provide particular services. Such an approach also permits the associated business processes to be conceived of as services.

Such a flexible design is then instantiated as a Service Oriented Architecture (SOA):

> Software architecture for a service-oriented approach (soa) implemented using a particular technology, e.g., CORBA, Web Services. SOAs enable domain-oriented operations through message exchange within a network of services. Services and messages have platform-independent formal definitions, based on open standards. They define standard interfaces and protocols to encapsulate information tools as services that clients can access without knowledge or control of their inner workings. Note that soa is a generic concept. SOA is that concept implemented using a particular technology, typically Web Services.
> http://www.e-framework.org/Resources/Glossary/tabid/642/filter/S/Default.aspx

Such an approach is able to address some of the key criticisms made about the architectural approach and provide a significantly more dynamic model of software development and selection. In particular it provides:

- a flexible approach to software combination and recombination (and consequently the ability to move beyond monolithic applications or fixed combinations of services)
- better opportunities for user engagement in the design and selection of software as particular functions or services can often be matched to planned business processes or other user activities.
- a better match between implemented practice and specification (as a result of the above selection and tailoring).
- a model that, although still using the underlying idea of architecture, is thought of in much more flexible and dynamic terms.

Although soa/ SOA offers a significantly richer approach than the traditional technical architecture we would suggest that it does not fully address some of the concerns we have raised about architectures and their use.[5] Even after the possibilities of soa have been considered, we contend that there is still a significant gap in how the repository community addresses issues around communicating and managing the success or failure of repositories and their interactions. These may be summarised as follows:

[5] It should be noted that it is the concept of soa that is being discussed. Issues surrounding the costs and effectiveness of its widespread deployment in practice are not part of this part of this analysis.

- Although soa takes account of non technical issues in the selection of services it is fundamentally about modelling software services and related business processes. It is not interested in representing non-technical interactions or constraints. Similarly soa isn't designed to include or communicate influences and cultural factors in its articulation of a system (however much they may be involved in the selection of suitable services or service implementations).
- Although soa / and SOA is able to select compatible services and is concerned with ensuring they can interoperate properly it is not clear to what degree they are interested in the articulation of the precise configuration of particular systems.[6]
- Although soa-influenced architectures are intended to involve input from non-technical managers and users in the process, the primary audience for the output of the process is the technical developer. The soa principles are meant to support good design and reusability but about technical architectures and communicated in a highly technical way.

Specific applications of architectures (the IE architecture) and soa/ SOA (the eFramework http://www.e-framework.org/) within the context of UK tertiary education will be discussed in more detail in section 12.

[6] This observation is influenced by the particular work of the eFramework which explicitly does not record information about instances– it stops at the level of recording particular arrangements of software (implementations in their terms) which support a given arrangement of services (a service usage model); counter examples/ corrections are welcome and will be incorporated in any subsequent revisions.

5 Towards the missing model

Thus far we have suggested that neither a traditional architectural approach nor a service-oriented approach are able to fully present the types of issues that need to be addressed by those attempting to describe, model, or manage repository and service interactions. Our critique of those approaches has by implication sketched an outline of what type of approach is missing. In this section this gap is examined in more detail, before, in the following section, one supplementary model is proposed.

When the interactions of a repository are being considered the primary function of an architectural model is to describe the component technical services, software, and standards in use for a particular purpose – detailed architectures might also note or reference the taxonomies used by the metadata standards.

There are of course other tools that can be used to describe some parts of these interactions. Some examples of these are: Service Usage Models (SUMs), Use cases, UML diagrams, Scenarios, and Domain models.

SUMs provide an application independent detailed view of a given arrangement of software services. A use case provides a view of a single interaction between an actor and a system. It delineates how the actor uses the system to accomplish a goal and records the different major decisions, paths, and outcomes as that interaction progresses. It originates in software design methodology and in that context is a tool which supports the derivation of system requirements. Use cases can be recorded in natural language but are also frequently recorded more formally in UML (or an equivalent). These tools all deal with hypothetical routes through the system.

Scenarios often provide the context of a use case or the source material from which use cases can be derived. They tell a story of how a user or users interact with a system. In doing this they not only engage with the system interactions but provide some context for the user and include what the user considers to be a successful use of the system. This contextualisation is intended to parallel a real setting or aggregate of a real setting. As such a scenario-based approach begins to address some of the types of information that we have noted are missing from other approaches. However, scenarios tend to be quite narrowly focused on particular interactions and are frequently closely tied in with requirements gathering and corresponding use cases. They are also intended to be ‘soft’ and not address technical issues in too much detail.

The idea of a domain model has been developed to expand on architectural model and as such it provides an interesting counterpoint to where we consider there to be a gap in the models that we, as a community, develop and use. The role of a domain model can be defined as follows:

> Domain models bridge the gap between the analysis of requirements (x-ref) and the production of design specifications. A domain model represents the common understanding of a key concept in the organisation. (http://www.jiscinfonet.ac.uk/InfoKits/creating-an-mle/mle-design/domain-models)

In a sense the domain model provides a view of a given context that allows a set of requirements or generic statements to taken and be developed for that context. For example, a domain model of assessment in higher and further education in the United Kingdom might present an agreed set of generic terms for this community to use to describe what they do. A domain model is a structured form of taxonomy or classification abstract from the actual practice of a community.

With these tools and approaches in mind, we consider that there is a missing approach that supports the description and potential analysis of the specific. In particular an approach is needed that is able to help capture and represent particular local practice and context alongside system details. That is able to take a view of users, repositories, and services in a particular context and articulate in selective detail what is actually going on in that context. Such a model would have to be able to consider and present technical systems and technical interactions alongside interactions in the user community (both within the community and with the system) and relevant cultural, legal, and social influences or constraints.

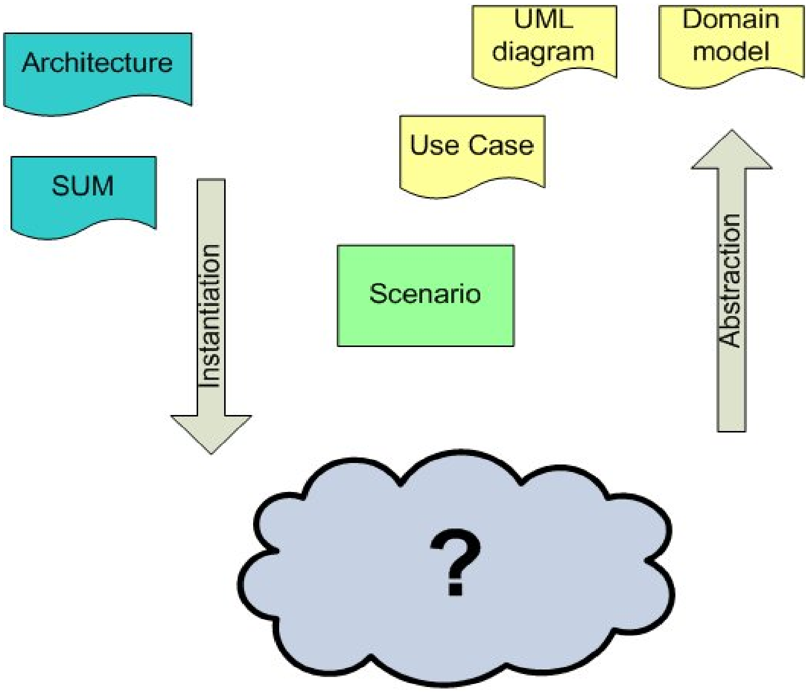


Figure 1 Describing repository interactions - what's missing?

Such a model should be able to select the key participants, interactions, and factors in a particular setting and present them in a way that enables better communication between technical and non-technical users and managers. It would also help articulate and possibly address particular types of questions about the success of failure of given repositories or service combinations and their use. Such questions might include issues like:

- How the publication tracking practices of academic departments interact with work of a repository.
- How the uptake of web2.0 technologies and institutional concerns about IPR interact with the use of the repository
- What dependencies a successful service has – on other services, community support, particular technical infrastructures, particular funding streams, or legal frameworks.

6 An ecological view of repository interaction

One method of developing a supplementary model of repository and service interactions is to review how other subject domains address similar questions of complex interactions of many different types in particular settings. One domain that addresses representing such complexity is ecology.

6.1 What do we mean by ecology?

Ecology may be defined as "the branch of biology dealing with the relations and interactions between organisms and their environment, including other organisms".[7] As such, it considers incredibly complex systems (often very selectively). It deals with many different scales (from the habitat of an organism to the habitat of a species). It takes account of many different types of interaction (e.g. chemical, physical, meteorological, social).

A classic example in ecology is a pond.[8]

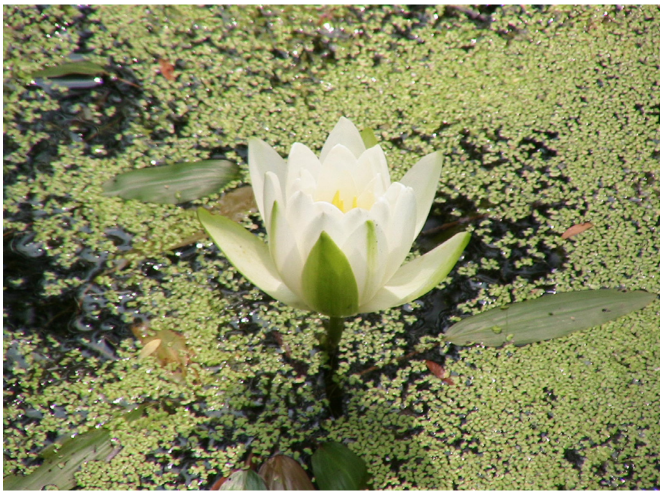


Figure 2 photo © Neal Singleton for http://9936.openphoto.net CC:PublicDomain

The ecosystem of a pond has a number of key features:

- it presents a number of different habitats (for example, the surface habitat and the shore habitat).
- it has a clear food chain with different trophic levels (for example, plant life (e.g. algae) sustains a small number of insects, which in turn sustain a smaller number of fish).[9]
- it is clearly influenced by both biotic and abiotic environmental factors (for example the presence of predators or the composition of the surrounding soil).[10]

---

[7] ecology. (n.d.). *Dictionary.com Unabridged (v 1.1)*. Retrieved June 08, 2007, from Dictionary.com website: http://dictionary.reference.com/browse/ecology

[8] Joe Lewis (n.d.) , Pond Ecology, Yale-New Haven Teachers Institute Curriculum Unit 92.05.07 http://www.yale.edu/ynhti/curriculum/units/1992/5/92.05.07.x.html

[9] "Relating to the feeding habits of different organisms in a food chain or web" trophic. (n.d.). *The American Heritage® Science Dictionary*. Retrieved June 28, 2007, from Dictionary.com website: http://dictionary.reference.com/browse/trophic

- it is affected by distinct biological and chemical processes (for example, the relative abundance of nitrogen will directly affect the growth of plant life and the abundance of plant life may have a direct effect on the sunlight present to animal life in habitat at the bottom of the pond.).
- within it all of these processes, species, and environmental factors interact with each other.
- it is a dynamic system that is constantly responding to changes.

Within this example it should be observed that ponds are often managed rather than entirely wild. They are cared for and specific action is taken to promote or hinder the growth and survival of particular components of the ecology.

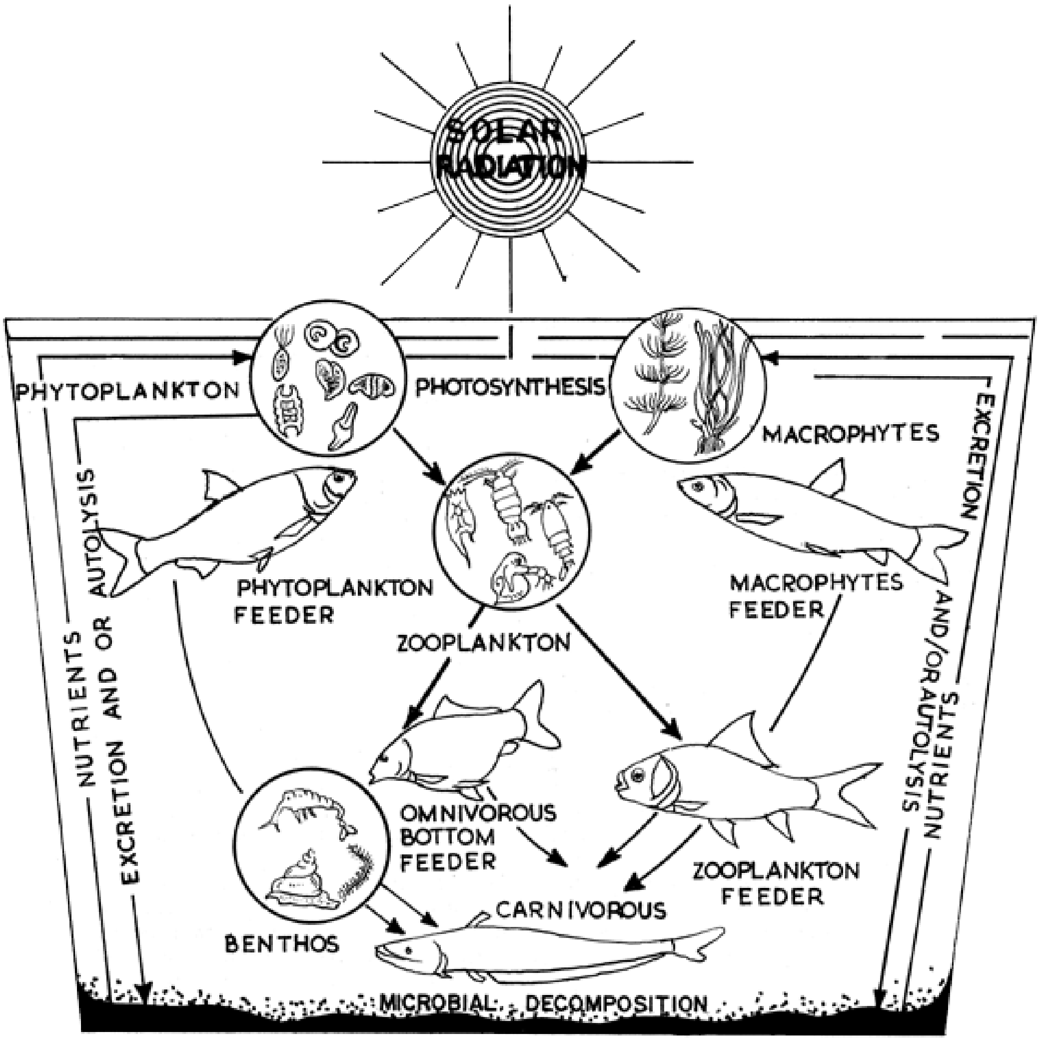


Figure 3 A Pond Ecosystem, from D. Kumar (1992) Fish culture in undrainable ponds: A manual for extension, FAO Fisheries Technical Paper No. 325. Rome, FAO http://www.fao.org/docrep/003/T0555E/T0555E00.HTM . Image © Food and Agricultural Organisation of the United Nations used by permission.

[10] Biotic "2. Associated with or derived from living organisms. The biotic factors in an environment include the organisms themselves as well as such items as predation, competition for food resources, and symbiotic relationships." biotic. (n.d.). *The American Heritage® Science Dictionary*. Retrieved June 28, 2007, from Dictionary.com website: http://dictionary.reference.com/browse/biotic ; "Not associated with or derived from living organisms" abiotic. (n.d.). *The American Heritage® Science Dictionary*. Retrieved June 28, 2007, from Dictionary.com website: http://dictionary.reference.com/browse/abiotic

6.2 Ecology and information systems

In the context of repositories and services there are a number of parallels between information systems and ecologies that suggest that an ecological approach may be both apt and useful – at least on a metaphorical level. Specific parallels will be addressed in more detail in section 8 but this section will observe a few general parallels.

Both ecosystems and information systems are complex networks involving many components. They exist in a dynamic changing environment and the interactions of the entities and the processes they create are much more significant than the isolated individual entities could be. A purposive ecosystem, as described in the example of a pond, is a closer parallel to the domain of repositories and services than a wild ecosystem would be.

An ecological approach to repository and service interactions allows a variety of types of information to be expressed. It can take a comprehensive view of repositories' contexts that addresses cultural, political, and financial influences as well as technical protocols. As an approach, ecology is aware that it is capturing a dynamic system, with continually evolving processes and with this awareness can try to indicate what and where change is occurring.

There is a risk that suggesting a way to use ecology as a metaphor may prove to be too complex and a further risk that, the process of defining the analogy may kill off the intuitive attraction of ecology. The authors are not ecologists; consequently in the development of this report we have developed our ecological view of repositories in dialogue with Chisholm and Schaider's handouts accompanying "Fundamentals of Ecology, Fall 2003", Lewis' "Pond Ecology" and through building on the previous work on information ecologies by Nardi and O'Day and by Davenport.[11]

6.2.1 Other work on information ecologies

The idea that ecological concepts might be a useful metaphor in the context of information systems or the interactions between people and ICT is not new.

The application of the idea of an information ecology to the world of corporate ICT use and systems has been examined by Davenport in *Information Ecology: Mastering the Information and Knowledge Environment*.[12] The interaction of people and ICT has also been explored by Nardi and O'Day in *Information Ecologies: using technology with heart*.[13]

Ecology as a metaphor or analogy for thinking about how people relate (especially through technology) crops up occasionally as people struggle to express the interactions of a group of people, the dynamics of that group, a location, and the 'tools' that enhance that interaction (be they chat software, coffee, or pool tables). A

---

[11] Chisholm and Schaider, (2003) "Fundamentals of Ecology, Fall 2003", Cambridge, MA: MIT, http://ocw.mit.edu/OcwWeb/Civil-and-Environmental-Engineering/1-018JFall2003/CourseHome/index.htm

[12] Davenport, T., & Prusak, L. (1997). *Information Ecology*. Oxford: Oxford University Press.

[13] Nardi, B., & O'Day, V. (2000). *Information Ecologies*. Cambridge: The MIT Press.

recent example of this intuitive expression is Dave Cormier's blog post "Building Ecologies - Making room for communities and networks" and various responses to it. [14] As part of that discussion George Siemens commented,

> I've stated previously that networks occur within something, that is, our capacity to form networks is influenced by a space or an ecology. Some ecologies are better suited for the formation of learning and knowledge networks than others. For example, certain corporate [settings] are structured to reward holding instead of sharing knowledge...a climate not conducive to the creation of new knowledge. Or, consider what we see occurring today with intellectual property - the climate of ownership of ideas at best only vaguely credible - creates an ecology in which the creation and dissemination of new knowledge is inhibited. Organizations - corporate, academic, or other - should be aware of the attributes of a successful ecology, namely one which allows the greatest probability for network formation. Dave Snowden stated in a presentation (can't remember where) that ecologies can't be engineered. While I don't know the exact scope and context of his statement, my first reaction is one of disagreement. Ecologies can at minimum be fostered (much like networks can be). Consider the economy as an ecology...or IP law...or innovation labs. In each instance, we can certainly influence how they develop and, to varying degrees, how they function. Anyway, the task is to explore what constitutes a desired ecology.[15]

There is also a significant body of work connected to the idea of an ecosystem funded by the European Commission under the 6th Research and Technological Development Programme. The Digital Business Ecosystems strand of projects focuses on exploring issues in the intersection of technology, economics, and social factors in order to support and develop the implementation and use of ICT infrastructure in connection with small and medium size enterprise in Europe.[16]

---

[14] Dave Cormier (2007) "Building Ecologies - Making room for communities and networks" Dave's Educational Blog http://davecormier.com/edblog/?p=107

[15] George Siemens (2007) elearnspace http://www.elearnspace.org/blog/archives/002930.html

[16] Digital Business Ecosystems. Edited by: F. Nachira, P. Dini, A.Nicolai, M.Le Louarn, L.Rivera Lèon. http://www.digital-ecosystems.org/book/de-book2007.html

7 The core concepts of repository ecology

7.1 Introduction

Given that an ecological approach may offer some useful insights into or perspectives on how repositories and services interact, the first step in using this approach is to sketch the basic features of ecosystems, some ecological methods and suggest some terminology that may be helpful in the conversation that follows.

Some of the basic ecological concepts that may be useful in themselves are defined by Nardi and O'Day and Davenport's work on information ecologies, others emerge from a further consideration of ecology proper. Most of the use of these ecological concepts by the repository community is likely to be unstructured and different users of the approach are likely to seize upon one or two concepts that help them to express a particular feature of their setting.

Beyond this way of using ecology however, it may be useful to consider some of the methods used in ecology. Two key approaches that will be briefly reviewed are the resource-tracking approach (analysing the movement of energy or nutrients) and the habitat approach (analysing the specific surrounding and community of a particular species or instance).

The following presentation of basic concepts (scale, entity, species, resource, environmental factor) is provided to support discussion and provide a common frame of reference for the use of an ecological approach. They are not intended to be prescriptive but descriptive of the shape of an ecological metaphor, and so to allow further consideration of the idea.

7.2 A dynamic (or evolving) system

One of the strengths of using an ecology-based approach is that it is understood intuitively that the system is inherently dynamic. This is one of the key differences between architecture as underlying metaphor and ecology as an alternative metaphor. When presented with a system architecture, we usually think in static terms, it is something that we build and is completed (or at least made public). Even if there is then substantial change to the system we consider it in terms of an add-on or extension, or rebuilding – there is a fixity to our model. Within an ecological model we understand that what we are presented with is a glimpse of a system that is undergoing constant change.

Implicit within this idea that a dynamic system is being described is the idea the components of the system change and adapt to their setting. Such adaptation may not occur at an even rate across the ecosystem in question and it may be appropriate to indicate particular communities or species that are rapidly changing. It should be noted that such change can be an intentional adaptation or a unexpected 'mutation' triggered by other species or environmental factors. This adaptation could be regarded as evolutionary – but the metaphor gets stretched slightly as the adaptation is generally both intentional and experimental

A species-based example of this sort of rapid change was the adaptation of institutional repositories from tools to support open access to also be tools to support RAE submissions – this also resulted in a substantial growth in the occurrence of repositories and the number of papers stored. A further example of a species that may be undergoing evolution is universities' course information systems – a proportion of the entities in this species are testing the deployment of the XCRI standard; if they are successful that species may change significantly.[17] This implies that when creating a ecological view of part of an ecosystem, the adaptability of the participants and the rate at which a given area is changing should be taken into consideration.

7.3 Scale

A basic principle of ecology is that the scale or the granularity of the ecosystem under discussion needs to be made explicit. Interactions or processes occur at particular granularities. Is the competition between two members of the same herd, two different herds, or two different species? Is the behaviour of a lecturer, a department or a university being described? Although processes at one level obviously affect other levels, a careful separation of levels is required to present them and understand how and why the work or fail.

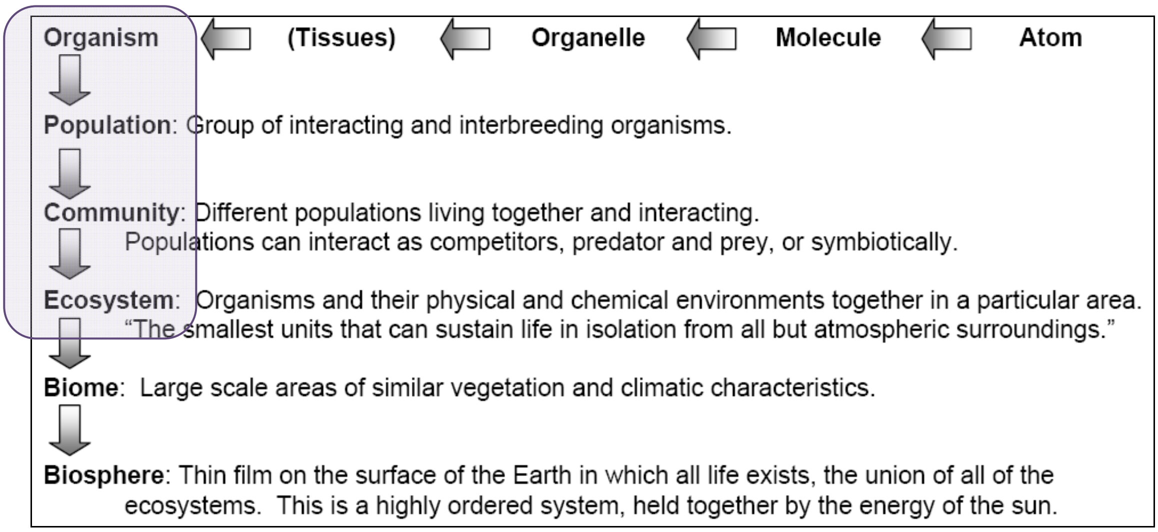


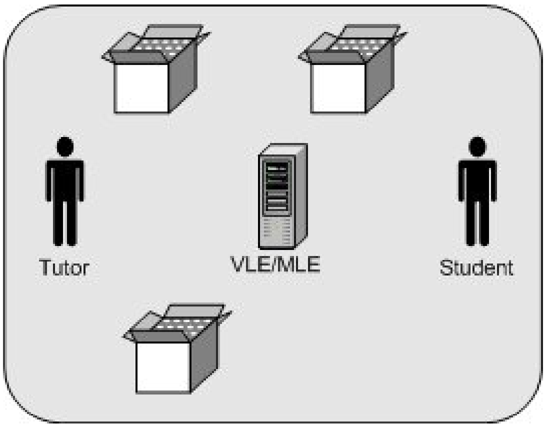
Figure 4 Scales at which ecology can be pursued. Figure from Fundamentals of ecology Fall 2003; lecture 1 http://ocw.mit.edu/OcwWeb/Civil-and-Environmental-Engineering/1-018JFall2003/LectureNotes/index.htm

The above diagram provides a hierarchy of scales (or levels) within an environment. Ecology constrains itself to thinking about the range from Organism to Ecosystem. There is a clear need to articulate what level of interaction being describe or model. Are you describing a microbe , a herd of giraffes, a river valley, or a city? Within the context of an ecology of repositories and services a comparable range 'Organism' to 'Ecosystem' may be the most useful parallel. Such a parallel may be mapped in different ways but the following makes some suggestions about how some of the different scales that an ecologically –influenced model might usefully distinguish and/ or comment on:[18]

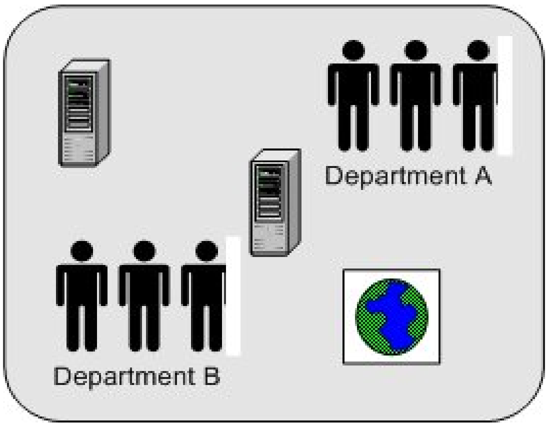
---

[17] Initial XCRI programme: http://www.jisc.ac.uk/whatwedo/programmes/elearning_framework/elfref_mmu.aspx ; Ongoing XCRI implementation and development work http://www.xcri.org/

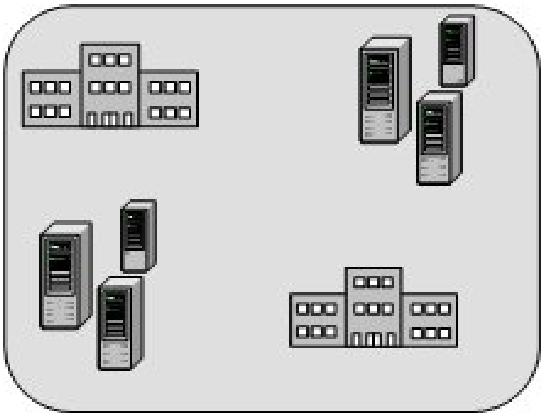
[18] It should be noted that in taking an ecological view of repositories and services they are regarded as 'living' entities within the ecosystem (more detail in the Entities section 8.4



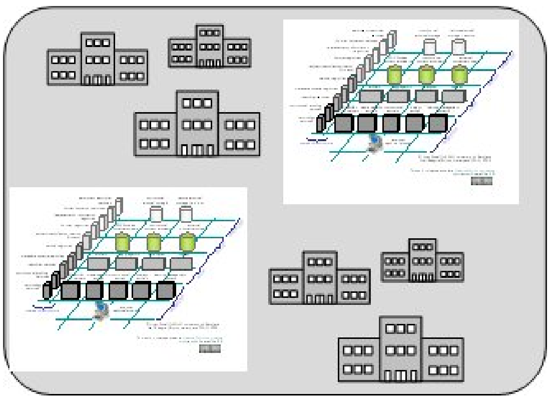
Individual users interacting with each other, a particular system, and assets within it. This is like a consideration of organisms.



Groups of users interact with different systems (local and external) which may also interact with each other. This is like a consideration of a population.



Insitutions deploy repositories and services which they use and which interact with each other. This is like a consideration of a community.



Large-scale communities create and operate information environments to support collaboration. This is like a consideration of an ecosystem.

Figure 5 Sample scales of repositories and people interaction.

Although most ecological approaches in the repository and service domain will focus on the community level of granularity (perhaps studying one entity, function, or habitat), it is important to remember that they are a particular localised view of the wider information environment (ecosystem) and will be influenced by events from other levels. Similarly a view of the interactions ongoing within the information environment will be influenced by the interactions occurring, for example, at institutional level. Articulating the influences of one level on another does however, need to be examined as part of the modeling process, as there can be no implicit hierarchy of influences (e.g. an agreement to enter into a national consortium does not necessarily change local repository practice and an institutional commitment to Open Access may not affect the perspective of particular academics). All of these may be influenced by processes and events completely outwith their ecosystem. It is suggested that these external influences that affect entire systems can be regarded as environmental factors (see section 8.7).

7.4 Entity and species

An entity is a tangible thing that exists within a ecosystem. In the development of an ecological view of repositories and services, repositories and services are regarded as 'living' participants in the ecosystem. This is perhaps the point at which the metaphorical usage of the concepts needs to be stressed the most. These instantiated systems with their software choices, customisations, user interests, and organisational commitments are being represented and summarised as living things. It is as reasonable to talk about the character of a repository as it is to talk about the character of a university or a city.[19] The most common types of entities are: users; repositories; services; objects; metadata records. In order to consider these things, in particular, repositories and services as 'living' participants it may be helpful to remember that they are social constructs (as well as bits of software).[20]

As suggested by the above consideration of scale entities are generally considered at the population level of granularity. For example repositories and services contain or act on a large or small number of digital objects (organisms) but also have a degree of corporate identity (i.e. it is meaningful to speak of the behaviour of a repository in the same way as one might consider the behaviour of a pack of wolves or herd of sheep). It should be noted that 'environmental factors' affecting entities or the connections between entities are not, in themselves, entities.[21] In order to consider repositories and services as 'living' participants it may be helpful to remember that they are social constructs.

Although this work could be expanded to include a consideration of the ecological approach at levels of granularity that focuses on objects and metadata records, this stage of the development of an ecological approach addresses provides a population-

---

[19] It should be noted that the choice of the word 'entity' is somewhat arbitrary. Its usage is not connected to databases. 'Participant' or 'thing' would be alternative choices. The reference to tangibility is somewhat redundant with the choice of 'entity' but it is to reinforce the following distinction.

[20] We're indebted to Wofram Hortsmann for suggesting the idea of a 'social construct' in his critique of this ecological approach as part of the workshop at ECDL2007

[21] An entity is therefore more narrowly defined than a W3C resource:, defined as "anything that has identity" 'Resource' definition mightinclude abstract concepts that might capture these factors (http://www.w3.org/TR/2002/WD-webarch-20021115/#glossary-resource).

based view of the articulation of a repository and services as we consider it to be the most directly useful and applicable to our immediate audience.

7.4.1 Repository

Within an ecological view a repository is an entity which supports a particular set of functions and meets certain user needs relating to the storage and management of digital assets for a defined group. Many repositories can also be considered as services or bundles of services (for example, an institutional repository may also be an OAI-PMH data provider). From a technical perspective a repository does not have to be a formal repository system but can be any other thing that supports this functionality. Examples of repositories include: arXiv; The Depot, Jorum, Daedalus, box.net, flickr; del.icio.us.

Further examples and a typology of repositories are available on the Repositories Research wiki at: http://www.ukoln.ac.uk/repositories/digirep/index/Typology

Although it is possible to define a repository purely as a data store with a set of associated services (whether formal web services, informal web-based services, or local services) when actual repositories and services are considered it is hard to discuss their success or failure without beginning to include something of their nature. We consider that presenting a particular repository as a living entity in an ecosystem is a useful thing to do and, furthermore, an approach not that dissimilar to existing definitions.

> In my view, 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 mangers, 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. [22]

7.4.2 Service

A service is an entity building on or otherwise interacting with a repository to offer added value to participants in the information environment in which it exists. A

---

[22] (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 ;

service may support technical interactions but it does not need to. Within an ecology-based approach a weekly conversation in a tearoom in which two colleagues consistently talk about good articles they've read is as much of a recommender service as a community forum, or an Amazon-style 'people who bought that..'. Examples of services include: The Information Environment Service Registry (IESR), the Digital Curation Centre, the Ethos service; the Repositories Research wiki; Pilot Engineering Repository Xsearch (PERX).

7.4.3 Species

In biology and ecology, a species is a collective term for a particular type of entity. Within a repository ecology similar groups can be discerned, such species might be: institutional repositories, aggregator services, library catalogues, blogs, students, teachers, system administrators.

Although the point of ecological descriptions or representations of repository ecosystems or communities is to focus on specific real entities, such descriptions may contain a combination of entities and species. For example, it may be useful to refer to an interaction with a particular student (a specific entity, an instance of the species), with a defined group of students (a specific group of entities, instances of the species), or with student as generic participants (a species).

The identification of the species, in this way, allows what is known about the behaviour of the species in general to be used to help understand the particular entity. As in biology or ecology, the observed characteristics of the entity are paramount but more general attributes of the species may illuminate the characteristics of the entity. For example, understanding the uptake in use of Comber University's repository (the entity) should be seen in the context of what is already known about the uptake of the use of institutional repositories in general (the species). It may digress from the standard pattern of the behaviour of repositories as a species, but should be considered in light of it.

7.5 Interactions

An interaction is a connection, relationship, or link between two or more entities or species in a population, community, or ecosystem. This interaction can have any nature; it may be a machine to machine technical interaction, an interaction between two people, or an interaction between a person/ people and systems. In a few circumstances it may be sufficient to note that there is some form of interaction or interactions between two entities or species that is important for the ecosystem but it will usually be useful to further elaborate on the nature of that interaction.

Defining interactions may simply involve indicating in some way what form of interaction is occurring. For example, interactions can include: harvests; talks to; emails; subscribes to rss feed from; edits objects; selects. Or it may also involve specifying any resources involved in the interactions. For example,

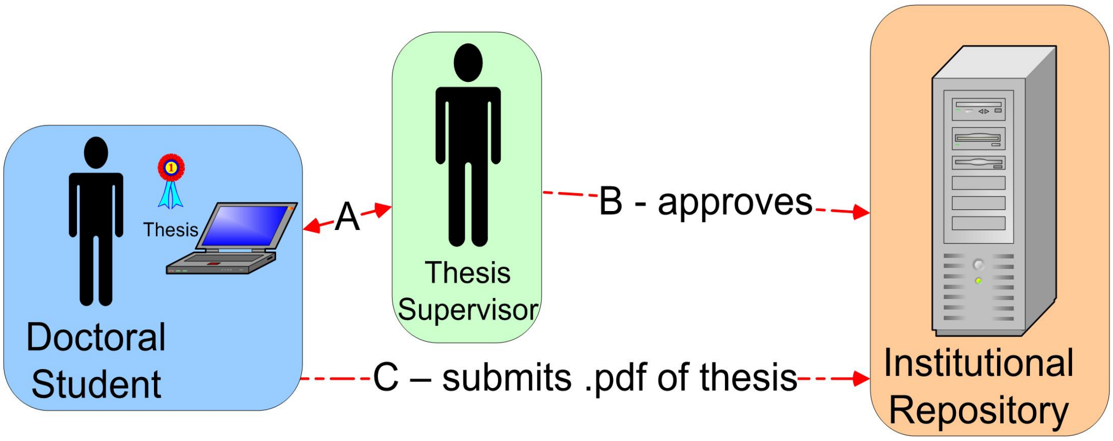


Figure 6 Ethesis interaction example

In the above figure the interactions (A) between the student and supervisor are unspecified (perhaps as they are too numerous to represent on this scale of representation or are secondary to the purposes of the diagram). The interaction B is clarified to show the interaction between the institutional repository and the supervisor (the supervisor approves the student's thesis deposit). The interaction C has a specified resource as well as an action involved. In the use of this concept we consider the approach of B or C, to be significantly more useful.

7.6 Resource

A resource is something that is passed from one entity to another as part of an interaction between them. This will normally be something that is essential to the wellbeing of the receiving entity. A resource may be modeled as an attribute of an interaction or as a nutrient - the choice will depend on the ecological approach chosen (implicitly or explicitly, this is explained in more detail in section 9.2). For example the DCC provides advice to the Edinburgh Research Archive about the sustainability of a mime type. This may be modelled either as:

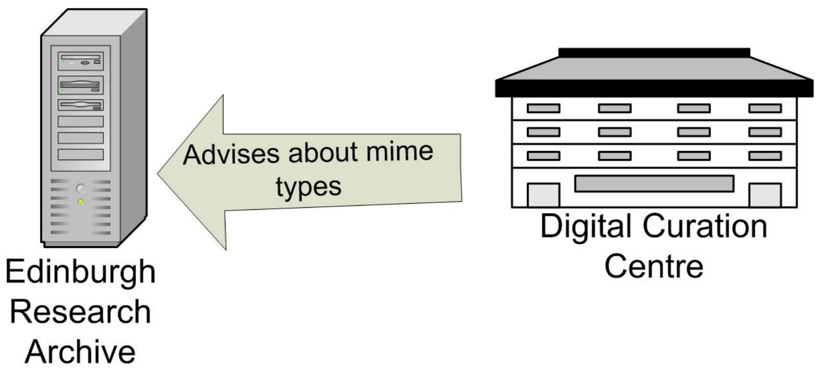


Figure 7 Resources as attributes of an interaction

or as

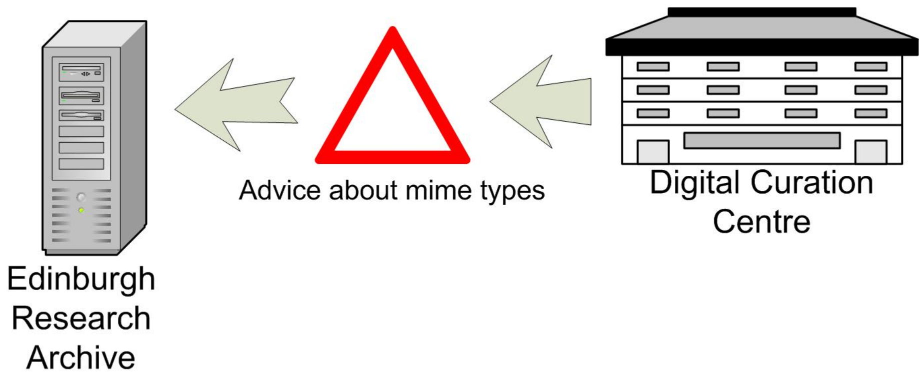


Figure 8 Resources as nutrients

In figure 6 the DCC advises the ERA, and the interaction 'advises' has the attribute 'mime types' which specifies the nature of the interaction. In figure 7 the same interaction is represented but the advice is represented as a nutrient being transferred rather than as an attribute of the interaction.

7.7 Environmental factors

An environmental factor is something that influences an entity, community, or ecosystem but is more general than an interaction between constituent entities. For example, a research council's mandate for Open Access deposit is likely to affect the number of papers available in that subject area in a given repository community and throughout the ecosystem(s) interacting with that subject area.

When considering what environmental factors are at work in an ecosystem, it is worth remembering that environmental factors may derive from both active participants in other environments and from the legal or cultural context of the wider environment.

In ecology effects on the community or ecosystem caused by active participants in other ecosystems (other entities or species not represented directly within the particular community or ecosystem being studied) could for example be the impact of migrating predators which are not directly part of the community being studied but have an ongoing effect on it. In the context of repository ecosystems such factors include: competition for resources (e.g. entities can't develop because available funding has been allocated outside of community) or suppression of competitors by other entities (e.g. political manoeuvring).

In ecology effects on the community or ecosystem not caused by other entities or species but rather by context and external forces are things like the effect of the weather or specific geological formations present in a given ecosystem. In a repository ecology such factors include: cultural drivers (e.g. the Research Assessment Exercise); the effect of ideals and concepts (the Open Access movement); the existence of funding (e.g. JISC funding streams), and legal constraints (e.g. copyright restrictions).

7.7.1 Using environmental factors

A number of issues should be noted about the use of environmental factors in describing a repository ecology. These are:

a) The influence of scales above the current one can be considered an environmental factor. For example, the influence of web trends or the influence a new rdf binding for Dublin Core. These could effect many parts of a system under consideration – both entities and interactions but are in themselves neither direct interactions nor new entities.

b) The effect of environmental factors should not be modeled as interactions. These are distinct things. An interaction is a direct connection between a number of participants in the ecosystem, an environmental factor is something that influences all or some of those entities and interactions – in visual terms an interaction is a line joining two entities; an environmental factor is a shaded region or a box listing things that affect the whole diagram.

c) It has been suggested that the idea of environmental factors in a description of repository and service interactions may have some commonality of intent with ideas present in Aspect-oriented programming (AOP).[23]

7.8 Other relevant features of ecologies

There are a number of other concepts from ecology that may prove to be of use as well. The following are a selection:

7.8.1 Keystone

Both species and entities may be qualified by the adjective 'keystone'. This designates a type of entity or particular entity that is essential to the growth or survival of a given community or ecosystem. For example institutional repositories are essential to the vision of an ecosystem set out by the JISC IE or The Depot[24] is an essential component of the maturing IE ecosystem (but – by design - not part of the mature one).[25] Such a designation will of course be somewhat subjective but the introduction of the concept supports the consideration of what entities or interactions within an ecosystem are essential to its success.

7.8.2 Biodiversity

It is inherent in ecological systems that more healthy ecosystems display greater biodiversity – i.e. there is a wider range of species within the system. Within the domain of repositories and services a metaphorical parallel may exist for this concept not only in the species or repository or service present but also with respect to the overlap between the species. The benefit of this can be seen in that there is not reliance on one educational institution, one software solution, one search service. Nardi and O'Day comment that "Monoculture –a fake, brittle ecology—gives sensational results for a short time, then completely fails".[26] One clear example of this within the information environment is if there was a dependency on a single software solution for repositories. Different repositories have different requirements and interfaces, and manage different types of objects. A single solution might allow rapid deployment but would then tie users into only being able to follow that software's

---

[23] AOP - This approach strives to address cross-cutting concerns in code design by treating them differently. for example -http://en.wikipedia.org/wiki/Aspect-oriented_programming

[24] The Depot (http://depot.edina.ac.uk/) provides an interim repository service for ePrints. It serves UK academics whose institutions do not yet have a repository and a redirection service for those whose do but are unaware of it. Once the information environment matures an interim repository should not be required.

[25] See also Nardi and O'Day, 53-4

[26] Ibid, 51. Note: as yet this contention that monocultures are unhealthy is not proven in the context of repositories and services.

development path.[27] It should be noted however, that there are, especially with the domain of ICT, instances where monocultures appear to be very successful and apparently quite sustainable. Consequently this particular metaphor should be used with great care.

7.8.3 Niche

Related to the above idea of evolution is the idea of a niche - a particular habitat, resource, or set of environmental factors that might exist within a given environment which would allow a particular species (or possibly a particular entity) to thrive. In the context of a repository ecology niches are likely to be created by a particular combination of institutional culture, external drivers, and available tools. In the example above - within some institutions the RAE created a niche for an institutional repository, within that set of institutions a primary focus on managing scholarly articles created a niche for particular software solutions that focus on that type of material. With a move to metrics-based research assessment in the future (irrespective of what that means), whichever software solutions are successful at exposing their content to the chosen metrics will thrive. Asking questions about what a particular set of conditions has made a community or population thrive is a key part of an ecological approach.

---

[27] See also Ibid, 51-2

8 Structured uses of ecological concepts

We have suggested that some basic concepts from ecology resonate with the situations and settings which participants an information environment are in, and that, the metaphor they offer is of use when they want to articulate the interactions occurring around their repositories or services. By treating technical systems, people, and organisations as entities in the same environment it is possible to selectively articulate the key interactions, dependencies, and influences that occur. As users, managers, implementers, or developers, this helps us to communicate and manage key dependencies and offers an alternative way to analyse and develop a service.

As much as these concepts are useful in provoking reflection, they may become considerably more useful when an approach that builds on their use through the application of particular ecological methods is employed. There are two particular methods used in ecology that, are easily accessible and may be of relevance; these are: resource tracking and habitat mapping.

8.1 Habitat mapping

Habitat or population mapping asks questions about the characteristics and nature of where things live. These questions can begin with a simple attempt to describe the setting a given entity finds themselves in, what other entities they interact with, what environmental factors they are exposed to, and what resources they have available. The detail of any attempt to fully describe either a 'real' ecology or a repository ecosystem can reach an unmanageable level of complexity quite quickly so it is crucial to consider the habitat and selectively present the key participants and factors.

This idea of thinking through and presenting the ecological context of a repository allows us to bring together insights from other types of analysis that we might carry out. Crucial issues or dependencies from other tools such as system design, business processes modeling, workflow analysis, management theory, and any other source can be brought together within the context of a habitat and complemented by any uniquely ecological insights. In the development of such a habitat-centred view, it is possible to capture the key entities, interactions, and environmental factors that influence a particular point – whether that point is a specific institutional repository, a lecturer, or metadata registry. We consider that the provision of such a metaphorical structure that allows the articulation of specific dependencies in their contexts is of benefit to the development community.

Within ecology, however, this approach's purpose is not only descriptive but also analytic and predictive. Figure 8 (following) is from a paper examining the efficacy of experimentation on food webs to model the environmental change on a species. The figure illustrates how the food web of a particular plant exists in a particular environmental setting and then locates settings for that population or community at different geographic scales.[28]

[28] Food-Web Models Predict Species Abundances in Response to Habitat Change Gotelli NJ, Ellison AM PLoS Biology Vol. 4, No. 10, e324 doi:10.1371/journal.pbio.0040324;

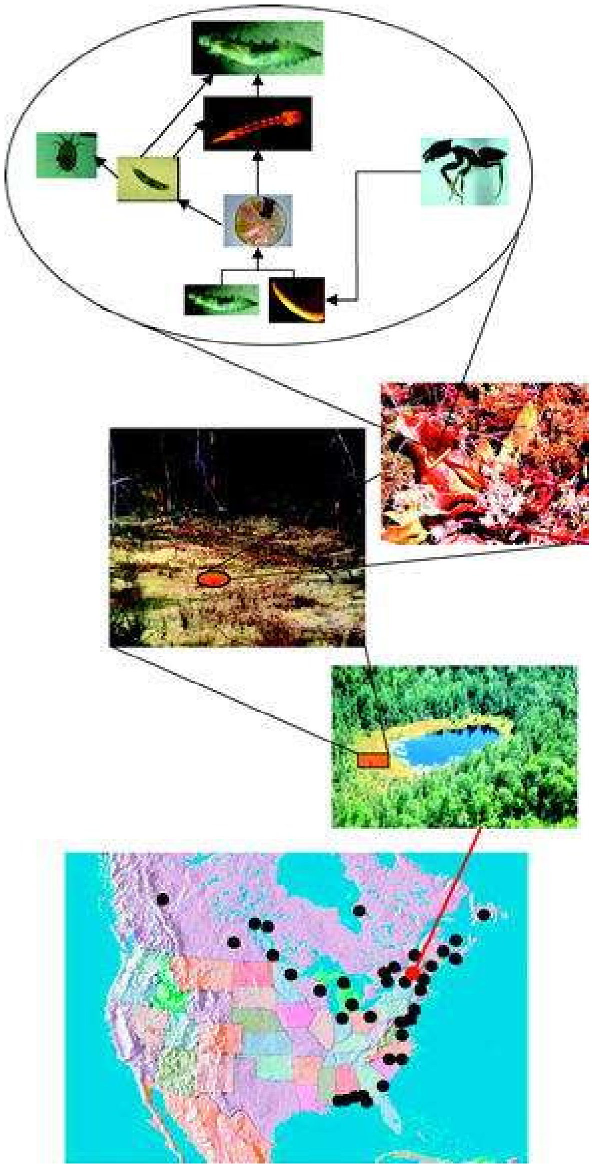


Figure 9 Gotelli NJ, Ellison AM, Food-Web Models Predict Species Abundances in Response to Habitat Change PLoS Biology Vol. 4, No. 10, e324 doi:10.1371/journal.pbio.0040324

Although we have been careful to stress the importance of local specifics when considering repository and service interactions, this example in ecology also points to how local specifics are often closely paralleled elsewhere. The bottom picture in the layer points to the occurrence of the preceding particular species food web within its habitats in North America.

We would suggest that habitat models within the repository ecology metaphor can, in a parallel manner, allow both a speculative analysis – what happens if this entity or environmental factor changes – and a comparative analysis – if we know that this system works or doesn't work in this concrete situation- are there similar settings that have these features? There is of course a tension in trying to find comparable settings, as we consider the ability to deal with particularity one of the great strengths of an ecologically-influenced approach. A useful guide to navigate this tension is to distinguish abstraction, which reduces the complexity of specific habitat by linking it to universal types or concepts, and which this approach wants to avoid, from selectivity, which reduces complexity though choosing to present only the most relevant specific features, and which already part of this approach to address issues of complexity.

The idea of a written of visual representation of a system's habitat being able to support speculative analysis is closely tied into the notion of using the ecological approach for the purpose of communication. There is something about the presented view of an ecosystem that lends itself to brainstorming and thinking about how the system might develop.

Although one of the possible developments of examining ecosystems is the potential to generate some form of mathematical models or simulations, this is not the intent of this body of work. Such mathematical models and simulations could enable developers or funders to model what might happen to the ecosystem in response to changes in the properties of organisms, species, populations and communities and their interactions. This area of modelling and simulating changes is a key area of ecology proper and it may have some application in the repository environment. It is however, not clear if this would offer anything beyond the more formal types of system modelling tools that already exist and are in use in other communities.

8.2 Resource tracking

In ecology the idea of resource tracking examines energy flow or nutrient flow within the community or ecosystem under consideration. The most well known example is probably that of the water cycle. Resource tracking examines how a particular resource moves around or is consumed.

The health or the growth of the ecosystem may be reflected by the supply of nutrients to entities within it. Figure 9 (below) illustrates nitrogen cycling - one form of resource tracking. In comparison with the idea of habitat-mapping , the resource tracking approach is much more focused on a particular resource or set of process that effect a given resource. It is seeking not so much to develop a broad view of a context but rather a broad view of view of a resource or resource process.

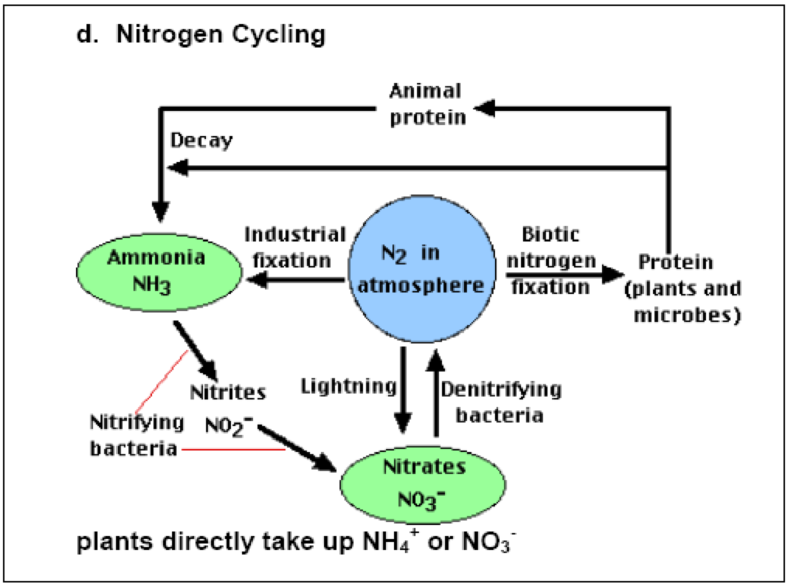


Figure 10 Fundamentals of ecology Fall 2003; lecture 9 http://ocw.mit.edu/OcwWeb/Civil-and-Environmental-Engineering/1-018JFall2003/LectureNotes/index.htm

In the context of repositories and services this approach may be of relevance in the consideration of the distribution of expertise, information, or money. Although there are problems with trying to apply this metaphor to these resources, it is of some use and is directly relevant when trying to consider the movement of digital objects as the one resource (intellectual work) may turn into many different things as it progresses around the system.

8.2.1 The use Food webs and food chains

Food webs and chains are one specific type of resource. The idea that particular species 'consumes' another species has a clear comparison with the idea of a data providers and service providers. Further, the numbers of particular consumer species may be dependent on the numbers of species they consume or that consume them. There may be a useful analogy with a trophic pyramid.

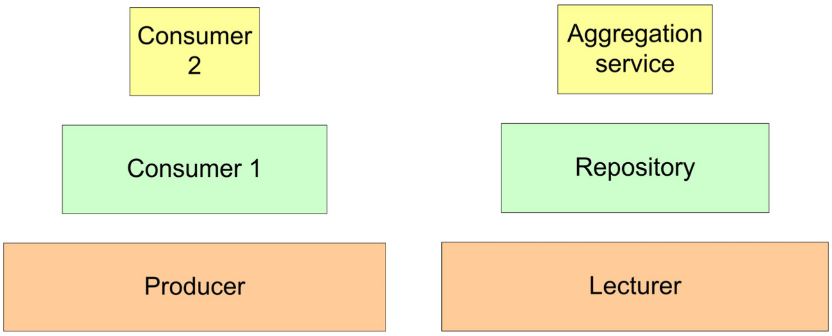


Figure 11 Trophic pyramids

There is however a difficulty with the metaphor here in that very few 'resources' within a repository ecology are actually used up by their 'consumption'. Limits on the growth of one species in an information environment are more likely to come from other members of that species (service A gets funding and B doesn't) or from unrelated environmental factors (available funding across the environment is decreased). The concept of a food web or chain can usefully illustrate dependencies between entities but the limits of the analogy should be remembered.

9 Illustrative examples of an ecological approach

9.1 Didet - the community of a learning environment

The DIDET project is a JISC and NSF funded project led by the University of Strathclyde, Stanford University and Olin College. It has developed strategies for collaborative classrooms in design-engineering supported by a software tool based on wiki technology. The classes offered by the participating departments allow "students to participate in global team-based design engineering projects that give them experience of working within multi-cultural contexts and enable them to develop global design team working skills."[29]

The project approach has been successful and is now integrated into the regular design engineering curriculum of the partners. During the development, however, issues about information literacy, copyright, and the relationship between the formal and informal digital spaces (the workspace and the library) proved unexpectedly complex. The following ecological view of DIDET attempts to capture some of this complexity and to note some questions that emerge from a consideration of this ecology.[30]

As an ecological view, a consideration of a supported design engineering course using the LauLima Learning Environment (LLE) and LauLima Digital Library (LDL) developed by DIDET is probably usefully considered at two different scales: initially as a small community and subsequently as a participant in a wider ecosystem. Its initial analysis at community-level presents a view of how DIDET works and the secondary analysis illustrates how this fits into the wider information environment.

The species within the community are students, lecturers, design coaches, and support staff, repositories with integrated services, and external collaborative services supporting collaboration (such as a video conferencing tool). The basic entities are specific students (A, B, C, D); lecturers (J, K), a coach (F), LauLima support (L), the LLE, the LDL, and other tools for collaboration (Polycom VC, Flashmeeting, etc.). Within this community all the species can be considered as keystone species – early plans to provide the functionality of the LLE and LDL within one system or to use the system without a human review of assets during the migration process between the LLE and LDL had to be revised in light of potential problems with Intellectual Property Rights (IPR) and with the need to enhance user-created metadata.

Some environmental factors present in the DIDET community are: the role of assessment in the collaborative design project (students who are assessed their design process, as presented in the LLE, as well as on their final design will interact with it differently); the support of senior university staff (Lecturers whose online contact time is counted as part of their overall contact time commitments may engage more extensively). Some other environmental factors are: the effect of finite project funding (sustainability issues have influenced the design of the community); the cultural

[29] http://www.didet.ac.uk/

[30] It should be noted that this section is the authors' perspective on DIDET and does not represent DIDET's view.

differences between the groups of students; the effect of working in different time zones on collaboration, and the effect of IPR legislation on collaboration.

The DIDET community exhibits a degree of biodiversity in the selection and use of tools for collaboration. Students are free to supplement the LLE as they wish, but resources created externally can then be referenced by or submitted to the LLE or LDL for more permanent use.

If the community developed by DIDET is considered within its wider setting, a number of questions emerge from an ecological model. The below figure considers the effect of IPR factors on the environment of the LLE and LDL and their users.

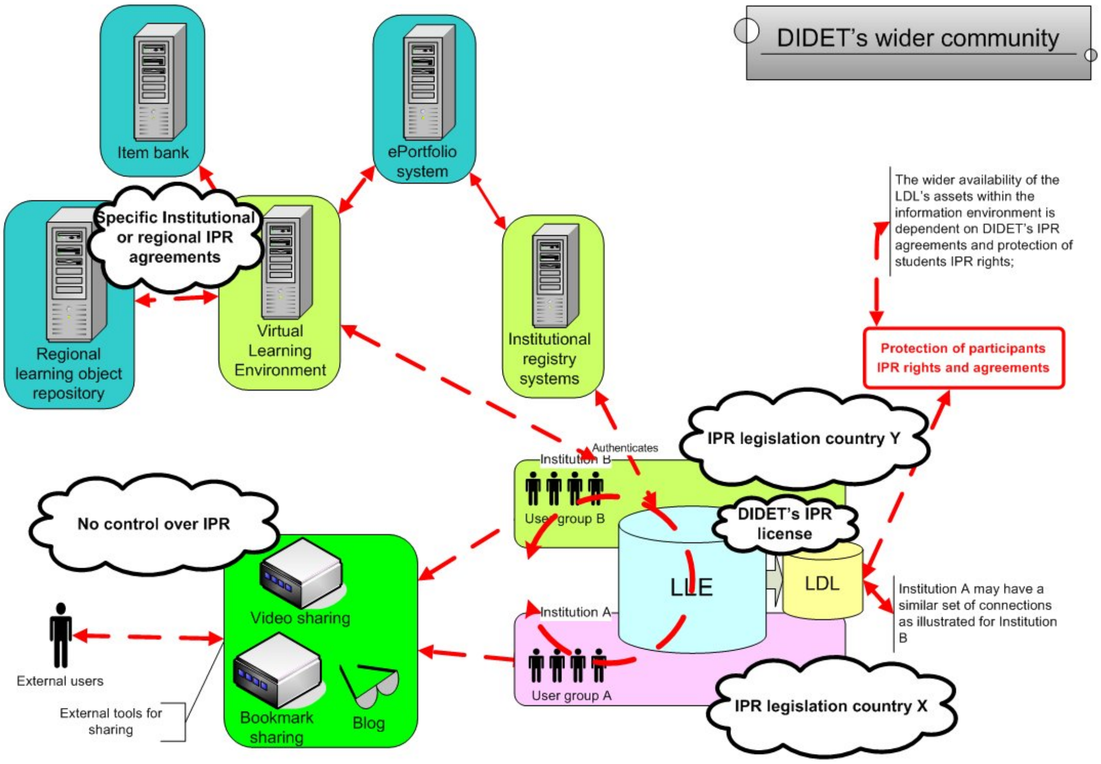


Figure 12 DIDET's wider community

DIDET has very successfully identified a niche and established a mechanism for multi-institutional collaboration within a discipline. Within a least one institution it has been able to integrate the use of local authentication tools – this adaptation to existing components of the one of the host communities makes participation in the LLE more straightforward.

IPR is a key environmental factor when the exposure of design assets generated in the course is considered. Both when students share them in an uncontrolled manner and when the LDL considers making some of these assets available to either future students within the LLE or to a wider student body. A further consideration of the influences that IPR has on the community might be illustrated by tracking resources (such as an assessment item or a textual excerpt) as they move through influence of the different IPR environments.

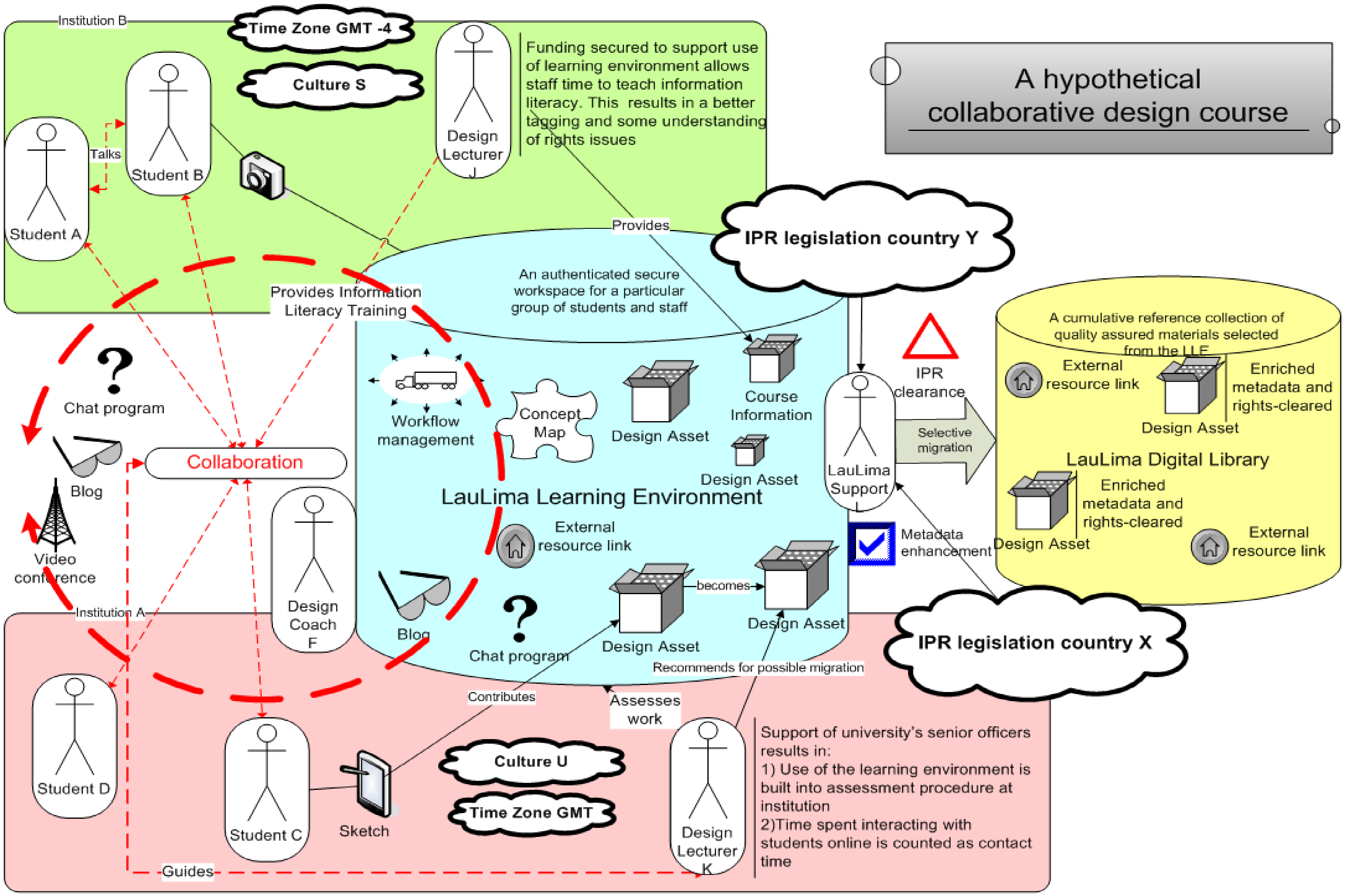


Figure 13 A DIDET community

An ecological view of DIDET presents the complexity inherent in a collaborative classroom. At a community-level the technological entities, pedagogical processes, and legal factors involved in the classroom are articulated together and when DIDET is considered as part of a wider ecosystem some of the challenges of managing learning materials in the context of a distributed classroom are highlighted.

9.2 An e-Crystallography dataset ecosystem

eBANK UK[31] is a JISC-funded collaborative project between UKOLN, the University of Southampton, Intute and the University of Manchester. It brings together an interdisciplinary team of chemists, digital librarians and computer scientists to explore the potential for integrating crystallography research datasets into digital libraries. It fits well into the scholarly knowledge life cycle by linking data, teaching, and research with publication. [32]

eBANK UK has successfully demonstrated the potential interactions of data-storing institutional repositories with data centres and commercial providers of data, especially in areas relating to the discovery and reuse of data. There are, however, very few examples of institutional repositories which are actually storing subject specific datasets in systematic ways and exposing their contents for discovery and reuse. The project is currently in its third phase:

Phase 3 of the project (2006 – 2007), is the preliminary scoping of a global network of data repositories - an eCrystals federation. This report is particularly interested in investigating what light an ecological perspective can shed on this potential 'dataset ecosystem'.

9.2.1 Identifying a scale

From an ecological perspective it is worth considering eBANK UK (Phase 3) on at least three different scales: population, community, and ecosystem.[33]

9.2.1.1 Population: the institutional dataset repository scale

An institution (or a department within it) makes a commitment to create a repository of datasets through the identification of resources (e.g. funding, human resources, equipment). The repository can then be populated through a number of means, e.g. as part of the workflow of an experiment, where the laboratory equipment is directly inputting data into the repository and collecting datasets from previous experiments.

9.2.1.2 Community: the institutional federation scale.

Institutional repositories containing datasets about crystallography federate their contents (through OAIPMH) and the institutions themselves work together to share resources and collaborate on issues such as training, curation, preservation, and policy development. Each participating institution requires sufficient resources and structures in place to support the population of their repositories with the relevant datasets and associated metadata *within* institutions. For example, each piece of

---

[31] http://www.ukoln.ac.uk/projects/ebank-uk/

[32] Lyon, L. eBANK UK: Building the links between research data, scholarly communication and learning. Ariadne July 2003, http://www.ariadne.ac.uk/issue36/lyon/

[33] There are, of course, other scales at which dataset repositories could be considered.

laboratory equipment deposits data and metadata directly into a repository; this data (or at least the metadata) is then shared *between* institutions.

9.2.1.3 Ecosystem: eBank UK – the national scale

The ecosystem of dataset repositories developed in eBANK UK captures the interactions of an institutional crystallography dataset repository federation with publishers and other services such as aggregators and other dataset repositories outside the federation. For example, eBANK UK has successfully demonstrated how the underlying data from an experiment can be made available in or through a publication. Such interactions also indicate that the connections that might exist between a data ecosystem and other ecosystems; for example, those in the print publications world. The three scales can be represented as follows,

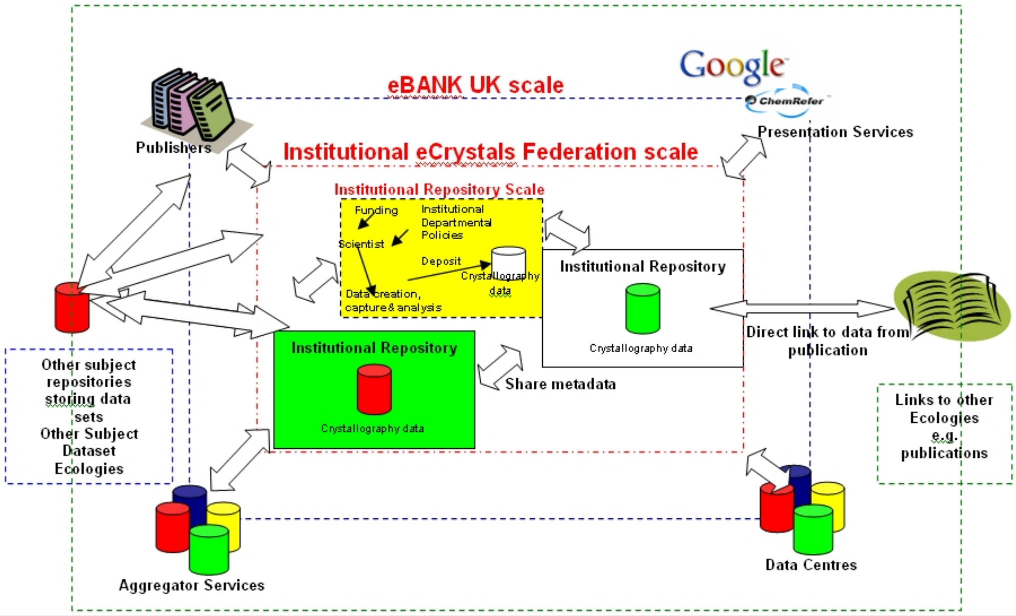


Figure 14 eBANK UK at different ecological scales

9.2.2 Dataset Species

At the Institutional Data Repository level, species are:

- Users
- Funders
- Researchers
- Scientists
- Metadata
- Relevant Institutional Stakeholders
- Institutional Repositories
- Datasets (Crystallography)

At the Institutional Data Repository Federation level, species are;

- Users
- Funders
- Researchers
- Scientists
- Metadata
- Policy makers
- Institutional Repositories
- Datasets (Crystallography)
- Aggregator services

At the eBankUK Model level

- Users
- Funders
- Researchers
- Scientists
- Datacentrers
- Publishers
- IR Federations
- Datasets
- Aggregator services
- Presentation / Portal services

9.2.3 eBank UK interactions

For each scale within the eBank UK ecosystem, real and potential interactions have been detailed below. It is useful to identify potential interactions to assess how they might help the population, community, or ecosystem flourish. For example, in eBANK UK the symbiotic relationship between institutions holding crystallography datasets and data centres is vital for the success of the ecosystem and needs to recognised and nurtured. The species and interactions of the dataset ecosystem are discussed in detail in the recent report 'Dealing with Data'.[34].

9.2.3.1 Interactions at the institutional dataset repository level

Species,	**Direction of interaction**	**Species,**	**Type of interaction**
Lab Machine	⇨	Lab Repository	Deposit
Lab Repository	⇨	Institutional Data Repository	Deposit and Validation Data Analysis
Chemistry Blog	⇔	Institutional Data Repository	Deposit, Share. collaborate
Scientist and Researcher	⇔	Lab Repository, Institutional Data Repository	Deposit, Discover, Re-use, administer
Funder	⇔	Scientist	Capture data, make data available
Dataset	⇔	Publication	Link

9.2.3.2 Interactions at the federation level

Species,	**Direction of interaction**	**Species,**	**Type of interaction**
Institutional Data Repository (A)	⇔	Institutional Data Repository (B)	Curate, Policy Development, Preserve, Develop Standards, Share Advocacy, Share Training
Institutional Data Repository (A)	⇔	Institutional Repository Dataset Federation	Validate, search, Harvest, Expose Records and Metadata
Users	⇨	Institutional Repository Dataset Federation	Discovery, Reuse, Linking, Citation
Presentation services (Google Scholar, CiteSeer, ChemRefer)	⇨	Institutional Repository Dataset Federation	Discovery, Reuse, Linking, Citation
Funder	⇔	Institutional Repository Dataset Federation	Capture data, make data available, make federation possible
Aggregator Service	⇔	Presentation Service	Data discovery, linking and citation
Publishers	⇔	Aggregator Services	Search and harvest
Subject Repository	⇔	Aggregator Services	Search and Harvest
Institutional Repository	⇔	Aggregator Services	Validation, search, harvest
Institutional Repository Dataset Federation	⇔	Library and Information Services	Harvest, expose, discover, citation
Institutional Repository	⇔	Publishers	Citation, Publish, discover

[34] Liz Lyon (2007) Dealing with Data
http://www.jisc.ac.uk/media/documents/programmes/digital_repositories/dealing_with_data_report-final.pdf

9.2.3.3 Interactions at the eBANK UK level

Species,	Direction of interaction	Species,	Type of interaction
Institutional Repository Dataset Federation	⇔	Library and Information Services	Harvest, expose, discover, citation
Institutional Repository	⇔	Publishers	Citation, Publish, discover
Institutional Repository Dataset Federation	⇔	Data Centre and Aggregator Services	Harvest, expose records and data
Institutional Repository Dataset Federation	⇔	Digital Curation Centre	Advisory

9.2.4 An eBank UK food web

The interactions with an eBank system could be viewed as simple food web.

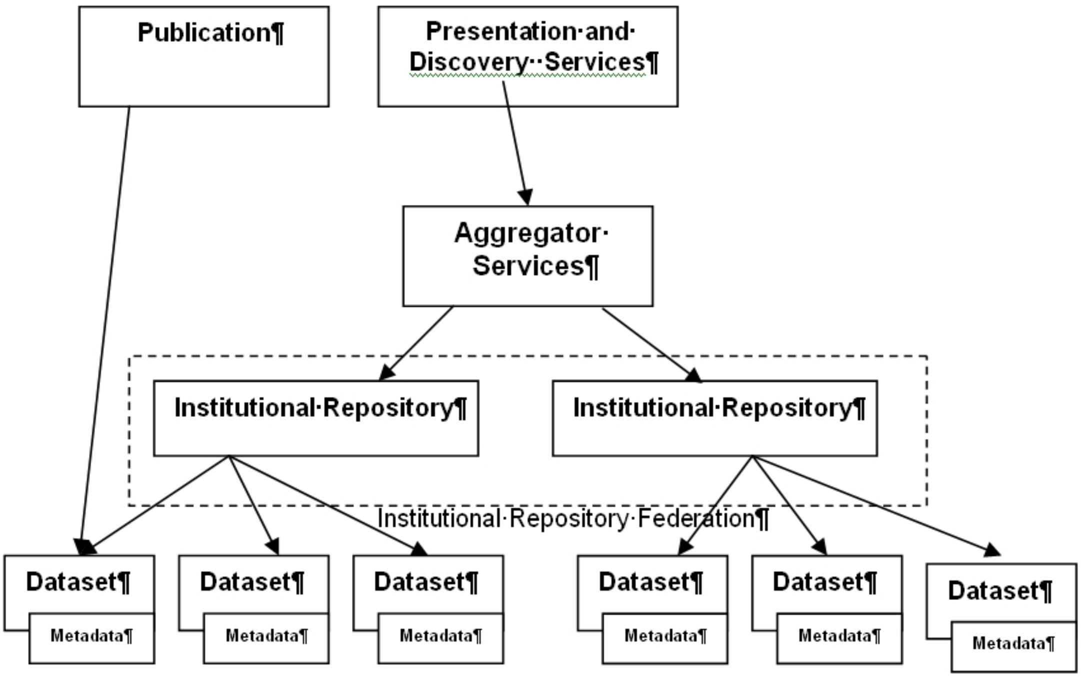


Figure 15 eBank food web

The food web shows how different species are ‘consuming’ others who are acting as nutrients. Although the figure is quite like an architectural representation of harvesting it should be noted that the interactions may also reflect processes other than technical ones (i.e. harvesting). There are also possible ‘chemical’ interactions between nutrients, e.g. data and metadata which could be analysed.

9.2.5 Environmental factors in the eBank ecosystem

Environmental factors which benefit eBANK UK include:

- ‘Will and motivation’ i.e. agreement amongst community that it should happen (i.e. right conditions)
- Bringing together of multidisciplinary team

- Homogeneity of datasets within the field of crystallography compared to the heterogeneity of datasets in other disciplines (e.g. physiology) which allows the easier implementation of dataset repository
- Funding
- the Research Assessment Exercise (stating that all publications submitted for the exercise make their underlying data freely available)
- the Open Access movement (arguing that all publications should make their underlying data freely available).
- the Guardian newspaper campaign 'Free our data' (arguing that publicly-funded data should be freely available- with particular (but not exclusive) focus on Geospatial data [35])

Factors potentially hindering the eBank UK ecosystem include:
- Competition for resources from eprint repositories (Resources such as :finance, storage space, infrastructure or expertise).
- Potential institutional immaturity in storing and exposing datasets locally.
- a lack of clarity about dataset ownership within institutions
- copyright restrictions,
- other legal or ethical constraints,
- embargo periods (researchers or others wanting to restrict access to their data for an initial period)

9.2.6 Biodiversity

It could be suggested that eBANK UK represents a monoculture. and, though very successful, might be fragile in the longer term. Consequently an important question, highlighted by an ecological approach is whether the conditions created to make eBank UK successful (e.g. funding, conditions, level and quality of interactions) can be sustained and replicated in other environments with the establishment of other subject specific institutional dataset repositories and federations. Does eBank UK resemble a monoculture because something within the design makes it so, or is it because it is the first such ecosystem to be developed (and is therefore a prototype for future institutional dataset repositories).

9.2.7 eBank UK: resource tracking

There are many resources that are passed between entities at different levels in the eBank UK ecosystem. They include the funding required to set up an institutional dataset repository, user training materials, the datasets themselves, their associated metadata.. If such resources are viewed as nutrients their movement around the ecosystem can be tracked. Such an approach facilitates an understanding of what has been or is needed to help the system thrive, through asking questions like "Has this repository entity benefited from resources provided by others? If so where did the resources come from? Are they available to a other entities in the ecosystem?"

It would be interesting to examine how datasets move and are transformed through an ecosystem, and what effect this may have, for example on their increased use perhaps, again this is something that could be modelled and simulated mathematically e.g.
- does the adding of quality metadata increase discovery and reuse

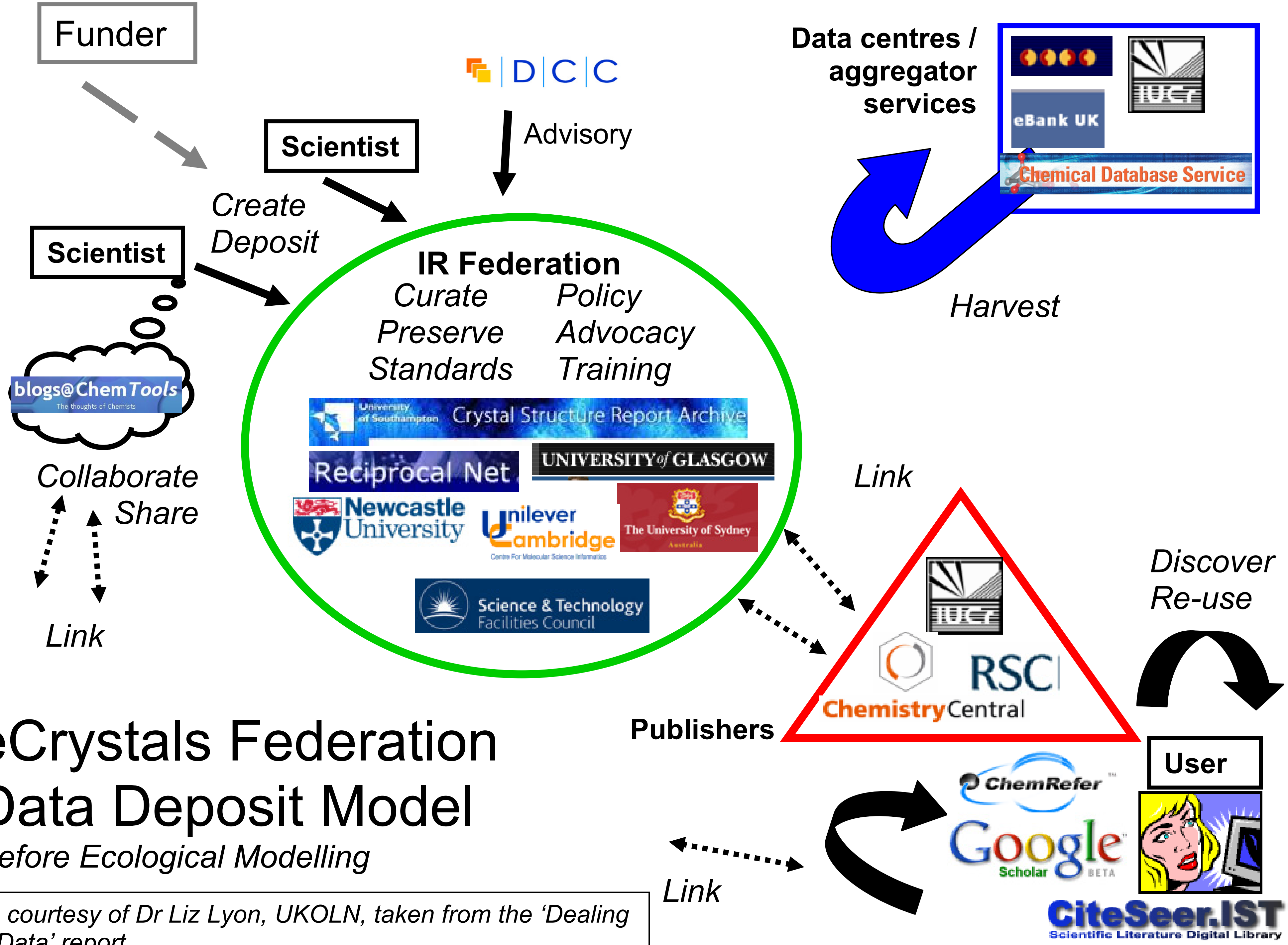
[35] http://www.freeourdata.org.uk/

- is analysed data more used in an ecosystem
- is the way datasets are disseminated important to the way it is discovered

Ecology can perhaps provide a fresh way of looking at existing systems and the interdependencies between entities.

9.2.8 Dataset Ecology Diagrams

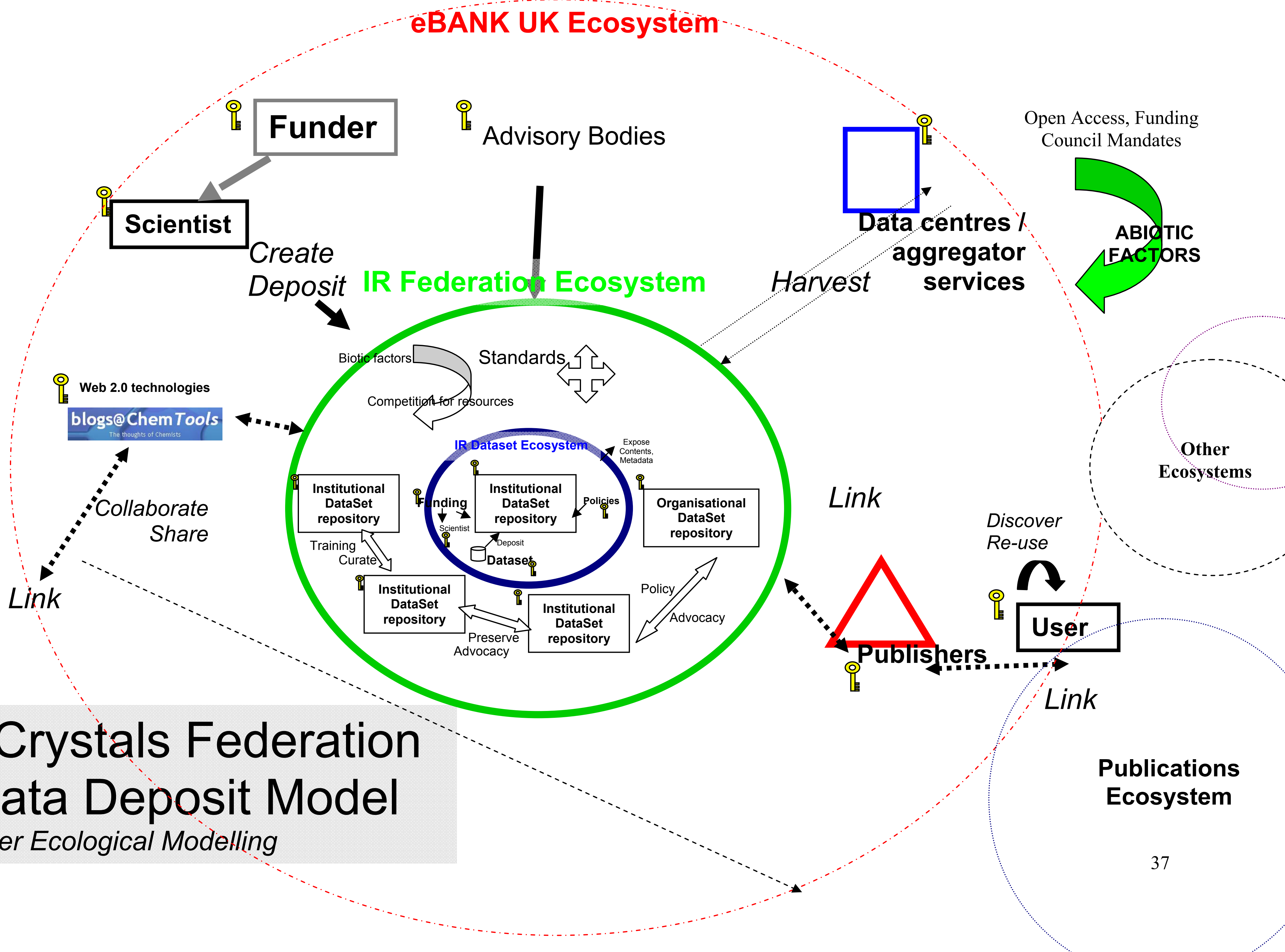
The final section of the eBANK UK model presents diagrams before and after an ecological approach has been applied.



eCrystals Federation Data Deposit Model

Before Ecological Modelling

Slide courtesy of Dr Liz Lyon, UKOLN, taken from the 'Dealing with Data' report



eCrystals Federation Data Deposit Model

After Ecological Modelling

10 An ecologically-influenced approach and existing work

The wealth of existing work within the repository space has created a degree of divergent approaches and terminological confusion in communicating how information and repository systems interact with each other, with other systems, with people, and of how the flow of information resources is managed. Knowledge and use of existing reference models, architectures and frameworks is often either geared towards members of specific communities, e.g. digital preservation, space science, e-learning or e-assessment or focused purely on technical architecture, without reference to the dynamics of a living system.

The ecological metaphor introduces a new set of terminology to this already crowded space. This section of the report aims to explore, briefly, how this ecological approach aligns with existing work and how it might be used in conjunction with them to support and enlighten, rather than duplicate. This exploration specifically engages with repository-related work ongoing in higher and further education in the United Kingdom.

10.1 The OAIS reference model

The Reference Model for an Open Archival System (OAIS) is aimed at information systems wishing to undertake long-term preservation in the UK it use in connection with repositories has been supported by many organisations, such as The Digital Curation Centre, the Digital Preservation Coalition, JISC, The British Library and The National Libraries of Scotland and Wales.

OAIS defines a reference model as "A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist" [36]. The OAIS reference model is very much an abstract framework and communication tool, working at a high level to specify functions and concepts, within the context of long-term preservation. Recent work [37] has shown that its applicability goes wider than that as it identifies an environment containing consumers, producers, repository functions and information flow common to many repositories, whilst bringing with it a heightened awareness of preservation issues. OAIS might prove useful for repositories wishing to communicate or audit their internal functions in relation to a detailed abstract model. Its focus, though, is on the 'archive' or 'repository' as a single system and it lacks support for defining and understanding the ways in which repositories interface with each other and with other systems.

[36] Consultative Committee for Space Data Systems (2002). Reference Model for an Open Archival Information System (OAIS). Blue Book, January 2002. Retrieved 2005-04-20, from: http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf

[37] For further information refer to Allinson, Julie, *OAIS as a reference model for repositories : an evaluation*, UKOLN, November 2006 http://www.ukoln.ac.uk/repositories/publications/oais-evaluation-200607/

There are some clear links between some of the intent of OAIS and some of the intent of the ecological approach – both are concerned with examining and expressing the specific practice of a repository across both technical and cultural factors. On the question of preservation both models are as concerned with the robustness of the institution as they are with the robustness of the software. It is also easy to see that issues like a mime type ceasing to be supported might be considered as an environmental factor. As such an ecological approach could inform a presentation of local practice which could then be assessed against OAIS's abstract reference model.

More interesting though is the possibility that an ecological approach could provide a way for repositories that have engaged with the OAIS model to consider their wider interactions and how they fit into regional, national and global ecosystems. The use of ecological terminology might easily be mapped to many of the concepts within OAIS. Looking ecologically at OAIS, we can already see a clear link between our notion of 'community' and the designated community concept in OAIS. The OAIS functional model already specifies a range of interactions within, and potentially between, systems and these could be presented and analyzed within a repository ecology.

10.2 An ecological approach and the e-Framework

Movement towards service-oriented approaches (soa) has impacted on the JISC strategy and IE and the increased awareness of the need to modularise service delivery and to integrate disparate systems and specifications has led to the development of the e-Framework for Education and Research. A joint initiative of the UK's Joint Information Systems Committee (JISC) and Australia's Department of Education, Science and Training (DEST), the primary goal of the e-Framework is "to facilitate technical interoperability within and across education and research through improved strategic planning and implementation processes" [38]. In this context, 'service' refers to technical services that are used to build true interoperability between systems, rather than the looser concept of a service used by the IE. Service Usage Models, a core element of the e-Framework and can be succinctly described as a model of how services (primarily software services) meet business needs.

The interaction between eFramework and ecological approaches is best considered in terms of how an ecological approach could inform a soa-based system. An ecological approach as has been outlined could form the articulation of a local context used by system designers in the selection of services and subsequent review of their suitability.

10.3 Domain models, and [eF:UL]

Recent work by the JISC building on the e-Framework has embraced the notion of domain models or maps. Domain maps (as discussed earlier) extend high-level informative abstractions, through process-oriented and human elements to low-level normative detail. Bill Olivier defines a domain as "a coherent area of practice" [39] and outlines various identifies the following elements of a domain model:

---

[38] The e-Framework for education and research http://www.e-framework.org/

[39] Olivier, Bill, 'Domain, Process and Service Usage Models', *JISC e-Learning Focus*, February 2007 http://www.elearning.ac.uk/resources/Domain%20Process%20and%20Service%20Models.doc/file_view

0 Boundaries of the domain
1 Related domains (context)
2 Domain stakeholders
3 Domain roles
4 Domain aims/purpose
5 Domain top level functions
6 Domain scenarios (human)
7 Domain practices and processes models (bridge)
8 Use cases (interaction with ICT)
9 Domain information models
10 Domain ICT System models [40]

JISC funded several projects to explore domain models (previously referred to as reference models). These have sought to develop an agreed abstract representation of particular domains from observed practice. This work will be continued in the form of the [eFramework Upper Layer] which seeks to capture domain information and link it relevant outputs of JISC projects (as outlined before the international eFramework records the services and SUMs, this work offers the possibility of capturing non-technical outputs in a similar structure).

An ecological approach provides a complementary mechanism for considering and framing many of these elements. Community or ecosystem diagrams can act as specific representations of part of a domain whilst concepts such as 'species' and 'entities' are useful when classifying and identifying both the human stakeholders, user agents and services within different systems. Relationships, workflows and service interoperability can also be usefully explored as interactions through the repository ecology. Although they both have other purposes and uses, on a given level an ecologically-influenced model of a system is an instantiated version of part of a domain map and a domain map in an abstraction of information from many modeled ecosystems.

10.4 A repository ecosystem and the JISC Information Environment

This report has already discussed why the architectural models set out as part of the development of information environments can benefit from supplementary models or approaches. This section reviews how an ecological approach can interact with the JISC Information Environment (IE).

An Information Environment is a mechanism for identifying and bringing together individuals, organisations and systems that ingest, store and disseminate information, along with the information resources themselves. For JISC, its own IE is identified as a core theme, aiming to "allow discovery, access and use of resources for research and learning [in the context of UK Higher and Further Education] irrespective of their location." [41]

---

[40] ibid

[41] JISC Information Environment http://www.jisc.ac.uk/ie

The JISC Strategy for 2007-2009[42] outlines three principles for technology development: community engagement, open standards and modularity. The IE supports these three principles by providing an abstract architectural view of the repositories and services in wider information landscape in which we are working, by helping to frame discussions, by supporting service development, by providing and managing funding, by providing guidance on the use of technical standards, and by engaging the community. Demonstrating a commitment to enhancing the user's experience of networked information in an educational context, the IE is ubiquitously illustrated by the diagram produced by Powell and Beagrie (figure 13). This diagram is however, only representative of a high-level view of the software architecture and it belies the complexity both of the developed instantiation of the IE, and the cultural, human, and conceptual factors that are intrinsic to it.

From this diagram it is clear that the IE is awash with 'species' (in particular types of software services), species that are gradually being realised as 'entities' such as the IESR, and the Depot. It is suggested that considering an ecological could help further conceptualise the IE for repositories and services and also help them navigate and understand this complex, unpredictable and changing landscape using well-defined ecological concepts.[43]

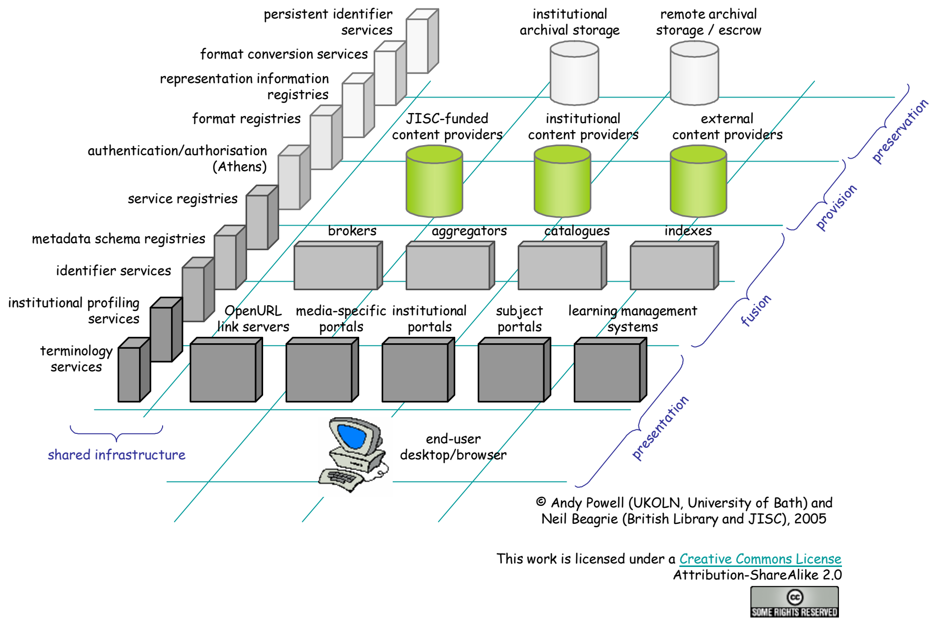


Figure 16 The JISC Information Environment

[42] JISC Strategy 2007-2009
http://www.jisc.ac.uk/media/documents/about_us/strategy/jisc_strategy_2007-2009.pdf

[43] If nothing else, this work highlights some of the problems inherent in the ongoing presentation of the IE only in terms of software architecture. The more of the IE that we build, the greater need we have for a presentation and understanding of it that is not just in architectural terms. The architecture provided and has driven the vision of the IE, but it is in danger of hitting all the problems outlined in section 4

10.5 An ecological approach and other work on repository interactions

Ecology offers a neutral metaphor for examining the complex interactions between people and systems. It can be used by repository managers, implementers, developers, funding agencies and users for communication and understanding. If it proves useful it might be applied to other aspects of the JISC Information Environment and beyond. There is, though, other activity which is working within a similar realm. There is not space in this report to cover these in detail. Indeed, the detailed sections above provide only a glimpse of the complexity and effort involved in developing architectures and models. But it is worth pointing to existing and ongoing work, as a placeholder for further effort to draw connections between these and an ecology-based approach.

Kerry Blinco's 'Wheel of fortune' [44] is one example of a successful visual aid for communicating the types and interlinking factors within a repository ecosystem that are combined within each individual repository. Work on repository typology in Heery and Anderson's 'Digital Repositories Review' [45] and since then within the Repositories Research Team [46] has attempted to further document these types and parameters, each of which could be viewed as 'environmental factors' or as the beginnings of a sub-set of typed ecologies.

Work on architectures like CORDRA and aDORE could be envisioned in ecological terms, acting as elements of an overall ecosystem. Other reference models, such as the DELOS manifesto for digital libraries provide additional frameworks for repository development and could impact on the structure of our ecosystem. The OAI-ORE project, funded by the Mellon Foundation in the U.S., is looking at developing specifications to support the exchange and re-use of compound information objects across systems. This offers a good example of an entity whose resulting outputs may have significant impact on objects and interactions within a repository ecosystem, just as OAI-PMH has had on the exchange of metadata between repository and other systems.

[44] Blinko, Kerry and McLean, Neil. 'A 'Cosmic' View of the Repositories Space (Wheel of Fortune)', 2004 http://www.rubric.edu.au/extrafiles/wheel/main.swf

[45] Anderson, Sheila and Heery, Rachel, *Digital Repositories Review*, AHDS and UKOLN, February 2005 http://www.ukoln.ac.uk/repositories/publications/review-200502/

[46] Repositories Research Team wiki (digirep) http://www.ukoln.ac.uk/repositories/digirep/

11 Next steps

The brief for this work emerged from the *Digital Repositories Review* and the *Repositories Roadmap*, which suggested the creation of a framework that could consider the relation between repositories and services, data flow between them and workflow issues. In the course of carrying out this work, it has become clear that such a framework needed to provide a conceptual context that would allow it to be both multifaceted (in addressing all types of interaction) and flexible. Consequently the developed ecological approach is a mixture of theory and method and is intended to support communication, management, and strategic thought in relation to these issues.

11.1 Who is this approach for?

In the development of this work the primary audience has been the development community, specifically, but not exclusively, the JISC development community and JISC programme and service managers.

We think that for those managing projects this approach is most likely to support their reporting process – as they seek to explain how what they have done fits in with their context and how that has influenced the ongoing uptake, success, or failure of their work. An example of the use of this sort of approach can be seen in the PROWE project, who developed a view of their ecology (based on the earlier work of Nardi and O'Day - see http://www.prowe.ac.uk/deliverables.htm#outputs).

It is possible that this work could be of use in indentifying opportunities for future development or for initial planning, but we consider that its strength for development projects is in supporting communication and analysis when the project is more developed.

We think that for programme managers, service managers, and analysts (including academics and consultants.) this approach additionally offers a way to assess the dependencies and vulnerabilities of services. In particular the habitat mapping and resource tracking approaches offer a way to begin to examine some of the questions raised in the digital repositories review and repositories roadmap.

Although there has been a positive overall response to the development of this work it should be noted that some people really like this approach and others really do not. Although some of this may be a communication issue, we think that it should be acknowledged that this, like any method or explanatory tool, is never going to suit everyone. A number of developers have questioned the need for a conceptual model or theory of repository interaction. The approach seems to have most resonated with those involved in the management of projects – especially those reporting about consortia.

11.2 What next?

The next stage for the development of the ecological approach is that it should be disseminated, used, and that further community feedback should be sought.

There are also, a number of key issues that have emerged from this work that we consider to be of relevance to the wider community irrespective of their view of an ecological approach. Perhaps the most important of these is the notion of environmental factors.

One of the distinctive features that has emerged from this work is the value of considering how to represent factors that do not fit within a modelled system but influence it. Within the ecological approach these are regarded as environmental factors. It has already been noted that in software design the aspect-oriented approach has been developed to address similar questions. We suggest as the community develops and plans repositories and services they take note of how their particular modelling tools take account of such cross-cutting or environmental factors.

11.3 Final thoughts

Ecology is not a new concept. As a metaphor it has the benefits of being instantly recognisable to everyone – our own environment and the ecologies within that are all around us. Its usefulness for the repositories space is in helping to make connections, to link technical complexity with non-technical complexity and offer a more neutral platform for understanding and communication. We believe it is a natural partner for ongoing work in developing information environments and offers one component in the process of documenting and developing repository provision across the world.

12 Further reading

- Bonnie A. Nardi, & 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 & 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
- Kerry Blinco & Neil McLean, "A 'Cosmic' View of the Repositories Space (Wheel of Fortune)", 2004, http://www.rubric.edu.au/extrafiles/wheel/main.swf
- Paul Miller "Interoperability. What is it and Why should I want it?" Ariadne Issue 24, 2000: http://www.ariadne.ac.uk/issue24/interoperability/intro.html
- F. Nachira, P. Dini, A.Nicolai, M.Le Louarn, L.Rivera Lèon. (Eds.) Digital Business Ecosystems. http://www.digital-ecosystems.org/book/de-book2007.html
- R. John Robertson, Anthony Troman, Paul Needham and Susan Copeland, "Repository Ecology: EThOS, the new UK e-theses service, national and institutional repository interaction", Presentation at JISC Conference, 13th March 2007, http://www.jisc.ac.uk/events/2007/03/event_conf_0307/abstract_ethos.aspx
- R. John Robertson, "Repository ecologies" Presentation at OAI 5, April 2007, http://indico.cern.ch/materialDisplay.py?contribId=11&sessionId=10&materialId=slides&confId=5710

12.1 Related work

- J. Barton & R.J. Robertson, "Developing a metadata lifecycle model" Workshop at CoLIS 5, June 2005, http://mwi.cdlr.strath.ac.uk/Colisworkshop.htm
- Anoush Margaryan, Sarah Currier, Allison Littlejohn, & David Nicol, Report on Learning Communities and Repositories, Community Dimensions of Learning Object Repositories, 2006, http://www.ic-learning.dundee.ac.uk/projects/CD-LOR/CDLORdeliverable1_learningcommunitiesreport.doc
- R. J. Robertson & 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

13 **Appendix 1: Worksheet on ‘The Ecology Mindset’**

Summary

Activity	Time Taken
Ecosystem Entities and Species Interactions Environmental Factors	20 minutes
Drawing your Ecosystem	10 minutes
Presenting your Ecosystem	20 minutes (5 minutes per pair)

Please record all your responses on the ‘Repository Ecology Sheet’
Please note that you will be working in **pairs**.

✍Task – Identifying your ecosystem (3-4 mins)

Together, think of an ‘ecosystem’ that has a digital repository in it (it may have many). It may be yours, it may be someone else’s, it may be an imaginary one e.g. a university has a repository for as many of its pre and post prints and various items of grey literature, such as presentations, monographs, articles from newspapers and magazines, from many academic departments.

✍Write the name of the ecosystem on the sheet provided (Section 1) and describe it briefly in one sentence.

✍Task - Identifying your entities and species and some of their characteristics (5-7mins)

First of all a quick reminder of what these terms mean

Entity

An entity is a tangible thing that exists within a repository ecosystem. (Please note that repositories and services are regarded as ‘living’ (in our ecology) and are largely treated in the same way as human participants in the ecosystem). Some example entities could be: users; repositories; services; objects; metadata records.

Species

A species within an ecosystem is a collective name for a particular type of entity. Example species are: institutional repositories, aggregator services, library catalogues, blogs, students, teachers, system administrators.

Why describe species?

When describing entities, the identification of the species allows us to find out what is known about the behaviour of the species in general and this can be used to help understand the particular entity. The observed characteristics of the entity are paramount but more general attributes of the species may illuminate the characteristics of the entity.

✍Task - List some of the entities and species in your ecosystem and describe the characteristics of one or two of them.
Please use the sheet provided (Section 2)

✍Interactions between entities and species
(5-7mins)

An interaction is a connection, relationship, or link between two or more entities or species in a population, community, or ecosystem.

Interaction can have any nature it may be:

- a machine to machine technical interaction
- an interaction between two people
- an interaction between a person/ people and systems

Interactions could be between entities, or entities and species or species to species.

In a few circumstances it may be sufficient to note that there is *some* form of interaction or interactions between two entities or species that is important for the ecosystem but it will usually be useful to further elaborate on the nature of that interaction.

Defining interactions may simply involve indicating in some way what form of interaction is occurring. For example, interactions can include:

- Harvests
- Talks to
- Emails
- Subscribes to rss feed from
- Edits objects
- Selects

Or it may also involve specifying any resources involved in the interactions.

✍**Please complete the table on the separate sheet by identifying the entities / species that are interacting, the direction of interaction (using arrows) and the type / description of the interaction (Section 3,4)**

✍**Environmental factors**

An environmental factor is something that influences an entity, species, community, or ecosystem but is more general than an interaction between constituent entities. For example, a research council's mandate for Open Access deposit is likely to affect the number of papers available in that subject area in a given repository community and throughout the ecosystem(s) interacting with that subject area. Other factors could be:

- competition for resources (e.g. entities can't develop because available funding has been allocated outside of community)
- suppression of competitors by other entities (e.g. political manoeuvring).
- cultural drivers (e.g. the Research Assessment Exercise)
- the effect of ideals and concepts (the Open Access movement)
- the existence of funding (e.g. funding streams)
- legal constraints (e.g. copyright restrictions)
-

✍**Task - Please complete the table on the separate sheet, filling in the column labelled 'Environmental Factors' (Section 3,4)**

✍**Drawing and presenting your Ecosystem**

You will have 10 minutes to draw and five minutes to present your ecosystem, use your ‘Repository Ecology Sheet’ to help you.

We would like you to draw your Ecosystem on the piece of paper, the pencils (with rubbers) and coloured pens provided. Please put your names and the title of the ecosystem in the top left hand corner of the sheet.

Tips for drawing your ecosystem:

- We have provided you an example of an overlay journal
- You do not have to draw all the entities and species in the ecosystem just a few indicating some interactions between the entities / species and some environmental factors influencing them.
- You may quickly sketch something in pencil (you can erase it if you make a mistake)
- **The final drawing should use the coloured pens provided**
  We suggest:
  - **Black** for entities and species
  - **Red** for interactions
  - **Blue** for environmental factors
- We will be taking a picture of your ecosystem, so ***please write clearly***!

Repository Ecology Sheet

Names of authors:

Name of Ecosystem:
Description:

1

Entities	Species	Characteristics of Entities / Species

2

Entity / Species	Type of Interaction	Entity / Species

Environmental Factors

3
4