LEarning About Digital Institutional Repositories

Creating an Institutional Repository: LEADIRS Workbook

By Mary R. Barton, MIT Libraries, mbarton [at] mit.edu
and Margaret M. Waters, consultant, mwaters [at] aya.yale.edu
with sponsorship from The Cambridge-MIT Institute (CMI).

LEADIRS II Workbook

Table of Contents

Chapter 1:
Building an Institutional Repository

Introduction

The Learning About Digital Institutional Repositories Seminars programme (LEADIRS) aims to describe and illustrate how to build an online institutional repository.

The LEADIRS series of seminars present specialists from the UK and abroad sharing their expertise and experiences in building institutional repositories.

This workbook book supplements the seminar presentations and offers practical advice as well as work sheets you can use to get started with your own repository programme. Where possible, we point you to real-world examples of planning aids or presentations used by university library teams in the UK and around the world.

The information in this book is as complete as possible at the time of writing. Because each institutional repository service will be unique to the institution where it is built, this information is meant to be helpful and to provoke discussion and exploration. It is not meant to be prescriptive. We cannot account for or anticipate the unique challenges and resources of your institution.

The authors are grateful to Margret Branschofsky and Julie Walker, colleagues at the MIT Libraries, for permission to adapt content from the DSpace Federation website and the DSpace at MIT project for this workbook.

How to Use this Material

LEADIRS materials are geared toward university librarians and senior level managers whose staff are building institutional repositories. Each chapter has information geared toward service managers, as well as domain-specific materials – in the technology chapter, for example.

In addition, financial managers will want to read the sections on cost modeling for institutional repository programmes.

Learning from others

Where possible, we provide links and references to universities that have already designed or built an institutional repository, so you can see real world examples of service plans, content policies, marketing materials, and so on.

Case studies

To highlight how specific universities in the UK have approached each phase of designing, building, and running an institutional repository, we present a series of case studies. These cases provide a starting point for discussion at the LEADIRS seminars, as well as among institutional repository teams at your university. See how peers at other universities face and solve similar problems to yours.

For further research

For each stage of building an institutional repository, we present the best websites, publications, and online resources to help you create your own service plan.

Note

Institutional repositories are a new breed of services and software, still in their nascence. The technology is constantly changing and new information on building and running these services is published all the time. The information we provide in the LEADIRS series and this book is meant to help you sort through this sea of information. Also, because your institutional repository service is unique to your institution, our information is not prescriptive. We offer no promise or guarantee of completeness in the guidelines and information we provide.

What is an Institutional Repository?

An **institutional repository** is a database with a set of services to capture, store, index, preserve and redistribute a university's scholarly research in digital formats.

The SPARC organization defines institutional repositories as follows:

- Institutionally defined
- Scholarly
- Cumulative and perpetual
- Open and interoperable

When we say institutional repositories are **open** and **interoperable** – it means they are OAI-compliant and allow open access to scholarly research.

Clifford Lynch defines IRs in the following way: "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."

Clifford A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" **ARL**, no. 226 (February 2003): 1-7.

Increasing momentum

Momentum has been building in the past several years for libraries to consider building repositories. Many of you will have read the House of Commons Science and Technology Committee report, released earlier this year, which recommends that "all UK higher education institutions establish institutional repositories on which their published output can be store and from which it can be read, free of charge, online." The report also suggests that government funding bodies "mandate their funded researchers to deposit a copy of all their articles this way".

As increasing amounts of research and scholarship exist in digital form, collecting and preserving this material serves multiple purposes. The House of Commons Report reinforces this point: "Self-archiving serves two main purposes: it allows authors to disseminate their research articles for free over the internet, and it helps to ensure the preservation of those articles in a rapidly evolving electronic environment."

How do People Use Institutional Repositories?

Universities and research libraries around the world use institutional repository in the following ways:

- Scholarly communication
- Storing learning materials and courseware
- Electronic publishing
- Managing collections of research documents
- Preserving digital materials for the long term
- Adding to the university's prestige by showcasing its academic research
- Institutional leadership role for the Library
- Knowledge management
- Research assessment
- Encouraging open access to scholarly research
- Housing digitized collections

Each university has a unique culture and assets that require a customized approach. The information model that best suits your university would not fit another campus.

Reference

See the PALS (Publisher and Library/Learning Solutions), report, "Pathfinder Report on Web Based Repositories," chapter 5, "Uses" for an excellent description of the varied uses for institutional repositories. (http://www.palsgroup.org.uk/)

Major Steps in Building an Institutional Repository

Broadly speaking, the following steps are the major milestones you will encounter in building an institutional repository. We present them here in logical order but realise that many of you will experience them differently.

- Learning about the process by reading about and examining other institutional repositories.
- Developing a Service Definition and Service Plan:
  - Conduct a needs assessment of your university.
  - Develop a cost model based on this plan.
  - Create a schedule and timeline.
  - Develop policies that govern content acquisition, distribution, and maintenance.
- Assembling a team

- Technology – Choose and install software platform
- Marketing
- Launching a Service
- Running a Service

Note that technology choices should reflect the requirements outlined in the service planning chapter. We address technology –software and hardware – in a separate chapter.

Use the Work Sheets at the end of each chapter to begin building your own service plan.

Most Common Challenges

The problems and hurdles which implementation teams face in building a repository include the following:

- Adoption rate by academics
- Providing for sustainability
- Developing policies
- Managing intellectual property rights
- University support
- Cost management
- Digital preservation
- Identifying key stakeholders

This workbook addresses these key challenges and points you to examples and references for further investigation. Also, see the Case Studies to learn how other implementation teams meet these challenges.

How Do I Get Started?

If you are just starting to design and build an institutional repository, focus on this chapter to get started planning what your service will offer and to learn about the decisions you need to make. Your service design is the foundation for all subsequent technology and budget choices.

If you have already started building a repository…

If you have already begun building an institutional repository, compare your existing plan to the material presented in chapter 2, to see if there are additional steps or decisions you may want to consider before proceeding.

Throughout the book, we include workbook pages you can use to keep track of pertinent questions and concerns, tracking your progress against the sample service plans we present.

Key References

In the past two years there has been an increasing amount of information published about institutional repositories. We have sorted through the literature and present the most useful links and references to help you find the most useful information quickly.

We consider the following articles and sites to be indispensable references:

- Digital Preservation Coalition
  http://www.dpconline.org/graphics/

- Crow, Raym. (2002) The Case for Institutional Repositories: A SPARC Position Paper, Washington, DC: Scholarly Publishing & Academic Resources Coalition.
  http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf

- Lynch, Clifford A. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" *ARL*, no. 226 (February 2003).
  http://www.arl.org/newsltr/226/ir.html

- SHERPA
  http://www.sherpa.ac.uk/documents/

- SPARC Europe
  http://www.sparceurope.org/index.html

See institutional repositories in action

Sometimes it's helpful to see how other universities and cultural institutions have organised or presented their repositories. The following sites link to a variety of online repositories and archives.

- ePrints list of UK repositories
  http://www.rdn.ac.uk/projects/eprints-uk/repositories/

- The Open Archives Forum, List of Repositories
  www.oaforum.org/oaf_db/list_db/list_repositories.php

- SPARC Europe, list of European institutional repositories
  http://www.sparceurope.org/Repositories/index.html#Europe

Next step

In the next section, we will guide you through the main steps to take to build an institutional repository, and the primary questions you need to answer along the way.

Chapter 2:
Planning Your Institutional Repository Service

Introduction

This section addresses the planning phases of building an institutional repository (IR), including service planning, staffing, marketing, and launching a service.

Subsequent chapters address technology choices, intellectual property issues and policies, as well as cost modeling.

Who should read this section?

The service planning material is geared toward university librarians and senior level managers whose staff are building institutional repositories. In addition, financial managers and hiring managers will want to read the sections on assessing existing resources and the staff skills required for operating an institutional repository.

Early-phase services

If you are just starting to design and build an institutional repository, focus on this chapter to get started planning what your service will offer and to learn some of the decisions you need to make. Your service design is the foundation for all subsequent technology and budget choices.

In the first chapter, we began by defining institutional repositories and outlining some of the many ways they are used. In this chapter, we address the planning phase and include a sample service plan you can customize to suit your service, showing the general milestones you will encounter.

Services already underway

If you have already begun building an institutional repository, compare your existing plan to those included here, to see if there are additional steps or decisions you need to take before proceeding.

We include workbook pages you can use to keep track of pertinent questions and concerns, tracking your progress against the sample plans we present.

Developing Your Service Model

It's important to define precisely how you intend to use the system and what type of services you will offer. For example, some universities build their institutional repository to hold only academic research. Others expand the service definition to include student theses, learning materials, or university records. Ideally, you want to decide this before you build the technical infrastructure of an institutional repository.

This section describes how to define your institutional repository service and then presents the major steps and decisions a team will encounter in planning an institutional repository service.

How do you define what your service offers?

An institutional repository is not defined solely by the software and database that contain your digital collections. It is a set of services – for those who deposit content, the academic and research communities you include, and for end users.

To develop a **service definition** for your Institutional Repository is to define what you will offer to all your users: those who deposit content in the repository and as well as end users. You determine your policy decision, the services the institutional repository will offer, the library's versus the content communities' roles, and the business plans of the service.

To create your service definition, you and your implementation team will answer a series of questions. For example, what kinds of content will you accept? Who can deposit content in the repository? Who will provide metadata?

For example, some institutional repositories accept only peer-reviewed material. Others, such as the University of California's Digital Library, accept prepublication materials. The Edinburgh University Library launched a repository for electronic dissertations and theses (EDTs) named Theses Alive! (http://www.thesesalive.ac.uk/ta_home.shtml). Others include learning objects, educational materials and other non-text items.

Creating Your Service Definition

To create the service model, your implementation team needs minimally to answer the following questions:

- What is the service's mission?
- What kinds of content will you accept?
- Who are the key users?
- Who are the key stakeholders?
- What services would you offer if you had unlimited resources?
- What can you afford to offer?
- Will you charge for services?
- What responsibilities will the library bear versus the content community?
- What are your top service priorities?
- What are the short-term priorities and long-term priorities?

Working with your colleagues on an implementation team, use the Work Sheet: Service Definition in this book to begin to articulate your service definition.

Institutional Repository Services

This section details the types of services an institutional repository can offer. Use this section to develop a detailed description of the support that your submitters and users can expect from the repository.

The service and support section leads to the resource allocation plan. A more extensive and complex service and support offering naturally leads to greater costs. One way to control costs is to offer one service and support level for free while reserving the option to offer other services on a for-fee basis. Managers may want to develop a separate matrix for those two categories for the purpose of developing the resource allocation plans.

There are many services that may be served by Institutional Repositories, some of which are currently supported by available technologies, others which will be developed in the near term, and still others that are merely in the 'identified needs' stage.

No doubt there are items that may fall into a fourth category of needs, which will be presented by early adopters of emerging digital research and teaching methodologies. While it is possible to plan for those services in the future – it is important to recognize the limitations of present systems.

Free vs. Fee-Based Services

The following table shows the range of services a library team might offer in an institutional repository. In this example, some services are available for free and others on a fee-based, cost-recovery basis.

Your service may offer some or most of these services – depending on how you structure the institutional repository and the needs of your community.

Whether you charge a fee for specific institutional repository services depends on your cost model, which we discuss in the Cost Modelling section of this book. In the example shown below, fees are charged to content communities that contract with the library to provide additional services.

Institutional Repository Services	
Core Services (free)	**Premium Services (fee-based)**
Setting up academic departments and other content communities in the institutional repository	
Metadata Services: Consultation	Metadata Services: Custom metadata creation
Training and user support for content submitters	Document services o Scanning o OCR o Reformatting files
Storage space allocation: basic	Extra storage space
Batch import of data: o Historic collections o Newly digitised collections	
IT Systems management	
Other:	

Reference

The University of Rochester offers a variety of core, or free, services, along with premium services to recover costs. See University of Rochester DSpace service, enumeration of core vs. premium services: http://www.library.rochester.edu/index.cfm?PAGE=1362 for more information.

Ways of Organizing Content

Each institutional repository service organizes content in a way to suit its university's unique culture and academic organizations. Many universities organize according to academic research centres or departments. This is by no means the only organizing principle.

One university (the University of Kansas) organizes its content using "hybrid communities" in the following ways:

1. **Formal Community** – Consists of departments, research centers, and groups already existing. Established submission guidelines and workflow. Example: Neuroscience Dept.

2. **Subject Community** – Open access, all academics can submit, or by proxy. Library staff review content before going online. Example: Policy Research Institute.

3. **Community of Interest** – An ad hoc group, crosses depts. Scholar-driven, membership limited to academic choice. Changes over time. Example: Social Science – cuts across departments.

How you organize content communities depends on the interests and allegiances of your academics. Most software systems call for customisations based on your content communities.

See the Technology chapter for discussion of current capabilities of the major available software systems.

Examples of Institutional Repository Service Definitions

Working with your colleagues, and using the Work Sheets in this book, you will begin to delineate a service definition for your institutional repository. Several institutional repository service teams have published information about their service models. These are valuable accounts of how others have solved some of the problems you face today.

Institution	***URL to learn about institutional repository programme***
University of Glasgow	http://www.lib.gla.ac.uk/daedalus/docs/eprintsleaflet.pdf
Queensland University	http://www.qut.edu.au/admin/mopp/F/F_01_03.html
University of Rochester	http://www.library.rochester.edu/index.cfm?PAGE=1362
Queens University in Canada	http://library.queensu.ca/webir/planning/q_space_planning_document.htm
Ohio State University	https://dspace.lib.ohio-state.edu/retrieve/335/KBRogers.pdf

Assembling a Team

Early on, you will want to identify staff responsible for each of the following tasks:

- Administering academic and staff surveys or interviews
- Conducting a needs assessment
- Synthesizing the results of surveys
- Developing your service model
- Developing a cost model
- Conducting resource assessments
    - Performing gap analysis
    - Developing requirements document
- Developing presentations for staff, academics and potential funders

These staff members constitute the **Implementation Planning Team**. Team representation should ideally include personnel from the libraries senior administration with budgetary responsibility, the archivist, and someone whose role is to interact regularly with academics.

The Implementation Planning Team should allow several months for the development of the Service Model, the associated costing and the implementation of the hardware and local customisation of the software.

Conducting a Needs Assessment Survey

A key foundation to your service plan is understanding the unique needs of your scholarly community.

How are scholarly works currently published or stored on campus? Who are the key stakeholders to include in your planning? What IT resources are available on campus? What are academic perceptions of the issues and problems of managing digital materials?

A successful needs assessment includes both informal and formal means of surveying the community:

- **Informal surveys** include face-to-face meetings with individual academics and administrators; email contacts, and monitoring existing web-based publishing services on campus.
- **Formal surveys** might include paper-based or online surveys of academics and staff, as well as formal presentations and Q&A sessions with departments and academic groups.

You can customize the Work Sheet: Academic Survey in this chapter to suit your service. Use it to understand how academics currently publish their research and how they might use an institutional repository at your university.

Reference

A team at the University of Rochester library embarked on such a study of academic needs: "Use a Shoehorn or Design a Better Shoe: Co-Design of a University Repository" by David Lindahl and Nancy Foster, available online at http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-13710/Participatory+Design+Conference+Paper+2004.07.31.pdf.

The Rochester team's website documents the study and results: http://docushare.lib.rochester.edu/docushare/dsweb/View/Collection-331

Creating a Service Plan

A typical institutional repository Service Plan is divided into three periods: Start-up, Growth and Maturity. These periods reflect milestones in phased development and allow for differing levels of resource commitment to meet those objectives.

Many institutions prefer to begin with a small, manageable service conducted alongside a targeted marketing and communication program to build awareness in the community. As momentum grows, they add content communities and reach additional academic departments.

Note: The activities and services listed below are meant to evoke ideas – not to comprise a definitive list.

Time Planning

Establish periods during which the activities will take place.

Typical Start-up period activities include:

- Identifying the service team and drafting the charge
- Identifying early adopters to begin adding content in a pilot, or beta, phase
    - Identifying historic collections to drive content
    - Identifying new research content
    - Identifying campus thought leaders
- Developing policies
- Choosing and identifying technologies
- Developing an advisory structure
    - Librarians
    - Academics
- Advocating for the service on campus
    - Senior administration
    - Academics

Staffing

Use the staffing work sheets in this book to track staff who contribute time and skills to the team. Whether or not you are able to hire new staff members – or if you rely on existing staff or some combination – you need to identify the skills already available to you in-house. In addition, you may need to contract out for additional resources, depending on the requirements of your service and the resources available to you.

Depending on the size and scope of your institutional repository service, you may need to hire, or allocate, resources to perform the following tasks:

Service and Support

- ❑ Assisting with community set-up
    - Web site design
    - Collection definition
    - Workflow definition (this may apply to those implementing DSpace only)
    - Batch loading of historic collections
- ❑ Supporting users
    - Telephone help line
    - Online help
    - Online documentation
    - FAQ pages

- ❑ Reviewing metadata
- ❑ Having library staff create metadata
- ❑ Developing customised metadata schemas
- ❑ Managing collections
- ❑ Consulting with communities and authors on preservation

These user support and service management tasks can be borne by existing staff or by hiring new staff. At some universities, an individual staff member devotes half or more of his/her time to the service, while at others, the tasks are divided among existing staff who retain responsibilities for other programmes as well.

User Support and Technology Roles

For a large or established service, there are two roles to be fulfilled: one primarily **technical** (IT systems/technology) and one **supporting users** and advocating for the institutional repository with academics and staff.

We also offer a work sheet that lists the skills needed for the two most prominent roles on the team. See the Work Sheet: Staffing in this book for sample job descriptions that outline the skills required for these positions.

- *User Support M*anager (or *Institutional Repository Service Manager* at some sites)
- *IT/Information Systems Specialist* to manage the technology

Again, depending on the scope of your service and the budget available to you, you may or may not hire new staff for these roles, or find/develop these skills among existing staff.

Running a Pilot or Early Adopter Programme

Each Institutional repository service is comprised of content collections – sometimes called **communities**. These are groups that contribute content to an institutional repository – either academic or administrative departments, colleges, centers, units, or labs, etc.

Many universities have found it helpful to run a **pilot programme** for their institutional repository service, showcasing a handpicked short list of early adopters who test the software and agree to join the programme early on.

This helps you to focus on adding one discrete group of content and users to the system, test the software, iron out procedures, and field test your policies and assumptions before launching the service to the entire university.

Once you identify a collection to join your pilot programme, choose someone as a liaison in that department to work with the repository's staff and implementation team.

Enlisting Content Communities

Finding communities at your institution to participate in a pilot or early adopter programme can take some legwork. Here are some suggestions for finding early adopter communities:

- Interview department heads and academics.
- Write an article for an academic newsletter or publication and ask for volunteers.
- Present the institutional repository service to writers and editors who publish content on campus.
- Contact departments, labs, research centers, etc. who publish content on the web.
- Post information about the institutional repository service to online tech discussion lists at your institution.

You may also want to survey academics to learn what they need from an institutional repository. See the Sample Academic Survey for ideas on how to quantify academics' needs.

Selection criteria

Selection criteria for early adopters might include the following attributes:

- A group or department friendly to your mission
- Diversity across discipline areas
- Diversity of content types or formats
- Including examples of different intellectual property rights management issues
- Collections of a manageable size; not an overwhelmingly large collection
- Campus opinion leaders
- Providing a content safety net – identifying content on the verge of being lost
- A community with enthusiasm for digital preservation and expertise in interpreting data

- Most important of all is a strong and reliable liaison within the community to work with the institutional repository service team

Selection criteria vary depending on the phase of development you're testing.

Finding a strong internal coordinator

When you are recruiting new early adopter communities, finding a strong internal coordinator to lead the community's work is half the battle. You might find this person in one of the following roles:

- Publications coordinator
- Assistant Dean
- Website manager
- Administrative staff

Note that it is usually not an academic who serves as liaison to the implementation team.

Work Sheet: Identifying Early Adopters

As you search for an academic department or research centre to join your institutional repository service or programme, you can use our sample work sheet to keep track of pertinent attributes of each department. This will help when setting your priorities in the service plan.

Marketing Your Institutional Repository Service

As you begin to build a service, it is critical to communicate how the service benefits the university community – in other words, to do some marketing to advertise the service on campus.

Some teams have been successful in targeting a handful of "thought leaders" on campus – getting them on board early to leverage their interest in the service. Another team contacted all the webmasters and writers at the university to raise awareness and generate leads among those who currently post scholarly content to university and department websites.

Potential Audiences

Academics

Some academics are already aware of the benefits of Institutional Repositories; others are less well versed in the new technology and the ways that it can help them manage digital assets. Still others are unaware that their work is at risk as it is currently managed.

Administration

Advocacy within the host institution is vitally important for sustainability.

External audiences

Thought leaders who avail themselves of the broader access possible from an institutional repository welcome recognition in the press and that coverage is correlated with an upsurge in both submitters and end-users in existing institutional repositories. The group includes both academics and central administrators.

- ❑ Institutional publications
- ❑ Coordination with university public relations office
- ❑ Web-site development
- ❑ User feedback sessions
    - o User discussion groups
    - o Usability sessions
- ❑ Academic advisory groups

Marketing Ideas

There are a variety of creative ways to market your institutional repository at your institution. You can use the ideas and tools we provide here to build awareness and adoption of your service.

Using a Top-Down Approach

- A top-down approach focuses on vice chancellors, deans, and administrators. Use word-of-mouth and direct influence. In this approach, you help create the institution's directive to build a digital repository, spreading the word to academics and staff.
- Garner institutional support by engaging influential academics and administrators before you launch the service.
- Develop a case demonstrating the value of the repository to the institution as a whole, showcasing university research and as a benefit to academic research.

Using a Bottom-Up Approach

- A bottom-up approach pitches the service to academics, staff, communities that publish on your university website, technical staff in departments, and groups dealing with publications, etc. In this approach, you prove the need for an institutional repository before requesting support at a higher level.
- Get academics interested in preserving their work for the long-term.
- Tap your Academic Advisory Committee to describe to their colleagues the benefits of using an institutional repository.
- Recognize that different departments have different cultures around scholarly communications, different digital needs. Your approach needs to take this into account.
- Look for academic acceptance in a wide range of disciplines, each with different cultures, and different publishing and digital needs.
- Approach academics who have publications on their department or faculty websites.
- Meet the editors, webmasters, and content managers on campus and present the service to them. They understand the challenges of online content management and preservation and can be great advocates for institutional repositories.

- Collaborate with other initiatives on campus for online content, courseware, etc.

Raising Awareness at the University

- Present your service in face-to-face meetings on campus – with communities, departments, individuals, by phone, in person, to staff, academics, IT departments, etc.

- Write a press release announcing the launch and distribute to all campus news outlets including faculty newsletters.

- Coordinate publicity at the department, library, and university level. Share marketing copy, posters, brochures with news office, websites, etc.

- Use printed brochures, posters, presentations and the university website to publicize the service.

- Plan events across campus and within content communities to publicize the launch of your service.

- Schedule a kick-off session for library staff to learn about your institutional repository service, ask questions, and build awareness.

- Build awareness of the institutional repository programme before you launch the service by running a pilot programme or early adopter programme.

- Do publicity both inside and outside the university. Some academics notice articles in the local newspaper and ask for more information.

- Listen to academics and end-users on campus, and remain flexible in your outlook as you gather requirements.

- Build interest in long-term preservation on campus.

- Offer presentations on topics of interest to academics and related to institutional repositories such as copyright, intellectual property rights in the digital age, etc.

Keeping in Touch with Content Communities

- Survey content communities annually to get feedback, gather new requirements, etc.

- Use an annual form to verify policy decisions.

- Run a Help line so content submitters and managers can reach the User Support Manager directly.
- Track bugs and enhancement requests.
- Share FAQs among content communities on campus.

See also the Marketing Lessons Learned from other institutional repository teams.

Adding Content to the Service

Among institutional repository teams we surveyed, the top priority and biggest challenge they face is getting academics and staff to submit content. Once they learn of the service and understand its importance, many academics are interested in the idea of a repository. But getting them to submit their content is the next challenge you face.

Enlisting Academic Participation

How you approach the problem of enlisting academic participation may include efforts on several fronts:

Marketing

Academics have to hear about your institutional repository service many times, over a period of time, and from several sources (print, online, in person). A good rule of thumb is that someone needs to have been exposed to your service seven times before they are fully aware of your service. Be sure to outline explicitly the benefits of your service to academics.

Technology

The user interface to your service ought to be attractive, easy-to-use, and well-documented.

Collaboration

Work with other web editors and websites on campus to see who's already posting academic research materials online and enlist their support and assistance. If someone on campus is working with online learning materials, you may need to work with them on content that is also to be housed in an institutional repository.

Intellectual Property/Policies

The easier it is for academics or departments to add content, the more likely they are to do so. You may need to offer guidelines, or even assistance, in clarifying rights issues.

Reference

Many institutional repository services face the hurdle of enlisting academic participation. Morag Mackie, project manager for advocacy at the University of Glasgow's DAEDALUS project, offers advice to her peers at other universities in an excellent article: "Institutional Repositories: Practical strategies from the DAEDALUS Project," by Morag Mackie, Ariadne, April 2004. http://www.ariadne.ac.uk/issue39/mackie/.

See also, David Prosser's presentation entitled "If You Build it, Will They Come? Filling an Institutional Repository," (2004) available at http://eprints.rclis.org/archive/00001005/01/Group7.pdf.

Promoting Your Service on Campus

Resource: Promotion and advocacy for your institutional repository

- Identifying 'champions' in academic departments who can encourage colleagues to take part is often the most valuable approach.
- Faculty will be more comfortable with providing content if they do not think that the e-prints movement will undermine the 'tried and tested' norms of scholarly communication. The fundamental message should be 'do not stop submitting papers to peer reviewed journals - but please deposit them in the e-prints archive *as well*.
- It is important whatever happens that e-print archives are run in such a way that they address the needs and working patterns of researchers. Things should be made as easy as possible for them to contribute.
- Set up a project web site that is linked to from the archive itself. This can act as a focus for developments and news.
- Publicize and promote the repository through university magazines, including the Library user newsletter; the distribution of literature about the value of institutional repositories, such as the SPARC *Create change* leaflet; and presenting at departmental meetings and university committees.

[Source: "A Guide to Setting Up Institutional Repositories" from CARL Institutional Repositories Pilot Project, Online Resource Portal. http://www.carl-abrc.ca/projects/ir/setting-up.htm]

Examples

The University of Glasgow's DAEDALUS programme offers a brochure to explain their service to academics: http://www.lib.gla.ac.uk/daedalus/docs/eprintsleaflet.pdf.

See the CARL Institutional Repository Project's online guide for academics and staff at Simon Fraser University Library and the Canadian Association of Research Libraries (CARL): http://www.carl-abrc.ca/projects/ir/. It lists their procedures and resources, as well as guidelines and policies for adding content to the system.

Queens University in Canada also publishes its guidelines and policies online: http://library.queensu.ca/webir/qspace-project/guidelines.htm.

Training and User Support

As you plan your service, consider the amount of training and support you will want to offer. Each of the varied user groups needs general exposure to the service –its features and how it is it used.

In addition, library staff who create metadata and add content need training specific to their jobs. Academics and their designated content contributors need training in adding content to the system and setting up content areas for departments or research centres.

- Library staff:
  - General procedures, understanding the service goals, etc.
  - User interface, adding content
  - Metadata procedures
  - Search methods
- Academics and Academic/Administrative Staff
  - General procedures, understanding the service goals, etc.
  - User interface, adding content
  - Metadata creation

Service Planning References

Collected below are service planning articles and guides to building an institutional repository.

We also include here lists of existing institutional repository services so you can see first hand how others organize and present their collections.

General references

Budapest Open Access Initiative:
http://www.soros.org/openaccess/

Campbell, Lorna M., Kerry Blinco, Jon Mason. Repository Management and Implementation. 2004:
http://www.jisc.ac.uk/uploaded_documents/Altilab04-repositories.pdf

CARL
Educause library
http://www.educause.edu/asp/doclib/subject_docs.asp?Term_ID=671

Digital Preservation Coalition
http://www.dpconline.org/graphics/

DSpace Project
Implementing DSpace section of DSpace.org website:
http://dspace.org/implement/index.html

Nixon, William J. *The evolution of an institutional e-prints archive at the University of Glasgow.* Ariadne Issue 32, July 8, 2002.
http://www.ariadne.ac.uk/issue32/eprint-archives/

Pinfield, Stephen; Gardner, Mike and John MacColl. *Setting up an institutional e-print archive*. Ariadne, Issue 31: April 11, 2002.
http://www.ariadne.ac.uk/issue31/eprint-archives/

Pinfield, Stephen. “Open archives and UK institutions: an overview”. *D-Lib Magazine*, 9, 3, March 2003.
http://www.dlib.org/dlib/march03/pinfield/03pinfield.html.

Pinfield, Stephen. “Creating institutional e-print repositories.” *Serials*, 15, 3, November 2002, pp. 261-264.
http://www-db.library.nottingham.ac.uk/ep1/documents/doc1/00/00/00/64/index.html.

Queens University Canada, QSpace project plan
http://library.queensu.ca/webir/planning/q_space_planning_document.htm

Rogers, Sally *A. Developing an Institutional Knowledge Bank at Ohio State University: From Concept to Action Plan*, 2003.
https://dspace.lib.ohio-state.edu/retrieve/335/KBRogers.pdf

Shearer, Kathleen. *A Step-by-Step Guide to Setting Up an Institutional Repository.* September 2002.
http://www.carl-abrc.ca/projects/ir/setting-up.htm

SHERPA tracks and posts links to relevant publications about creating eprints repositories:
http://www.sherpa.ac.uk/documents/

SHERPA Project Proposal, Version 2, October 2002.
http://www.sherpa.ac.uk/documents/proposal.pdf

SPARC Europe:
http://www.sparceurope.org/index.html

SPARC Europe FAQ:
http://www.sparceurope.org/Repositories/index.html

SPARC Institutional Repository Checklist and Resource Guide:
http://www.arl.org/sparc/IR/IR_Guide.html

Directories of Institutional Repositories

ePrints list of UK repositories:
http://www.rdn.ac.uk/projects/eprints-uk/repositories/

PALS report on Institutional Repositories (PDF), page 43 available from the Projects page of the PALS site: http://www.palsgroup.org.uk/.

SPARC Europe, list of European institutional repositories:
http://www.sparceurope.org/Repositories/index.html#Europe

SPARC Institutional Repository Checklist and Resource Guide:
http://www.arl.org/sparc/IR/IR_Guide.html#appendix

Information on How to Enlist Faculty Participation

Mackie, Morag. “Filling Institutional Repositories: Practical strategies from the DAEDALUS Project.” Ariadne. April 2004.
http://www.ariadne.ac.uk/issue39/mackie/

Prosser, David. "If You Build it, Will They Come? Filling an Institutional Repository," Presentation. (2004).
http://eprints.rclis.org/archive/00001005/01/Group7.pdf

Service Planning References

JISC Biannual Project Report, including list of service planning steps and items:
http://hairst.cdlr.strath.ac.uk/documents/HaIRST-FAIR-BR0204.pdf

Queens University: Q-Space Project Plan:
http://library.queensu.ca/webir/planning/q_space_planning_document.htm

Software Project Management Templates and Tools
(Primarily subscription based):
http://www.projectconnections.com/knowhow/template_list/newformat/planning-scope.html

Washington State University. *Cost Factors in Digitzation*,
http://digitalwa.statelib.wa.gov/newsite/projectmgmt/costfactors.htm

First Hand Accounts of Building an Institutional Repository

Beaudoin, Patsy and Margret Branschofsky, "MIT's DSpace Experience: A Case Study." (2004)
http://www.dspace.org/implement/case-study.pdf

The Glasgow ePrints Service at the University of Glasgow describes its service in a pamphlet for faculty and staff:
http://www.lib.gla.ac.uk/daedalus/docs/eprintsleaflet.pdf

Hubbard, Bill. 'Building Repositories of eprints in UK Research Universities'. in "Capturing Edinburgh's Research" event, University of Edinburgh Library, Edinburgh, 23 October 2003. PowerPoint presentation available at
http://www.thesesalive.ac.uk/archive/SherpaEdinburgh.ppt

Sally Rogers of the Ohio State University DSpace team writes about the planning process for their institutional repository, called KnowledgeBank:
https://dspace.lib.ohio-state.edu/retrieve/335/KBRogers.pdf

Queensland University in Australia posts the service definition of its ePrint repository: http://www.qut.edu.au/admin/mopp/F/F_01_03.html

Queens University in Canada publishes its IR service model and project plan:
http://library.queensu.ca/webir/planning/q_space_planning_document.htm

The University of Rochester publishes the service model for its DSpace institutional repository project:
http://www.library.rochester.edu/index.cfm?PAGE=1362

Vanderbilt University Library staff share their plans to include an institutional repository project into the university's strategic plans:
http://staffweb.library.vanderbilt.edu/strategicplan/diglib/recommendations.htm
and
http://staffweb.library.vanderbilt.edu/strategicplan/diglib/report/reporttosteeringcommittee.doc

Service Planning Workbook Sections

This section offers a series of work sheets and planning aides to help you start planning your institutional repository service.

- Work Sheet: Project Planning Timeline
- Work Sheet: Institutional Repository Services: Free or Fee-Based
- Work Sheet: Service Model Definition
- Work Sheet: Staffing
- Work Sheet: Identifying Early Adopters
- Project Planning: Sample Academic Survey
- Project Planning: Lessons Learned
- Marketing: Lessons Learned
- Marketing: Sample Communications Plan

Work Sheet: Project Planning Timeline

This section presents the major steps and decisions a team faces in building an institutional repository service.

Project Planning Timeline

	Early weeks	*Beginning phase*	*Before launch date*	*Launch*	*Long-term events*
Define service definition					
Assemble a team					
Choose technology					
Early Adopter programme					
Marketing the service					
Launch events					
Run the service					

Detailed Project Planning Steps

Developing a Service Definition
- Survey academic needs

Assembling a team
- Assess current staff skills and talents
- Hire and reassign staff, as needed

Technology
- Research and choose a software platform
- Install hardware and software systems

Marketing
- Develop marketing materials
- Develop web site, brochures, communications vehicles

Early Adopter programme
- Batch load existing collections
- Training, setting up new content collections

Launching a Service
- Launch events
- Marketing/PR campaign
- Training courses for submitters

Running a Service
- Long term tasks – running the system, growing the service, etc.
- System support
- Help line

Work Sheet: Institutional Repository Services – Free or Fee-Based

Your implementation team can use this work sheet to determine which services you might offer in your institutional repository and whether you will offer them for free or on a fee-based, cost-recovery basis.

Consider whether your institutional repository service will offer each of these services and if so, note whether you intend to charge an additional fee.

This list is not exhaustive – it is intended to cover the typical services institutional repositories are asked to provide to content submitters and users. Your service is unique and may have fewer or more services depending on your community's needs.

Reminder: Use a Phased Approach

When planning the services you will offer in your institutional repository, it is helpful to brainstorm a broad list of services at the start. You will not need to offer all of these services at the outset. You can offer new services in phases – startup, growth, and maturity – as staff skills increase and experience grows.

Institutional Repository Services	Core service (free)	Premium service (fee-based)
Setting up academic departments and other content communities in IR		
Metadata consultation		
Custom metadata creation		
Training content submitters		
User support for content submitters (troubleshooting, etc.)		
Document services o Scanning o OCR o Reformatting files		
Basic storage allocation		

Institutional Repository Services	**Core service (free)**	**Premium service (fee-based)**
Extra storage space		
Batch import of data: o Historic collections o Newly digitised collections		
User reporting		
IT Systems management		
Other:		
Other:		

Work Sheet: Service Model Definition

In order to write a detailed service model definition for your institutional repository service, your implementation team ought to consider the following issues and questions.

This list is not exhaustive, merely suggestive. The unique needs of your institution will dictate the parameters of your service definition. Note additional issues and questions in the space provided below.

- What is the service's mission?
    - Increase impact of academic research
    - Raise visibility/prestige of institution
    - Create an institutional leadership role for the Library
    - Showcase university's research output
    - Prepare for RAE
    - Manage institution's IT costs
    - Capture the Institutional record
    - Provide vital services to academics
    - Help Libraries to meet the challenges of the digital realm
    - House digitised collections
    - Manage learning materials
    - Encourage Open Access
    - Other _____
    - Other _____
- What kinds of content will you accept?
    - Published, peer-reviewed literature
    - Pre-Prints
    - Datasets
    - Research Materials
    - Learning Materials
    - Institutional Records
    - Theses
    - Conference Proceedings
    - Electronic Journals
    - Other _____
    - Other _____
- Who are the key users?
    - Academics
    - Library Staff
    - Students
    - Administrators
    - Internal Research Staff

- External Researchers
- Other _____
- Other _____

- Who are the key stakeholders?

  - Academics
  - Library Staff
  - Students
  - Administrators
  - Internal Research Staff
  - External Researchers
  - Other _____
  - Other _____

- What services would you offer if you had unlimited resources?

- What can you afford to offer?

- Will you charge for services?

- What responsibilities will the library bear versus the content community?

- What are your top service priorities?

- What are the programme's short-term priorities and long-term priorities?

Additional issues relevant to your institution:

_____
_____
_____
_____
_____

Work Sheets: Staffing

As you assemble a team to develop your institutional repository, you will look for certain skills sets you want to include – either by including existing employees, perhaps splitting responsibility among several by allocating their time, or with new hires.

Staffing User Service and Support Roles

Use this table to note staff members who can contribute skills and time to the institutional repository service.

User Service and Support Roles	Staff Available	% Time Allocated
❑ Setting up content collections (or communities)		
o Web site design		
o Collection definition		
o Workflow definition		
o Batch loading of historic collections		
❑ Supporting users		
o Telephone help line		
o Offering online help		
o Creating online documentation		
o Writing FAQ pages		
❑ Reviewing metadata		
❑ Having library staff create metadata		
❑ Developing customised metadata schemas		
❑ Managing collections		
❑ Consulting with communities and authors on preservation		
❑ _____		
❑ _____		

Work Sheet: Staffing

Depending on your existing staffing and the funding available to you, you may or may not be able to hire new staff for an institutional repository initiative.

In this work sheet, we offer two way of tracking staff and skills needed to run your service: tracking staff members and their skills, or enumerating skills needed to run the service and then finding or hiring staff who bring these skills to the team.

Tracking Staff and Skills

Staff Member	Skills	% of Time Dedicated	Employment Category
ex: Joe Smith	*Programmer*	*50%*	❑ On staff ❑ Contractor ❑ Permanent ❑ Temporary

Locating Staff to Provide Skills Needed for the Service

Skills Needed	Staff Member	% of Time Dedicated	Employment Category
ex: Programming	*Joe Smith*	*50%*	❑ On staff ❑ Contractor ❑ Permanent ❑ Temporary

Sample Job Descriptions

The following job descriptions show the two primary roles on an IR team:

- A User Support Manager
- An Information Systems and Technology Specialist

You may decide to hire/designate one individual to each role, or you may find these skills divided among a few existing staff members. What's important is that these skills are dedicated to the service, not that they are embodied in one staff member.

Note: In these sample job descriptions, we purposely do not enumerate the traditional library roles and skills that contribute to institutional repository services. These skills are well understood. We focus instead on the need for new skills and roles occasioned by an institutional repository implementation.

Sample Job Description: User Support Manager

Reporting to the library's Assistant Director for Technology director, the User Support Manager has primary responsibility for managing the communication with and support of the institutional repository's users, and particularly those users adding content to the system. This position requires a knowledgeable, enthusiastic, and self-motivated individual.

Responsibilities

- Take primary responsibility for all aspects of the system's user management.
- Perform user training for library staff and content contributors.
- Provide expertise and assistance to the library's public services staff in their support of IR end-users.
- Coordinate and manage the definition and setup of new IR content groups, and coordinate and communicate with library subject specialists.
- Perform outreach to university community, including site visits to academics and open training sessions.
- Make recommendations on new functionality to IT systems programmers based on feedback from academics, submitters, and library staff.
- Work with public relations organizations to publicize the institutional repository service.

- Coordinate importing of historical collections with the System Manager, including collection assessment, metadata consulting, conversion referral, and developing metadata crosswalks if necessary.

- Provide consultation on university policies and legal and regulatory issues related to intellectual property and sponsored research as they relate to the institutional repository service.

- Work on projects and teams with library and other groups at university who are closely aligned with IR services (such as a Metadata Advisory Group, online education initiatives, etc.).

- Chair the Service Advisory Group and participate in the Policy Committee.

Additional responsibilities

_____

_____

_____

Qualifications

- Master's degree in library science, or equivalent experience

- Experience with using, and helping others in an academic setting to use web-based software

- Extensive knowledge of library practices and goals, especially with regard to technology

- Working knowledge of web-based publishing tools and practices (such as HTML)

- Excellent written and oral communication skills and interpersonal skills

- Understanding of library mission and ability to communicate system mission and functionality clearly to key library staff and users at the university

Additional qualifications

_____

_____

_____

Other experience desired

- Experience supporting complex library systems
- Knowledge of the university community and research interests

Sample Job Description: Information Systems and Technology Manager

To run an institutional repository system you need a technologist who can take primary responsibility for the technology.

Reporting to the library's Assistant Director for Technology, the Information Systems and Technology Manager has primary responsibility for all aspects of the technical management of the institutional repository. The position requires a knowledgeable, enthusiastic, and self-motivated individual.

Responsibilities

- Hold primary responsibility for all aspects of the system's technical management.
- Coordinate related work by the Libraries' Systems Office and IS staff, including training.
- Contribute bug-fixes and other enhancements to the systems developer if applicable. (For open source systems, primarily.)
- Perform system monitoring, testing, and debugging.
- Provide system administration.
- Monitor and upgrade utility programs and middleware.
- Develop approved system enhancements.
- Manage hardware contracts and system administration tasks for IR servers, documenting operational issues
- Participate in projects and teams working on activities related to the IR service.

Additional responsibilities

_____

_____

_____

Qualifications

- Master's degree in computer or library science, or equivalent experience

- Experience programming (generally speaking, Java) and managing code written by others
- Understanding of network (especially web) development issues
- Experience with Unix systems and basic system administration skills

_____

_____

_____

Other experience desired

- Experience with Open Source development projects and procedures by which source of a project is shared with the community. (For open source projects only.)

_____

_____

_____

_____

Work Sheet: Identifying Early Adopters

These are some helpful categories for choosing content collections or communities to include in a pilot programme for early adopters.

Community	Formats	IP issues	Collection size	Metadata	Liaison help w/data	Typical users	Opinion leaders	Attract wide audience
University Press	Books	X	X					X
Statistics	DVI, teX, Postscript					X		
Botany	Images, Datasets, Audio			FGDC, Dublin Core		X		
Computer Science	Software, Code	X	X			X	X	
Theses Online Project	Some multiple formats		X	MARC				X
Medical School	Medical images, Datasets					X	X	
Classics	Text (multiple fonts)		X	MARC				
Business School			X	MARC			X	X
GIS Datasets	ASCII, DBtables, ESRI, raster			FGDC				

Project Planning: Sample Academic Survey

As part of the campus needs assessment step, you will want to develop a detailed picture of academics' current practices and expectations for publishing and distributing their research.

Reference

This survey is adapted from a survey used by the DSpace team at MIT. In a separate study, a team at the University of Rochester library embarked on a study of academics' needs: "Use a Shoehorn or Design a Better Shoe: Co-Design of a University Repository" by David Lindahl and Nancy Foster, available online at http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-13710/Participatory+Design+Conference+Paper+2004.07.31.pdf.

Sample Needs Assessment Questions

We have developed a series of questions you can customize and supplement to survey your university's academics about your institutional repository plans.

1. Which college or faculty are you affiliated with at the university?
2. How long have you been at the university?
3. Are you a tenured academic?
4. How important to you are the following statements about the benefits offered by an institutional repository. "An institutional repository would be a valuable tool if it could…"
    - Make preprint versions of my research available to a worldwide audience.
    - Make my research available faster than the traditional publishing process.
    - Make available types of materials that have not been made available through the traditional publishing process, including large datasets and rich media formats such as audio, video, and graphic images.
    - Make my research available with very little effort on my part and without my having to maintain a website of my own.

- ❑ Provide long-term preservation of my digital research materials.
- ❑ Make it easy for other people to search for and locate my work.
- ❑ Allow me to search for the most current findings of my colleagues throughout the university.
- ❑ Preserve the research of the university in a convenient, central place.

5. What digital formats do you use to create your research materials, conference materials, or other scholarly communication?

   - ❑ Data formats
   - ❑ Source code
   - ❑ Binary formats
   - ❑ BinHex format
   - ❑ Postscript formats
   - ❑ Video formats
   - ❑ Graphical image formats
   - ❑ Audio formats
   - ❑ Text formats
   - ❑ Page description formats
   - ❑ Microsoft Office Suite formats
   - ❑ Other

6. Which formats would you likely submit to an institutional repository?

   - ❑ Data formats
   - ❑ Source code
   - ❑ Binary formats
   - ❑ BinHex format
   - ❑ Postscript formats
   - ❑ Video formats
   - ❑ Graphical image formats
   - ❑ Audio formats
   - ❑ Text formats
   - ❑ Page description formats
   - ❑ Microsoft Office Suite formats
   - ❑ Other

7. If you distribute preprint articles, how do you distribute them?

   - ❑ Post them to my own website.
   - ❑ Post them to my department's website.
   - ❑ Post them to a discipline-specific preprint site.
   - ❑ Send them out by email.
   - ❑ Mail out paper copies.

- ❑ Other, please specify: _____

8. How many peer-reviewed articles, conference papers, datasets, or other types of scholarly communication do you typically author or co-author annually?

   - ❑ 0-1 per year
   - ❑ 2-4 per year
   - ❑ 5-7 per year
   - ❑ 8-10 per year
   - ❑ Greater than 10 per year

9. Which of the following typically apply when you submit an article or other work for publication?
   (Select up to three of the most common requirements you have.)

   - ❑ I surrender the copyright of all submitted materials to the publisher.
   - ❑ I pay a flat fee to submit materials.
   - ❑ I am required to pay page charges.
   - ❑ Neither the article nor any part of its essential substance, tables, or figures may be published or submitted elsewhere before submission to the publisher.
   - ❑ Preprints may be posted on recognized preprint servers if the server is identified to the editor upon submission of the paper, or with other minor restrictions.
   - ❑ I retain the right to post the final work on the web.
   - ❑ I retain the right to use the published materials in the courses I teach.
   - ❑ I don't typically get engaged in the submission process, so I am unfamiliar with its requirements.
   - ❑ Other

10. Overall, how interested are you in using an institutional repository when it becomes available at this university?

    - ❑ Extremely interested
    - ❑ Somewhat interested
    - ❑ Neither interested nor uninterested
    - ❑ Not very interested
    - ❑ Not at all interested

11. My concerns about submitting to an institutional repository include: (Select your top three concerns.)

    - ❑ I worry it might constitute prior publication and prevent me from submitting my work to journals.
    - ❑ I am hesitant to submit my work to a repository that does not have a formal review policy or other quality control process.
    - ❑ I prefer that only my formally published works be available for public consumption.
    - ❑ I am hesitant to assign distribution rights for my scholarly works to the university.
    - ❑ I would be worried about the risk to the patentability of my ideas.
    - ❑ I am concerned that works submitted to an institutional repository will not have citation value and will not count towards tenure.
    - ❑ I am uncomfortable using electronic resources such as word processors, spreadsheets, the Internet, and email.
    - ❑ I already submit to a preprint server.
    - ❑ Other

12. The university library is considering expanding its basic institutional repository service to include some custom and consultative services. Please indicate your interest in using the following services:

Personalized information services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Customized reporting services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Publishing services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Digital conversion services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Reformatting services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Reformatting consulting	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use

Collection administrative services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Metadata consulting	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Metadata services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use
Batch import services	Would not use	Probably would not use	Might or might not use	Probably would use	Definitely would use

13. Who in your department typically makes the purchasing decision for services such as those listed in the question above?

    - ❑ Individual academics
    - ❑ Head of the lab, center, or department
    - ❑ Department administrative officer
    - ❑ Other

14. How do you get most of your information about university programmes or initiatives?

    - ❑ University newspapers
    - ❑ Student newspapers
    - ❑ University website
    - ❑ Faculty newsletters
    - ❑ Departmental newsletters
    - ❑ Local newspaper
    - ❑ Other, please specify: _____

Project Planning: Lessons Learned

Here we present the top 'lessons learned' – in essence, what these teams wish they'd known before embarking on their institutional repository service. Because each institutional repository service is different, these ideas may not all apply all to your service.

Teams that are already implementing institutional repositories report that the following strategies and tactics were helpful in planning their respective IR services.

Management

- It's important to do business and operational planning in parallel with your technical installation effort.

- When speaking with deans and administrators, be prepared to answer questions about cost and copyright, and to make the case for enhancing the institution's reputation.

- Be flexible in defining your service model – it may take several iterations, and you'll revisit some decisions during the implementation phase.

- The academic calendar influences the service's progress enormously.

- Communicate your service's development and progress to Library staff along the way to generate awareness and acceptance.

- Cross-functional teams among library staff are very important for business planning and proved invaluable for

    - Analysing staffing impact
    - Creating a representative management structure
    - Developing the communications plan
    - Envisioning a future set of premium services
    - Integrating the repository service into all corners of the library's daily operations

- Stress the importance of managing through uncertainty.

- Professional development for library staff is critical.

- Many sites found that the Library's research focus helped to establish the library as an active participant in academics' research processes.

- Your IR Policies need clear explanations and examples.

Technology

- Some institutions quickly exceeded their initial storage. More content was submitted, larger files, etc. Keep this in mind when planning for capacity.

- It is helpful to have a test server in addition to the development and live servers. This lets you try out new collections or communities in a testing environment without impacting the production system.

Content Acquisition

- Some services find it politically advantageous to add the Dean's research materials to the service early on, to help spur academic acceptance of the service.

- Academics generally don't submit their own content – dept staff, web editors, or administrative staff do.

- When loading large content collections, some services find it helpful to get 'seed funding' for one time loading of large collections – the university press collection, for example.

- Seeding a collection with existing content also helps attract more end users and other content collections.

Note: See also the Marketing Lessons Learned section of this chapter.

Marketing: Lessons Learned

In marketing their institutional repository service, implementation teams have learned a variety of lessons about how to communicate the benefits and advantages of using an IR. They share here the most important lessons they learned in marketing their IR services.

Getting the Word Out

- Many institutional repository teams stress the importance of **educating academics** about why an institutional repository is important. Faculty customs and culture need to change.

- Prepare an **"elevator speech"** to describe your service and its benefits in a brief speech. You'll be glad to have a quick description of the service benefits when you spot a key administrator or academic waiting for the lift.

- **Ongoing communication** on campus is vital. Go back to communities often, use newsletters, personal follow up, and phone calls to answer questions, remind community liaison to get new content added to the repository.

- **Timing** is critical. Summer is a great time to market your service, if academics are around. Also, summer is a good time for staff to learn something new. During the beginning and end of the academic term, your message can be lost.

- Keep in **close contact** with existing communities. Staff at one institution take turns contacting communities to post content and move the service forward. Having several team members contact the community avoids the feeling of nagging your community liaisons.

- Use **success stories**, quotations from academics, and time lines for how long it takes to start a community to encourage other communities to get going.

- Work with the university's **Grants Office** to reach academics who need to demonstrate in grant proposals how their work will be distributed and preserved.

- **Invest** your time and money in getting a community going – jump-start a successful service as a community test case.

- **Different disciplines think differently**, and have different content and publishing needs. Address them specifically.

- **Educate academics** about issues of digital scholarship, preservation, etc. This makes it much easier to persuade them to add their research to an institutional repository.

- **"Easy"** sells – that is, showing individual academics how easy it is to submit and find content.

- Having a **persistent identifier** for their research is the single **best selling point** for an institutional repository when talking with academics.

- **Word-of-mouth** among academics and end-users is invaluable. When content contributors and users start using an institutional repository, word can spread quickly around your institution.

- Success doesn't always follow immediately after you publish an article, make a presentation, or persuade an academic to preserve his/her work. Your **marketing efforts pay off eventually**, even if it doesn't feel that way immediately.

- Use the term **"visibility"** in addition to "open access" – it's a softer term and for some, more easily understood.

Recruiting Content Collections/Communities

- There's a **long process** to launch a new content community in an institutional repository.

- The rhythm of the **academic calendar** is significant.

- There's no such thing as too much **publicity**.

- A Community's **policies** are set at the highest level of the community. For example, the chair of an academic department would most likely drive decisions about content acceptance in the department's institutional repository community.

- The concept of open access can be a tough sell. Providing **exposure for their content and digital preservation** is easier to sell to academics.

- The institutional repository team can take a **personal approach** to finding pilot programmes. Go out to talk with faculty groups, individual

academics.

- **Criteria** for selecting early adopter communities change as you go along. As you build communities, the size of collection and diversity of formats become more important than at the start.
- Library **heads of collections** are most helpful in recruiting early adopter participants.
- Each year, **survey your communities** to see who's still there, who's in charge of submitting content, and so on. This keeps your records clean and provides a way to get back in touch with all communities annually.
- Community **staff** change often.

Getting Content Submitted

- Consider launching new collections of content in **bi-annual "releases"** in fall and spring. This provides motivation to finish new collections and garner publicity. It also sets expectations about when new content collections will appear.
- Submitters need **more training** in entering metadata properly.
- Once academics and staff are trained to submit content, their **metadata is surprisingly rich**. But there are some problems with the submissions, including lack of authority control.
- **Authority control** is desperately needed for metadata – for example, using dropdown menus with author names, dept. names, degrees offered, etc.
- If you find that academics are having trouble submitting content, you might try **training graduate students** in submitting content, dealing with metadata, finding content to upload, etc.

Top Selling Points for Signing Up Early Adopters

- Persistent identifiers for content.
- Uniform presence for the university's research.
- Community and author control

- The self-archiving clause from key publishers. Emphasize this to academics who are concerned about the first publication clause in publishing contracts.

Marketing: Sample Communications Plan

To launch an institutional repository service at your university, you need to raise awareness and understanding of your service. This sample communications plan illustrates the many ways you can reach your audience and build interest among academics, staff, and end users.

We offer here a sample communications plan adapted from the DSpace team at MIT. You can customize to suit your institutional repository.

Note that the activities and documents you include in a communications plan for your service will depend on the specifics of your system and your university. This plan aims to reach all potential target audiences including library staff, academics and researchers, end users of the content, alumni, and the general public.

Sample Communications Plan

The first table below identifies those communication activities that will be ongoing, while the second one describes communication events that are tied to the service's launch date.

Ongoing Communications Activities				
Event	**Purpose**	**Person(s) Responsible**	**Target Audience**	**Repeat**
IR Website	Provide general information	User Support Manager	World	Continuous
Online Newsletter	Disseminate news about IR service	User Support Manager	Library staff, Advisory groups, Early Adopters	Quarterly
Meetings with Academics and Community Groups	Acquaint prospective communities with IR service	User Support Manager	Academics, Content Communities	Continuous
FAQ	Provide concise responses to commonly asked questions	System Support Manager & User Support Manager	World	As needed
Update Libraries Publications Packet	Incorporate IR into library's services	University News Office	Library's targets	Annual
Links from library website: (home page, departmental pages, subject	Incorporate IR into library's services	Webmaster	Web site audience	As needed

Ongoing Communications Activities				
Event	**Purpose**	**Person(s) Responsible**	**Target Audience**	**Repeat**
pages)				
Hands-on training sessions	Educate library staff	System Support Manager & User Support Manager	Library Staff	Continuous
Articles, Press Releases & Publicity Events	Raise awareness	University News Office	University, World	Continuous
Orientation for new academics and staff	Inform academics about IR, benefits	User Groups	New Academics and Staff	Annual
Press Kit	Provide Information about IR service	University News Office	Press	Annual

Institutional Repository Launch Events and Activities

Communications Events and Activities			
Event	**Purpose**	**Person(s) Responsible**	**Target Audience**
Spotlight on University Home Page	Advertise launch events, Raise awareness of IR service	University News Office	Campus
All Staff Meeting	Raise staff awareness of IR, answer questions	Planning Team	Libraries Staff
Articles in Campus Magazines and Newspapers	Promote participation; raise awareness on campus	User Support Manager and University News Office	Campus, Academics, Alumni
Live Demo for Library Council	Build understanding of IR service	IR System Manager, User Support Manager	Library Council
IT Partners Conference	Raise awareness on campus	User Support Manager	IT Staff
All Staff Meeting	Share final report of planning committee	IR Planning Team	Library staff
Alumni Events	Fundraising	University News Office, Development Office	Alumni
Printed Brochure	General information about IR service.	University News Office	Public
All Academics, Electronic Mailing	Promote launch event	University News Office	Academics
Press Releases for Launch	Announce launch of IR service	University News Office	General public, Higher education, Research community, Library community
Launch Event	Celebrate launch of IR service & fundraising	University News Office	University community, donors

Chapter 3:
Choosing an Institutional Repository Software Platform

Once you have determined the need to create an institutional repository and begun planning your service offering, it is time to examine the available systems closely to choose which one matches your needs.

Library directors need to balance the need for innovation in managing digital archives with available resources and budget constraints. Each Institutional Repository platform has unique strengths. This document outlines the technology choices and the software platforms available to you. Our purpose here is not to recommend any single system to you. Your choice of an Institutional Repository system depends entirely on the unique needs of your institution.

One of the most important steps is to visit other online IR initiatives to get a feel for the scope, quality and ease of use for the end user. See the list of university libraries using each of the institutional repository software systems outlined in the *Institutional Repository Software Platforms* section.

Potential Uses

The Institutional Repository platforms presented below can potentially serve a variety of uses, some with customisation:

- Pre-print and e-print archives
- Online theses
- Educational materials
- Digital libraries materials delivery
- University records management (in the future)
- Alternative publishing platforms (in the future)

Components of an Institutional Repository System

The essential components of an institutional repository are:

- Interface for adding content to the system
- Interface for searching/browsing/retrieving content
- Content database for storing content
- Administrative interface to support collection management and preservation activities

Additional features may include integration with other university systems, including online courseware, etc.

Choosing a Software Platform

To choose a software platform for your Institutional Repository, it is a good idea to assemble a team consisting of university library administration and staff along with information technology staff. Each member contributes expertise on how the system should operate and the features required – whether service features (metadata, submission workflow, content types, etc.) or underlying server issues (operating system, databases, search mechanisms, etc.).

This document outlines the issues your team might consider, features to look for, and the strengths of the top IR platforms currently available:

- Basic technology building blocks
- Product features to look for
- Technology product models
- Other technical aspects of running a service
- Implementation steps
- Cost considerations
- Major IR software providers
- Feature checklists

We also point you to sites already using each platform, so you can see for yourself how other universities have implemented these systems.

Basic Technical Building Blocks of an Institutional Repository

An institutional repository system consists of the following technology building blocks:

- Windows or Unix/Linux Servers
- Web Server, such as Apache and related web application tools
- Database, such as MySQL, DB2, Oracle, Postgres, SQL Server
- Institutional Repository Software

Note: You may need several servers for the service – for each of these phases: development, testing, and production.

Product Features to Consider

When examining a software platform, look for the following features:

- File formats supported: text, images, datasets, video, audio, etc.
- Metadata standards (descriptive, technical, preservation, rights)
- Interoperability: OAI compliance, Z39.50, SRW, etc.
- Permanent item address or locator (e.g., persistent URL)
- Search/browse of metadata
- Full-text search
- Workflow, submission for content approval
- User authentication and authorisation:

    - Back-end: content contributor, editor, administrator, metadata editor
    - Front-end: end-user access to content
- Customisation: API (application programming interface) for customising the software, extending features as needed

Software Developer/Distributor

- Free vs. commercial software (licence, subscription fees)
- Open Source vs. proprietary
- Technical support available:
    - for fee vs. free
    - by phone
    - by email
    - via online forums

Technology Product Models

Institutional Repository software platforms are available in several different licensing and distribution models.

- **Proprietary Software**
  You typically pay for the software and, optionally, any additional subscription or consulting fees. You own the use of the software and, with a subscription, get software upgrades. With a programming interface, or API, you can customise the software, but the software vendor owns, creates, and maintains the source code.
- **Open Source Software**
  You download the software platform, in most cases for free, and your IT staff can examine, customise, and enhance the source code. A central governing body manages the source code, but it is open for changes and enhancements from the development community (for example, CDSware, DSpace, EPrints, Fedora, Greenstone)**.**
- **Software Service Model**[*]
  A software vendor owns and distributes a software platform, or also hosts and manages your data for you. In this model, the software vendor provides additional services for a fee, and also controls and updates the software source code (for example, Open Repository or bepress). The three options are:
    - Run and managed locally
    - Run locally, managed remotely by vendor
    - Run and managed remotely by vendor

Which software model suits you best depends entirely on your needs and resources. For example, are you comfortable using an open source software platform that is open for community development, or would you prefer a more

[*] In the future, consulting companies may build on top of open source products, building services and offering hosting along with the free, open source software.

conventional software vendor-customer relationship, where you can pay for technical support, upgrades, and consulting as needed.

Be aware of hidden costs in all product models. It might be beneficial to speak with other university librarians who have built institutional repositories with the systems you are evaluating.

Technical Issues Once a Service is Running

Depending on the software platform you choose, your technical staff may manage the following aspects of the service delivery:

- Service availability (24/7)
- Scalability (growth)
- Backup and recovery
- System maintenance
- Extensibility: access to other university resources, systems
- Customisation
- Internationalisation/multilingual support
- Data loading

Implementation Steps

Implementing a software system consists of the following steps:

- Examining service needs and requirements
- Choosing a software platform
- Assembling and setting up necessary hardware, servers
- Installing and configuring software
- Creating a test/demo version of your system
- Customising the interface as needed
- Training staff
- Creating content approval workflows: accept, edit, reject, etc.
- Loading content
- Testing system

Cost Considerations

The costs of software and hardware platforms are generally predictable. However, the costs of overall planning, implementation, training staff, and running the system depend on your service plan.

Several Institutional Repository platforms presented here are free and open source. This means the software itself is cost-free, and your technical development staff can modify the source code and enhance its features.

Technology cost considerations

- Software costs (one-time and ongoing)
- Hardware, servers, etc.
- Operations staff
- Programming staff (if necessary)
- Backup and recovery
- Preservation

Planning for the Long Term

When choosing a software platform for your Institutional Repository, consider not just your current needs but try to envisage how your service might look ten or fifteen years from now.

- What types of content do you think you might need to host?
- How broadly will your faculty and staff adopt the service?
- How might the volume of submissions grow as the adoption curve rises?
- What preservation initiatives will you apply to stored files and data?
- What is the data migration or data export strategy were you to move to a new system?
- How vital to the institution is the content you are storing/preserving?

You may not have the answers to these questions today. But consider these questions as you investigate software solutions to meet your current and anticipated needs.

Digital Preservation

Along with open access to research material, digital preservation is an important motivation for building institutional repositories – to ensure digital research materials are available and accessible in the long term. While institutional repositories cannot *do* digital preservation, they are an important tool to reach this objective. Several key initiatives are addressing digital preservation issues: SHERPA and the Digital Preservation Coalition most prominently.

> Digital information is lost when it is left unattended while hardware, software and media continue to develop. Without intervention, an e-print may be subject to media degradation within a few years. Even if the e-print is securely backed-up, a few more years will see the e-print's content become inaccessible as software and hardware change.
> Pinfield, Stephen and Hamish James. *The Digital Preservation of e-Prints*. D-Lib Magazine (September 2003).

Digital preservation strategies

An excellent site for learning about digital preservation is the Preserving Access to Digital Information (PADI) site published by the Australian National University. They offer a concise guide to the most often discussed **strategies** for digital preservation – migration, adherence to standards, encapsulation and emulation – along with additional resources for in-depth information: http://www.nla.gov.au/padi/topics/18.html. Also, see their "Information Trails" to get started: http://www.nla.gov.au/padi/trails/index.html

There are a number of preservation strategies employed to offer short or long-term preservation. This list is by no means complete or detailed, but offers a springboard to discussion and further research.

Preservation Strategies

- **Bitstream Copying** – or backing up data, where you make a duplicate of the digital object.
- **Durable, Persistent Media** – where you preserve the physical media, or CD, on which the object is stored.
- **Migration** – where you copy data from one technology to another to avoid obsolescence of both the physical media and the data format.
- **Standards** – relies on recognized, long-term standards over proprietary formats.
- **Emulation** – process of reproducing software and hardware environments to translate code from one computing environment to run on another.
- **Encapsulation** – as part of an emulation strategy, where objects and metadata are grouped together to help decode and render object later.
- **Preservation Metadata** – describes the software, hardware and requirements of the digital object to use in preserving the object.

Sources: PADI: http://www.nla.gov.au/padi/topics/18.html
Cornell University Library:
http://www.library.cornell.edu/iris/tutorial/dpm/terminology/strategies.html.

Learning more about digital preservation

The Digital Preservation Coalition (http://www.dpconline.org) offers a wealth of information and practical guidance on digital preservation. There are two particularly useful guides for library staff building institutional repositories that have recently been published:

- Directory of Digital Repositories and Services in the UK
  http://www.dpconline.org/docs/guides/directory.pdf
- Contracting Out for Digital Preservation Services: Information Leaflet and Checklist, by Duncan Simpson
  http://www.dpconline.org/docs/guides/outsourcing.pdf

An important guide from the DPC is Paul Wheatley's "Institutional Repositories in the context of Digital Preservation" available from the DPC site at http://www.dpconline.org/docs/DPCTWf4word.pdf. He discusses the requirements and challenges of providing for long-term accessibility of archived objects.

As Paul Wheatley notes, there has yet to develop a consensus for best practices for long-term digital preservation. Technologies and strategies are still emerging. See the Resources section of this chapter for further information on researching digital preservation and keeping on top of this evolving area.

Institutional Repository Software Providers

The following are some of the more well known software developers/vendors offering Institutional Repository software. This is not an exhaustive list but you might examine these when choosing the system that best suits your needs:

- Archimede
- bepress
- CDSware
- CONTENTdm
- DSpace
- EPrints
- Fedora
- Greenstone
- Open Repository

Refer to the section, *Institutional Repository Software Platforms*, in this book, for information on each of these featured software systems.

Reference

For information on how institutional repository platforms are evolving, see Alan McCord's article, "Institutional Repositories: Enhancing Teaching, Learning, and Research" published by Educause Evolving Technologies Committee (October 2003), http://www.educause.edu/ir/library/pdf/DEC0303.pdf.

Technology References

General technology references

Campbell, Lorna M., Kerry Blinco and Jon Mason. Repository Management and Implementation: A White Paper for alt-I-lab 2004. Prepared on behalf of DEST (Australia) and JISC-CETIS (UK) (2004).
http://www.jisc.ac.uk/uploaded_documents/Altilab04-repositories.pdf

Crow, Raym. *The Case for Institutional Repositories: A SPARC Position Paper*. SPARC: Scholarly Publishing & Academic Resources Coalition (2002).
http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf

Johnson, Richard K. *Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication*. D-Lib Magazine, (November 2002).
http://www.dlib.org/dlib/november02/johnson/11johnson.html

McCord, Alan. *Institutional Repositories: Enhancing Teaching, Learning, and Research.* Educause Evolving Technologies Committee (October 2003)
http://sitemaker.umich.edu/dams/files/etcom-2003-repositories.pdf
or http://www.educause.edu/ir/library/pdf/DEC0303.pdf

Pinfield, Stephen, Mike Gardner, and John MacColl. *Setting up an institutional e-print archive* Ariadne 31 (2002)
http://www.ariadne.ac.uk/issue31/eprint-archives/intro.html

RLG DigiNews, Editor's Interview with Clifford Lynch, August 15, 2004.
http://www.rlg.org/en/page.php?Page_ID=19481&Printable=1&Article_ID=1601

Surveys and Comparison of Software Systems

Budapest Open Access Initiative, *A Guide to OAI-compliant Institutional Repository Systems* (2004).
http://www.soros.org/openaccess/software/.

CARL Institutional Repository Pilot Project list of tools and technology,
http://www.carl-abrc.ca/projects/ir/tools-e.htm

Crow, Raym. Institutional Repository Checklist and Resource Guide, SPARC: Scholarly Publishing & Academic Resources Coalition (2002).
http://www.arl.org/sparc/IR/IR_Guide.html

DiLauro, Tim, *Choosing the components of a digital infrastructure*. First Monday, volume 9, number 5 (May 2004).
http://firstmonday.org/issues/issue9_5/dilauro/index.html

Jones, Richard D. *DSpace vs. ETD-db: Choosing software to manage electronic theses and dissertations.* Ariadne Issues 38 (January 2004).
http://www.ariadne.ac.uk/issue38/jones/intro.html

Nixon, William J. *DAEDALUS: Initial experiences with EPrints and DSpace at the University of Glasgow.* Ariadne Issue 37 (October 2003).
http://www.ariadne.ac.uk/issue37/nixon/intro.html

Digital Preservation

Cornell University Library. *Digital Preservation Management:* Implementing *Short-term Strategies for Long-term Problems (2004)*
http://www.library.cornell.edu/iris/tutorial/dpm/index.html

Digital Preservation Coalition. *Directory of Digital Repositories and Services in the UK (*2004).
http://www.dpconline.org/docs/guides/directory.pdf

ERPANET, *Ingest Strategies Guidance Document*, 2004.
http://www.erpanet.org/guidance/index.php#technology

Jones, Maggie and Beagrie, Neil. *Preservation Management of Digital Materials.* Digital Preservation Coalition (2002).
National Library of Australia. Preserving Access to Digital Information. *PADI Thesaurus: Costs*
http://www.nla.gov.au/padi/topics/5.html

PADI *Digital Preservation Strategies*.
http://www.nla.gov.au/padi/topics/18.html

PADI Trails. Information Trails on digital preservation topics.
http://www.nla.gov.au/padi/trails/index.html

Pinfield, Stephen and Hamish James. *The Digital Preservation of e-Prints*. D-Lib Magazine (September 2003).
http://www.dlib.org/dlib/september03/pinfield/09pinfield.html

Simpson, Duncan. *Contracting Out for Digital Preservation Services: Information Leaflet and Checklist*. Digital Preservation Coalition (2004).
http://www.dpconline.org/docs/guides/outsourcing.pdf

Wheatley, Paul. *Institutional Repositories in the context of Digital Preservation* DPC Technology Watch Report (2004).
http://www.dpconline.org/docs/DPCTWf4word.pdf

Work Sheet: Requirements Document for IR Software Systems

This specification lists some features to look for when evaluating institutional repository software systems.

Product Distribution

- Open Source or Proprietary
- Type of purchase plan (free or commercial, one-time or renewable licence)
- Is remote hosting available?

Programming and Customisation

- Documented API (application programming interface) which allows the university's programming staff to customise and extend the software platform's features (applies to Open Source systems and proprietary)
    - What features have API's (e.g., user interface)?
- If Open Source, what programming language?

File Formats Accepted

- Text (documents, theses, books)
- Images
- Datasets
- Video
- Audio
- Computer programs
- CAD/CAM
- Databases
- Complex/multi-part items

Technical Features

- Persistent identifier (CNRI handles, etc.)
- Workflow for content approval, submission
    - Flexible
    - Multi-step
    - Roles-based
    - Centralised or decentralised submissions
- Versioning (over-time, different formats)
- Search engine/full-text search
- Metadata browse, search and sort features
- Content management features (e.g., preservation)

- User authentication and authorisation:
  - Back end - content contributor, editor, administrator, metadata editor
  - Front end – content to end users
- Support for multiple languages in search, user interface
- Bulk importing and exporting of data
- Single submission consisting of several related digital objects
- Real-time updating and indexing of deposits
- Rendering of non-web formats

Metadata Standards

- Descriptive
- Technical
- Preservation
- Rights
- _____

Interoperability

- OAI-compatible
- Z39.50 or SRW
- METS (or other packaging standards such as IMS CP)

System Administration

- User Management
- Adjustable user permissions
- Supports user authentication (x.509 or LDAP)
- Registration, roles-based security, authentication, authorisation, etc.
- Reporting features
- Logging features
- Scalability
- Clustering with automatic fail over
- Backup and recovery

System Configuration/Constraints

- Operating Systems supported
- Database(s) supported
  (Specify if database is included with system)
- Other 3rd party software tools required? (especially commercial products)
- Recommended hardware configuration
- Is client software required other than web browser?

Technical Support

- Paid/non-paid
- By phone
- By email (direct)
- By open email distribution list
- Are paid consultants available from software development organization to help with implementation?

Technical Documentation

- Full Documentation available
- Online Help
- System documentation
- Programming documentation, if applicable
- Are sample configuration files supplied with system?

Additional Factors to Consider

- _____
- _____
- _____
- _____

Reference Sites

List URLs of sites currently running the software system:

- _____
- _____
- _____
- _____

Institutional Repository Software Platforms

The landscape of software platforms for building institutional repositories is constantly changing. We outline below the major features and benefits of the leading institutional repository software platforms:

- Archimede
- Bepress
- CDSware
- CONTENTdm
- DSpace
- EPrints
- Fedora
- Greenstone
- Open Repository

In addition to brief profiles of each platform, we offer a work sheet to help you define the feature list and requirements you will use to decide which platform is best for your institutional repository service.

Note: This list of software platforms is not exhaustive. There are several other software platforms that libraries choose to use. The information in this section is as accurate as possible at the time of writing. Software standards and releases change constantly, so be sure to research current offerings using this information as a starting point.

Resource

An excellent resource for researching institutional repository software is the Budapest Open Access Initiative's *Guide to Institutional Repository Software*, available at http://www.soros.org/openaccess/software/.

Archimede

URL: http://www1.bibl.ulaval.ca/archimede/index.en.html

Description: Developed at Laval University Library, Archimede is open source software for building institutional repositories. It has been developed with a "multilingual perspective," offering English, French and Spanish interfaces. With a focus on internationalisation, the software interface is independent and not embedded in the code. This allows you to develop additional language-specific interfaces without re-coding the software itself. It also lets users switch from language to language "anywhere and anytime" while searching for and retrieving content.

Availability

- Free, open source software, delivered under the GNU general public licence.
- Download Archimede software from SourceForge: http://sourceforge.net/projects/archimede

Features

- Inspired by the DSpace model, using communities and collections of content.
- The search engine is based on open source Lucene, using LIUS (Lucene Index Update and Search), a customized framework developed at Laval by the library staff.
- OAI compliant.
- Uses a Dublin Core metadata set.

Technical support: http://sourceforge.net/projects/archimede/

Example site

- Laval University Library

bepress

URL: http://www.bepress.com/repositories.html

Description: Developed by the Berkeley Electronic Press, bepress builds and hosts their customers' repositories. Cost includes software, custom implementation, infrastructure, training, hosting, offsite backup, technical support, and software upgrades.

Recently announced software partnership with ProQuest Information & Learning called Digital Commons@ which combines bepress with ProQuest's library of electronic theses and dissertations (ETDs).

Availability: Commercial software, paid licence and subscription fees

- Flat rate: between $8k and ~$50k per year for unlimited series, unlimited papers, cost relates to size of repository, anticipated usage. Includes full training, documentation, technical support, customer service, and software upgrades.
- Variable rate: $4-5k licence fee and usage fees per content series, per posted paper

Features

- Offers EdiKit client software for entering content to repository
- OAI compliant
- XML data exporting
- Customisation through API, templates
- Full text search

Technical support: Available through paid software licence.

Example sites

- Boston College
- University of California's eScholarship Repository (http://repositories.cdlib.org/escholarship/)
- Florida State University
- New England Law Library Consortium

CDSware (CERN Document Server Software)

URL: http://cdsware.cern.ch

Description: Developed by CERN, the European Organization for Nuclear Research, based in Geneva, CDSware is designed to run an electronic preprint server, online library catalogue, or a document system on the web.

Availability

- Free, open source software distributed under the GNU General Public Licence
- Latest version: CDSware v0.3.3
- Download location: http://cdsware.cern.ch/download/

Features

- OAI compliant
- MARC 21 metadata standard
- Full text search
- Database: MySQL
- Extensibility: API available
- Powerful search engine with Google-like syntax
- User personalization, including document baskets and email notification alerts

Technical support

- Free email support at cds.support@cern.ch or through mailing list: project-cdsware-users@cern.ch
- Paid technical support is also available.

Example site

- CERN document server: http://cdsweb.cern.ch/
  At CERN, CDSware manages over 400 collections of data, consisting of over 600,000 bibliographic records, including more than 250,000 full text documents.

CONTENTdm™

URL: http://contentdm.com/

Description: Digital Collection Management Software by DiMeMa, Inc. providing tools for everything from organizing and managing to publishing and searching digital collections over the Internet. CONTENTdm also offers scalable tools for archiving collections of any size.

Availability: Commercial software. Pricing based on collection size.

- Rate between $7,000 for a maximum of 8,000 stored items and $40,000 with no limit on stored images.
- Additional cost of $1,000 - $6,000 for an annual maintenance agreement fee, which is included in the fee for the first year.
- One day of installation assistance and on-site training is available for $2,500 plus travel expenses.
- Free 60-day trial available

Features

- All content types accepted
- OAI compliant
- Dublin Core metadata
- XML data export
- Z39.50 compatible
- Multiple-collection searching
- Automatically add collections to WorldCat
- Product includes several components: CONTENTdm Server, Acquisition Station software (can be copied and installed on up to 50 Windows machines for distributed use), search client and web templates, PowerPoint plug-in

Technical support: Available through an annual maintenance agreement fee, between $1,000 - $6,000. Installation support also available for a fee.

Example sites

Full list of organisations using CONTENTdm at http://contentdm.com/customers/customer-list.html including:

- University of Arizona
- University of Iowa
- University of Oregon
- University of Washington Libraries
- Oregon State University
- Colorado State University Libraries
- Brigham Young University

DSpace

URL: http://www.dspace.org

Description: DSpace is a digital library system designed to capture, store, index, preserve, and redistribute the intellectual output of a university's research faculty in digital formats. Developed jointly by HP Labs and MIT Libraries.

Availability

- Free, open source software jointly developed by MIT and Hewlett Packard Labs.
- Latest version: DSpace 1.2.1
- Distributed through the BSD open source licence
- Download at http://sourceforge.net/projects/dspace/

Features

- All content types accepted
- Dublin Core metadata standard
- Customisable web interface
- OAI compliant
- Workflow process for content submission
- Import/export capabilities
- Decentralised submission process
- Extensible through Java API
- Full text search using Lucene or Google
- Database: PostgreSQL, or SQL database that supports transactions, such as Oracle, MySQL

Technical support

- DSpace-Tech mailing list for technical questions, discussions: http://www.dspace.org/feedback/mailing.html

Example sites

- Cambridge University
- Cranfield University
- Drexel University
- Duke University
- University of Edinburgh
- Erasmus University of Rotterdam
- Glasgow University
- Hong Kong University of Science & Technology Library
- Massachusetts Institute of Technology
- Université de Montréal (Erudit)
- University of Oregon

EPrints

URL: http://software.eprints.org

Description: GNU EPrints is free, open source software developed at the University of Southampton. It is designed to create a pre-print institutional repository for scholarly research, but can be used for other purposes.

Availability

- Current version: GNU EPrints 2.3.6
- Distributed under the GNU general public licence
- Download software at http://software.eprints.org/download.php
- Demo server: http://software.eprints.org/demo.php

Features

- Any content type accepted
- Archive can use any metadata schema.
- Web-based interface
- Workflow features: content goes through “moderation process” for approval, rejection, or return to author for amendment.
- MySQL database
- Extensible through API using Perl programming language.
- Full text searching
- RSS output

Technical support

- EPrints-tech mailing list: http://software.eprints.org/maillist.php
- General announcements and “underground” discussion list also available at http://software.eprints.org/maillist.php.
- EPrints wiki: http://wiki.eprints.org/w/

Example sites

Full list of 141 known sites at http://software.eprints.org/archives.php, including:

- California Institute of Technology
- CogPrints Cognitive Science Eprint Archive
- Digitale Publikationen der Ludwig-Maximilians-Universität München
- Glasgow ePrints Service
- Institut Jean Nicod - Paris
- National University of Ireland (NUI) Maynooth Eprint Archive
- Oxford EPrints
- Psycoloquy
- University of Bath
- University of Durham
- University of Southampton

Fedora

URL: http://www.fedora.info/index.shtml

Description: Jointly developed by University of Virginia and Cornell University, Fedora (Flexible Extensible Digital Object Repository) serves as a foundation for building interoperable web-based digital libraries, institutional repositories, and other information management systems. It demonstrates how you can deploy a distributed digital library architecture using web-based technologies, including XML and Web services.

Availability

- Free, open source
- Distributed under the Mozilla open source licence
- Information available on future release of Fedora Phase 2: http://www.fedora.info/documents/fedora2_final_public.html
- Download the current release, Fedora 1.2.1 at http://www.fedora.info/release/1.2/

Features

- Any content type accepted
- Dublin Core metadata
- OAI compliant
- XML submission and storage
- Extensibility: APIs for management, access, web services
- Content versioning
- Migration utility

See a full list of Fedora features at http://www.fedora.info/.

Technical support

- Free online support through mailing list: https://comm.nsdlib.org/mailman/listinfo/fedora-users
- Fedora WIKI: http://www.fedora.info/wiki/bin/view/Fedora/WebHome

Example sites

- Indiana University
- Kings College, London
- New York University
- Northwestern University
- Oxford University
- Rutgers University
- Tufts University
- University of Virginia
- Yale University

Greenstone

URL: http://www.greenstone.org/cgi-bin/library

Description: Developed by the New Zealand Digital Library Project at the University of Waikato, Greenstone is a suite of software for building and distributing digital library collections. Greenstone was developed and distributed in cooperation with **UNESCO** and the **Human Info NGO**.

Availability

- Free multi-lingual, open source software
- Distributed under the GNU General Public Licence
- Current version: Greenstone v2.51
- Download location: http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=download

Features

- Multilingual: Four core languages are English, French, Spanish and Russian. Over 25 additional language interfaces available
- Includes a pre-built demonstration collection
- Offers an "Export to CDROM" feature

Technical support

- Online support: http://www.greenstone.org/cgi-bin/library?e=p-en-download-utfZz-8&a=p&p=support
- Technical email list: https://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-devel
- General user discussion list: https://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-users
- Commercial support is available for a fee.

Example sites

Full list at http://www.greenstone.org/cgi-bin/library?e=p-en-support-utfZz-8&a=p&p=examples shows a great variety of applications of the software platform, including:

- Books from the Past/ Llyfrau o'r Gorffennol
- Gresham College Archive
- Peking University Digital Library
- Project Gutenberg at Ibiblio
- Texas A&M University: Center for the Study of Digital Libraries
- University of Applied Sciences, Stuttgart, Germany

Open Repository

URL: http://www.openrepository.com/default.asp

Description: Open Repository is a new offering from BioMed Central, providing a commercial software service for building institutional repositories. Based on the DSpace software, it offers multiple editions with varying levels of support and service.

Availability

- Commercial services ranging in price from £5,000 to £10000 for set up and an additional £2,500 to £25,000+ in maintenance fees
- Three editions currently offered: Standard, Silver and Gold
- Product information available at http://www.openrepository.com/Products.html

Features

- Accepts wide variety of content formats
- Conversion utility to create PDFs
- OAI-based metadata
- Feature list available at http://www.openrepository.com/Open.Rep.Sales.Flyer.pdf

Technical support

- Full technical support available

Chapter 4: Legal and Regulatory Environment and Policy Development

This section addresses a variety of legal and intellectual property issues you may encounter as you develop an online institutional repository.

These materials will guide you through the relevant intellectual property rights issues (IPR) and point you to resources to help develop policies for running your institution's service.

Our goal is to suggest issues you may need to consider and to offer policy guidelines based on best practices of library and university staff who are currently running institutional repositories. We also refer you to recent legislation on intellectual property rights issues.

Caveat

This seminar and related materials point out some of the legal issues you may encounter and suggests sources for further research. This information is correct and current to the best of our abilities but does not constitute legal advice.

You are strongly advised to seek legal counsel to consider your particular needs and environment.

Also, each locality and each university faces unique legal and institutional regulations and practices. Just as Scottish law differs from English law, so too will the Universities of Edinburgh and Cambridge devise different policies.

Getting Started

Library staff and researchers routinely deal with intellectual property policies and publishing rights issues for printed matter. Online materials are similarly governed by government copyright, legal deposit laws, and publishing contracts.

Each university that develops an institutional repository needs to research and write policies and regulations for its collections. We include a checklist of policy issues to consider as you design and build a service, and then populate it with electronic content.

Because each institutional repository is unique, and regional laws differ, your policies will be unique to your service. The publications and websites listed below offer valuable information on the legal and regulatory environment affecting institutional repositories, particularly in light of the recent changes in the UK copyright laws.

Who should read this section?

The material in this section is appropriate for a readers just embarking on an institutional repository project as well as those already building a service. Generally speaking, all members of the team developing an institutional repository need to understand applicable IP government regulations, and be conversant with the customs and practices at your university.

But ultimately, responsibility for setting and monitoring IP issues falls to the **project manager**, working with the university's **copyright expert** and **legal counsel** (if available).

If you create a **Policy Advisory Group** to set rules for the repository, this group's charge includes writing the repository's rules for copyright ownership and licences for the service (both for depositing content and accessing content). A Policy Advisory Group typically includes the following roles:

- Library staff
- University administrators
- Faculty
- Legal counsel or copyright expert
- Archive staff
- Information systems staff
- Other staff as needed

The *Policy* section below provides additional details and worksheets for creating policies for your service. Also, see the *Policy Worksheet* for a list of items for the committee to address in its meetings.

If you are in the early stages of an IR project…

If you are just beginning, or in the early stages of building a repository, read this section to familiarise yourself with the whole landscape of legal issues and policies to consider. It will likely prompt questions you may want to use as a starting point in talks with your university's copyright expert (or legal counsel's office).

Next, use the checklist for policy development to set tasks for a Policy Committee of library, academic, and administrative staff to write policies for your repository. We include links to other universities' policies, which may help you to get started writing your own.

IP laws and customs vary from place to place, and from university to university, so you'll most likely want to customise the policies we include as examples.

If your IR project is underway already…

If you have already begun building an institutional repository, you may already have a list of policies or established practices, or at least a list of issues you need to address. Compare your list with those listed in the Policy Worksheet in this section to see if there are additional policies you need to create.

As they write a list of policies for your service, your Policy Committee may be interested to see how other universities address similar concerns. See the list of *Sample Policies* in this chapter.

You may also need to audit your existing procedures and licences to ensure your service complies with regulations and policies that affect your service.

Refer to the sections below on *copyright*, *licensing*, and *recent legislation* to for pointers to recent legislation and literature in these areas.

Outline: Legal and Regulatory Environment

Understanding Intellectual Property Rights (IPR) for Institutional Repositories
- Copyright
- Licensing
- Rights Management

Relevant Legislation
- 2003 Copyright Changes
- Legal Deposit
- Freedom of Information Act

Policy Guidelines for Institutional Repositories
- Policy Worksheet
- Sample Policies
- Memorandum of Understanding

Resources for Further Research
- Government Offices and Acts
- Intellectual Property Rights Guides
- Copyright and Licensing for Scholarship

Understanding Intellectual Property Rights (IPR) for Institutional Repositories

This section describes the primary **Intellectual Property Rights** (IPR) issues you may encounter when creating an institutional repository at your university, including copyright, content licensing, and rights management.

Intellectual Property Rights refer generally to the ownership and rights over work produced and distributed both online and in print.

Copyright and Content Licensing

Copyright offers protection to content creators to control how their material can be used and distributed. For details on the specifics of how copyright is defined and enforced, see the UK Patent Office's excellent website describing all the terms and conditions of copyright: http://www.patent.gov.uk/coy/indetail/basicfacts.htm. Understanding copyright issues is vital to the success of an institutional repository project. Your university's copyright officer is the best person to interpret how these laws affect your university and your institutional repository.

Institutional repositories deal with **copyright issues** on two fronts: in **collecting content** from scholars, by which they must secure the rights to distribute and preserve the content, and in **distributing content** to end users, by which they must balance the tenets of open access with copyright protection.

As you work with faculty who want to submit their content to an institutional repository, you might want to encourage them to **retain copyright** to their work or at least retain rights to publish their work electronically when they publish their papers. We offer details and guidelines below.

Content Licences

Content licences are the legal agreements by which content can be distributed. Typically, an institutional repository might have these two licences:

- **Deposit licence:** An agreement between the creator (or copyright holder) and the institution giving the repository the right to distribute and preserve the work.
- **Distribution licence:** An agreement between the author or creator or copyright holder and the end user governing the uses that can be made of the work.

Each university has unique copyright needs, as well as existing copyright agreements with its faculty. Your institutional repository's policies and licence

needs are unique. This will inform your choice of licences. Consult your institution's legal counsel or copyright officer to draft this licence to suit your service needs and concerns.

For sample deposit licences, see Gareth Knight's "Report on a deposit licence for E-prints" (2004) published through SHERPA (Microsoft Word format). http://www.sherpa.ac.uk/documents/D4-2_Report_on_Deposit_Licenses.doc

Creative Commons Licence

The Creative Commons Licences offer content creators and distributors a variety of licences, letting the content creator stipulate conditions for using the licenced content. See the Creative Commons site at http://creativecommons.org/ for information on the licences offered and tools for content creators/distributors.

The Creative Commons site also offers excellent background information on the legal concepts of fundamental intellectual property concepts: http://creativecommons.org/learn/legal/.

Copyright Guidelines for Scholars

Scholars who place their research in institutional repositories may need additional information on copyright issues. Several organisations provide excellent information and guides to understanding copyright for scholarly research.

The Creative Commons group offers important information on content licensing for faculty, researchers, and authors (http://creativecommons.org/learn/licenses/).

Publishers usually will agree to an author's request to retain rights to post content to a website or institutional repository. Faculty should be encouraged to retain these rights before and after publishing their work so they can contribute their content to online repositories.

Project RoMEO offers excellent guidance for scholars interested in self-archiving. Their website provides valuable information on negotiating content agreements with publishers, with a guide to how publishers commonly licence content from faculty.
http://www.lboro.ac.uk/departments/ls/disresearch/romeo/.

The EPrints project publishes extensive information and guidance on self-archiving and open archives, as well as a glossary of terms in this area (http://www.eprints.org/glossary/) and links to the most important sites for research http://www.eprints.org/self-faq/.

Rights Management

The terminology around managing copyright in the digital age can be somewhat confusing. This section describes how rights management can be implemented in an institutional repository.

Digital Rights Management generally refers to the software used to manage rights in an automated way. There are a number of research initiatives currently addressing these issues.

Rights Metadata projects address the means to represent rights information in metadata. Two notable standards have emerged: XRML, a private initiative, and ODRL, an open standard. Many open source projects use the open ODRL specification.

- XrML: http://www.xrml.org/
- ODRL: http://odrl.net/

For institutional repositories, **Rights Management** generally refers to how content is distributed in accordance with copyright rules and to indicate who owns the copyright for the content. Institutional repositories usually aim for open access. However, there may be instances where access needs to be restricted, such as information related to patentable materials.

The software platform you use to build your service may offer technology to manage access rights. Here is an example where the policies you set and the technology features you use need to work together. For more information on how different software systems handle access control, see the *Technology* chapter of this document.

At MIT, for example, the DSpace access control system allows restricted access to material related to technology patents and to other content that needs to be restricted for a period of time.

Recent Legislation

A number of recent legislation changes in the UK affect the way online digital repositories address intellectual property rights:

- Copyright Changes
- Legal Deposit
- Freedom of Information Act

We address each one here, pointing you to the acts themselves as well as additional resources to help determine how your projects may be affected.

Recent Copyright Changes

Changes in 2003 to the UK Copyright Law work to bring UK laws into closer harmony with EU copyright law. The most notable changes affect the definition of 'fair dealing' and libraries' ability to make copies for entities conducting commercial research. Such research is no longer covered as an exemption from copyright law.

For example, libraries are held responsible for checking whether a user's copy request is for commercial purposes or not.

How these changes affect your project depends on the population you serve and policies your library follows.

Copyright Resources

The Copyright Licensing Agency Ltd. (CLA) outlines the major changes to the law: http://www.cla.co.uk/directive/index.html.

The UK Patent Office offers guidance on complying with new copyright rules: http://www.patent.gov.uk/copy/notices/2002/guidance2.htm and http://www.patent.gov.uk/copy/notices/2003/copy_direct3.htm.

See the text of the new provisions here: http://www.legislation.hmso.gov.uk/si/si2003/20032498.htm.

Legal Deposit

In 2003, the UK's legal deposit law was extended to include electronic, non-printed materials, in the Legal Deposit Libraries Act 2003.

This act requires publishers to deposit with the British Library (and the other statutory legal deposit libraries), copies of all newly published electronic works. This has broad implications for how libraries store and manage digital

collections and provides further impetus for building and managing institutional repositories across the UK.

Legal Deposit Resources

See the text of the Legal Deposit Libraries Act 2003: http://www.hmso.gov.uk/acts/acts2003/20030028.htm.

Her Majesty's Stationery Office publishes a list of Frequently Asked Questions (FAQ) about the act: http://www.hmso.gov.uk/faqs.htm.

The British Library publishes its policies on Legal Deposit: http://www.bl.uk/about/policies/legaldeposit.html.

Freedom of Information Act

The Freedom of Information Act 2000 gives individuals to access to information held by public government agencies. There are numerous implications for libraries and other public record holders. It directly affects the need to handle and distribute institutional records efficiently and quickly.

In addition, all public authorities must adopt and maintain approved publication schemes.

Freedom of Information Resources

The text of the Freedom of Information Act 2000 is available here: http://www.hmso.gov.uk/acts/acts2000/20000036.htm.

Note that Scotland has its own Freedom of Information Act and Information Commissioner. For more information, see http://www.scotland.gov.uk/government/foi. The Freedom of Information (Scotland) Act 2003 is available online: http://www.hmso.gov.uk/legislation/scotland/acts2002/20020013.htm.

The Information Commissioner's Office sponsors a website discussing local authorities' response to the FOI: http://www.informationcommissioner.gov.uk/cms/DocumentUploads/FOI%20Survey%20Findings%20for%20Local%20Authorities.pdf.

The Lord Chancellor's site: http://www.dca.gov.uk/foi/foidpunit.htm.

The Guardian publishes a special section on the Freedom of Information Act: http://www.guardian.co.uk/freedom/0,2759,178243,00.html.

The UK government has issued guidance on how the passage of the Freedom of Information Act affects the Data Protection Act of 1998: http://www.dca.gov.uk/consult/foi/dpsaresp.htm.

The text of the Data Protection Act 1998 is available here: http://www.hmso.gov.uk/acts/acts1998/19980029.htm. Scotland does not have separate provisions for data protection and use those of the 1998 act.

Confer with Your University's Copyright Officer

To learn how these recent changes in legislation affect your university, locale, and library projects, you will want to consult with your university's copyright officer or an expert in the field of intellectual property law.

Remember that this seminar and associated materials do not constitute legal advice.

Policy Guidelines for Institutional Repositories

Each university that develops an online institutional repository needs to research and write policies and regulations for its collections. This section offers guidelines for crafting your institutional repository's unique policies, and includes the following topics:

- Creating Policy Guidelines
- Forming a Policy Advisory Group
- Issues to Consider
- Technology Implications
- Sample Policies
- Using a Memorandum of Understanding
- Worksheet: Creating Policies for an Institutional Repository

Creating Policies for Your Service

Because each locality and university has unique laws and customs, and because different functional roles have unique perspectives, each university that builds an institutional repository can gain efficiencies by convening a policy advisory group to examine its unique needs and create policies to govern its service.

There are three kinds of policies:

1. Policies that your project team can resolve internally – for example, a list of supported formats.
2. Policies related to library policies – such as collections or access to collections.
3. Policy decisions related to the university's policies – user authentication and identification, privacy policies, theses, etc.

Forming a Policy Advisory Group

Most successful institutional repository projects form a Policy Group to advise on all policy decisions. The Policy Group can help to determine your institutional repository's policies on content submission and distribution, privacy and licensing issues, and other policy guidelines.

The Policy Group's Role

The Policy Group generally makes decisions related to institutional repository services, standards, and functionality. When the group's decisions have significant financial, service, or public relations impact, the issues may be referred to the library's governing body, made up of senior library administrators, for endorsement.

Tip: Keep your university's copyright officer informed as you make decisions about legal issues and licences.

Makeup of the Group

Each university's needs and customs will determine the membership of its Policy Group. Because the group's decisions may have implications for staffing and cost models, it is most important that a wide group of decision makers be involved.

A typical Policy Advisory Group is a cross-functional team that might include the following library and university staff members:

- Associate Director of Technology
- Associate University Librarian for Public Services
- Head of Collection Management Services
- Head of Document Services
- Head of Archives
- Head of one Divisional Library
- Information System Manager
- User Support Manager
- Information Systems Manager
- University Press representative

Naturally, the makeup of your Policy Group may differ according to your needs.

Policy Issues to Consider

Policies fall into three broad categories:

- **Content** – formats, kinds of content you'll collect, etc.

- **Collections** – what constitutes a collection, how collections are managed and administered, if you're organising your content by department or clusters, etc.

- **Copyright** – intellectual property agreements and rights issues.

Policy Checklist

Your Policy Group may address some or all of the following issues:

- What **types of materials** will be accepted into the repository?
- **Whose work** can be included in the repository?
- What are the **categories of content** that need policy statements, such as theses, educational materials, etc.
- Are **Student Projects** accepted, or only faculty-created content?
- Criteria for determining **what constitutes a collection** in the repository. Who determines, sets, and authorises membership?
- How is your **repository structured** – around individual faculty or authors, or by department, research division, etc. Are collections of content built around an academic department or an individual?
- Do you have **contingency plans** if a department or research centre on which a collection is built, ceases to exist?
- **General rights and responsibilities** of libraries and those who create collections of digital content.
- **Content guidelines** for submission and organisation.
- **Privacy policy** for registered users of the system.
- **Theses**. Will you collect and preserve online theses? Who owns the copyright to the theses at your institution?
- **Fee vs. Free Access**. Are there areas of your system that you will charge users to access?
- **Restricting Access**. Patent issues may require that some theses be unavailable for a period of years. Other contributors may wish to limit access to their content.
- **Downtime**. What level of downtime is acceptable for your system?
- **Licensing**. You may need to consult your university's copyright officer on content licensing issues.

- **Preservation formats**. Which formats are supported, and to what degree?

- **Withdrawal of items**. Can items ever be deleted, or only hidden?

- **Metadata**. Who is authorised to enter metadata? Only library staff or faculty and contributors?

Assessing Your University's Existing Policies

As you begin to assemble policies for an institutional repository at your institution, you may need to reference your institution's existing policies with faculty. There may be existing policies on copyright ownership, for example.

Each university or institution has different policies for faculty intellectual property rights, based either on explicit, written policies or by tradition. For example, Cambridge's policies are different from Oxford's, which are different from Edinburgh's.

Technology Implications of Policy Decisions

Depending on how your digital repository is structured, there may be technology considerations for changing policies after you launch the system. Build your service flexibly to accommodate policy shifts where feasible.

For example, you may decide to organise your repository according to academic departments or research centres. To create a digital collection based on the work of an individual scholar may be feasible in your system, but at some stage, you may need to provide a way to aggregate individuals' collections into a group – if for example, a pioneering professor starts his/her own section in the repository, which will later be included in the department's collection.

Revisit Ongoing Policy Issues

The Policy Group's role is ongoing. They may need to revisit issues after the system launches and new issues arise. The Policy Group's decisions have important implications for how the service fits into the library's overall mission and the university's goals.

For example, several years ago, the notion of an "electronic signature" was not universally accepted. Now that it's gained wider acceptance, policies governing such signature would have to adjust to current practices.

As your repository grows, new needs will arise within the institution and new demands will be made of the service. Be prepared to change your policies as the repository service matures.

Sample Policies

Several universities have posted their policies online. It is instructive to see how they run their online services.

- Archive of European Integration
  http://aei.pitt.edu/

- California Digital Library (CDL) at the University of California
  http://repositories.cdlib.org/escholarship/policies.html

- DSpace at MIT
  http://libraries.mit.edu/dspace-mit/mit/policies/index.html

- Hong Kong University of Science & Technology
  http://library.ust.hk/info/repository.html

- National University of Ireland
  http://eprints.may.ie/faqs.html

- Open University of the Netherlands
  http://dspace.learningnetworks.org/index.jsp

- E-Print Repository at Queensland University of Technology
  http://www.qut.edu.au/admin/mopp/F/F_01_03.html

- University College London IPR policies
  http://www.ucl.ac.uk/Library/scholarly-communication/ipr.htm

- Theses Alive! project at Edinburgh University Library
  http://www.thesesalive.ac.uk/ta_submitters_faq.shtml

Using a Memorandum of Understanding

A **Memorandum of Understanding** is a document that describes the rights of parties involved in a transaction or agreement.

Some institutions use a Memorandum of Understanding to delineate the legal rights and responsibilities of each party who uses an institutional repository: the sponsoring institution, its content collections, authors, and/or end users.

These Agreements are Optional

It is not essential that you provide a memorandum of understanding for your university's institutional repository. Some universities choose to offer them to their content collections, departments, or faculty. Other universities determine that their institution's existing policies and agreements with faculty are sufficient for these purposes.

Example Text

The following example offers sample text for such a document. Your university's copyright officer will of course want to review and customise the text to suit your institution.

Memorandum of Understanding for an Institutional Repository

This Memorandum of Understanding is made by and between the _____ content collection of the UNIVERSITY NAME ("Collection") and the NAME OF INSTITUTIONAL RESPOSITORY of the UNIVERSITY NAME ("Provider") and is entered into this _____ day of _____, 2004.

The Collection acknowledges that it has read and is familiar with the following documents (which are appended hereto and are made a part hereof) and agrees to abide by the policies, terms and conditions therein:

- Content Guidelines
- Collection Policies
- Collection Startup Procedures
- Institutional Repository Distribution Licence
- Privacy Policy
- Format Support Policy

The Collection has appointed _____ to be the LIAISON for this collection with the understanding that this person will fulfil the following duties:

- Communicate with the Institutional Repository Director regarding content and technology issues on an as needed basis.
- Provide the information needed to set up and maintain collections.

The Collection agrees to provide confirmation information concerning the its status on a yearly basis as requested by the UNIVERSITY Libraries.

Provider agrees to distribute and preserve the collections entrusted to it by Collection and to provide the services described the Institutional Repository Collection Policies document referred to above.

Collection	**Provider**
Signature:*	Signature:
Title:	Title:
Date:	Date:

* Should be signed by head of content collection.

Policy Work Sheet

This section outlines the questions your Policy Group faces in developing the policies for your Institutional Repository. We present considerations to keep in mind and questions you need to answer as you create your service policies.

Your decisions are informed by the unique circumstances and regulations of your institution. We present a range of issues, and you can note below how you might handle each one.

- Content and Collection Policies
- Submission Process
- Copyright and Licences
- Metadata
- Privacy Policies
- Service Policies

Content and Collection Policies

How you organise and regulate content for your service will depend on the institution's culture as well as faculty requests and expectations. This section contains questions and guidelines to help in the process of crafting your policies.

Defining Collections

- How will your collections be organised?
  For example, will content be grouped by academic department, by subject, by type (theses, etc.)?

- What constitutes a collection?

- Who determines and authorises submitters?

- What are your contingency plans if a department ceases to exist?

Content Guidelines

- What types of content will the repository accept?
  - Technical Reports
  - Working Papers
  - Conference Papers
  - Preprints, "Postprints"
  - Books
  - Theses
  - Datasets

- Learning objects
- Digitised historical objects

- Who can submit content: faculty, staff, students, etc.?

- Must the work be education or research-oriented?

- Does the work have to be in digital form?

- Will the repository accept peer-reviewed content only, or is non-peer-reviewing content allowed?

- Does the work have to be in finished form, ready for distribution?

- Does the author or owner have to grant the service the right to preserve and distribute the content?

- If the work is part of a series, must other works in that series be contributed as well?

- Which document types will you accept?
  - Text
  - Images
  - Audio
  - Video

- Which document formats will you accept? Will you offer preservation for any of these formats?

MIME type	Description	Extensions
application/marc	MARC	marc, mrc
application/mathematica	Mathematica	ma
Application/msword	Microsoft Word	doc
application/octet-stream	Unknown	(anything not listed)
application/pdf	Adobe PDF	pdf
application/postscript	Postscript	ps, eps, ai
application/sgml	SGML	sgm, sgml
application/vnd.ms-excel	Microsoft Excel	xls
application/vnd.ms-powerpoint	Microsoft Powerpoint	ppt
application/vnd.ms-project	Microsoft Project	mpp, mpx, mpd
application/vnd.visio	Microsoft Visio	vsd
application/wordperfect5.1	WordPerfect	wpd

MIME type	Description	Extensions
application/x-dvi	TeXdvi	dvi
application/x-filemaker	FMP3	fm
application/x-latex	LateX	latex
application/x-photoshop	Photoshop	psd, pdd
application/x-tex	TeX	tex
audio/x-aiff	AIFF	aiff, aif, aifc
audio/basic	audio/basic	au, snd
audio/x-mpeg	MPEG Audio	mpa, abs, mpeg
audio/x-pn-realaudio	RealAudio	ra, ram
audio/x-wav	WAV	wav
image/gif	GIF	gif
image/jpeg	JPEG	jpeg, jpg
image/png	PNG	png
image/tiff	TIFF	tiff, tif
image/x-ms-bmp	BMP	bmp
image/x-photo-cd	Photo CD	pcd
text/html	HTML	html, htm
text/plain	Text	txt
text/richtext	Rich Text Format	rtf
text/xml	XML	xml
video/mpeg	MPEG	mpeg, mpg, mpe
video/quicktime	Video Quicktime	mov, qt

Submission Procedures

- Is there an approval process for content being submitted?
- Are submitters notified of an item's progress in the submission process?
- Are there content size limits for individual items, individual faculty members, or collections?

Intellectual Property Rights (IPR)

- Must content submitters own copyright for submitted content?

- What policies do you need for author permissions and licensing terms?
- Do you require copyright transfer for submitted items, or do you want only a non-exclusive right to distribute the work?
- Who is responsible for ensuring compliance with publisher copyright issues?
- At your university, who holds the intellectual property rights for faculty research, course materials, etc.?
- What are your existing intellectual property rights agreements with faculty?
- Who owns the copyright to theses at your university?

Note: As you develop content policies, you will want to consult your university's copyright officer on content licensing issues.

Metadata

- Which metadata standards will you use or support? (This may depend on the software platform you use.)
- Who is authorised to enter metadata?
  Only library staff or faculty and content contributors?
- Who determines if the metadata meets the service's quality standards?
- Who can correct mistakes or errors in metadata?
- Is there an approval process for metadata?

User and Privacy Policies

- Will you have a user agreement with end-users of the system?
- Will you institute a privacy policy for those who register with the system?

- Will you authenticate users of the system?
- Will you allow limited access to certain items?

Additional Service Policies

Preservation Formats

- Which formats are supported, and to what degree?

Withdrawal of Items

- Will you provide for withdrawal of items?
- Does withdrawal mean deletion?
- Are there circumstances that would warrant deletion of an item from the repository?
- If you allow withdrawal from public view (without deletion), how will you handle the metadata? Will there be some information for end-users saying that the item was withdrawn?
- The following are potential provisions for withdrawing items:
    - Removed from view at request of the author.
    - Removed from view at the university's discretion.
    - Removed from view at the library's discretion.
    - Removed from view by legal order.

General

- Define the general rights and responsibilities of libraries and collections in the service.

Paid Access

- Fee vs. Free Access? Are there areas of the system that users need to pay to access?

Backup and Recovery

- What level of downtime is acceptable for the system?
- How secure must content be?
- What guarantee do you offer to content submitters regarding backup and recovery?

Government Resources

These resources outline the current state of copyright law and other intellectual property issues.

UK Links

The **Copyright Licensing Agency** (CLA) offers news and regulatory guidelines, including an introduction to the New Copyright Law: http://www.cla.co.uk/.

The **UK Patent Office** presents a comprehensive set of links and information related to intellectual property issues: http://www.patent.gov.uk/links/index.htm.

The **UK Patent** Office also publishes a website about intellectual property rights: http://www.intellectual-property.gov.uk/.

The **Scottish Law site** provides comprehensive coverage of Scottish legal issues: http://www.scottishlaw.org.uk/.

Research the **Freedom of Information Act** (2000) at the **Information Commissioner's Office** site: http://www.informationcommissioner.gov.uk/eventual.aspx.

The **Library and Archives Copyright Alliance** (LACA) monitors and lobbies the government and the EU on issues of copyright law that affect their members: http://www.cilip.org.uk/committees/laca/laca.html. The LACA also publishes a very useful list of web links on these topics: http://www.cilip.org.uk/committees/laca/laca3.html.

The **Regional Development Agencies** in England aim to further regional economic development and efficiency. There are RDAs in each of in eight regional areas plus London:

- Government Office for London
  http://www.go-london.gov.uk
- Government Office for the East Midlands
  http://www.go-em.gov.uk
- Government Office for the East of England
  http://www.go-east.gov.uk

- Government Office for the North East
  http://www.go-ne.gov.uk
- Government Office for the North West
  http://www.go-nw.gov.uk
- Government Office for the South East
  http://www.go-se.gov.uk
- Government Office for the South West
  http://www.gosw.gov.uk
- Government Office for the West Midlands
  http://www.go-wm.gov.uk
- Government Office for Yorkshire and The Humber
  http://www.goyh.gov.uk

EU Links

The **European Bureau of Library, Information, and Documentation Associations**: http://www.eblida.org/ecup/.

For information on Irish intellectual property law, see the **Information & Communication Technology Law in Ireland** site at http://www.ictlaw.com/ip.htm.

The **Intellectual Property Rights Helpdesk** offers IP-related assistance for European researchers: http://www.ipr-helpdesk.org.

The **SURF** site from the Netherlands links to a variety of European sources in its **Copyright Management for Scholarship** website: http://www.surf.nl/copyright/info/resources.php.

US Links

The **United States Copyright Office** at the **Library of Congress** is the best central source for US copyright issues: http://www.loc.gov/copyright.

Intellectual Property Rights

Recent Publications

The House of Commons Science and Technology Committee recently published **Scientific Publications: Free for all?** – a survey of the state of scientific research publishing, with important findings and discussion of

institutional repositories. It advocates that all UK libraries establish institutional repositories and raises key issues for implementation and direction: http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf.

The Wellcome Trust published an influential report on open access, called **Costs and Business Models in Scientific Research Publishing.** It is available online at http://www.wellcome.ac.uk/doc_wtd003185.html.

IPR Guides

One of the best papers on Intellectual Property Rights is the **Cedars Guide to Intellectual Property Rights**, which explains all the main concepts and recent developments in IPR issues in the UK: http://www.leeds.ac.uk/cedars/guideto/ipr/guidetoipr.pdf.

The **CEDARS Bibliography, Copyright and Intellectual Property Rights Issues** by Catherine Seville lists hundreds of resources on these issues: http://cedars.bodley.ox.ac.uk/cedars/bibliog/CS.cfm.

The **Center for Intellectual Property and Copyright in the Digital Environment** (CIP) has up-to-date information, seminars, and mailing lists on associated topics: http://www.umuc.edu/distance/odell/cip/.

The **European Bureau of Library, Information and Documentation Associations** lobbies on behalf of library organisations and offers a wealth of resources at its site: http://www.eblida.org/.

To learn about digital rights management, see the **IEEE** paper, “Towards a Digital Rights Expression Language Standard for Learning Technology”: http://ltsc.ieee.org/meeting/200212/doc/DREL_White_paper.doc.

The **IP Mall** website lists dozens of intellectual property websites and resources and is one of the best sites on the Internet for researching IP issues: http://www.ipmall.info/fplchome.asp.

The **Joint Information Systems Committee (JISC)** published results from its May 2003 Intellectual Property Rights workshop: http://www.ariadne.ac.uk/issue36/iprws-rpt/.

Praxis offers an excellent collection of background information on intellectual property issues: http://www.praxistech.org.uk/links/bg_info_on_ip.asp.

The **World Intellectual Property Organization** (WIPO) is a specialised agency of the United Nations, dedicated to intellectual property rights: http://www.wipo.int/.

Rights Management

For information on **Metadata and Rights Management**, see the Open Archives Initiative's OAI-Rights programme at http://www.openarchives.org/news/oairightspress030929.html.

A **whitepaper** describing how rights should be described in the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a joint effort of staff from Project RoMEO and the Open Archives Initiatives group is available at http://www.openarchives.org/documents/OAIRightsWhitePaper.html.

Copyright, Licensing and Preservation for Scholarship

Dr. Theo Andrew published a valuable briefing paper through JISC on the intellectual property issues raised by putting electronic theses in an institutional repository: *Intellectual Property and Electronic Theses*. JISC Legal Information, (22 September 2004). http://www.jisclegal.ac.uk/publications/ethesesandrew.htm#author.

The **CAMiLEON** project develops and evaluates technologies and strategies for long term preservation of digital scholarship: http://www.si.umich.edu/CAMILEON/about/aboutcam.html.

Charles Oppenheim, "Information Ownership, Copyright and Licences" http://www.zbmed.de/eahil2002/proceedings/oppenheim-proc.pdf

Copy Own is a resource on copyright ownership for the higher education community: http://www.inform.umd.edu/copyown/.

Copyright Management for Scholarship: http://www.surf.nl/copyright/

Creative Commons group's information on content licensing for scholars: http://creativecommons.org/learn/licenses/. You can also download its freely available licence text and for general licensing information and related technology: http://creativecommons.org/.

The **Open Archives Initiative** (OAI) investigates issues and standards for interoperability among institutional repositories: http://www.openarchives.org/index.html.

Project RoMEO (Rights Metadata for Open Archiving) is an excellent resource for library staff building an institutional repository: http://www.lboro.ac.uk/departments/ls/disresearch/romeo/. It investigates the rights issues around 'self-archiving' of academic research in institutional repositories.

SHERPA investigates issues pertaining to scholarly communication and publishing: http://www.sherpa.ac.uk/index.html. Its site offers an excellent list of articles on related topics: http://www.sherpa.ac.uk/documents/.

The **JISC Legal Information Service** has a comprehensive list of links and resources posted about intellectual property rights: http://www.jisc.ac.uk/legal/index.cfm?name=lis_helpsites_iprights. They also offer a substantial **glossary** of terms: http://www.jisc.ac.uk/legal/index.cfm?name=lis_glossary.

The **SURF** site also publishes the **Zwolle Principles** and related information on copyright ownership and rights management: http://www.surf.nl/copyright/keyissues/scholarlycommunication/principles.php.

The **Zwolle Group's information on copyright agreements** is particularly helpful and explains copyright terminology very clearly: http://www.surf.nl/copyright/keyissues/scholarlycommunication/agreements.php.

Chapter 5: Guidelines for Cost Modelling for Institutional Repositories

Introduction

One of the first questions about an Institutional Repository programme is "How much will it cost?" Unfortunately, there is no one simple answer to this question. Clearly, it depends on the scope of your service requirements and the resources available to you. But we can help you to clarify the factors affecting your budget and to identify cost categories.

In this section, we describe the primary cost factors and issues to consider when building a budget or cost model for your institutional repository service. Your cost categories will depend on the size and scale of the service. We include cost modelling information for a variety of scenarios – for those building a small-scale service using existing staff, as well as more complex services.

While there is no set formula to determine how much it will cost to build a repository at your university, we offer tools to help you create a realistic cost model.

Who should read this section?

This section is geared toward senior managers and finance staff who will create staffing plans and budgets for your institutional repository service. But all senior staff must be familiar with these concepts and cost inputs for a successful service to account for all direct and indirect costs.

This way, senior staff will be aware of the options available and end-goals to keep in mind. For example, you may start off using existing staff, unable to hire new, dedicated resources. And, as the service grows, you may need to account for existing staff doing their jobs differently. Otherwise, you can miss the need to account for staff contributions to the repository.

Without budgeting properly for all resources involved in the service you may run the risk of underestimating total costs leaving the programme in a deficit, and thereby risking its success and sustainability.

Getting Started

As you read through this section, make note of the cost areas you need to address for your institution. Even if you are starting with a more basic model in mind, read through all the cost categories listed, so you will understand the whole landscape as your service grows into a more complex model.

You may start off using only existing staff, with no new hires for the service. Keep in mind the next, more complicated models – having existing staff do the same jobs in a different way, for example. The goal is to grow with the end-goal in mind, not haphazardly as staff pitch in their time and expertise to make the service a success.

Note

The guidelines presented here assume your institutional repository service will be established within an existing organisation as an additional service. The operational costs are measured from the perspective of that organisation and do not include costs for other departments (for example, the time authors spend submitting items to the repository). It captures the opportunity costs of foregone activities of existing staff by including costs for all staff whose time will be redirected significantly to work on the new service.

Strategies for Building an Institutional Repository

There are several operational models for running an institutional repository:

- **Build your own** using open source software.
  In this model, you download the software and manage the service in-house. Your staff do local customisations to the code and participate in an open source community of developers.

- Join other universities in a **consortium** to build and run institutional repositories on a central technology platform.
  In this model, a group of universities, typically organised by geographic proximity, share hardware with varying local customisation. Shared resources help to manage costs. For example, the White Rose Consortium has one user support manager for three universities.

- **Outsource** the tasks of building and running the hardware and software to a commercial service provider.
  In this model, an outside service provider hosts your content and may additionally provide other outsourced services such as user support, training, etc.

- Use a **hybrid** approach that combines some of the above methods.

Each method has its benefits – no one model works for everyone. After determining your service definition and the features you need to support it, you can choose the method or strategy that works best for you.

No Easy Answers

Note that whichever method you choose, there are no shortcuts or "turn key solutions" to building an institutional repository. You still need to design a service, apply the proper technology platform, create policies, recruit content communities, enlist faculty participation, and market the service to your users.

Each Unique Service Has Unique Costs

Each university's service is unique. Each unique service has unique costs. These costs will change over time – an early stage project faces different costs than a mature service running at capacity.

Aspects that affect your service's **direct costs** include the following:

1. Content
2. User communities
3. Existing resources
4. Service size and scale
5. Service maturity – startup, growth, and maturity phase each bear different costs.

The service maturity has a tremendous impact on costs. Most notably, support activities decrease over time. For example, over time, community set up will be a smaller and smaller part of the User Support Manager's role.

Your service's **indirect costs** ought to account for these items:

- Strategic planning – ongoing development of the service may or may not be directly related to feature development in the system.
- Support staff – the need for library staff to work as faculty liaison may also develop over time.

Rather than using a formula, or relying on any one else's budget model, you will want to develop your own cost model.

Cost by Activities

One method to account for the costs of building an institutional repository is to factor costs based on activities you and your staff need to perform to implement the service. In this model, you might account for staffing not by individual head counts, but by activities those staff members perform:

- Marketing
- System support
- Content acquisition
- Training, and so on

See the work sheet in this chapter titled *Separating Costs by Activity Categories* to learn how to account for the activities needed to build a service.

How Technology Choices Impact Costs

The technologies, hardware, and software you use to build an institutional repository are a significant cost factor for your service. Whether you build the service in house or buy software and services from an outside vendor, the cost of building and maintaining a complex service will be central to your cost model.

In the *Technology* chapter of this book, we outline the various software platforms you can use to build an institutional repository.

Whether you use open source software (such as eprints or DSpace), buy commercial software, or contract out for software and services, you still need to develop a service definition, create policies, market the service at your university, work with faculty and submitters, maintain quality control, and do long-range planning for the system.

For open source software, you may need to customise the user interface and provide other programming and software development efforts to set up and run the software.

Example: Contracting for Software and Services

For example, if you contract for software and services, you may get the following services:

- System support and maintenance
- System equipment
- Software systems
- Technical training for library staff and content submitters
- Batch ingest and bulk loading
- Metadata creation
- Technical support

However, depending on the services offered and the pricing model, you still need to perform the following services in-house:

- Strategic planning
- Marketing:
    - Developing marketing materials
    - Marketing activities, in house and external
    - Public relations and communications
- Policy Development

Additional Considerations

As you consider whether to build an institutional repository in house or contract out for development and maintenance, note that there are some tradeoffs:

- Library visibility in the community
- Loss of customisation
- Exit strategy:
  - How will the content be managed
  - Switching costs may be high

Library Visibility

Libraries have been struggling to maintain the awareness of members of the university community. With a digital library, users working online from home, or another off-campus location, may not even realise that the library provides the service they are using.

Outsourcing services such as marketing your institutional repository, or content hosting may have the same effect.

Institutional repositories can be used to showcase the library on campus as much as they can for showcasing the university's research to the world.

Loss of Customisation

With collaboration or outsourcing, you may lose the ability to make the best decisions for your community. Only you can decide the costs and benefits to those options.

Exit Strategy

At the start of any major technology project it is wise to consider a potential exit strategy should the technology platform, or the project itself, fail to reach maturity.

For example, when you choose a software platform, you might consider the potential costs and complexity should you need to migrate or close your service at some point. If you choose a software service provider, you might consider strategies in the event your service provider ceases operation or shifts its direction or business operations away from the current platform.

Although many are familiar with moving content, such as a library catalogue, from one provider to another, or from one version to an upgrade, there is an additional consideration with institutional repositories. If in the future you opt to move to an open source software platform, you may not want to be on that technology learning curve with a fully operational, mature service.

Benefits to Developing a Repository On-Site

While some universities may benefit from outsourcing the development of their repositories, for others, the level of customisation they require – and the presence of in-house programmers – may lead them to develop a repository on-site using either open source or commercial software systems.

Among the benefits of developing an IR in-house include the following:

- Increased visibility to the Library in the university community
- Complete customisation of policies and user interface
- Responsiveness to local user needs and preferences
- Increased contact with constituents – your team helps them build online collections/communities, not outside consultants
- Continuity of development, management, maintenance
- Control of content
- Ownership of system

The development process reaches beyond creating and maintaining the software system itself, including associated training, planning and customising for the institutional repository service.

Budget Inputs

It is important first to define your institutional repository's service model in order to estimate costs, although this may necessarily be an iterative process. You may start out assuming "We will do only what we can afford to do." Then, as service requirements and staff commitments grow, you may need to seek additional funding. It may be impossible to plan a service without knowing how much it will cost. As with other planning processes, creating a cost model is an iterative process.

Specific data that you need to gather include the following cost categories:

- Staffing
- Overhead
- IT Systems: Hardware and Software
- Services

Staffing

Staffing can be the single biggest cost of running an institutional repository.

We segment staff categories, separating those directly responsible for the provision of the new service and those whose involvement is significant to the service. In this approach we consider all of the staff that will experience any significant change in the way that they perform their daily routines due to the implementation of the new service. This includes the need to be trained on

and stay abreast of changes in the institutional repository in order to serve their primary constituents, be they end-user researchers or the faculty who are submitting their works. Capturing these costs allows us to fully account for all of the costs associated with the service, and to understand how they may change over time as the system grows and the impact on staff changes. It also provides a mechanism to measure the contribution that the host organisation brings to the new enterprise. We will talk more about that in the section on Budget Impact.

Staff may logically be categorised as having direct responsibility for the daily operation of the service, having a job that will significantly be altered because of the new service, or no impact. For staff that will have the primary responsibility for the service, both the number of staff and the level of commitment (full or part-time) will be defined by the service model. For others for whom you want to capture the opportunity cost of foregone activities, interviews may be a more appropriate method to gather the data for allocation. Support staff may be considered separately or as part of an estimated overhead multiplier that would include office expenses as well.

Position	Allocation	Salary	Benefits Rate	Total Costs
User Support Manager	100%	£50,000	20%	£60,000
Librarian	20%	£40,000	20%	£9,600
Senior Manager	5%	£60,000	20%	£3,600
Total Staff Cost				£73,200

Resource

See the Service Planning chapter worksheets on *Staffing* to track roles and skills you need to add to the service team.

Overhead or Indirect Costs

Once all your staff costs have been identified, you can begin to consider the *indirect costs* associated with their employment. These costs would include office space, equipment, human resources staff time, etc. Some organisations may have already developed an overhead rate that is routinely applied in the budget process and that is of course a suitable substitute in many cases. Some costs to keep in mind include ongoing costs as well as one-time expenses:

- Office space
- Utilities
- Supplies
- Professional development expense
- Training materials
- Marketing materials

System Equipment

The systems and equipment costs you might face in building an institutional repository service include the following:

- Software costs (one-time and ongoing)
- Hardware, servers, etc.
- Operations staff
- Programming staff (if necessary)
- Backup and recovery
- Preservation

The cost of the equipment can meet or exceed the staff costs. This will naturally depend on many factors. What resources already exist with which to start a pilot programme, how the pilot is scoped, and how quickly content is submitted will affect costs.

It is important to have a scaling strategy in place if new equipment will be purchased. To properly account for this category it is likely that you will want to think in terms of an escrow account for system equipment purchases over time in order to capture the full annual cost of the service.[1]

The main input for this section of the cost model is the service definition and associated requirements you determined in the Service Planning chapter, which you can use to develop a Request for Proposals (RFP) from the IT experts within your organisation. Once existing resources are evaluated, then you can approach outside vendors for additional purchases or estimates of future costs.

[1] Remember that storage costs have been falling dramatically over time and that trend is expected to continue.

Services

Finally, it may be necessary or advantageous to contract with other services on campus for some support aspects of the service. For example, the server on which the repository is run may reside outside of the library in an information services or campus computing department. In that case, it is likely that there would be a fee charged for the space and some other support such as tape backups, or system monitoring. Normally, this will vary across institutions.

Cost-Recovery Services

To balance the costs of running an institutional repository, some services offer premium services on a cost-recovery basis. In this model, basic services of setting up a content collection might be free for users, but adding additional storage or metadata services might cost extra.

Your service may offer some or most of these premium services – this depends on how you structure your institutional repository service and what are the needs of your community. In the example shown below, fees are charged to content communities that contract with the library to provide additional services.

Institutional Repository Services	
Core Services (free)	**Premium Services (fee-based)**
Setting up academic departments and other content communities in the institutional repository	Extra storage space
Metadata Services: Consultation	Metadata Services: Custom metadata creation
Training and user support for content submitters	Document services • Scanning • OCR • Reformatting files
Storage space allocation: basic	Extra storage space
• Batch import of data: • Historic collections • Newly digitised collections	
IT Systems management	

Reference

The University of Rochester's institutional repository offers a variety of core, or free, services, along with premium services to recover costs. See University of Rochester DSpace service, enumeration of core vs. premium services: http://www.library.rochester.edu/index.cfm?PAGE=1362 for more information.

Budget Impact

Understanding the cost of managing an Institutional Repository is important, but often what you really seek to understand is how much additional funding will be necessary to make it work today.

Another way to think about the costs is to categorise them according to how they will actually affect the budget of the host organisation. Obviously, the opportunity cost of foregone activities by staff is not a cost that will need to be funded directly but rather seeks to capture the true cost of running the system including hidden costs.

Similarly, unless there are a large number of new staff hired to run the service administrative costs will not be affected. At MIT, for example, the team classified costs in distinct categories. The **incremental costs** were essential to delivering the service in the way the team envisioned it at MIT. **Principal costs** are those which are considered enhancements above a bare-bones implementation, and so include things such as travel and professional development, as well as additional time spent by different members of the library staff to increase understanding and awareness of the service. Comprehensive costs are meant to capture all of the costs associated with contributions from existing staff and to demonstrate the value associated with building the service within an existing organisation such as the library.

Example: Types of Costs

The table below gives examples of the types of costs that might be classified in this way. Of course, what one organisation may see as a luxury, another may see as an absolute necessity, so these are merely suggestions. The point is to be able to determine exactly what additional funding will be required to run the service as defined by each organisation.

Incremental	Principal	Comprehensive
New staff within host organisation	Redirected staff	All staff affected
Office space, supplies and equipment [2]	Administrative support, travel, professional development	
System equipment purchase	Additional storage capacity	
Support fees paid to other departments		
Total Needed to Fund Basic Service	**Total Needed to Fund Enhanced Service**	**No Affect to Existing Budget**

Budgeting Over Time

As you determine your cost model, you may need to revise the model as you move through the following phases for the service:

- Start-Up
- Growth
- Maturity

Completing at least a skeletal plan for each stage can help manage the costs over time – for example by avoiding the purchase of new equipment that will be rendered inadequate long before it is obsolete. The cost model should be viewed as a work in progress with regularly scheduled reviews to track the validity of the growth scenario.

Cost Outlook

Several factors may influence the costs of running an institutional repository over the coming years. Institutional repositories are still in the early stages of

[2] An overhead rate could be applied if available.

development and adoption. As they become more common, standards will develop for service development, policies, digital preservation, and content development. As standards develop, it may be come easier to predict associated costs.

For example, over the past two years, we have gained a better understanding of the many uses for institutional repositories, such as learning object repositories or the role of institutional repositories in managing theses and dissertations. In the case of theses, this has led to a better understanding of the associated IPR issues The institutions that have worked through them are sharing their findings with the community.

Resource

Dr. Theo Andrew published a valuable briefing paper through JISC on the intellectual property issues raised by putting electronic theses in an institutional repository: *Intellectual Property and Electronic Theses*. JISC Legal Information, (22 September 2004). http://www.jisclegal.ac.uk/publications/ethesesandrew.htm#author.

Work Sheet: Separating Costs by Activity Categories

Use this work sheet to separate according to activities you may perform in house or by contract with an outside consultant or agency.

Activities or tasks	Internal Staffing Costs			Outsource	
	% of effort	Salary or pay level	Costs	Provider	Costs
System support: ▪ 24/7					
System maintenance					
Software upgrades					
User interface development					
Marketing activities:					
▪ Internal					
▪ External					
Marketing materials:					
▪ Internal					
▪ External					
Support:					
▪ Library staff					
▪ Submitters					
▪ End users					
Technical support					
Metadata help					
Help develop/market content communities					
Develop and maintain website					
Training					
▪ Library staff					
▪ Submitters					
▪ End users					
▪ User surveys					
Travel					
Professional development					
Fundraising					
Admin. support					

Work Sheet: Tracking Revenue for Services

You may decide to generate revenue for your institutional repository by charging a fee for certain activities. For example, you may decide to charge a fee for library staff assisting in metadata creation. Use this table to identify potential sources to provide each service – whether among your staff or from outside contractors or service providers.

Activities or tasks	Potential revenue item in fee-based service	Potential sources for this service
System support: ▪ **24/7**		
System maintenance		
System equipment		
Software upgrades		
User interface development		
Marketing activities:		
▪ **Internal**		
▪ **External**		
Marketing materials:		
▪ **Internal**		
▪ **External**		
Support:		
▪ **Library staff**		
▪ **Submitters**		
▪ **End users**		
Technical support		
Metadata help		
Help develop/market content communities		
Develop and maintain website		
Training		
▪ **Library staff**		
▪ **Submitters**		
▪ **End users**		
Develop policies		
Batch loading		
Strategic planning		
▪ **New service development**		

▪ **User surveys**		
Travel		
Professional development		
Fundraising		
Admin. support		

Work Sheet: Budgeting for an Institutional Repository

The worksheets you completed in earlier chapters of this book will determine the parameters of your cost model. The answers to these questions depend in large part on you define your service in the Service Definition worksheet, the Service Policies you create, and the Staffing model you plan.

Key Questions for Cost Modelling

You may find it helpful to answer as many of these questions as possible when preparing your cost model.

1. Will we need additional administrative assistance to support new personnel?

2. Will we need specialised space or equipment to support the programme?

    a. Help line phone capabilities

    b. Special power supply/security for server room

    c. Special equipment to alert if server fails

3. If yes, what will be the installation availability and costs?

4. What pay grade is necessary to support specialised skills of new employees?

5. What training does existing staff need?

6. What impact will the programme have on existing library staff?

7. What impact will the programme have on existing support staff?

8. How can we account for overhead?

9. If we are to offer for fee services – what are our revenue projections for those services?

10. What changes would we need to make if the service grows faster than we expect?

11. How can we reduce costs if the service grows more slowly that we expect?

12. How can we account for preservation expenses?

13. What activities or services are we willing to forego in order to support this service with existing staff?

14. What costs can we share with other universities?

15. What costs can we share with other university communities or departments?

16. What outside resources/services might we want to contract for?

For Collaborators:

17. What costs will we incur in the event that our collaborators decide to discontinue the relationship?

18. What costs will we incur if we decide to "go it alone?"

For Commercial Services

19. What costs will we incur if the service is changed or discontinued?

20. What costs will we incur if we are forced to suffer a break in service while we transition?

Resources

Andrews, Dr. Theo. “Intellectual Property and Electronic Theses.” *JISC Legal Information* (22 September 2004).
http://www.jisclegal.ac.uk/publications/ethesesandrew.htm#author

Barton, Mary R. and Julie Walker. “Building a Business Plan for DSpace, MIT Libraries’ Digital Institutional Repository.” *Journal of Digital Information* (2003).
http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Barton/

The House of Commons Science and Technology Committee, *Scientific Publications: Free for all?*
http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf

PADI: Preserving Access to Digital Information at the National Library of Australia – Counting the Costs of Digital Preservation: Is Repository Storage Affordable by Stephen Chapman
http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Chapman/chapman-final.pdf

PALS (Publisher and Library/Learning Solutions), report “Pathfinder Report on Web Based Repositories,” page 24 on budgets.
http://www.palsgroup.org.uk/

QSpace – Queens University Repository Project Plan,
Contains some v. useful info on budgeting, sample project costs table, last page:
http://library.queensu.ca/webir/planning/q_space_planning_document.htm

Rogers, Sally *A. Developing an Institutional Knowledge Bank at Ohio State University: From Concept to Action Plan*, 2003.
https://dspace.lib.ohio-state.edu/retrieve/335/KBRogers.pdf

University of Rochester DSpace service, enumeration of core vs. premium services: http://www.library.rochester.edu/index.cfm?PAGE=1362.

Washington State University. *Cost Modelling for Library Digitization Projects*
http://digitalwa.statelib.wa.gov/newsite/projectmgmt/costfactors.htm

Wellcome Trust, Costs and Business Models in Scientific Research Publishing:
http://www.wellcome.ac.uk/doc_wtd003185.html

