



# DSpace: Planning and Design of an Institutional Repository



**III International Conference on  
University Libraries**

Margret Branschofsky  
MIT Libraries  
Oct. 28, 2004



# Scholarly Communication in Transition

# Advances in Technology

- Methods of communication and distribution no longer restricted to formal publishing enterprises
- In some disciplines Online Discussion Lists are the primary method of exchanging ideas
- Author websites, Departmental websites

# Experimentation within Disciplines

- Preprint archives
  - “arXiv” - physics, mathematics, computer science, & quantitative biology
- E-journals launched independently of publishers
- New journal formats - interactive features

# Variety of digital research and educational materials

- Bio-medical images produced by instruments
- Acoustical soundings of the ocean floor
- Video simulations of physical phenomena
- Computer programs
- Statistical data sets
- Geo-spatial data sets

# Purpose of Scholarly Communication

- *Build archive* for future generations of scholars
- *Responsibility* to maintain archive
  - Publishers, Scholarly Societies, Libraries?
- *Obligation* to share with other researchers
  - Recent reports from funding agencies in the U.S. (NIH) and U.K. (House of Commons Scientific Technology Committee) have proposed requirements to make publicly funded research data openly accessible

# Libraries Can no Longer Guarantee Access to E-journals

- Access limited by license agreement
  - Who
  - Length of contract
- Means to access content (system) is controlled by publishers
- Publishers may stop service, go out of business

# Economics of Scholarly Communication

- Increased awareness due to rising costs
- Library organizations supporting alternative models (SPARC)
- Academic authors increasingly aware of power to promote change
  - Role in the publication process (editorial work, peer review)
  - Copyright – no longer automatic transfer of all rights to publishers
- Open access movement – some success

# DSpace Conceived to Address Problems

- Enables libraries to take active role in distributing & preserving work produced by their institutions
- Architecture designed to serve needs of a multidisciplinary organization

# Institutional Repositories

- Institutionally-based
- Scholarly material in digital formats
- Cumulative and perpetual
- Open and interoperable

# Possible Content

- Preprints, eprints
- Published books, articles
- Technical Reports
- Working Papers
- Conference Papers
- E-theses
- Images (visual, scientific, etc.)
- Audio/Video files
- Datasets
  - (statistical, geospatial, scientific)
- Databases
- Websites
- Learning Objects
- Digitized library collections

# DSPACE

# DSpace is...

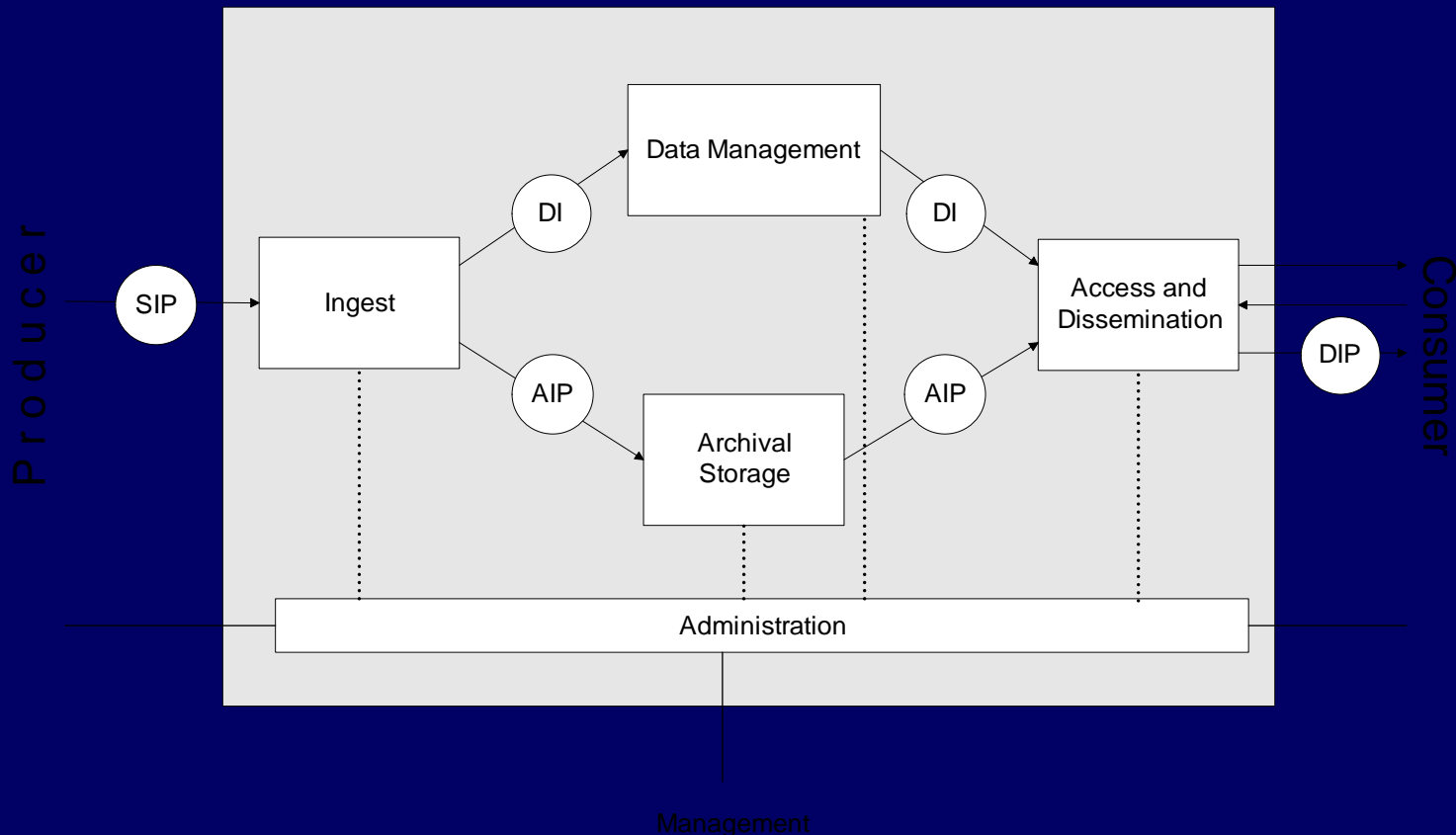
- An open source **digital asset management system**
- A technology platform for **Institutional Repositories**
- A **federation** of digital repositories across multiple academic research institutions
- A **production service** of MIT Libraries to its local research community

# DSpace

- **Captures**
  - Digital research material in any formats directly from creators
- **Describes**
  - Descriptive, technical, rights metadata
  - Assigns persistent identifiers
- **Distributes**
  - Searches metadata
  - Delivers via Web, with necessary access control
- **Preserves**
  - Large-scale, stable, managed long-term storage

# DESIGN

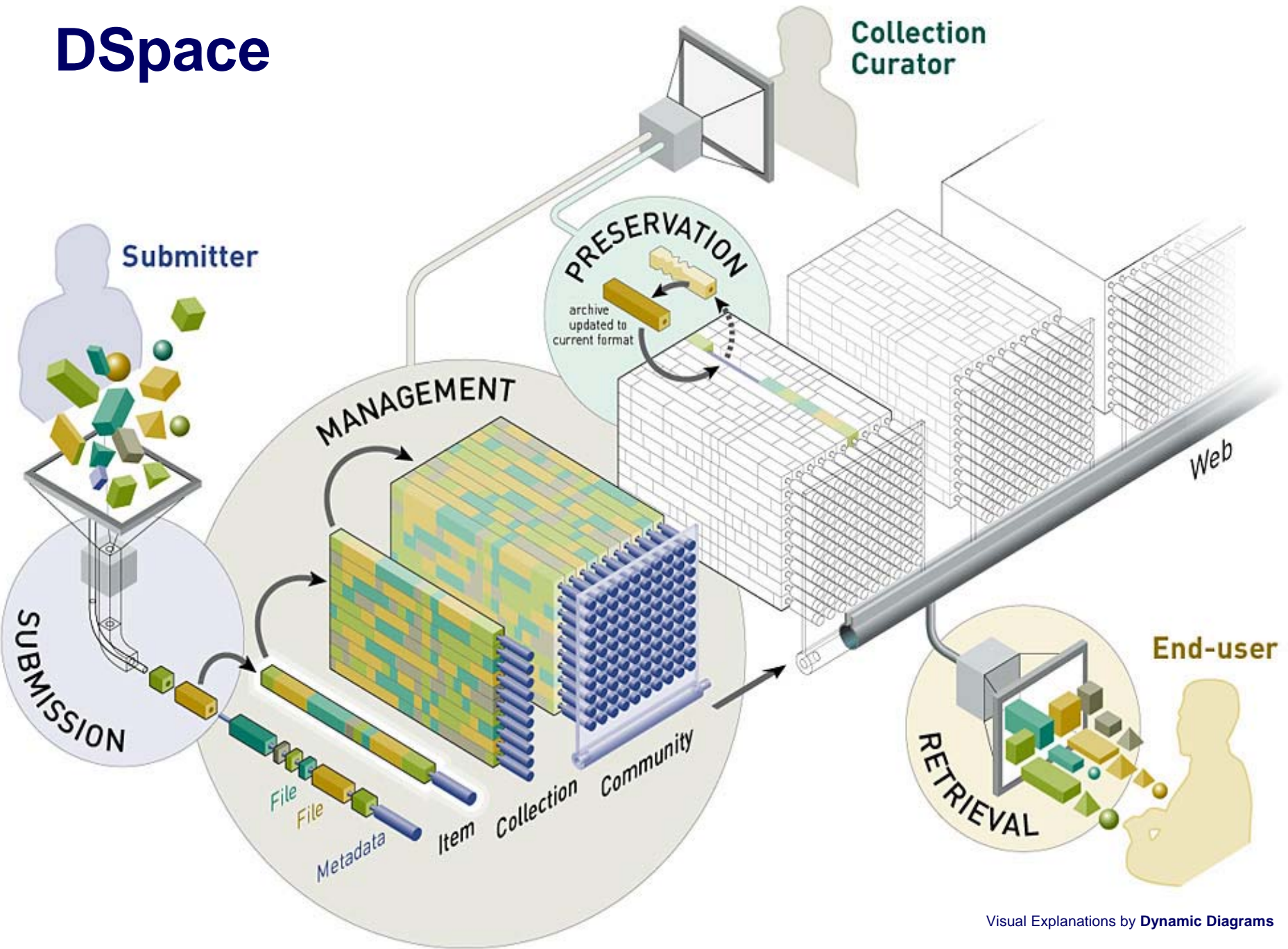
# Open Archive Information System (OAIS) Reference Model



# Information Model

- **Communities & Sub-communities**
  - Research units of the organization
  - Accommodate different policies
- **Collections**
  - Arbitrary groupings of items
  - Different workflow options for submission
- **Items**
  - Logical units of content
  - Receive persistent identifiers
- **Bitstreams**
  - Individual digital files
  - Receive preservation treatment

# DSpace



**Search DSpace:**   
[Advanced Search](#)[→ Home](#)**Browse**

- [Communities & Collections](#)
- [Titles](#)
- [Authors](#)
- [By Date](#)

**Sign on to:**

- [Receive email updates](#)
- [My DSpace](#)  
authorized users
- [Edit Profile](#)
- [Help](#)
- [About DSpace](#)

DSpace at MIT &gt;

**DSPACE VERSION 1.1**

New features include Advanced Search capability with fielded searching, more powerful keyword searching, improved performance and bug fixes.

[More news...](#)**Search**

Enter some text in the box below to search DSpace.

 **Communities in DSpace**

Select a community to browse its collections.

- [Center for Global Change Science](#)
- [Center for Innovation in Product Development \(CIPD\)](#)
- [Center for Technology, Policy, and Industrial Development \(CTPID\)](#)
- [Department of Ocean Engineering](#)
- [Laboratory for Information and Decision Systems \(LIDS\)](#)
- [MIT Press](#)
- [Singapore-MIT Alliance \(SMA\)](#)
- [Sloan School of Management](#)

**What can you find in DSpace?**

MIT Research in digital form, including preprints, technical reports, working papers, conference papers, images, and more.

**Is this all of MIT's research?**

No. DSpace is limited to digital research products. For items in print, go to [Barton: MIT Libraries' catalog](#). DSpace is young and growing rapidly. Check back often.

Search DSpace: [DSpace at MIT >](#)

   
[Advanced Search](#)

[Home](#)

**Browse**

- [Communities & Collections](#)
- [Titles](#)
- [Authors](#)
- [By Date](#)

**Sign on to:**

- [Receive email updates](#)
- [My DSpace](#)  
authorized users
- [Edit Profile](#)
- [Help](#)
- [About DSpace](#)

## Center for Global Change Science

Community home page



In:

Search for:

or browse

The Center for Global Change Science (CGCS) seeks to better understand the fundamental processes and mechanisms controlling the global environment. The interdisciplinary Center, founded in January, 1990, utilizes theory and observations of oceanic, atmospheric, and terrestrial systems, to improve the ability to accurately predict global environmental change. Two major activities of the CGCS are the Climate Modeling Initiative, and the Joint Program on the Science and Policy of Global Change. Visit the [CGCS website](#) for more information.

### Collections in this community

- [Joint Program on the Science and Policy of Global Change Reports](#)

Report Series intended to communicate research results, and provide useful reviews and commentaries on various aspects of the global climate change issue.

### Recent Submissions

[Induced technical change and the cost of climate policy](#)

[Uncertainty in climate change policy analysis](#)

[Description and validation of the MIT version of GISS 2-D model](#)

[Responses of primary production and total carbon storage to changes in climate and atmospheric CO<sub>2</sub> concentration](#)

[Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model](#)

### CGCS Website

Visit the CGCS home pages for more information:

- [Joint Program on the Science and Policy of Global Change](#)

**Search DSpace:** [Advanced Search](#)[Home](#)**Browse**[Communities & Collections](#)[Titles](#)[Authors](#)[By Date](#)**Sign on to:**[Receive email updates](#)[My DSpace](#)  
authorized users[Edit Profile](#)[Help](#)[About DSpace](#)[DSpace at MIT >](#)  
[Center for Global Change Science >](#)

## Joint Program on the Science and Policy of Global Change Reports

Collection home page



In:

Search for

or browse

Subscribe to this collection to receive daily e-mail notification of new additions

The Joint Program on the Science and Policy of Global Change is an interdisciplinary organization that conducts research, independent policy analysis, and public communication on issues of global environmental change. Established in 1991 as a shared activity of the Center for Global Change Science (CGCS) and the Center for Energy and Environmental Policy Research (CEEPR), it brings together scientific, economic, and policy aspects of global change issues, with a particular focus on climate. Visit the [Joint Program's website](#) for more information.

**Recent Submissions**[Induced technical change and the cost of climate policy](#)[Uncertainty in climate change policy analysis](#)[Description and validation of the MIT version of GISS 2-D model](#)[Responses of primary production and total carbon storage to changes in climate and atmospheric CO<sub>2</sub> concentration](#)[Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model](#)[Global Change Joint Program Website](#)[CGCS Website](#)[CEEPR Website](#)



# TECHNOLOGY

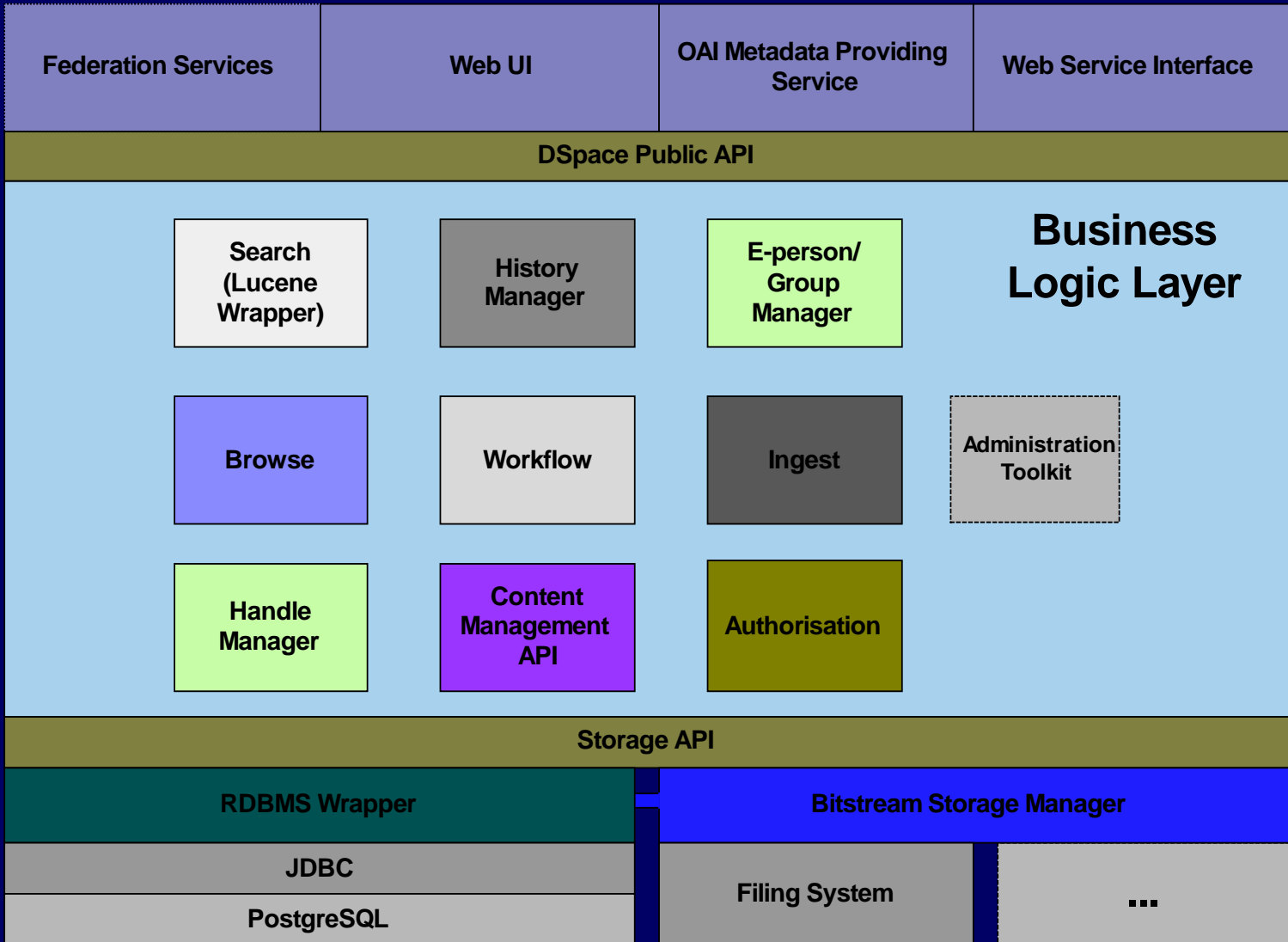


# Open Source Software

- Code available to all, free of charge
  - Download from SourceForge.net
  - Over 100 academic institutions using/testing DSpace
- No formal support from HP or MIT
  - Helpful discussion list – [dspace-tech.mit.edu](mailto:dspace-tech.mit.edu)
- Shared responsibility for software enhancement and evolution
- Safety in numbers (e.g. large community of adopters with vested interest)
- Governance structure in development

# Standards-based

- Standards Compliant
- OpenURL
- OAI-PMH
- Z39.50 and SRW
- CNRI Handles – persistent identifiers
- X.509 digital certificates for access control



# Open Software

- Java (hardware and OS platform independent)
- Apache, Thomcat
- PostgreSQL database manager
- Lucene search engine



# CHALLENGES

# Challenges

- Faculty Acceptance
- Policies
- Sustainability
- Digital Preservation


# Faculty Acceptance

- Build trust by *adding value*
  - Google indexing, preservation
- Accept variety of content
  - preprints and publications, digital research material, educational material
- Respect discipline differences
  - access control, review process, copyright and IP attitudes
- Develop institutional support
  - broad advocacy, mission relevance



# Institutional Repositories

*unlike traditional libraries*

- Not necessarily “published” material
  - Institutionally-based selection criteria
  - Shared selection, processing responsibility
  - Shared curatorial control
- 

# Institutional Repositories

*unlike traditional archives*

- Acquisition at point of creation
- Shared selection, processing responsibility
- Shared curatorial control
- Includes non-document material

# Policy Definition

- *New services need new policies*
- *Institutional*
  - Intellectual property, privacy
  - Ownership of theses, research articles, teaching material, etc.
- *Library*
  - Collections, e.g. student work? teaching material?
  - Access, e.g. restrictions? fees?
  - Use of Creative Commons licenses

# Sustainability

- Startup funded with grant money
- Business plan for MIT
  - Built operating cost models
  - Developed revenue options
- Unique to every organization
  - Tied to service model, local resources



# DIGITAL PRESERVATION



# Digital Preservation

- Repositories don't "do" preservation
- Preservation operations are defined by
  - Digital collections in hand
  - Cost/benefit tradeoffs
  - Local policy

# Digital Preservation

- MIT Philosophy
  - Lots of digital material *is already lost*
  - Most digital material is *at risk*
  - Preserving bits better than nothing
  - Capture as much information as possible
  - Evaluate cost/benefit tradeoffs over time

# Digital Preservation

## ■ *Supported*

- Provides for future content usability
- Migration for texts, images, audio, etc.
- Emulation for software, multimedia, etc.

## ■ *Unsupported*

- Bit preservation at minimum
- Migration when possible  
e.g. commercial conversion services

# Digital Preservation

## ■ MIT Policy

### ■ *Supported* formats

- e.g. TIFF, SGML/XML, PDF

### ■ *Known/unsupported* formats

- e.g. Microsoft Word, PowerPoint (common)
- e.g. Lotus 1-2-3, Visicalc, WordStar (less common)

### ■ *Unknown/unsupported*

- Highly complex and rare formats
- e.g. one-of-a-kind software programs

# Digital Preservation

- Emerging international preservation infrastructure, e.g.

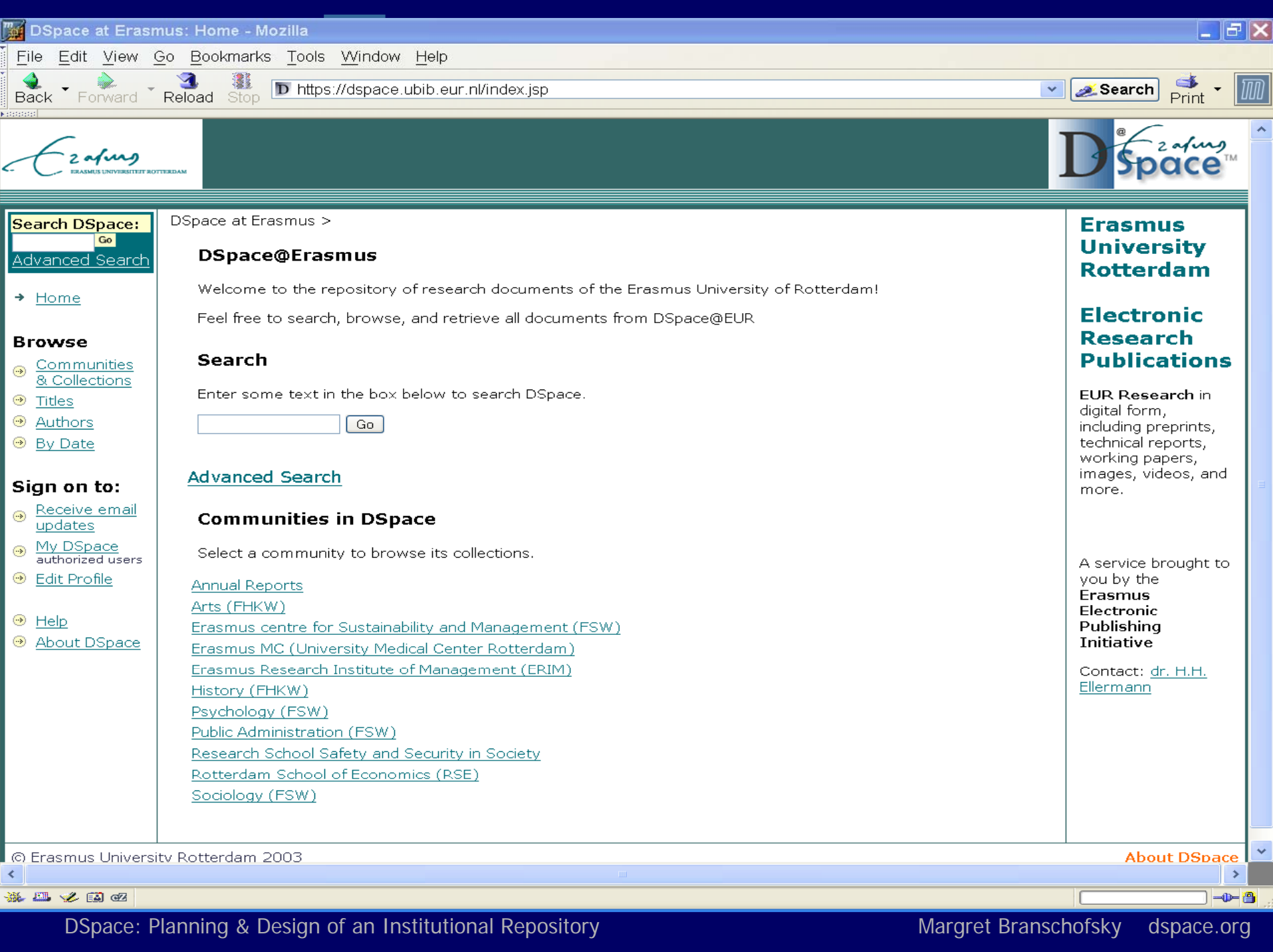
Global Digital Format Registry

UK Digital Curation Centre

UK National Archives PRONOM

# Conclusion

- Digital Libraries require stable infrastructure
- Commercial products are expensive, functionally lacking
- Open source software has great potential, presents challenges for our community



**Search DSpace:**  
  
Go  
[Advanced Search](#)

DSpace at Erasmus >

### DSpace@Erasmus

Welcome to the repository of research documents of the Erasmus University of Rotterdam!  
Feel free to search, browse, and retrieve all documents from DSpace@EUR

### Search

Enter some text in the box below to search DSpace.

### [Advanced Search](#)

### Communities in DSpace

Select a community to browse its collections.

- [Annual Reports](#)
- [Arts \(FHKW\)](#)
- [Erasmus centre for Sustainability and Management \(FSW\)](#)
- [Erasmus MC \(University Medical Center Rotterdam\)](#)
- [Erasmus Research Institute of Management \(ERIM\)](#)
- [History \(FHKW\)](#)
- [Psychology \(FSW\)](#)
- [Public Administration \(FSW\)](#)
- [Research School Safety and Security in Society](#)
- [Rotterdam School of Economics \(RSE\)](#)
- [Sociology \(FSW\)](#)

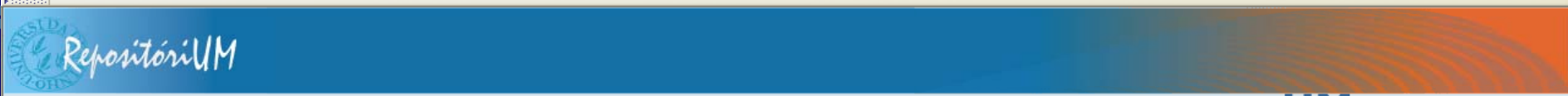
**Erasmus University Rotterdam**

**Electronic Research Publications**

EUR Research in digital form, including preprints, technical reports, working papers, images, videos, and more.

A service brought to you by the **Erasmus Electronic Publishing Initiative**

Contact: [dr. H.H. Ellermann](#)



**Pesquisar RepositóriUM:**  
  
[Pesquisa Avancada](#)

**Percorrer:**

- Comunidades & Coleções
- Títulos
- Autores
- Por Data

**Entrar:**

- Subscrever
- Meu RepositóriUM  
utilizadores autorizados
- Editar Conta
- Ajuda
- Acerca do RepositóriUM

Universidade do Minho >

**Bem-vindos ao repositório institucional da Universidade do Minho!**

Para ter mais informações sobre este novo serviço (por exemplo, como depositar conteúdos no caso de ser membro da U.M., qual a relação do RepositóriUM com *Declaração de Berlim sobre o Acesso Livre ao Conhecimento nas Ciências e Humanidades*, etc.), [clique aqui](#).

**Pesquisa**

Introduza um texto na caixa em baixo para pesquisar no RepositóriUM.

**Comunidades no RepositóriUM**

Seleccione uma comunidade para percorrer as suas colecções.

- [Biblioteca da Universidade do Minho](#)
- [Centro de Engenharia Biológica](#)
- [Departamento de Engenharia de Polímeros](#)
- [Departamento de Sistemas de Informação](#)
- [Núcleo de Estudos em Gestão](#)



**O que pode encontrar no RepositóriUM ?**

Investigação da UM em formato digital, incluindo teses e dissertações, artigos, relatórios técnicos, etc. O RepositóriUM está a dar os primeiros passos e a crescer constantemente. Visite-nos com frequência.

**É Mestre ou Doutor pela Universidade do Minho ?**

O RepositóriUM pode preservar e divulgar a sua tese ou dissertação! [Clique aqui](#), para mais informações.

## DÉPÔT DE DOCUMENTS

Chercher dans le dépôt :  
  
Lancer  
[Recherche avancée](#)

Érudit | Dépôt de documents (PROTOTYPE) >  
**Bienvenu dans le serveur de dépôt de documents**  
Une brève description...

- Accueil
- Naviguer dans les :**
  - communautés & collections
  - titres
  - auteurs
  - dates
- Me connecter pour :**
  - mes abonnements
  - ma zone personnelle utilisateurs autorisés
  - mettre à jour mon profil
- Aide
- À propos...

### Communautés dans le dépôt

Choisissez une communauté pour naviguer dans ses collections.

- [Centre d'analyse des politiques publiques](#)
- [Centre de recherche en économie agroalimentaire](#)
- [Centre de recherche sur les politiques et le développement social](#)
- [Centre études internationales et mondialisation](#)
- [Centre interuniversitaire de recherche en analyse des organisations \(CIRANO\)](#)
- [Centre interuniversitaire de recherche en économie quantitative](#)
- [Centre interuniversitaire de recherche sur la science et la technologie](#)
- [Chaire de Recherche du Canada en Mondialisation, Citoyenneté et Démocratie](#)
- [Collectif de recherche sur l'itinérance, la pauvreté et l'exclusion sociale](#)
- [École de service social](#)
- [Groupe de recherche en économie de l'énergie, de l'environnement et des ressources naturelles \(GREEN\)](#)
- [Groupe de recherche sur la cohésion sociale](#)
- [Groupe d'étude et de recherche sur la sécurité internationale](#)
- [Institut d'études européennes](#)
- [Programme de publication et de diffusion numériques des thèses](#)
- [RISQ & CIRASST](#)

### Plate-forme de tests de Dspace

Veillez noter que vous accédez à un **prototype**. Les documents qui seront ajoutés dans cet environnement de test et de développements ne seront pas conservés, de même que les numéros d'identification des documents déjà stockés. L'application peut être mise en arrêt à tout moment. L'interface utilisateur est en cours de modification.

Le serveur de prépublication d'Érudit utilise l'application [DSpace](#).

[Commentaires](#)

# DSpace in Spanish

- Universidad de Navarra, Spain
- [dspace.unav.es](http://dspace.unav.es)



# More Information

- 
- [dspace.org](https://www.dspace.org)
  - [dspace.mit.edu](https://dspace.mit.edu)

