



MIT's DSpace

A good fit for ETDs

Margret Branschofsky

Keith Glavash

MIT LIBRARIES

What it is

- **Institutional Repository** for MIT faculty's digital research and teaching materials
- **MIT Libraries - Hewlett Packard Research Labs** collaborative development project
- **Open Source** system

What it does

- **Captures**
 - Digital research materials in various formats
 - Directly from creators (e.g. faculty)
- **Describes**
 - Descriptive, technical and rights metadata
- **Distributes**
 - Via WWW, with necessary access controls
- **Preserves**
 - Various levels based on digital file format submitted

Possible content

- **Theses/Dissertations**
- **Audio Files**
- **Video Files**
- **Datasets**
- **Learning Objects (OCW)**
- **Images**
- **Technical Reports**
- **Working Papers**
- **Conference Papers**
- **Preprints/Articles**
- **Books**
- **Text**

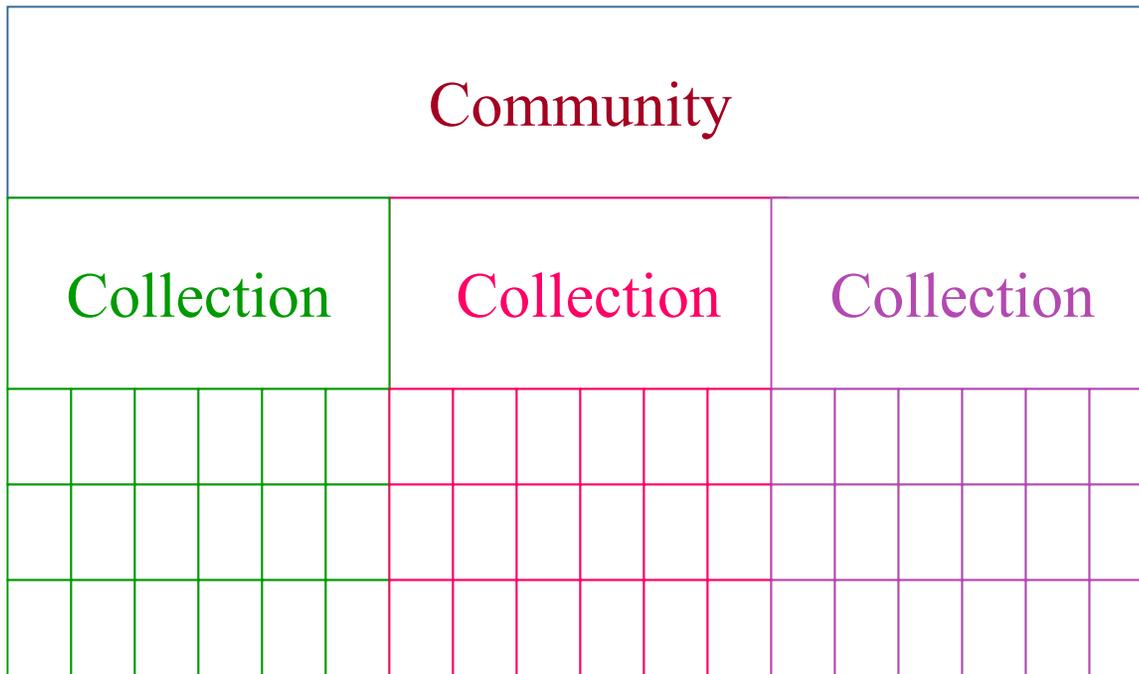
DSpace communities

- **Departments, Labs, Research Centers, Schools...**
- **Localized policy decisions determine...**
 - **Who contributes content**
 - **Who will have access**
 - **Submission workflow for collections**
 - **Submitters, Reviewers, Editors, Coordinators**
- **Some customization of look & feel**
- **Communities supply metadata**

Information model

- **Communities**
- **Collections**
 - Distinct groupings of like items
 - Separate workflows
- **Items**
 - Qualified Dublin Core metadata record
 - Can appear in multiple collections/communities
- **Bitstreams (files)**
 - Complex objects – multiple files
 - Multiple formats - same content

Information Model



Items

Files

10001110000001111000101000010001000101000100010111000100101010101000010011111100010100
 10001110000001111000101000010001000101000100010111000100101010101000010011111100010100
 10001110000001111000101000010001000101000100010111000100101010101000010011111100010100
 10001110000001111000101000010001000101000100010111000100101010101000010011111100010100
 10001110000001111000101000010001000101000100010111000100101010101000010011111100010100
 10001110000001111000101000010001000101000100010111000100101010101000010011111100010100

Preservation

- **Format registry**
- **Supported file formats**
 - **Confident that preservation techniques possible**
 - **Migration or emulation**
- **Known file formats**
 - **Batch migration where possible**
- **Unsupported file formats**
 - **Bit preservation at minimum**

Standards-based

- **Open Source**
- **CNRI “handles” for persistent identifiers**
- **Open URL linking/SFX-enabled**
- **OAI-PMH for exposing metadata**
- **METS profile**

Intellectual Property Rights

- Click-through **license** during submission
- Grants DSpace **non-exclusive right** to acquire, manage, preserve, distribute the item
- Does ***not*** grant DSpace **copyright**
- Copy of **license stored with item**

Future – Federation (1)

- **Goals of federation**
 - Share technical and other resources
 - Promote standards and interoperability
- **Hundreds of institutions have downloaded the software and expressed interest**

Future – Federation (2)

- **Six universities have joined MIT in a formal DSpace Federation**
 - **Columbia University**
 - **Cornell University**
 - **Ohio State University**
 - **University of Rochester**
 - **University of Toronto**
 - **University of Washington**

MIT's ETDs

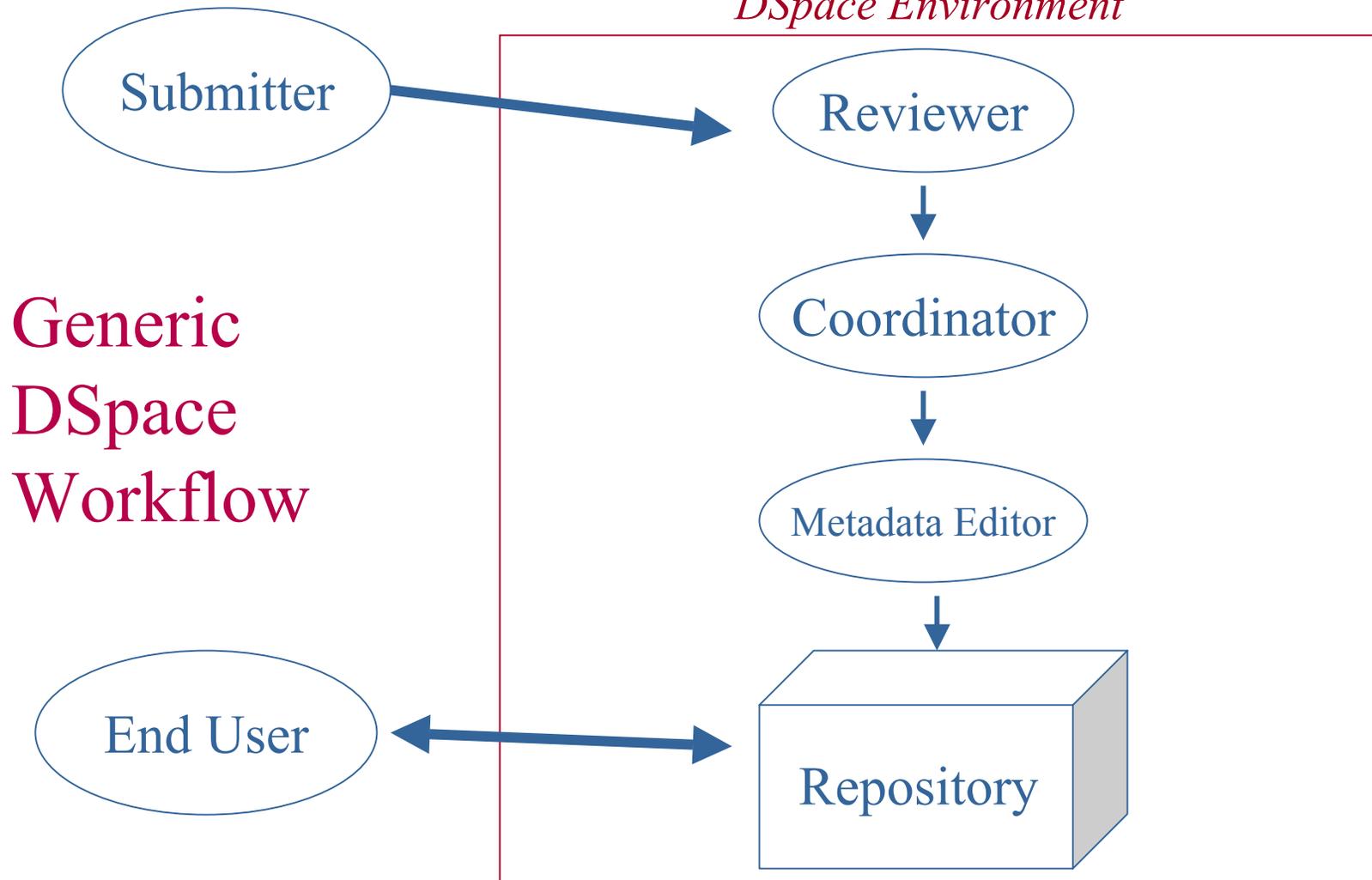
- **Since 1996** have used a locally-developed system based on Dienst software
- Collection consists of **~250 born-digital** and over **8,000 scanned** documents
- **<http://thesis.mit.edu>**
- Currently **migrating to DSpace** in two phases
 - design/implement workflow for new ETDs
 - shift existing ETDs and allow more scanned ETDs to be added

Generic DSpace Workflow

1. **Submitter** – author
2. **Reviewer*** – content verifier/editor/advisor
3. **Coordinator*** – administrative sign-off prior to release
4. **Metadata editor*** – cataloging

***optional**

DSpace Environment



Generic DSpace Workflow

ETD Stakeholders @ MIT

- **Student authors**
- **Faculty advisors**
- **Department or Grad School administrators**
- **Library/Archives (collect & manage)**

Prototype ETD steps (1)

- **Graduating student**
 - Submits thesis file(s)
 - Supplies basic metadata

Prototype ETD steps (2)

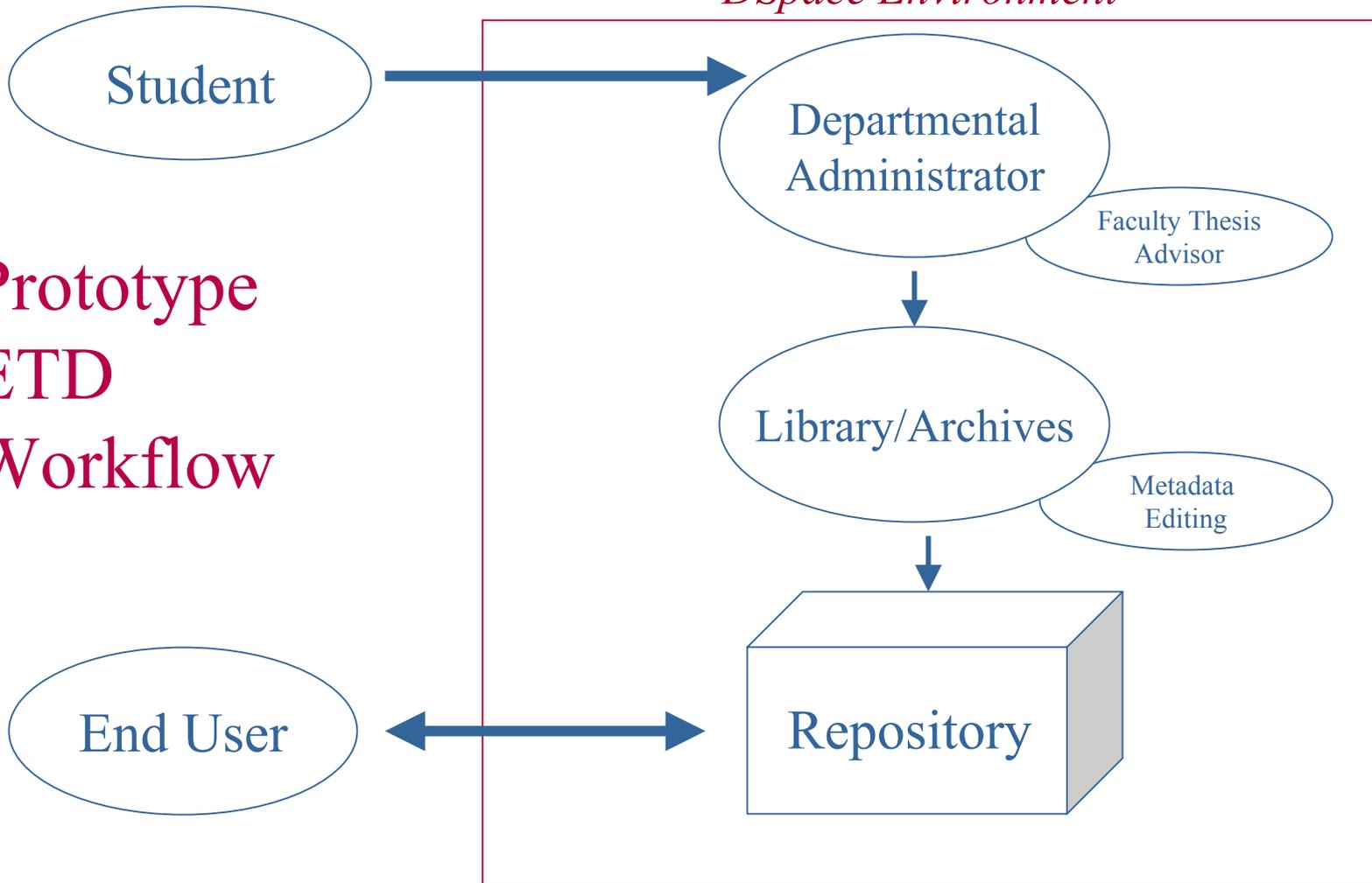
- **Department administrator**
 - **Receives email notification of thesis submission to DSpace**
 - **Faculty thesis advisor**
 - **Verifies document is authentic thesis for which degree is being granted**
 - **Confirms patent status**
 - **Approves submission on DSpace allowing advance to Library/Archives**

Prototype ETD steps (3)

- **Library/Archives**
 - Checks against patent hold database
 - Edits metadata (e.g. name authority)
 - Approves submission allowing thesis to be released and viewed online, OR
 - Temporarily holds release

DSpace Environment

Prototype ETD Workflow





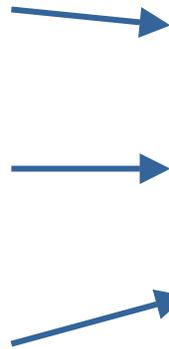
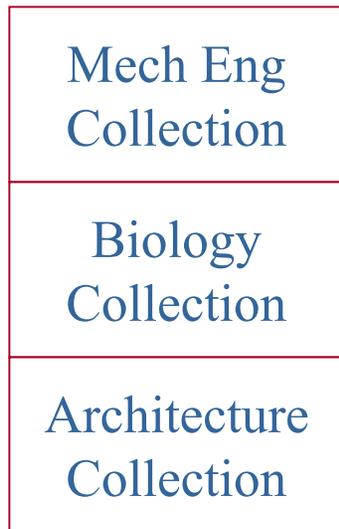
ETD
Community
(owned by Library)

Mech Eng
Collection

Biology
Collection

Architecture
Collection

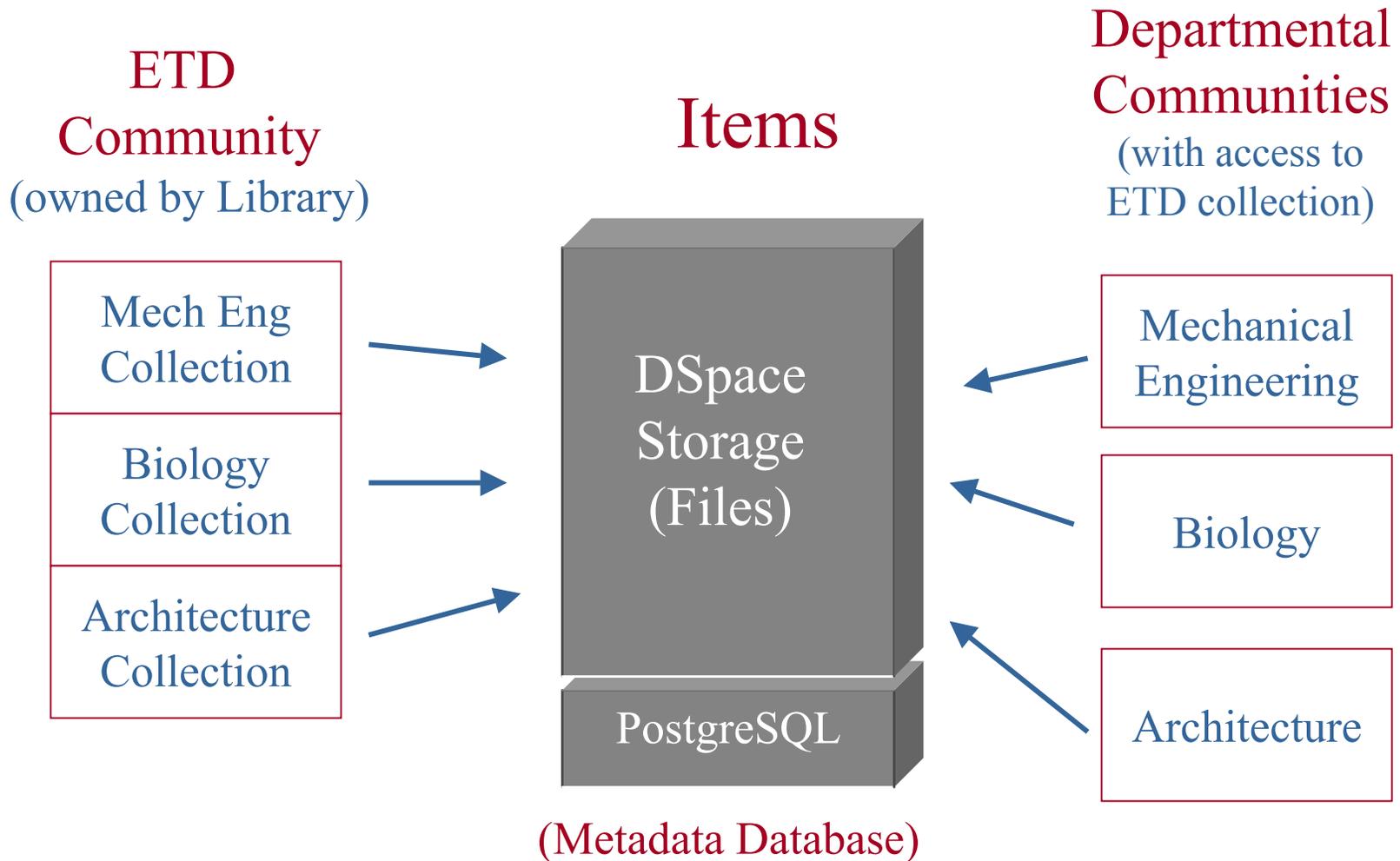
ETD
Community
(owned by Library)



Items



(Metadata Database)



Development possibilities for ETDs (1)

- **Virtual workspace** for thesis authors & advisors
- **Additional workflow configuration** possibilities
 - Re-ordering steps
 - Simultaneous reviewers
- **Adding functionality** to any step (e.g. printed forms)

Development possibilities for ETDs (2)

- **Preloaded student information**
- **Automated name authority control**
- **Automated metadata generation from submitted document**
- **Presentation & access tools**

More information

DSpace:

www.dspace.org

MIT Installation of DSpace:

web.mit.edu/dspace

MIT ETD's:

thesis.mit.edu

Margret Branschofsky

DSpace User Support Manager

MIT Libraries

margretb@mit.edu

Keith Glavash

Head of Document Services

MIT Libraries

kglavash@mit.edu