VIVO and Scholarly Repositories: Synergistic Opportunities
What is an Institutional Repository?

An Institutional Repository is an software platform for collecting, managing, preserving, and sharing the intellectual output of a research institution in digital formats.
What is an Institutional Repository?

- Supports *all formats* (text, media, data, etc.)
- Provides *discovery and delivery* services for content, on the Web and via Google and other search engines
- Supports *persistent* linking and citation to research material
- Includes a *service model*, e.g. for Open Access and/or digital archiving, and
- *set of policies* for governing digital research asset
Institutional Repository Snapshot

• Many platforms available
  • Open Source (DSpace, Eprints, Fedora, Greenstone, Archimede, Invenio, IR+)
  • Commercial (Digital Commons, DigiTool, CONTENTdm)

• ~2000 registered repositories world-wide (more in reality)

• Containing >2m research “items”

• Collectively rival PubMed, arXiv and other large centralized archives
What Does DSpace Support?

- DSpace (and most scholarly Institutional Repositories) is focused on the management, long-term archiving and Open Access sharing of research and teaching output of the local institution.

- It’s not just metadata or an index of content, it’s the content itself.

- Common content types include:
  
  | Preprints, eprints | E-theses |
  | Published articles, books and chapters | Images (visual, scientific, etc.) |
  | Technical reports and working papers | Audio/Video files |
  | Conference papers | Websites |
  | Databases | Learning Objects |
  | Datasets (statistical, geospatial, scientific) | Digitized library collections |
Repository Problem

• Repositories manage many data types

• metadata has diverse models, representations

• current XML/RDBMS support doesn’t scale
Exhibit UI

Exhibit
Publishing Framework for Data–Rich Interactive Web Pages

Exhibit lets you easily create web pages with advanced text search and filtering functionalities, with interactive maps, timelines, and other visualizations like these:
317 Publications sorted by: pub-type; then by... •  grouped as sorted article (84)

- **Staged Self-Assembly: Nanomanufacture of Arbitary Shapes with O(1) Glues**
  Erik D. Demaine, Martin L. Demaine, Sándor P. Fekete, Mashhood Ishaque, Eynat Rafalin, Robert T. Schweller, and Diane L. Souvaine
  *Natural Computing* (2008)

- **Ordinal Embeddings of Minimum Relaxation: General Properties, Trees, and Ultrametrics**
  Noga Alon, Mihai Budiu, Erik D. Demaine, Martin Farach-Colton, MohammadTaghi Hajiaghayi, and Anastasios Sidiropoulos
  *ACM Transactions on Algorithms* (2008)

- **Refolding Planar Polygons**
  Hayley N. Iben, James F. O'Brien, and Erik D. Demaine
  *Discrete & Computational Geometry* (to appear)
  This paper describes an algorithm for generating a guaranteed-intersection-free interpolation sequence between any pair of compatible polygons. Our algorithm builds on prior results from linkage unfolding, and if desired it can ensure every edge length changes monotonically over the course of the interpolation sequence. The computational machinery that ensures against self-intersection is independent from a distance metric that determines the overall character of the interpolation sequence. This decoupled approach provides a powerful control mechanism for determining how the interpolation should appear, while still assuring against intersection and guaranteeing termination of the algorithm. Our algorithm also allows additional control by accommodating a set of algebraic constraints that can be weakly enforced throughout the interpolation sequence.
  URL: [http://hdl.handle.net/1721.1/41878](http://hdl.handle.net/1721.1/41878)
  [View/Open] | application/pdf | 333.2Kb

- **Algorithmic Graph Minor Theory: Improved Grid Minor Bounds and Wagner's Contraction**
  Erik D. Demaine, MohammadTaghi Hajiaghayi, and Ken-ichi Kawarabayashi
  *Algorithmica* (2007)

- **Linearity of Grid Minors in Treewidth with Applications through Bidimensionality**
  Erik D. Demaine, and MohammadTaghi Hajiaghayi
  *Algorithmica* (2007)
VIVO and DSpace Together

- Link publications listed in VIVO to DSpace PIDs for full-text items
- Link DSpace items of all types to appropriate VIVO entities (e.g., people, organizations, research projects, etc.)
- Using DSpace as a long-term archiving platform for relevant VIVO entities
VIVO and DSpace Together

- *Mine* research topics/keywords from DSpace items to link them to appropriate VIVO entities (i.e., people, organizations, 'concepts')

- Incorporate DSpace full-text into VIVO search
  DSpace supports OpenSearch (e.g. link keywords to an OpenSearch query across federated repositories for syndicated results)

- Streamlined, multi-system deposit/update, UI (Exhibit?)
VISION

- Registry of (active) researchers including minimal profile info
- Assigns unique, persistent IDs
- Name disambiguation to identify correct ID
- May include bibliographic information, other activities
STATUS

• Non-profit corporation, jointly governed by publishers, universities, libraries, researchers

• Principles include Open Access to registry data

• Scheduled to launch late 2011 – early 2012
A Scholarly Data Ecosystem

Centralized MIT Research Profiles (VIVO)

Bibliographic Acquisition & Management

Full-text DBs

DOI

Institutional Repository

External Bibliographies (Local)

HR data

Other Local Data
VIVO + DSpace + Exhibit + ORCID

- Full-text DBs
- Harvesters (ORCID)
- DOI
- Institutional Repository
- External Bibliographies (Local) [DOI, Dspace URI, ORCID]
- Web browser
- Data Warehouse (HR and ORCID data)
- Other Local Data

Centralized MIT Research Profiles (VIVO)

Bibliographic Acquisition & Management

Exhibit
Key Infrastructure

- Repositories and tools (VIVO, DSpace, Exhibit)
- *Standard identifiers* (URIs) for entities of interest: people, institutions, documents, datasets, etc.
- Common *content models and ontologies* for metadata/content – shared registry?