GKR - The GALILEO Knowledge Repository

Sara Fuchs

COMO XXII – Athens

October 15, 2010
The Problem

- Much of USG’s intellectual product never appears in a permanent printed form
- Exists as disorganized pockets of digitally born objects & media scattered among individual hard drives, departmental servers, and removable storage media across an institution
- While USG’s digital scholarly output reflects substantial investment of resources, assets and effort, it
  - Lacks curatorial stewardship;
  - May be inaccessible;
  - Exists on unsustainable hardware, software, or individual support;
  - Needs future-proofing migration strategies.
The scholarly publishing system

• Cost of print and electronic journals increasing at rate that greatly outpaces the general rate of inflation putting enormous pressure on university budgets

• Publishers charge exorbitant subscription fees for research conducted with public funds

• Scholars limit access to their own research by granting publishers a monopoly on their work
The changing landscape

• Wide dissemination of scholarly materials is on the rise and challenging publisher business models
• Grant funding agencies are increasingly requiring public access to output and researchers are responsible for compliance
• Increased awareness among faculty of how publisher contracts can prohibit use of materials for teaching purposes and furthering research
The changing landscape

• Some faculty are working alone to increase access to their research
• Some faculty are choosing to act together to increase access to their research and to take back control of their intellectual property from publishers
• Wide dissemination of scholarly materials is often made possible through open access journals and repositories
Open Access (OA) defined

• “Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions... OA is compatible with copyright, peer review, revenue (even profit), print, preservation, prestige, career-advancement, indexing, and other features and supportive services associated with conventional scholarly literature... OA literature is not free to produce or publish.”

   – Peter Suber, http://www.earlham.edu/~peters/fos/overview.htm
Open Access channels

• Journals
  - 5000+ DOAJ peer-reviewed journals
  - Variety of business models
    o Advertising, Institutional subsidies, membership dues, publication fees, fund-raising, volunteers, etc.
    http://oad.simmons.edu/oadwiki/OA_journal_business_models
  - BioMed Central, PLoS
Open Access channels

• Repositories/archives
  ▪ 1650+ OA/OAI repositories
    o data, disciplinary, institutional (IR); 372 IRs in the U.S.
      http://www.opendoar.org/
  ▪ arXiv.org (e-prints rep), Dryad (data rep), DLESE (science educ materials), DLNET (math and eng LO)
  ▪ Some IR content peer-reviewed
  ▪ OA policies require deposit in repositories
Public access to research

  Public Access Policy requires research be made openly available in PubMed Central

- National Center for Atmospheric Research [http://opensky.library.ucar.edu/policy/](http://opensky.library.ucar.edu/policy/)
  NSF sponsored lab requires scientists make their research openly available in their institutional repository

  “require its scientists to publish their original research articles in scientific journals that allow the articles and supplementary materials to be made freely accessible in a public repository within six months of publication.”
Public access to research

  Act last introduced in the Senate 6/09 requiring nonclassified research arising from grants funding by 11 government agencies be made openly available, including NSF, DoD, DoE, and NASA.

  Open Government Initiative - “all federal research agencies should require that papers published by the investigators they support be made freely available to the public as soon as possible”

• NSF Data Management Plan
University OA policies

- Harvard Faculty of Arts and Sciences, Feb 2008
- Stanford School of Education, June 2008
- Boston University, Feb 2009
- MIT, Mar 2009
- Oberlin College, Nov 2009
- University of Virginia, July 2010
- Rollins College Faculty of Arts and Sciences, Feb 2010
- Duke University, Mar 2010
- San Jose State University, Apr 2010
“When I write a paper, I put it on my Web site and it might get 5 million hits. If it’s really good and I send it to a leading journal in my field that prints 16,000 copies and I’m supposed to take it off my Web site. In what universe does that make sense?”

http://www.arl.org/sparc/innovator/harvardfas.shtml
Goals of the library

• Provide stewardship for digital scholarship and research content
• Enhance access to the faculty’s intellectual output
• Raise awareness of and advance changes in scholarly communication on campus
Reasons to submit to an IR

• Increase access to your intellectual product: Multiply the ways in which it may be used to support education
• Publicize your work: Open access scholarship is cited 40% more often (Harnad)
• Know more about the impact of your work: Use statistics on number of downloads, etc.
• Get more life out of works-in-progress, conference papers, research reports, and other material that may not see publication in a peer-reviewed journal
• Generate renewed interest in previously published material
• Collect and share works for use in evaluation and accreditation
Types of Content for IRs

- Electronic theses and dissertations
- Research and technical reports
- Pre-print research
- Post-print research
- Working papers
- Conference papers
- Open access journals
- Service publications
- Research proposals

- Research data sets
- Organizational documents
- College and university archives
- Organizational reports
- Institutional planning and evaluation documents
- Learning objects
- Other resources that reflect the quality and diversity of the institution
GALILEO Knowledge Repository (GKR)

• Building a system-wide approach to Institutional Repositories
• Concept developed by Regents Advisory Committee on Libraries (RACL) in August 2004
• Federal grant (IMLS National Leadership) awarded 2009 to seed the project
• Leads: Georgia Institute of Technology & University of Georgia
Institutions / Libraries involved

- Albany State University
- Georgia State University
- College of Coastal Georgia
- Georgia Southern University
- Georgia Institute of Technology
- Kennesaw State University
- Medical College of Georgia
- University of Georgia
- Valdosta State University
- North Georgia College and State University (SC survey)
GKR Project Components

- Build meta-searching site to connect scholarly content across USG
- Establish a repository hosting service for discrete IRs at four USG institutions - Albany State, Coastal Georgia, Medical College of Ga., Georgia Gwinnett College
- Establish other IR-related services: copyright research; digitization; content submission; and preservation
- Partner with NGCSU to conduct an assessment survey of the USG faculty’s usage and perceptions of IRs
- Document and make available the GKR organization and technical model to others interested in establishing statewide IRs
Timeline

2009: Hire GKR Manager and Developer
2010: Create GKR metadata repository; Establish IR-Related Services
2011: Provide IR-Related Services
2012: Dissemination (guide and workshop)
Disciplines in GKR

Select a community to browse its collections.

- Agriculture, Agriculture Operations, and Related Sciences [0]
- Architecture and Related Services [589]
- Area, Ethnic, Cultural, and Gender Studies [0]
- Biological Sciences [0]
- Business, Management, Marketing, and Related Support Services [0]
- Computer and Information Sciences and Support Services [0]
- Education [15]
- Electronic Theses and Dissertations [99]
- Engineering [100]
- English Language and Literature/Letters [0]
- Family and Consumer Sciences/Human Sciences [0]
- Foreign languages, literatures, and Linguistics [3]
- Health Sciences [37]
- History [0]
- Legal Professions and Studies [0]
- Liberal Arts and Sciences, General Studies and Humanities [0]
- Library Science [0]
- Mathematics and Statistics [0]
- Multi/Interdisciplinary Studies [0]
- Natural Resources and Conservation [0]
- Parks, Recreation, Leisure, and Fitness Studies [0]
- Philosophy and Religious Studies [0]
- Physical Sciences [0]
- Psychology [0]
- Public Administration and Social Services [0]
- Social Sciences [136]
- University Archives [41]
- Visual and Performing Arts [0]
Mapping Tool: browsing by discipline

GALILEO Knowledge Repository Mapping Tool

Use the GALILEO Knowledge Repository Mapping Tool to map individual institutional repository collections to central GALILEO Knowledge Repository collections.

Select "GALILEO Knowledge Repository (gkr)" to view and edit the central communities and collections.

Select one of the individual institution links to view and edit that institution's communities and collections, and to map that institution's collections to the central GALILEO Knowledge Repository collections.

- Albany State University (asurame)
- College of Coastal Georgia (ccga)
- Georgia Institute of Technology (gatech)
- Georgia Southern University (georgiasouthern)
- Georgia State University (gsu)
- Kennesaw State University (kennesaw)
- Medical College of Georgia (mcg)
- North Georgia College and State University (ngcsu)
- University of Georgia (uga)
- Valdosta State University (valdosta)
Mapping Tool

- Engineering
  - Aerospace Engineering
  - Bioengineering
  - Biomedical Engineering
  - Chemical and Biomolecular Engineering
  - Civil Engineering
  - Computer Engineering
  - Electrical Engineering
  - Engineering Science
  - Environmental Engineering
  - Industrial Engineering
  - Materials Science and Engineering
  - Mechanical Engineering
  - Nuclear and Radiological Engineering
  - Nuclear Engineering
  - Polymer, Textile, and Fiber Engineering
Mapped collections

Communities in SMARTech
Select a community to browse its collections.

- Center for Experimental Research in Computer Systems (CERCS) [152]
- CERCS Technical Reports [152]
- Center for the Enhancement of Teaching and Learning (CETL) [15]
- The Classroom [15]
- College of Architecture (CoA) [2973]
  - Architecture Centennial Lecture Series [14]
  - Architecture Program Lecture series [7]
  - City and Regional Planning Book Series [2]
  - City and Regional Planning Program Invited Speakers [3]
  - City and Regional Planning Program Studio Final Presentations [1]
  - College of Architecture Books [1]
  - College of Architecture Brochures [2]
  - College of Architecture Faculty Publications [30]
  - College of Architecture Invited Speakers [2]
  - College of Architecture Newsletters [12]
  - College of Architecture OSP Research Reports [124]
  - College of Architecture Research Forum [16]
  - College of Architecture Theses and Dissertations [1032]
  - College of Architecture Undergraduate Projects [1518]
  - Design and Technology for Healthy Aging Initiative at Georgia Tech (DATHA) [4]
  - Douglas C. Allen Annual Lecture [1]
  - Georgia Tech Solar Decathlon 2007 [1]
  - Industrial Design Lecture Series [10]
  - PhD Focus [8]
  - T. Gordon Little Lecture Series in Imagination [1]
  - Urban Connections [4]
- Center for Assistive Technology and Environmental Access (CATEA) [34]
  - CATEA Fact Sheets and Reference Guides [27]
  - CATEA Invited Speakers, Seminars, Symposia, and Workshops [1]
  - Center for Assistive Technology and Environmental Access OSP Research Reports [6]
- Center for Quality Growth and Regional Development (CQGRD) [75]
  - Center for Quality Growth and Regional Development (CQGRD) Research and Studies [10]

Disciplines in GKR
Select a community to browse its collections.

- Agriculture, Agriculture Operations, and Related Sciences [9]
- Architecture and Related Services [589]
  - 24th National Conference on the Beginning Design Student [40]
  - Adopting Building Information Modeling and Integrated Practice Symposium [8]
  - Architecture Centennial Lecture Series [14]
  - Architecture Program Lecture series [7]
  - Atlanta BeltLine Decision Support Tool: Strategic Planning Session - 2008 [8]
  - CATEA Fact Sheets and Reference Guides [27]
  - CATEA Invited Speakers, Seminars, Symposia, and Workshops [1]
  - Center for Assistive Technology and Environmental Access OSP Research Reports [5]
  - Center for Quality Growth and Regional Development (CQGRD) Research and Studies [10]
  - City and Regional Planning Book Series [2]
  - City and Regional Planning Program Invited Speakers [3]
  - City and Regional Planning Program Studio Final Presentations [1]
  - College of Architecture Books [1]
  - College of Architecture Brochures [2]
  - College of Architecture Faculty Publications [30]
  - College of Architecture Invited Speakers [2]
  - College of Architecture Newsletters [12]
  - College of Architecture OSP Research Reports [100]
  - College of Architecture Research Forum [15]
  - College of Architecture Theses and Dissertations [99]
  - College of Architecture Undergraduate Projects [100]
  - Design and Technology for Healthy Aging Initiative at Georgia Tech (DATHA) [4]
  - Douglas C. Allen Annual Lecture [1]
  - Georgia Tech Solar Decathlon 2007 [1]
<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down in the dumps : uncovering new meaning in the landscape of landfills</td>
<td>Brandes, Lauren Marie (University of Georgia, 2005-03-31)</td>
<td></td>
</tr>
<tr>
<td>Current problems of international taxation of electronic commerce</td>
<td>Kerimov, Nuran Gudrat (University of Georgia, 2005-03-31)</td>
<td></td>
</tr>
<tr>
<td>Reconstructing paradigms of expression [electronic resource] / by Christopher Murphy.</td>
<td>Murphy, Christopher Eric. (2009-08-26)</td>
<td></td>
</tr>
<tr>
<td>The use of land trusts to preserve abandoned graveyards in the American Southeast</td>
<td>Smith, Jason Oliver (University of Georgia, 2005-03-31)</td>
<td></td>
</tr>
<tr>
<td>This is not a thesis: a thesis on david markson</td>
<td>Underwood, Samuel Benjamin (University of Georgia, 2005-03-31)</td>
<td></td>
</tr>
<tr>
<td>Developing a Modular Hydrogeology Ontology Extending the Sweet Ontologies</td>
<td>Tripathi, Ajay (Digital Archive @ GSU, 2005-08-08)</td>
<td></td>
</tr>
</tbody>
</table>
IR-related Services

• To reduce barriers to recruiting scholarly content
  ▪ Copyright research
  ▪ Digitization
  ▪ Content submission
  ▪ Preservation

• Services resulted from the USG-wide GKR stakeholder meeting of November 30, 2007
Copyright Research Services

• Provide guidance and assistance in resolving intellectual property concerns regarding IR submissions

• Emails sent to the GKR Manager; Georgia Tech and University of Georgia will assist with requests
Digitization Request Form

<table>
<thead>
<tr>
<th>Current use:</th>
<th>Daily</th>
<th>Several times a week</th>
<th>Once or twice a month</th>
<th>Several times a year</th>
<th>Once a year</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often is the material accessed in analog form?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Potential use:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>One a scale of 1-5, 1 highest and 5 lowest, to what extent would your institution benefit from access to the material in digital format?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Barriers:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>One a scale of 1-5, 1 highest and 5 lowest, to what extent does the analog form limit use of the material?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context:</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the collection part of a larger digital collection?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Would scanning of these materials complete an existing digital collection?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are there related digital collections online?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Content Submission Services

Interpreting/applying GKR Metadata Guidelines (modified Dublin Core):
• Resolving content submission issues in DSpace
• Submitting content and metadata when needed

Valdosta MLIS: 2-4 student interns as content submitters:
• Students earn 3 credit hours, gain valuable work experience
• GKR / VSU MLIS relationship unique in statewide repositories
GALILEO Knowledge Repository Metadata Guidelines by Priority

Status abbreviations: M = Mandatory; MA = Mandatory if applicable; Rec = Recommended; RA = Recommended if applicable; Opt = Optional

<table>
<thead>
<tr>
<th>MANDATORY</th>
<th>Element</th>
<th>Qualifier</th>
<th>Status</th>
<th>Scope Note</th>
<th>Instructions / example</th>
</tr>
</thead>
<tbody>
<tr>
<td>date</td>
<td>date</td>
<td>accessed</td>
<td>M</td>
<td>Date DSpace takes possession of item.</td>
<td>Provided by DSpace. Example: 2005-01-04T18:08:50Z</td>
</tr>
<tr>
<td>date</td>
<td>date</td>
<td>available</td>
<td>M</td>
<td>Date or date range item became available to the public.</td>
<td>Provided by DSpace. Example: 2005-01-04T18:08:50Z</td>
</tr>
<tr>
<td>date</td>
<td>date</td>
<td>issued</td>
<td>M</td>
<td>Date of publication or distribution.</td>
<td>Example: 2001-01</td>
</tr>
<tr>
<td>identifier</td>
<td>identifier</td>
<td>uri</td>
<td>M</td>
<td>Uniform Resource Identifier</td>
<td>Provided by DSpace. Example: <a href="http://hdl.handle.net/123456789/60">http://hdl.handle.net/123456789/60</a></td>
</tr>
<tr>
<td>description</td>
<td>description</td>
<td>Provenance (approved)</td>
<td>M</td>
<td>History of object.</td>
<td>Provided by DSpace.</td>
</tr>
<tr>
<td>description</td>
<td>description</td>
<td>Provenance (available)</td>
<td>M</td>
<td>History of object.</td>
<td>Provided by DSpace.</td>
</tr>
<tr>
<td></td>
<td>description</td>
<td></td>
<td></td>
<td>Draft standard language regarding use</td>
<td>Draft standard language regarding use</td>
</tr>
</tbody>
</table>
Digital Preservation Services

- Built on LOCKSS (supports “distributed digital replication” approach)
- Closed Archive (no direct public access)
- Automated format emulation tools
- Low Cost (planned minimal expense, low barriers to adoption for mid-size institutions)
- Flexible, adaptable multi-inst. model
- LC / NDIIPP partnership (1 of 8 initial)

http://metaarchive.org/
Education for Hosted Sites

• Site visits planned for each hosted site
  GKR staff meet with library staff and research faculty to introduce them to the GKR project, IRs and scholarly communication

• DSpace training via Wimba collaborative learning software
  Repositories, DSpace structure, creating communities and collections, item submission and searching

• Metadata training via Wimba
  Dublin Core best practices, applying GKR metadata guidelines in DSpace
USG Faculty IR Interest Survey

• GKR sponsoring a USG-wide faculty survey to assess perceptions, experiences w/ IR use, author’s rights issues, OA publishing activity
• Survey by Jennifer Campbell-Meier (NGCSU Library).
  ▪ Ph.D. work, Communication & Information Sci., Univ. of Hawaii
• Results analyzed with GKR Outreach and Evaluation Committee
  ▪ Used to improve GKR technologies, services, marketing
• Learn about perceptions held by academicians about IRs and Open Access models and build business strategies to address them
• No surveys of IR use at statewide level exist currently
  ▪ Will inform GKR’s development / Serve as model survey
Dissemination of practices

• Document the GKR organizational model
• Document the GKR technical model:
  ▪ Guidelines for searching, harvesting repository metadata
  ▪ Technical specifications for GKR technologies
• Communicate, disseminate to statewide and consortial organizations via:
  ▪ GKR web site
  ▪ Presentations at meetings and professional conferences
  ▪ Articles and announcements through library publications
Conclusion: GKR’s impact

- Increase visibility to USG’s digital scholarship and research
- Promote information sharing and discovery of research from a single web site
- Improve access to learning for the citizens of Georgia at large
- Create outlet for new forms of instructional media and scholarship
- Provide stewardship for the least permanent (i.e. non-published) elements of the USG’s intellectual works
- Demonstrate effectiveness of IRs to USG faculty
- Advance scholarly communication
Links

GKR Project Site:
http://www.library.gatech.edu/gkr/

Training and Staging Site:
http://gkr-dev.library.gatech.edu/
Questions?

Sara Fuchs, Digital Initiatives Librarian, Georgia Institute of Technology
sara.fuchs@library.gatech.edu

Marlee Givens, GKR Manager
marlee.givens@library.gatech.edu

Keith Gilbertson, DSpace Technical Lead
keith.gilbertson@library.gatech.edu