

Integrating Scholarly Repository Services into Consortial Organizations and Statewide University Systems

Tyler Walters

Presented by: Sara Fuchs & Keith Gilbertson

Georgia Institute of Technology
Atlanta, Georgia, USA

The Problem

- ❑ Much of the University System of Georgia'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.



Types of Content

- 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
- 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 USG



GALILEO Knowledge Repository

- GALILEO: **GeorgiA** **LI**brary **LE**arning **O**nline
- Proposing a system-wide approach to institutional repositories as well as a collaborative strategy for promoting open access to scholarly information
- Concept developed by Regents Advisory Committee on Libraries (RACL) in August 2004

GKR: The Project Basics

- ❑ 3-yr project to finalize several components toward establishing a comprehensive statewide repository
- ❑ \$1,664,650 project total: \$857,005 (IMLS) & \$833,946 (institution share)
- ❑ Leads: Georgia Institute of Technology & University of Georgia
- ❑ Participants: Georgia State Univ., Georgia Southern Univ., Kennesaw State Univ., Medical College of Georgia, Valdosta State Univ., Albany State Univ., College of Coastal Georgia
- ❑ Scholarly Communication Survey: North Georgia College and State University

GKR Project Components

- Build a repository of standardized metadata harvested from IRs within the USG (GT and GKR Technical Committee)
- Hosting service for IRs at three USG institutions (GT)
- Establish IR-related services: copyright research (GT/UGA); digitization (UGA); content submission (GT/VSU); and preservation (GT)
- 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 model to others interested in establishing statewide IRs
- Develop and implement a statewide and consortial repositories symposium and workshop



Timeline

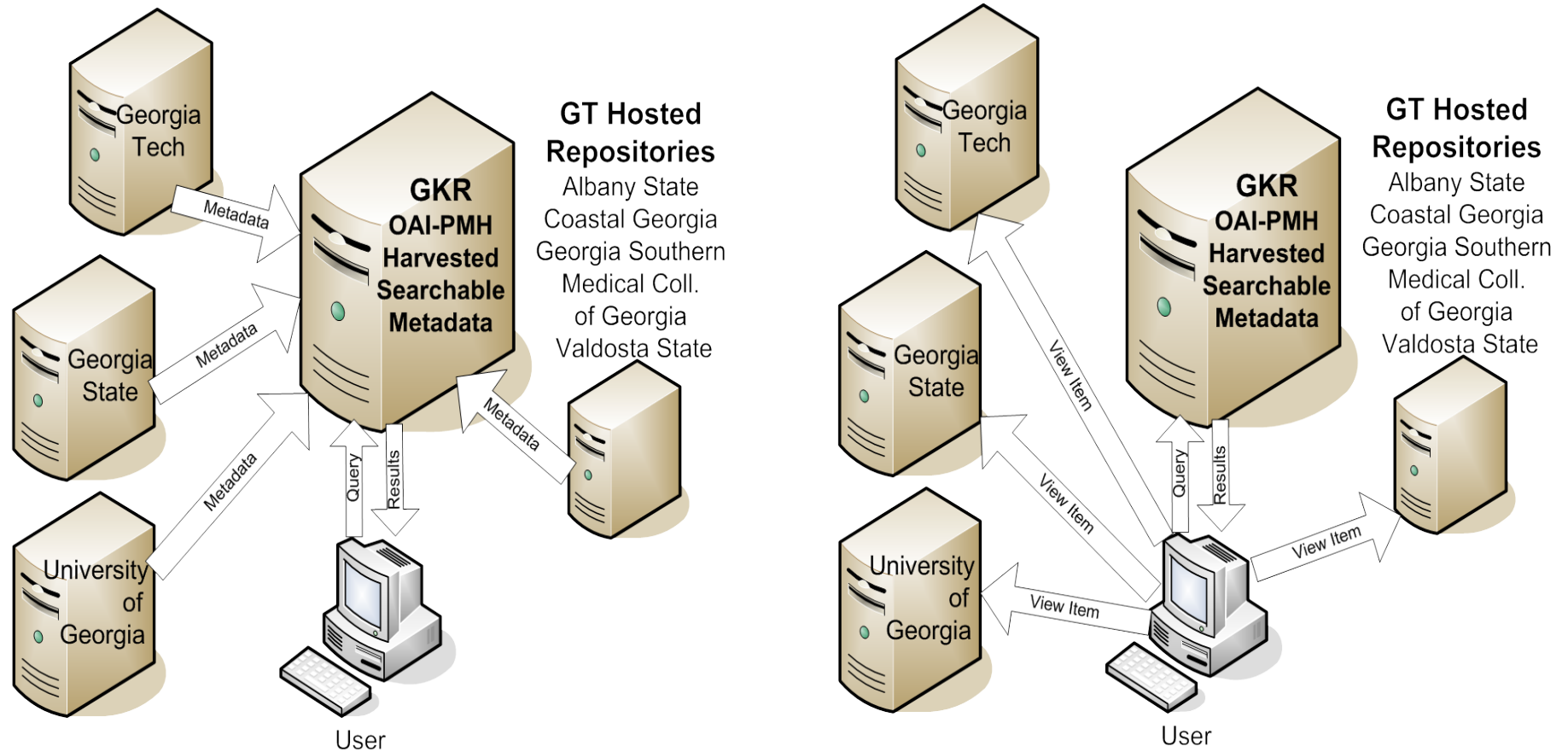
2009: Hire GKR Manager and Developer

2010: Create GKR metadata repository;
Establish IR-Related Services

2011: Provide IR-Related Services

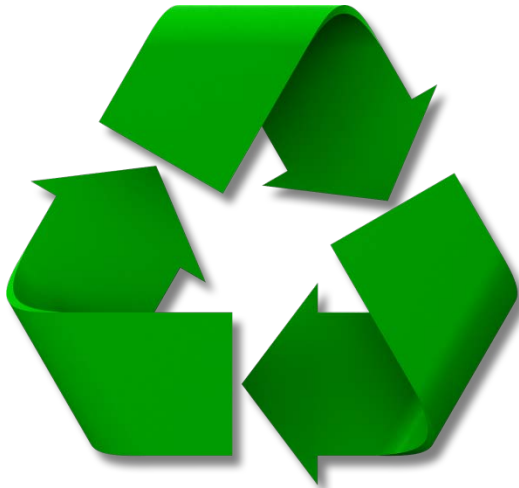
2012: Dissemination (guide and workshop)

GALILEO Knowledge Repository



Technology Goals - Reduce, reuse, recycle

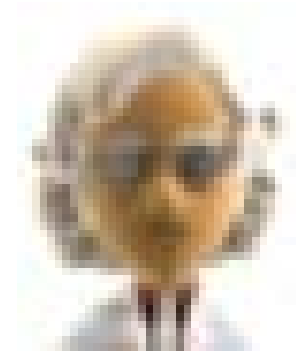
- Certain technological problems come up when creating a consortial repository system
- These problems have already been solved by others
- Let others do the work!



Work the Technical Lead hasn't done...



Thanks, Georgia Tech and USG Faculty and Staff!





Current Infrastructure

Two servers at Georgia Institute of Technology

- 5+ terabytes of storage
- 24 Gigabytes of RAM/system
- Powerful CPU
- DSpace Repositories hosted at other Institutions
- Digital Commons Sites hosted on other servers

Thank you repository software developers!



Harvesting Process (Original Prototype)

Metadata from member repositories were pulled into the GKR using:

- PKP Harvester component
- MySQL data dump of harvested metadata
- Shell and awk scripts to format dumped metadata
- Command Line DSpace Import (check for errors and repeat if necessary)

Current Process (DSpace 1.6 Harvester)

Edit Collection: University of Georgia - New Version

[Edit Metadata](#) [Assign Roles](#) [Content Source](#)

Content source: This is a standard DSpace collection
 This collection harvests its content from an external source

Harvested Collection Location

OAI Provider: http://tolson.libs.uga.edu.:8080/oai/request

OAI Set id: all

Metadata format: DSpace Intermediate Metadata

Content being harvested: Harvest metadata and references to bitstreams.

Last Harvest Result: Harvest from http://tolson.libs.uga.edu.:8080/oai/request successful on 2010-06-11 11:06:54.122

Current harvest status: Collection ready for harvest

[Change Settings](#) [Import Now](#) [Reset and Reimport Collection](#)

[Save](#) [Return](#)

Click the button!

Thank you Texas A&M!

Mapping Tool

GALILEO KNOWLEDGE REPOSITORY 
Digital Scholarship from the Institutions of the University System of Georgia [Login](#)

[Home](#)




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.

[« hide](#)

- [GALILEO Knowledge Repository \(gkr\)](#) 
- [Albany State University \(asurams\)](#) 
- [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 (gkr 43 »)
 - Aerospace Engineering (gkr 44 »)
 - Bioengineering (gkr 45 »)
 - Biomedical Engineering (gkr 46 »)
 - Chemical and Biomolecular Engineering (gkr 47 »)
 - Civil Engineering (gkr 48 »)
 - Computer Engineering (gkr 49 »)
 - Electrical Engineering (gkr 50 »)
 - Engineering Science (gkr 51 »)
 - Environmental Engineering (gkr 52 »)
 - Industrial Engineering (gkr 53 »)
 - Materials Science and Engineering (gkr 54 »)
 - Mechanical Engineering (gkr 55 »)
 - Nuclear and Radiological Engineering (gkr 57 »)
 - Nuclear Engineering (gkr 56 »)
 - Polymer, Textile, and Fiber Engineering (gkr 58 »)

Quickly Getting Content into Medical College of Georgia

MCG has Open Access articles in other databases

- [SWORD submission test from BioMed Central](#)
- Worked beautifully!
- MIT has SWORD publisher workflow project
- SWORD workshop 2:15 – 4:30 PM Friday July 9

Thank you SWORD, BioMed Central, and MIT pioneers!



PubMed Central Open Access Import

Import full-text articles from PubMed Central

Innovation by Scott Lapinski at Harvard Medical School Library

Find and download articles

Limit to open access records only

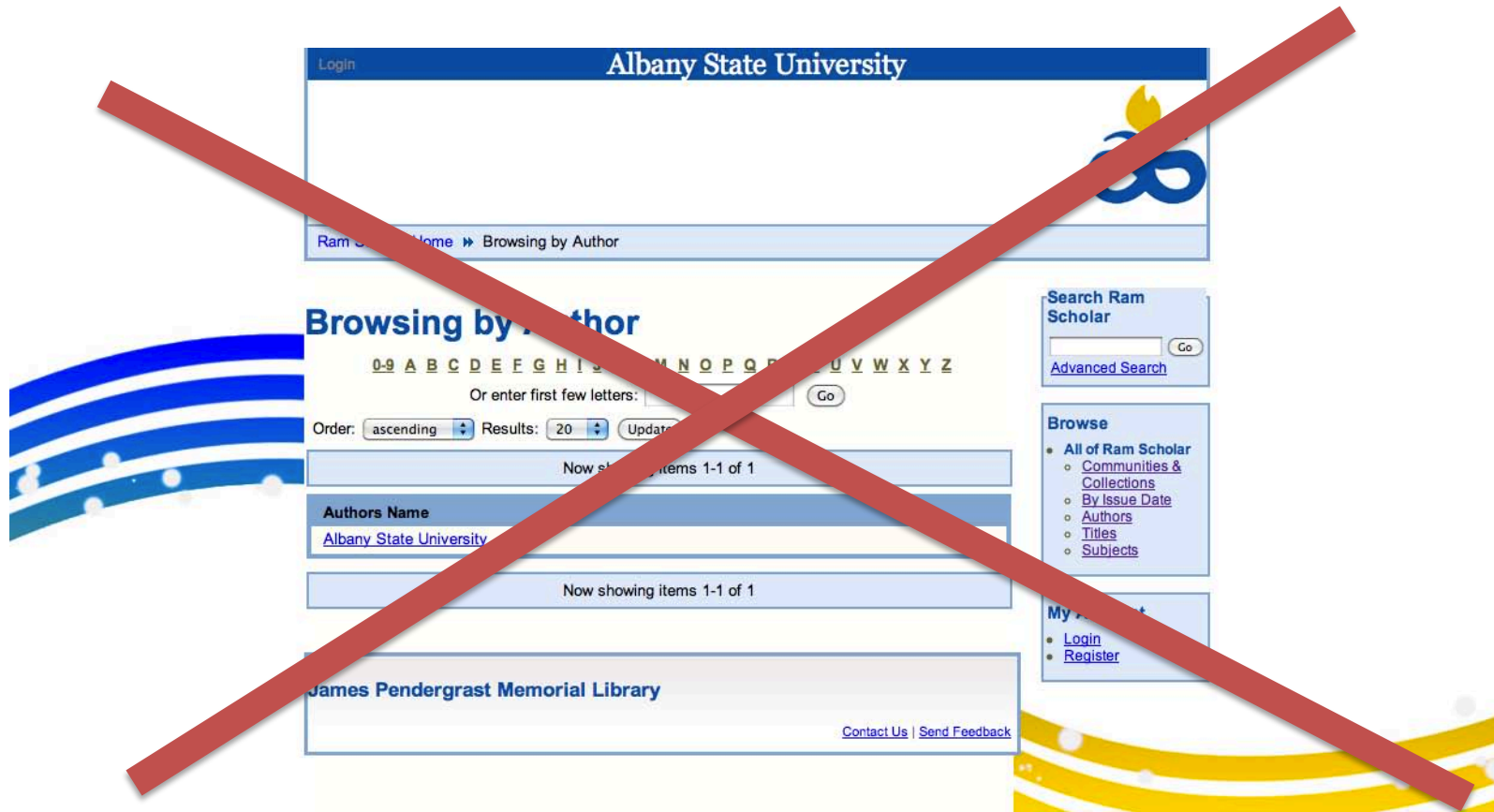
Convert metadata into DSpace format and import items

Thank you Scott Lapinski! (Harvard Medical School)



Theming with the DSpace xmlui ("zoom-louie")...

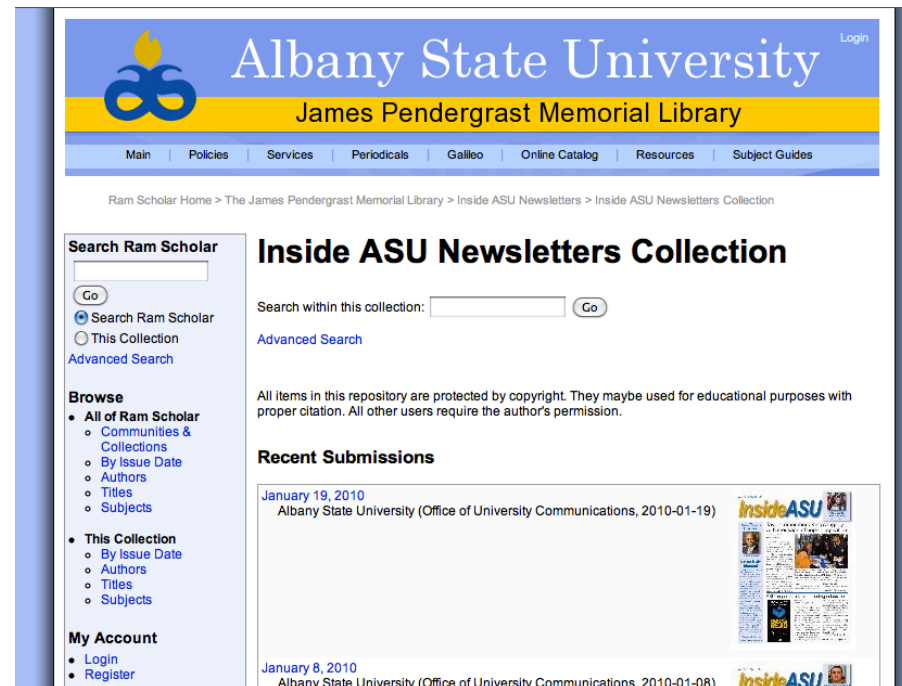
The Big, Bubbly Ribbon



“New” Albany State Theme



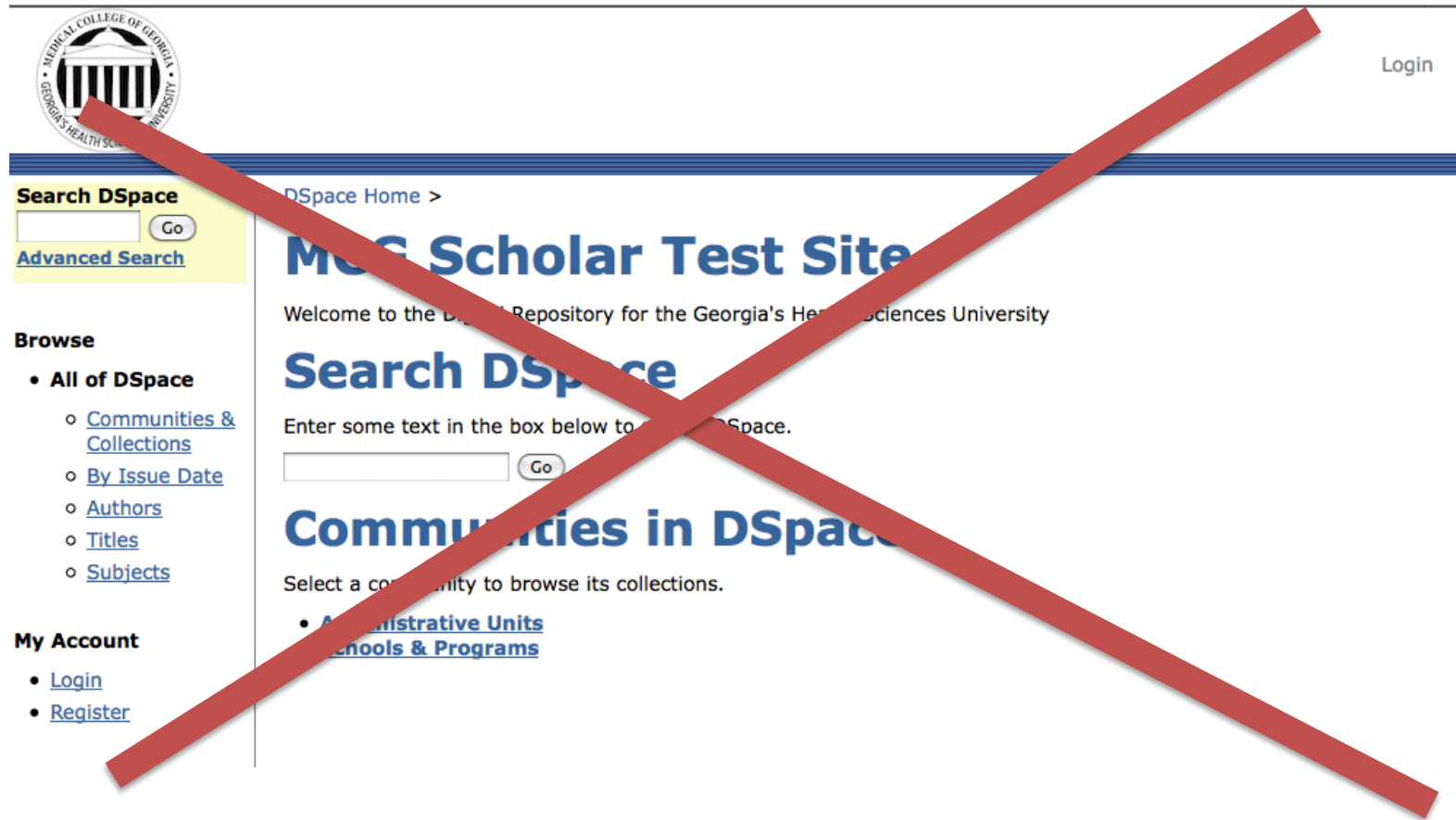
The screenshot shows the Kent State University Libraries website. The header includes the Kent State University logo (1910-2010) and the text "UNIVERSITY LIBRARIES". A navigation menu contains links for Home, Research, Services, About Us, Help, Contact Us, Search, Site Index, and Personal. The breadcrumb trail reads: "DRC Home > University Libraries > Special Collections and Archives > Kent State Centennial Collection". The main heading is "Kent State Centennial Collection". Below it is a search bar with a "Go" button and radio buttons for "Search DRC" (selected), "This Collection", and "Advanced Search". A "Browse" section lists "All of DRC" (Communities & Collections, By Issue Date, Authors, Titles, Subjects) and "This Collection" (By Issue Date, Authors, Titles, Subjects). A "My Account" section includes "Login" and "Register". The main content area describes the collection, provides copyright and license information, and lists recent submissions: "Franklin Hall Construction" and "McGivrey Hall Construction", each with a thumbnail image.



The screenshot shows the Albany State University website. The header includes the Albany State University logo and the text "Albany State University James Pendergrast Memorial Library". A navigation menu contains links for Main, Policies, Services, Periodicals, Galileo, Online Catalog, Resources, and Subject Guides. The breadcrumb trail reads: "Ram Scholar Home > The James Pendergrast Memorial Library > Inside ASU Newsletters > Inside ASU Newsletters Collection". The main heading is "Inside ASU Newsletters Collection". Below it is a search bar with a "Go" button and radio buttons for "Search Ram Scholar" (selected), "This Collection", and "Advanced Search". A "Browse" section lists "All of Ram Scholar" (Communities & Collections, By Issue Date, Authors, Titles, Subjects) and "This Collection" (By Issue Date, Authors, Titles, Subjects). A "My Account" section includes "Login" and "Register". The main content area includes a copyright notice, a "Recent Submissions" section with two entries: "January 19, 2010" and "January 8, 2010", each with a thumbnail image of an "Inside ASU" newsletter.

Thank you Jeff Craft (OhioLINK), Kent State,
Bill Anderson (Georgia Tech)!

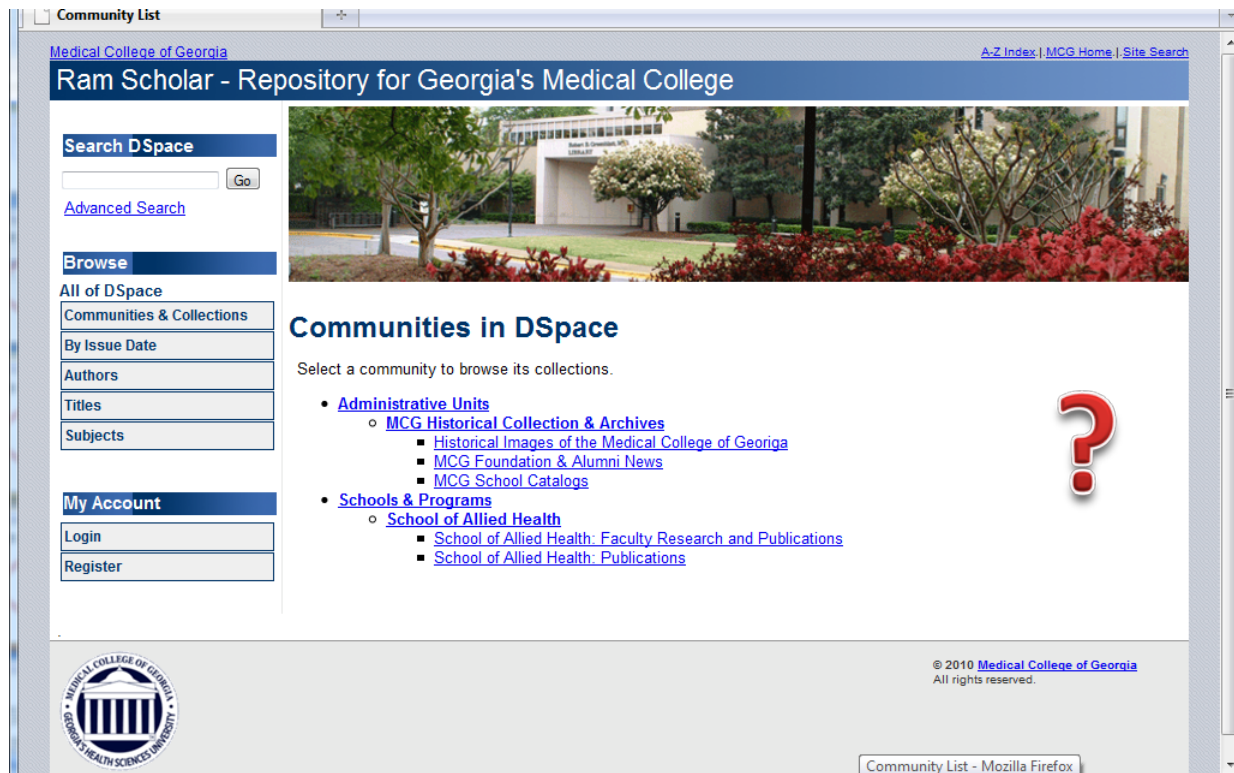
Medical College of Georgia, Strike 1



The screenshot shows the DSpace interface for the Medical College of Georgia. A large red 'X' is drawn across the entire page, indicating that the content is blocked or unavailable. The visible elements include:

- Top Left:** Medical College of Georgia logo.
- Top Right:** Login link.
- Search DSpace:** A search bar with a 'Go' button and a link to 'Advanced Search'.
- Browse:** A list of navigation options:
 - All of DSpace
 - Communities & Collections
 - By Issue Date
 - Authors
 - Titles
 - Subjects
- My Account:** Links for Login and Register.
- Main Content:**
 - DSpace Home >
 - MCG Scholar Test Site**
 - Welcome to the Digital Repository for the Georgia's Health Sciences University
 - Search DSpace**
 - Enter some text in the box below to search DSpace.
 - Communities in DSpace**
 - Select a community to browse its collections.
 - Administrative Units
 - Schools & Programs

Medical College of Georgia, Strike 2



Community List

Medical College of Georgia [A-Z Index](#) | [MCG Home](#) | [Site Search](#)

Ram Scholar - Repository for Georgia's Medical College

Search DSpace

[Advanced Search](#)

Browse

All of DSpace

Communities & Collections

By Issue Date

Authors

Titles

Subjects

My Account

Login

Register

Communities in DSpace

Select a community to browse its collections.

- [Administrative Units](#)
 - [MCG Historical Collection & Archives](#)
 - [Historical Images of the Medical College of Georgia](#)
 - [MCG Foundation & Alumni News](#)
 - [MCG School Catalogs](#)
- [Schools & Programs](#)
 - [School of Allied Health](#)
 - [School of Allied Health: Faculty Research and Publications](#)
 - [School of Allied Health: Publications](#)

© 2010 [Medical College of Georgia](#)
All rights reserved.

Community List - Mozilla Firefox



MCG Theme

Medical College of Georgia wants the DSpace repository to be viewed as the *Institutional* Repository:

Theme will *not* look like the library website; it will look more like the main institution website

Overheard at OR2010

☐ Institutional Repository software could benefit greatly from the work of design and interface professionals

Work that the Technical Lead is still waiting for others to do...



Backup and Recovery Plan

Currently in development:

- Necessary for credibility as a trusted service
- Rescue from hardware problems, developer mistakes, natural disasters, unnatural disasters
- Must be tested!

Early thank you to DuraCloud developers and anyone who sends me a sample backup plan – please?



Inline Media Playing on Remote Repositories

- Some IRs now allow the user to play embedded multimedia content from the browser (example: @mire A/V streaming module for DSpace)
- We wish to allow playback of audio and video files stored in remote repositories (Example: Play a University of Georgia file from the GKR metadata site)



Full-Text Searching of Remote Repositories

DSpace Harvester includes OAI-ORE support

Trick 1: Import Thumbnails for Harvested Records

Trick 2: Allow full-text search of articles for which GKR would have had only the metadata records



IR-Related Services

To reduce barriers to recruiting scholarly content:

- Copyright Research Services
- Digitization Services
- Content Submission Services
- Digital Preservation Services



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

Significance

Current use: *

	Daily	Several times a week	Once or twice a month	Several times a year	Once a year
How often is the material accessed in analog form?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Potential use: *

	1	2	3	4	5
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?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Barriers: *

	1	2	3	4	5
One a scale of 1-5, 1 highest and 5 lowest, to what extent does the analog form limit use of the material?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Context: *

	Yes	No
Is the collection part of a larger digital collection?	<input type="radio"/>	<input type="radio"/>
Would scanning of these materials complete an existing digital collection?	<input type="radio"/>	<input type="radio"/>
Are there related digital collections online?	<input type="radio"/>	<input type="radio"/>

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

MANDATORY

Element	Qualifier	Status	Scope Note	Instructions / example
date	accessioned	M	Date DSpace takes possession of item.	Provided by DSpace. Example: 2005-01-04T18:08:50Z
date	available	M	Date or date range item became available to the public.	Provided by DSpace. Example: 2005-01-04T18:08:50Z
date	issued	M	Date of publication or distribution.	Example: 2001-01
identifier	uri	M	Uniform Resource Identifier	Provided by DSpace. Example: http://hdl.handle.net/123456789/60
description	Provenance (submitted)	M	History of object.	Provided by DSpace.
description	Provenance (approved)	M	History of object.	Provided by DSpace.
description	Provenance (available)	M	History of object.	Provided by DSpace.
				Draft standard language regarding use

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 the 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
 - Will inform GKR's development / Serve as model survey

The GKR Model

Last deliverable:

- 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

IMPACT: GKR

- Increase USG visibility to its 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/>

Thank You!

Tyler Walters

tyler.walters@library.gatech.edu

Sara Fuchs

sara.fuchs@library.gatech.edu

Keith Gilbertson

keith.gilbertson@library.gatech.edu

GKR Manager: marlee.givens@library.gatech.edu