

iCAMP:

Curate, Archive, Manage, Preserve

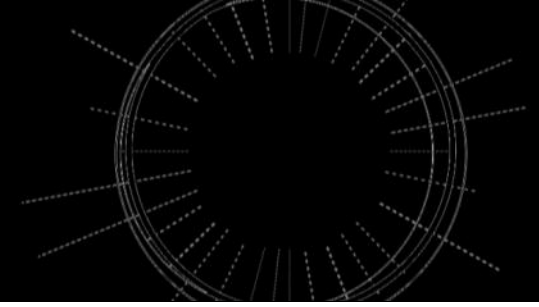
A review of digital curation
curriculum development
efforts at UNT and other institutions



Martin Halbert
UNT Dean of Libraries
CNI Fall 2012 Membership Meeting
Monday, December 10, 2012
Washington, D.C.

UNT[®]
UNIVERSITY
OF NORTH TEXAS™

Overview



1. Research data management curricula gaps
2. Summary of other key projects & findings
3. iCAMP project aims and research methodology
4. Review of iCAMP results
5. Questions raised for the future



Wordle™ of NEH-ODH Data Management Policy

UNT®

UNIVERSITY
OF NORTH TEXAS™

The (Now Obvious) Gap

There are too few people, too little preparation, and too little institutional infrastructure available to manage too much data being generated from too many sources.

"By 2018 the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know how to use the analysis of big data to make effective decisions." –McKinsey Global Institute, Big data: The next frontier for innovation, competition, and productivity (2011)

"How to make sense of all these data? People should be worried about how we train the next generation, not just of scientists, but people in government and industry." –Johns Hopkins researcher, reported in Economist report Data, Data Everywhere

While 75% of the information in the digital universe is generated by individuals, enterprises have some liability for 80% of information in the digital universe at some point in its digital life. –IDC 2011 Digital Universe Study



How do we bridge a Gap like this?

- Our institutions and associated educational curricula are not yet able to effectively organize and manage the research data being generated
- Massive data loss is occurring and will continue to occur for some time
- It will take years of effort to build up the infrastructure and expertise required to administer the zettabytes of information now being created by society



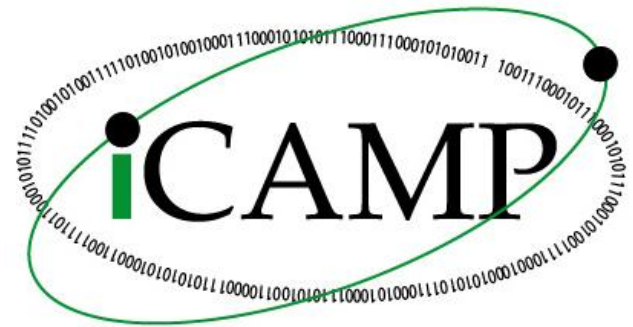
Note: you don't cross the Grand Canyon in one day, you do it in stages. You camp down in the middle at the end of the first day and hike up the other side the next day.

UNT[®]

UNIVERSITY
OF NORTH TEXAS™

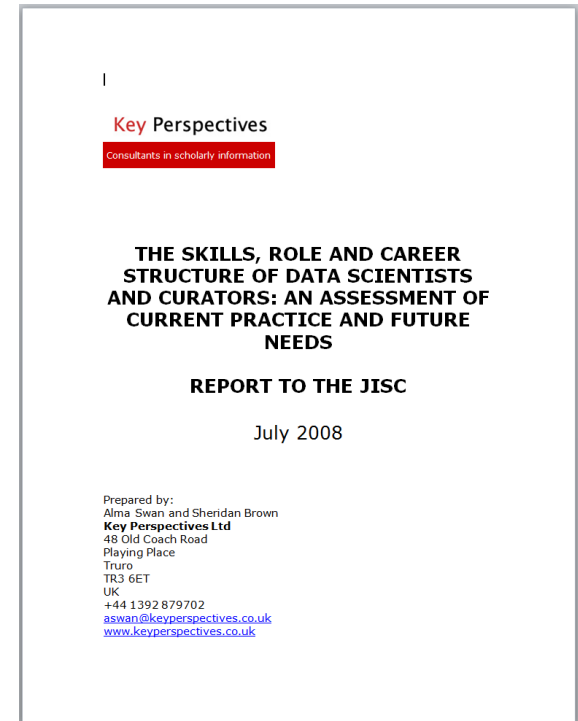
iCAMP Project: Driving Question

- *How can LIS education be changed and enhanced*
- *to result in well-prepared information professionals, scientists, and scholars*
- *that can take on the challenges and problems of digital curation, data management, and digital preservation,*
- *incorporating extensive training and practical experience*
- *in the context of contemporary distributed learning and web-based courses?*



2008 JISC Report: Skills, Role and Career Structure of Data Scientists and Curators (Key Perspectives Ltd)

- Distinguished four roles: *"data creator, data scientist, data manager and data librarian"*
- *"There is no defined career structure for data scientists and this is a major problem"*
- *Little agreement on what training/preparation is appropriate (Formal? On the job? Continuing ed?)*
- *There is a "continuing challenge in remaining properly skilled up" in data science skills*
- *"The role of the library in data-intensive research is important and a strategic repositioning of the library with respect to research support is now appropriate" (however libraries are poorly situated for this repositioning)*



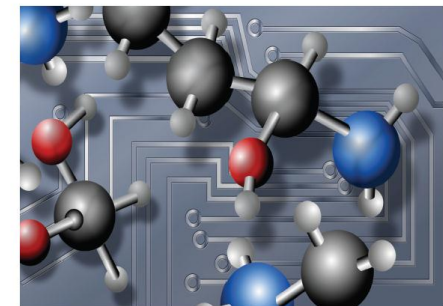
2010 ARL E-Science and Data Support Services Report



- 87% indicated that support efforts were collaborative, involving data centers (disciplinary or central academic computing services), research groups, library data and institutional repositories”
- The a majority of libraries report a mixed response strategy of reassigning existing staff, providing training to existing staff, or hiring new staff
- “many libraries are uncertain about the kind of expertise needed”, but of the 65 positions studied, 72% had degrees in library and information science at the master’s or PhD level for current or planned positions
- Whatever the educational background, the ability to become a trusted colleague of researchers is key
- Most ARL institutions report that e-science support efforts at their universities are hampered by a lack of institutional consensus and resources

E-Science and Data Support Services A Study of ARL Member Institutions

August 2010



Catherine Soehner, University of Michigan
Catherine Steeves, University of Guelph
Jennifer Ward, University of Washington

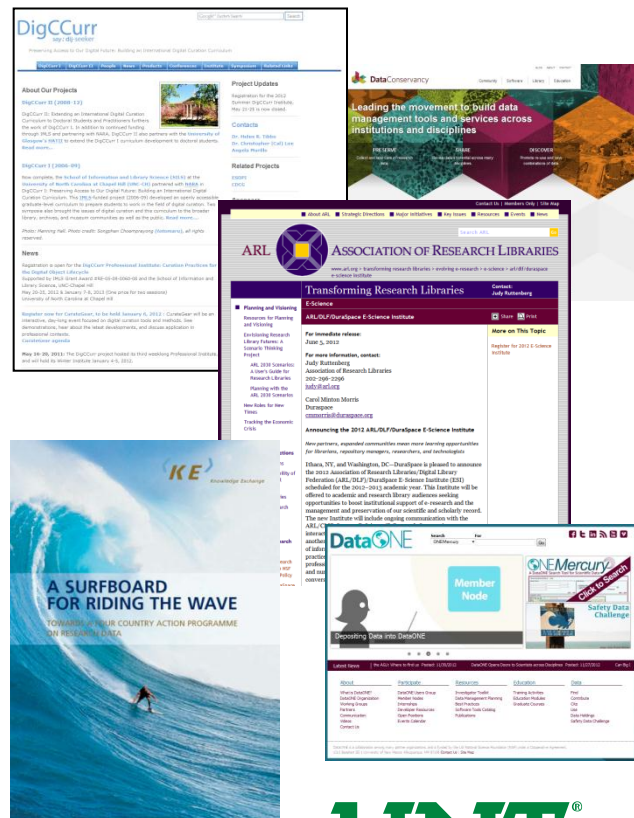
Surveys of the Data Science Trend, and the Gap



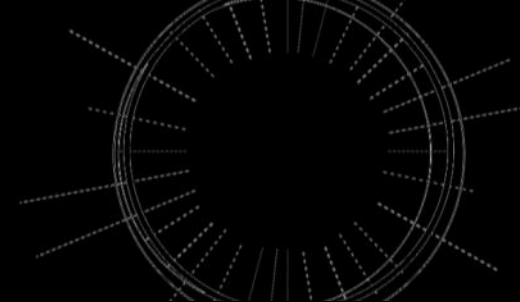
- 2008 Swan Report to JISC on Data Skills
- DigCCUR I and II (UNC Chapel Hill)
- E-Science and Data Support Services (ARL)
- Research Data Alliance
- ARL Institute
- Academic Librarians and Research Data Services: Preparations and Attitudes (University of Tennessee)

Additional relevant projects

- DigCCUR I and II projects by UNC Chapel Hill
- ARL E-Science Institute
- 2011 European Knowledge Exchange report in response to EC Riding the Wave report
- Data Conservancy and DataOne
- Projects underway at individual iSchools in response to Big Data trend and data management planning

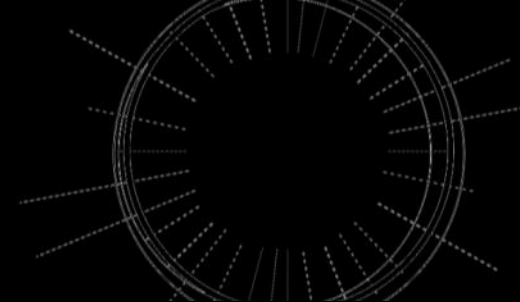


iCAMP Project



- 3-year IMLS 21CL capacity building project (9/1/2011 through 8/31/2014)
- 2011 IMLS 21CL \$624,663 grant (paired w/ 2011 21CL \$226,786 DataRes grant)
- Collaboration between
 - UNT College of Information (Texas Center for Digital Knowledge and Dept. of Library and Information Sciences)
 - UNT Libraries

iCAMP objectives



- Develop four courses (web-based) in digital duration and data management (*curricular content will be freely available to others*)
- Deploy a robust technical infrastructure to support student learning, practical engagement, and training
- Explore the concept of a “teaching library”
- Implement a virtual teaching environment for students’ active experimentation and discovery learning (*tools will be open-sourced and freely available*)

iCAMP Primary Beneficiaries



- Graduate students majoring in Information Science
- Post-masters information professionals retraining for digital curation jobs
- Discipline-specific graduate students who may be responsible for managing data

Terminology Variations Encountered

- Digital curation
- Digital preservation
- Digital archiving
- Data management
- Data science
- eScience
- Data analytics



iCAMP Focus on Competencies



- **Competency** (working definition):

A cluster of applicable knowledge, skills, abilities, and attitudes that can be improved with training and professional development, affect the job, and correlate with performance on the job; Measurable against set standards

- **iCAMP Competency:**

Knowledge, skills, abilities, and attitudes to carry out a wide range of professional functions in support of digital curation and data management responsibilities.

Methodology for Competencies Analysis



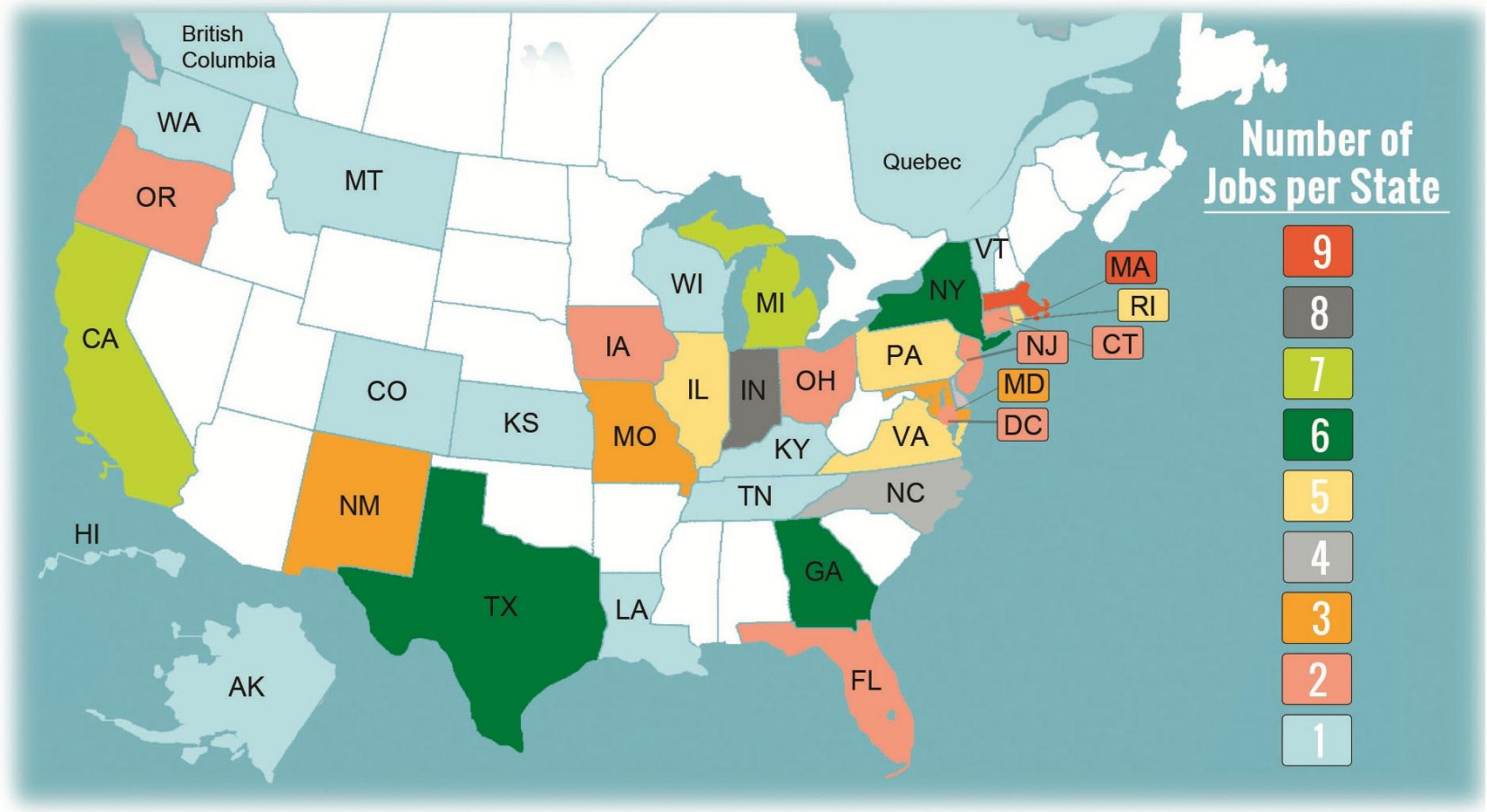
- Sources of job announcements were monitored for several months, including *ALA JobLIST*, *ARL's Job Announcements*, *LIS Jobs*, and *Digital Curation Exchange Jobs* that cover all regions of the United States and Canada
- Primary search terms used to locate relevant job advertisements included any derivative terms for curate, preserve, archive, and manage – e.g., curation, curator, preservation, archiving, archivist, management, manager, etc.
- A second set of terms was then used in combination with the primary terms to narrow search results; those terms included data, information, digital content, digital object, digital asset, digital information, digital collection, digital data, research data, and scientific data

Methodology for Competencies Analysis (cont.)



- An initial set of 110 job advertisements (collected between October 2011 and March 2012) comprised the corpus for this analysis
- Data were entered into NVivo qualitative analysis software, and content analysis coding techniques were used
- Data were categorized into five dimensions: position title, institution, degree, experience, and skills and knowledge, and then determined frequencies and patterns of occurrence of specific job characteristics and requirements.

Job Advertisements Analysis



N = 110

UNIVERSITY
OF NORTH TEXAS™

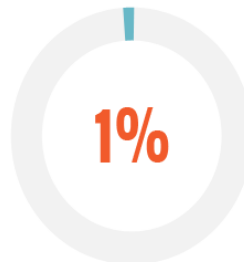
Job Advertisements Analysis: Education Requirements



85% of the job ads required or preferred an ALA-accredited Master's degree as an educational qualification for the job

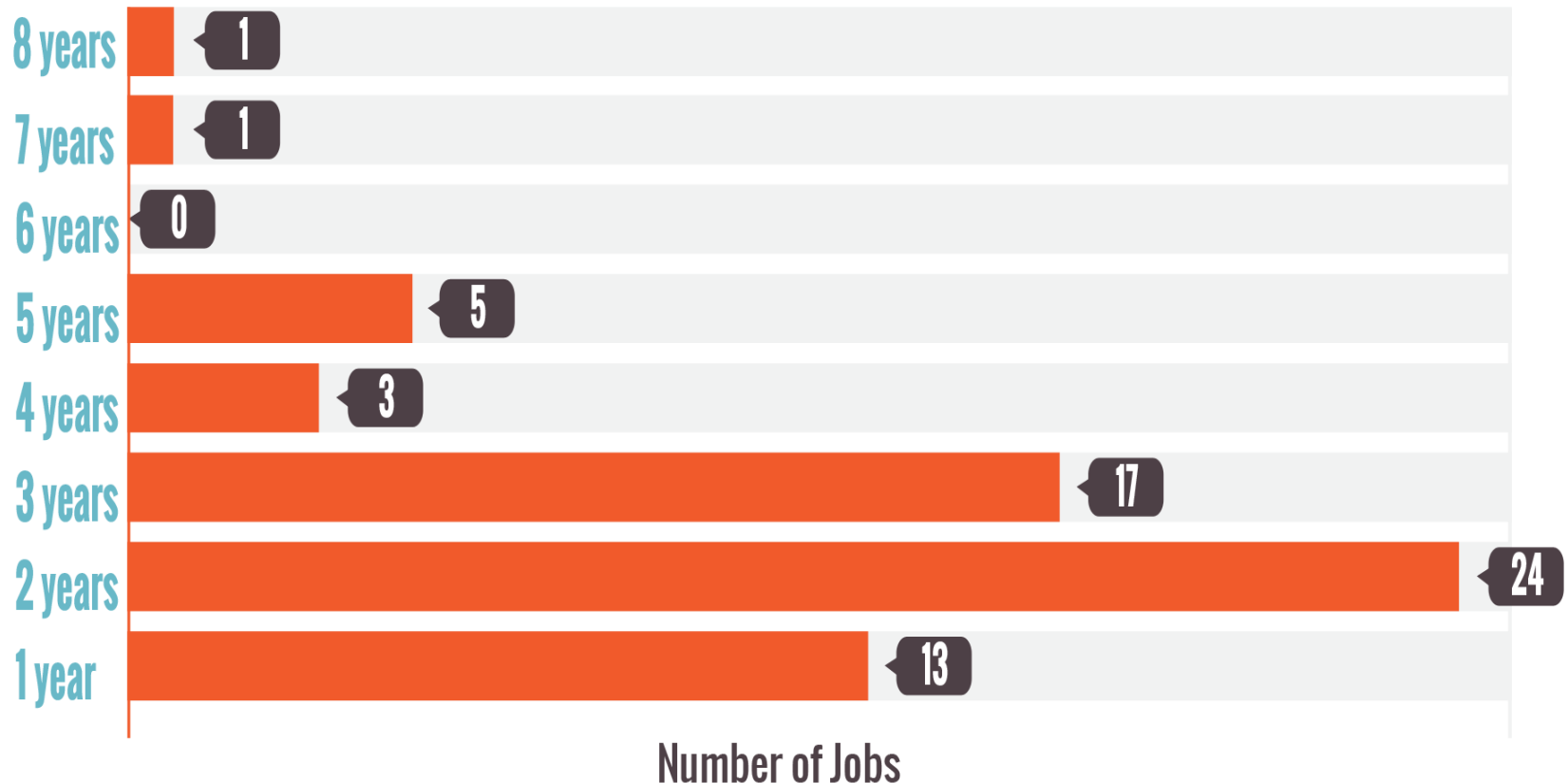
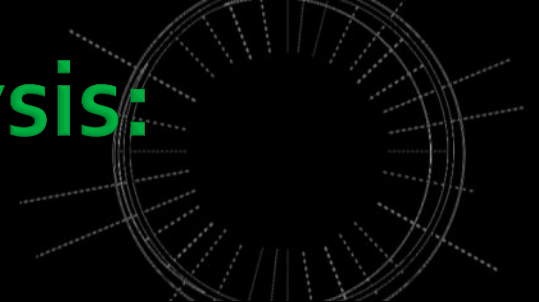


28% of the job ads referenced the requirement or preference for a degree in other disciplines (in addition to or instead of a degree in LIS)



1% of the job ads required a PhD

Job Advertisements Analysis: Experience



Job Advertisements Analysis: Skills and Knowledge



Areas	No. of Job Ads	Percent
Working in an Information Technology Intensive Environment	64	58%
Standards and Specifications	60	55%
Project Management	49	45%
Personal and Interpersonal Skills	47	43%
Research and Trends	41	37%
Tools and Applications	35	32%
Liaison and Support	34	31%
Functional Skills for Curation	27	25%
Working Knowledge for Curation	25	23%
General Library/Archive Skills	24	22%
Professional Development	17	15%
Other Domain Knowledge	4	4%

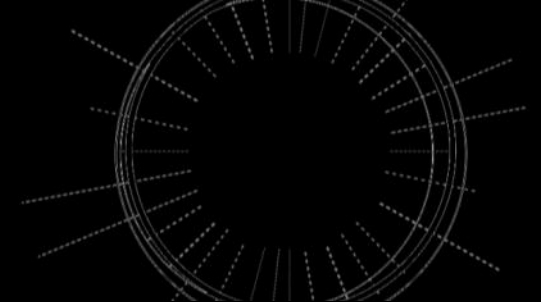
Source: Kim, J., Warga, E., & Moen, W. (2012). *Digital curation in the academic library job market*. Presented at the ASIST 2012 Annual Meeting

iCAMP Competency Groups



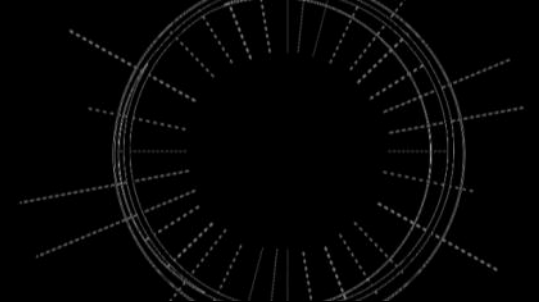
- Communication & Interpersonal
- Curating and Preserving Content
- Curation Technologies
- Environmental Scanning
- Management, Planning, and Evaluation
- Services
- Systems, Models, and Modeling

iCAMP Courses



1. Digital Curation and Data Management Fundamentals – Spring 2013
2. Tools, Applications, and Infrastructure for Digital Curation – Spring 2013
3. Managing and Preserving Digital Data and Information – Summer 2013
4. Digital Curation and Data Management Seminar – Fall 2013

iCAMP Technology



- Traditional learning management system: Blackboard learn
- Virtual lab: Drupal Islandora Fedora platform
- Sandbox: UNT Libraries technology environment

iCAMP Desired Outcomes



- Increase in students' preparedness by enhancing practical training
- Positive change in discipline-specific graduate students' knowledge and perceptions
- Positive effects on distributed, online LIS education
- Improved LIS faculty and UNT librarian's abilities and attitudes for collaboration in LIS education

Questions and concerns for the future raised by this project

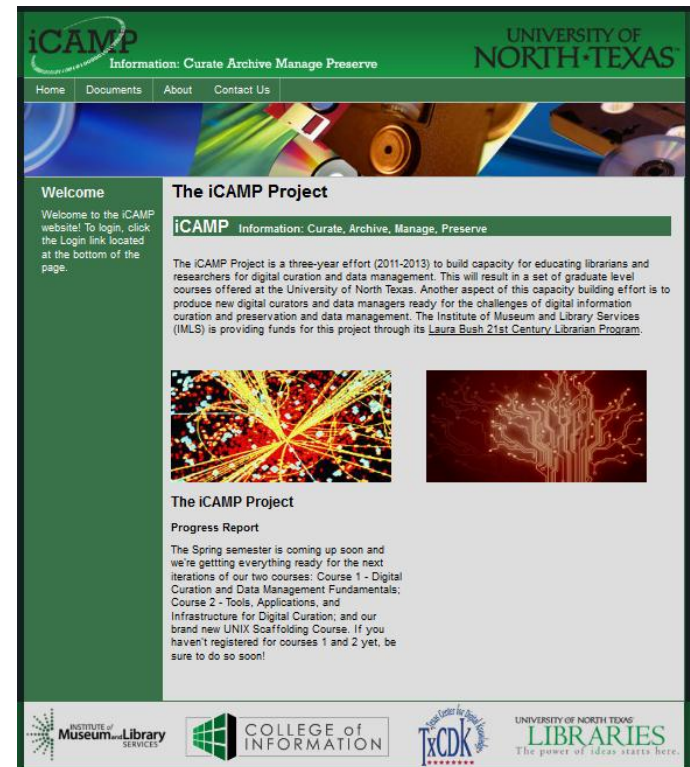


- Our analysis primarily examined recruiting in the LIS sector; to what extent will libraries be engaged in research data management?
- Libraries continue to mostly hire LIS graduates for these roles; will LIS programs be able to adapt to these new needs and provide effective preparation for such roles?
- Even if LIS programs adapt, will they be agile enough to continue to adapt to meet new needs as the field keeps changing rapidly?

iCAMP Project URL and Staff

<http://icamp.unt.edu>

- Principal Investigators
William E. Moen, College of Information
Jeonghyun Kim, College of Information
Martin Halbert, UNT Libraries
- Professional Staff
Ana Krahmer, UNT Libraries
Mark Phillips, UNT Libraries
- Graduate Research Assistants
Brenda Cantu, Paulette Lewis,
Jacqueline Salter, Edward Warga,
Library and Information Sciences
Joseph Helsing, Computer Science and
Engineering



UNT[®]
UNIVERSITY
OF NORTH TEXAS™

Questions and Answers

