The Library eResearch Task Force was established in April 2012 with the charge of recommending: a) the scope of Library-supported eResearch activities, b) organizational responsibilities for such work, and c) resource requirements. The Task Force is performing a local, as well as a national, environmental scan and has devoted time to identifying current services and activities that the Library/campus are pursuing, as well as gaps and future needs. The group’s charge asked us to perform several tasks, each of which will be addressed in a report that is due on December 31. These tasks include:

- Assessing immediate needs for expanding cyberinfrastructure support, training workshops, and data management plan support;
- Identifying training and re-tooling needs among Library faculty and staff, and other individuals who are candidates for training;
- Reviewing existing programs in the Scholarly Commons and other Library units;
- Recommending organizational responsibilities for research data services provided by the Library.

As way of background, we would like to note the following operating definition for eResearch:

Eresearch refers to the concept of research using digital technology (e.g., computing, networks, digital data) in fields including science, social science, and the humanities. Eresearch can be described as collaborative, computing and data intensive, and often interdisciplinary. Libraries are uniquely qualified to manage, curate, and steward research data. Research data are the numerical, descriptive, or visual facts, observations or experiences on which an argument, theory or test is based. These may be raw or analyzed, experimental or observational. While archivists and librarians recognize that research data represent a portion of a broader information context, the interests of scholars, funding agencies, and academic institutions have focused on preserving and managing the specific output of the research activities in which scholars engage. Funding agencies, institutions, and researchers have a high interest in ensuring that research data is well-managed and accessible for validation and reuse of research results. Libraries (including archives) because of their accessibility on many levels, and their expertise in the provision of access, information management, and long term preservation are well positioned to support these activities.

At this time, we would like to report several of the gaps that we have identified in the Library’s eResearch/data management programs, in order to initiate a discussion about how the Library might address these needs.

**Gap Area: Current Library Support for Research Data Creation, Acquisition, Use, and Curation**

At present, data services and eResearch support are distributed throughout the Library. The Task Force recognizes and applauds the efforts of the Library to develop, fund new, and re-shape existing roles to support eResearch. At the same time, the Task Force intends to recommend that these efforts to articulate research data services in Library professional roles be stepped up considerably. Several data services positions have recently been appointed. Units and individuals have responded independently to this emerging service need.
Data management support has recently been included in the core competencies for subject specialists, and recent public service faculty position announcements include data service needs in the position descriptions. However, the overall hiring activity does not indicate a strong a programmatic commitment to this emerging campus-wide need, in which the Library is expected to play a key role. The hiring approach has not been as focused or strategic as that of peers (e.g., Purdue, University of California San Diego, University of North Carolina at Chapel Hill, University of Virginia, Penn State) in organizing to address eResearch services and functions.

While the initial approach to eResearch support was to encourage services to develop at the unit level, both the Library and the campus have come to the point where the need now requires the Library to ensure a consistent baseline of facility in all units with research data support. A highly decentralized approach to any kind of support services that have a common thread ends up foundering if it doesn't have some basic organizational foundation. There are bound to be gaps in service (i.e., researchers who do not know how to get assistance to acquire, manipulate, or develop data management plans in their subject area) in the current arrangement.

**Gap Area: Professional Development, Training, Communication, and Cohesive Public-facing Services**

While support for research data services cannot be confined to one unit, it should have a cohesive base, including professional coordination, and it needs to be distributed throughout subject and important functional areas. Demand increases daily for public-facing programs, internal communication, and one-on-one consultation with campus researchers who are creating and using data, but who do not have access to appropriate services. The Task Force has identified a number of “immediate” needs for public-facing programs, ongoing monitoring of research data developments across funding agencies and a variety of subject fields, and professional development programs focused on supporting librarians and other Library professionals to obtain training and re-tooling to focus their work on research data support:

- **Training & Instruction**
  - Researcher-focused workshops on a variety of topics related to research data management, curation, and use. Coordinate them Library-wide through the Scholarly Commons Savvy Researcher workshop series.
  - Subject specialists and functional specialist training workshops on digital scholarship tools, principles of data curation, description of data services infrastructure, and user services for data management; external, internal, and with GSLIS/CIRSS;
  - Sponsor external presenters for Research Data Management Workshops/meetings, spring 2013 (already planned with UIC, NUL, UChicago).

- **Communication & environmental scanning**
  - Establish tight coordination between Scholarly Commons and unit-level data management web and other virtual services;
  - Gather and analyze researcher requirements;
  - Working with the OVCR, invest in ongoing analysis of NSF and other agency data management plans;
  - Invest more in development of the DMPTool — web-based data management planning wizard;

---

• Identify support to expand existing Scholarly Commons research data services (web documentation, data manipulation, instruction, and training);
• Establish Library active participation in local, regional, national & international professional groups and activities (BRDI, RDAP, DCC, IDCC, etc.);
• Monitor and communicate about funder requirements, changes in research policy and requirements at the campus/university levels, and policy and best practices for data management and curation in the LIS field and across subject domains;
• Track Big Data and other eResearch funding Programs and Proposals, and communicate with Library and other campus groups about local opportunities and impact;
• Engage in campus-wide researcher profile service conversations, including expansion of support for BibApp/Connections researcher profile service.

• Consultation in & support of research data use
  ▪ Establish a core group of subject and functional experts with operational activities that contribute to the Research Data Service program;
  ▪ Develop and sustain a virtual service catalog of research data services available to Urbana campus researchers, on and off-campus;
  ▪ Develop research data consulting service guidelines for subject liaisons;
  ▪ Expand the Library’s Data Purchase Program;
  ▪ Develop and publicize best practices for data management and curation within and across disciplines, working with experts in discipline-specific needs, funder requirements, and those who specialize in preservation and metadata best practices;
  ▪ Identify ways incorporate specific needed expertise from elsewhere on campus—e.g. IGB; NCSA, Beckman, OVCR, etc.

The Library should be able to provide both basic and in-depth research data services as needed by the campus. Through the Task Force’s examination of the work of peer institutions across the U. S., Canada, and in the United Kingdom, it is readily apparent that our peers are doing this with greater consistency than Illinois, and they have made significant shifts in resource allocation to support it (see programs at Purdue, UC San Diego, Columbia). The University of Illinois is one of the largest public “research intensive” universities in the U. S., and the Library ought to make the resource commitments to play an active role in supporting researchers to maintain their competitive edge in this area of strategic importance to the institution.

Gap Area: Research Data Management, Curation, and Preservation Service

While the needs identified above address user-facing needs, there is a further unmet need on campus and in the Library for research data curation services--repository services, data publishing, and data transformation, migration and management. This service would work hand-in-hand with librarians and other specialists. It also is required for the Library to be able to participate in a campus-wide service; by doing so we can help ensure that data repositories and publishing services operate in a way that will facilitate the identification and preservation of materials having long term research value, and can work out pathways to ensure that preservation as a campus or Library-managed resource. To that end, the Task Force intends to recommend that the Library establish a Research Data Service program that provides support to Urbana campus researchers in the use, management, preservation, and access to digital data that researchers create, acquire, or manipulate.

This service is proposed to be a core component of a campus-wide research data services program provided in partnership with the OVCR, CITES, NCSA, GSLIS, and the Graduate College. This service could be modeled after
the relationships established through the IDEALS Repository and Scholarly Communications Program. At the request of the Vice Chancellor for Research, the Library, OVCR, CITES, NCSA and GSLIS are preparing a proposal for a suite of research data curation and cyberinfrastructure services for consideration by the Provost and Chancellor. This “component” of the Library’s eResearch services would be considered a core part of the campus-wide support for research data management and long-term preservation.

It will be critical for the Library portion of the proposed campus –supported Research Data Service to be situated in the Library (as are IDEALS and Medusa), and to work seamlessly with the Library’s subject specialists, data services librarians, archivists, and functional specialists to ensure that a consistent level of research data services are available to researchers on the Urbana campus. These services will include metadata support, technology training, repository development, digital preservation/curation, and archival appraisal, accessioning, arrangement, description, and access. The required elements of the Research Data Service, that are not currently supported in the Library, or anywhere on campus, are the following:

- **Data curation services**
  - Assistance in the development of data management plans (in collaboration with subject specialists and the Scholarly Commons)
  - Data transfer and ingest support services
  - Metadata creation support
  - Data preservation services
  - Convenient access to data replication, and integrity checking services

- **Data discovery, integration, and publication services**
  - Data set registration with DataCite
  - Data set provenance and reuse description
  - Data submission process support
  - Data set availability alert services
  - Ontology development
  - Portal development (access to data and services)
  - Linking of data sets to resulting publications
  - Linking of data sets to resulting analysis, mining results, visualization
  - Archival appraisal, arrangement, description and long term preservation (in collaboration with University Archives, IDEALS, and Library Preservation)

- **Data analysis and visualization services**
  - Data set transfer and transformation
  - Database creation and management
  - Data analysis support services (in collaboration with the Scholarly Commons)
  - Data mining support services (in collaboration with the Scholarly Commons)
  - Data visualization support services (in collaboration with the Scholarly Commons)

Libraries are uniquely qualified to manage, curate, and steward research data. While archivists and librarians recognize that research data represent a portion of a broader information context, funding agencies, institutions, and researchers have a high interest in ensuring that research data are well-managed and accessible for validation and reuse of research results. Libraries and archives, because of their expertise in the
provision of information management, preservation, and access services are well positioned to support campus activities to manage and curate research data.

The Library eResearch Task Force:

Susan Braxton
Harriett Green / Caroline Szlowicz
Tom Habing
Karen Hogenboom
Chris Prom
Mary Schlembach
Sarah Shreeves
Sarah Williams
Beth Sandore Namachchivaya (ch.)

Data Management RSS feed:

http://illinois.edu/lb/iList/5001

eResearch Task Force Committee page:

http://www.library.illinois.edu/committee/eresearch_task_force/eresearch_task_force_charge.html