University of Manitoba Libraries Data Management Strategic Agenda

Prepared by:

Vera Keown, Head, Sciences and Technology Libraries, University of Manitoba Libraries

Carell Jackimiek, Manager, Digital Initiatives, University of Manitoba Libraries

Jeffrey Marcus, Canada Research Chair, Department of Biological Sciences, Faculty of Science

Gary Strike, Liaison Librarian, Elizabeth Dafoe Library, University of Manitoba Libraries

Revised on April 17, 2014

I. Introduction

Scientific discovery can be characterized by three paradigms, experimental, theoretical and computational science. In 2007, Dr. Jim Gray, a database software researcher at Microsoft, gave a presentation where he described how the exponential growth in computing power was driving a fundamental change in the way scientific research is conducted. He called this shift the "fourth paradigm". Gray's fourth paradigm describes the new era in scientific research characterized by data-intensive, collaborative and multidisciplinary scientific inquiry. These are the driving forces behind the need for e-science and data management initiatives. Such volumes of data will require specialized tools and skills in order to be acquired, managed, analyzed, visualized, shared and preserved. The term e-Science (and its related term e-Research, which covers all disciplines) is used to describe the development of the tools and technologies required to support this fourth paradigm of research. "E-science departs from well-established experimental and theoretical methodologies with its large-scale, data-driven, and computationally intense characteristics. E-science fundamentally alters the ways in which scientists carry out their work, the tools they use, the types of problems they address, and the nature of the documentation and publication that results from their research. E-science requires new strategies for research support and significant development of infrastructure"[1].

Research is an expensive and time consuming endeavor. However, with vast amount of data being created and stored digitally, it is now possible, in theory, for researchers to more easily manage, share, and re-use data, allowing researchers to build upon existing research, verify results, and combine data and information within and across disciplines to form new research outputs. However, in order to do so, data must be created and maintained throughout the research life-cycle in a manner that allows for its long-term preservation. The data must also be documented and described in order to be searchable. "All too often the results of expensive and time-consuming research as represented by rich data sets are lost due to the absence of sound data management plans. Redundant research is undertaken because the previous research data is no longer available. Opportunities for analysis of data across time are lost along with the historical data sets. Even when data has been properly stored and preserved it benefits no one if it isn't easily discovered, retrieved and repurposed".[2]

Funding agencies are also calling upon researchers and institutions to make research more publically accountable and accessible. Funding agencies in Europe and the US are now requiring researchers to submit data management plans along with their grant applications. These plans must stipulate how a researcher will gather, store, manage, and share their data publically in order to be eligible for public research funding. It is expected the Canadian funding agencies will follow suite in the near future.

---

[1] **Association of Research Libraries.** 2007 Evolving E-Research: ARL Joint Task Force on Library Support for E-Science [Internet]. Washington, DC. [Cited 2013 January] Available from: http://www.arl.org/rtl/eresearch/escicharge.shtml

[2] Johnston, Wayne. "Ontario Agri-Environmental Research Data Repository." Funding application. University of Guelph, 2011. Accessed January 14, 2013 https://journal.lib.uoguelph.ca/index.php/perj/article/viewArticle/2014/2638

Data preservation and sharing is a relatively new concept for most researchers. Institutions generally lack formal data management processes, tools, and training to help researchers successfully manage their data. Libraries, with their experience managing, preserving and providing access to information are seen as the ideal centres for supporting academic researchers' data management needs. "E-science has the potential to be transformational within research libraries by impacting their operations, functions, and possibly even their mission"[3]. The Association of College & Research Libraries' Research Planning and Review Committee recently identified "library involvement in data curation, including collaboration with their research communities, as one of the 2012 top ten trends in academic libraries".[4]

However, data management is a massive undertaking that will require collaboration amongst many institutional players. The role of the library in any institution's e-Science initiative can be summed up by Lewis' quote on "Libraries and the management of research data":

> *"Perhaps the starting point for any discussion about libraries and research data is to ask whether managing data is actually a job for university libraries. The answer to this question is a straightforward yes and no. Yes, in the sense that data from academic research projects represents an integral part of the global research knowledge base, and so managing it should be a natural extension of the university library's current role in providing access to the published part of that knowledge base. No, because the scale of the challenge in terms of infrastructure, skills and culture change requires concerted action by a range of stakeholders, and not just university libraries."*[5]

"While libraries may have little immediate engagement in these processes, clearly e-science has the potential to be transformational within research libraries by impacting their operations, functions, and possibly even mission"[6].

What follows is a strategic agenda for the University of Manitoba Libraries (UML) which outlines a set of coherent and aligned strategies which can serve as the UML's platform for its e-Science strategic plan. The strategic agenda will guide implementation over the course of a few years as it is translated into the organization's strategic plan.

---

[3] Association of Research Libraries. 2007 Evolving E-Research: ARL Joint Task Force on Library Support for E-Science [Internet]. Washington, DC. [Cited 2013 January] Available from: http://www.arl.org/rtl/eresearch/escicharge.shtml

[4] ACRL Research Planning and Review Committee, "2012 Top Ten Trends in Academic Libraries," *College & Research Libraries News* (June 2012): 311-320, http://crln.acrl.org/content/73/6/311.full.pdf+html

[5] Lewis, M. (2010) "Libraries and the management of research data", Envisioning Future Academic Library Services: Initiatives, Ideas and Challenges, pp145-68

[6] Association of Research Libraries. 2007 Evolving E-Research: ARL Joint Task Force on Library Support for E-Science [Internet]. Washington, DC. [Cited 2013 January] Available from: http://www.arl.org/rtl/eresearch/escicharge.shtml

II. Background

The University of Manitoba Libraries (UML) is in the very early stages of defining its role in e-Science and e-Research support services. Both the UML and the University are in the exploratory stage of e-Research/e-Science initiatives and services. Relationships, collaborations, responsibilities, goals and objectives, and strategy all need to be investigated and defined.

There is strong support from various stakeholders such as the Vice-President of Research's office, IST, and faculty, for developing an e-Science strategy at the University of Manitoba. They are keenly aware of the need to develop tools and processes for managing research data and agree that the UML would be the department best suited to the lead planning and implementation of an institutional data management program. A critical factor for success is the development or adoption of tools that allow researchers to perform E-science tasks quickly, easily, and cost-effectively so that they want to use the E-Science infrastructure being created. All parties acknowledge that additional funding would be required by the libraries and others in order to both start-up and sustain a data management program.

Information Systems and Technology (IST) does not currently have a strong mandate to support research computing needs. As a result, researchers have devised their own strategies for managing, storing, and sharing data, many of which do not satisfy the legal (e.g. FIPPA) and ethical requirements as set out by university policies on ethics and research. The present institutional IT infrastructure managed by IST cannot support the data storage and sharing requirements of researchers, nor can it be easily expanded to do so.

The Research Quality Coordinator (Office of the VP Research) has been working on getting support from various departments for the development of a Collaborative Research Platform in order to satisfy the funding and privacy requirements of safely and securely storing and transmitting research data. Such a platform would provide researchers with a secure environment in which to store, analyze, and share their working data.

While the UML does not currently have library staff with research data management experience, it does have staffs that possess related skills sets and experience in digital asset management, and data and GIS products and services. The UML supports and manages an institutional repository, MSpace, which is dedicated to collecting, storing, preserving and making accessible digital versions of the intellectual output of UM community members. The UML is currently developing a Digital Asset Management system that, with customization, could be used to store and share research data. The Libraries' subscribes to a number of statistical data sources and provides dedicated staff to support social sciences and statistics teaching and research. The Libraries' also provides access to Geographical Information System (GIS) tools such as ARC GIS, AutoCAD, Geomatica, and Google Earth Pro, as well as support via the GIS Librarian.

The Office of Research Services had concerns that researchers would not welcome the intrusion into their data management by other units within the organization. They felt that unless mandated by the funding agencies, researchers would resist internal policies and procedures for the formal management of data though the research life-cycle. They felt that since the Libraries was already seen as a repository, that the researchers would be more acceptable of UML's involvement in managing published data (though not raw data).

Finally, while Canadian funding agencies have yet to introduce data management policies, it is expected that they will follow American, European, and Australian agencies who require researchers to submit and follow through data management plans in order to be awarded funding. By committing to developing data management programs and processes now, the University of Manitoba will be well positioned to respond to these requirements in the near future.

Strengths:

- UML is well-respected across campus and generally have good relationships with faculty and the university administration.
- Skill development potentially relevant to e-Science/data management
    - Metadata skills were developed in some staff in managing the institutional repository, digital asset management, GIS, and sourcing secondary data.
    - Some staff members have advanced degrees in disciplinary fields.
- UML has strong consortial ties with Canadian academic libraries who are also working on developing data management programs and expertise. UML is currently working with other western university libraries to put forth a proposal to COPPUL for the development of shared storage and data preservation facilities and resources.
- One of the goals in the Libraries Strategic Plan is to integrate librarians into the research process.
- UML is committed to supporting open access which also encompasses open data.

Weaknesses:

- Most librarians are not involved in the research process with faculty. Some may not be fully committed to e-research support as an area of emphasis for the library.
- The current IT infrastructure is not adequate to support data sharing and storage requirements.
- UML has no in-house data management expertise.
- UML cannot support data management programs with existing resources (human and financial).
- UML may decide not to take the lead on data management for the University and support will wane.

Opportunities:

- Support from various institutional stakeholders for the development of data management programs and process and to have UML lead these.

- Funding agencies are likely to introduce data management policies in the near future which would require the University to invest in data management resources.
- There is a new CIO of IST who is keen to support researchers.
- The University's mission to "Create, Preserve, and Communicate Knowledge" speaks to data management.

Threats:

- Faculty are wary of new, centralized systems
- No new funding will be made available to support data management

III. Potential Opportunities that Align with Strategic Priorities

A. Data Curation

- Hire a data management specialist.
- Develop inventories of skill sets required by existing library staff to become proficient in data management and existing skill sets. Raise staff awareness of data management skills. Develop a training program for UML staff in metadata, data curation, data management planning, and data sharing. Host a data management bootcamp for library staff.
- Investigate existing tools for developing data curation, e.g. Dataverse. Investigate the technologies and infrastructure platforms.
- Update Collection Development and Institutional Repository policies to include data curation.
- Partner with a research department or centre on a pilot project that would enable the libraries to learn more about the nature of the work involved in supporting e-science through a low-stakes endeavor.

B. Metadata

- Investigate types of data, file types, and metadata standards.
- Train librarians and technical services staff members in metadata standards and creation.

C. Support

- Include information on discipline-specific external data repositories and/or local data repositories on subject LibGuides.
- Provide information and links to data management planning tools, templates, examples, funding agency requirements, intellectual property issues, etc., via the subject guides (similar to the Open Access LibGuide).
- Support researchers in the development of metadata to for data sets that are being published alongside articles.
- Provide links to external expertise.
- Search for data management policies from other institutions and draft recommendations for related UM policies.

D. Outreach

- Host a data management day to educate faculty and administration and communicate what we are doing.

E. Facilities

- UML could work with IST and/or its consortial partners to develop a Virtual Research Environment and/or local data repository that meets the needs of researchers.
- Create a physical e-Science lab with specialized technology and software (e.g. analysis and visualization software), and trained personnel to provide support and consultation services.

F. Other

- UML would play a leadership role in bringing together internal and external stakeholders to develop a broader data management plan for the university.
- Continue to expand and refine the Libraries roles within data management through planning and policy development.
- Participate on a national data management working group, e.g. Research Data Canada.
- Investigate collaborations on data management infrastructure with COPPUL or CARL Libraries.

IV. Risk Assessment

A. Opportunity Cost

- Research data is at risk. The University already recognizes that institutional research data is at risk. The lack of institutional tools, resources, and support to faculty and researchers in the secure management, curation, and sharing of their research data has forced them to try to develop individual solutions. A vast majority of these methods rely on researchers purchasing and maintaining their own computers, servers, external and/or off-site back-up. Sharing of data with collaborators is equally concerning as this is often done using external hard drives, thumb drives, or cloud services (e.g. Google docs, Dropbox, etc.). There have already been reported cases of data being stolen, misplaced or lost due to a disaster (e.g. the Duff Roblin fire)

- Funding agencies will mandate data management and sharing and the university will not be able to quickly respond to the requirements. The greatest consequence of not being able to satisfy the funding agency requirements is the loss of funding opportunities.

- If UML decides not to lead data management initiatives at the University they will lose an important opportunity to become an essential partner and leader in the research process and to develop a new line of business. Someone else will control the organization and access to information at the university, thus diminishing the libraries' role. Researchers and other research supporters are looking for guidance for the present and future and will be receptive to this effort. If the library does not participate in this work, they will cede leadership in this effort to other campus entities.

- If UML do not lead the data management initiatives at the University, it runs the risk of being viewed as not being an innovative or progressive library which will hinder our ability to attract skilled staff. We could hinder our ability to collaborate with internal and/or external stakeholders because we won't have anything to offer. We could lose credibility with other academic libraries

- Existing library staff will not have the opportunity to expand their skills and expertise into this important new area for libraries, which will have a negative effect on their career.

B. Risks of Failure

- There are two risks in pursing services and initiatives in data management: a risk that we will be unable to accomplish what we set out to do, and the risk that we will be too successful and overwhelmed. This can be mitigated through careful resource planning and a phased approached to product and service implementation.

- A pilot project would be high profile and results actively reviewed. If UML fails to launch or complete a successful pilot, it will lose credibility. This risk can be mitigate by having a well-defined and supported plan, defined goals and outcomes, good communication, and by making good use of partnerships.

- There risks are involved within UML itself. Reorganization or reassignment of duties may be required. Traditional services may need to be changed or phased out in order to provide new services. Some staff may be resistant to change. Well defined change management can help mitigate some of these risks.

- There are risks to partnering with internal and/or external providers. The reputation of some providers could undermine the Libraries' credibility with stakeholders. There is a risk that partner organizations may not be able to deliver the expected results which would reflect negatively on the UML. Stakeholders could lose confidence in the whole data management initiative. In order to mitigate these risks, partners must be chosen carefully, and there must be commitment for the project from the senior administration.

- There is a risk that the UML may invest considerable resources into the development of a data management initiative only to see it be taken over by another group on campus. In order to mitigate this risk, UML will need to continuously work to communicate the value of the unique expertise, skills, and experience of the library and how they contribute to the initiative's success.

V. Organizational Implications

A. Staff Capacity

The Library has determined that additional expertise will be required in order for the Libraries to support institutional data management initiatives. At the very least, a Data Management Specialist will need to be hired. Training for liaison librarians and technical services staff members is required. Additional staff, such as developers, will be required to support the development of a Virtual Research Environment and/or the customization of the DAMS to support data curation and sharing.

An inventory of skills and capacities currently available or accessible to the libraries should be undertaken and reviewed in light of what skills and capacities will be needed to achieve the opportunities identified. Several librarians already possess expertise that may be leveraged towards developing initiatives and services for e-Science (information architecture, policy design, archives and discovery and delivery services for metadata, technology infrastructure, etc.) although it is unclear how much additional skill sets (including those beyond what a librarian is likely to possess) may need to be hired, brought in from other units or brought in as contract workers for a fixed term.

UML will need to foster a change in organizational culture that promotes an entrepreneurial mindset and encourages taking risks. The recent reorganization of the UML has positioned the Discovery and Delivery unit to take on the technical aspects of the E-Science initiative. Liaison librarians will need to continue to shift their focus away from traditional book and journal collection development and management to collecting and managing data. There will be a need to focus on providing value added data services.

B. Tools and Resources

UML needs to assess the extent to which it can provide massive digital storage for research data and provide access to discoverability tools. If the choice is to move in one or both of these directions, a significant capital outlay may be required that UML has not previously considered.

Similar to staff capacity above, UML will need to be able to determine what tools and resources are currently available or accessible to the libraries to apply towards accomplishing the opportunities that we have identified. Some tools and resources identified may already be assigned to existing services and so discussions of their possible reallocation may need to be initiated, with an eye on maximizing the libraries' impact on achieving the university goals. Plans will need to be made to obtain needed tools and resources that are not readily accessible. These plans could include forming partnerships with internal or external agencies (see below) or pursuing grants.

Library and University policies will need to be reviewed and revised to incorporate data management practices.

Primo would need to be customized to accommodate data records.

UML does not have the financial resources to support data management. A budget will need to be developed and written into the annual budget requests. UML may consider applying for grants to fund some aspects of the initiative.

C. Institutional Partners

Many if not all of the identified opportunities will require that the libraries enter into formal or informal partnerships with other agencies at the university, including university administration, IST, ORS, and faculty researchers. The primary institutional partners will be IST, Office of the VP Research, ORS, and Deans, faculty, students, and research centres.

These partnerships represent an opportunity for the Library to help to provide services without needing to purchase them outright. The Library might provide consulting services to researchers via the Office of Research Services, who would provide the grant planning and submission process, or to the grant writers employed by the various faculties.

D. External Partners

The Library will develop new relationships and strengthen existing relationships with other Canadian academic libraries and consortiums in order to collaborate on a regional or national data repository initiative as well as to share experience and expertise.

The Library, in collaboration with regional or national partners, must contribute to the operation of a sustainable, distributed preservation architecture. This may mean being responsible for off-site storage in a closed, cloud configuration.

Different libraries will possess different strengths and capabilities. Once gaps in staffing and resources have been identified, an exploration of how we might address these gaps should be conducted. UML may attempt to leverage their existing relationships with libraries within their consortiums, or develop new relationships with libraries that are known to possess the strengths required.

VI. Next Steps

- Present the strategic agenda to libraries' administration.
- In collaboration with the UML management team, prioritize potential opportunities and identify 2-3 for immediate implementation.
- Establish a cross-departmental e-Science team at the UML to initiate the projects approved by the UML management team.
- Determine what the scope of the libraries involvement in supporting e-Science at the university should be.
- Begin or continue discussions with collaboration partners to ensure that they are supportive of project goals during the planning stages. Identify other libraries that are active in this

area, especially those from institutions who are similar to UML, and make connections to learn more about their experiences and approaches.

- Report on the UML e-Science activities through existing communication channels (external blog, emails to library staff, etc.)
- Conduct additional interviews (formal or informal) of key stakeholders to gain additional insight and understanding of needs for e-Science support at the institutional level.

Appendix A: Glossary of Terms

Cloud Computing

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation.

COPPUL

COPPUL stands for the Council of Prairie and Pacific University Libraries. It is a consortium of 23 university libraries located in Manitoba, Saskatchewan, Alberta and British Columbia. Member libraries cooperate to enhance information services through resource sharing, collective purchasing, document delivery, and many other similar activities. COPPUL strives to be a cohesive and collaborative organization which provides leadership in the development of solutions that meet the academic information resource needs of its member institutions.

Cyberinfrastructure

Cyberinfrastructure (or CI) describes research environments that support advanced data acquisition, data storage, data management, data integration,data mining, data visualization and other computing and information processing services distributed over the Internet beyond the scope of a single institution. In scientific usage, cyberinfrastructure is a technological strategy for efficiently connecting laboratories, data, computers, and people with the goal of enabling novel scientific theories and knowledge. The term "cyberinfrastructure" was coined in the U.S. and other countries have different terms for this type of technological infrastructure. Cyberinfrastructure now often includes systems for managing, archiving and preserving data, in addition to data processing, and so can include digital libraries and archives and the software and hardware to support them.

DataVerse

The Dataverse Network is an open source application to publish, share, reference, extract and analyze research data. It facilitates making data available to others, and allows you to replicate others work. Researchers, data authors, publishers, data distributors, and affiliated institutions all receive appropriate credit.

A Dataverse Network hosts multiple dataverses. Each dataverse contains studies or collections of studies, and each study contains cataloging information that describes the data *plus* the actual data and complementary files.

Data Curation

Data curation refers to the value-added activities and features that stewards of digital content engage in to make digital content meaningful or useful. The data portion of this term sometimes refers specifically to research data (the outcomes of conducting research) and sometimes to digital content of any kind.

Digital Curation

Digital curation includes *digital* preservation data *curation*, as initially defined by the Digital Curation Center of the UK when it was founded. This term encompasses the full lifecycle of digital content management: selection, preservation, maintenance, collection and archiving of digital assets. Digital curation is generally referred to the process of establishing and developing long term repositories of digital assets for current and future reference by researchers, scientists, historians, and scholars.

Data Life Cycle

The data lifecycle – and data lifecycle management – deals with tracking, managing, and understanding data and metadata as it flows through organizations.

Data Management Plans

Data Management Plans describe how a research project will manage, dissemination and share research results. This includes descriptions of the types of data, samples, physical collections, software, curriculum materials, and other materials produced by a project; the standards used for data and metadata; policies for access and sharing data, including privacy protection of privacy and intellectual property, and security; policies and provisions for re-use, re-distribution, and the production of derivatives; and methods of archiving data, samples, and other research products, and for preservation of access to them. Funding bodies increasingly require grant-holders to produce and maintain Data Management Plans, both at the proposal stage and during the project.

Data Set

A set of files containing both research data - usually numeric or encoded - and documentation sufficient to make the data re-usable.

Data Sharing

Data sharing is the practice of making data used for scholarly research available to other investigators.

Data Visualization

Data visualizations are visual representations of data, or abstract information. In the context of e-Science data visualization is closely related to scientific visualization, an interdisciplinary branch of science primarily concerned with the visualization of three dimensional phenomena (architectural, meteorological, medical, biological, etc.), where the emphasis is on realistic renderings of volumes, surfaces, illumination sources, and so forth, perhaps with a dynamic (time) component. Visualizations and simulations are a key part of scientific communication in the digital era, and require sophisticated software to execute (i.e. the visualizations are often not static files that can be captured and preserved like a digital image).

Digital Asset Management System (DAMS)

A DAMS (Digital Asset Management System) infrastructure can ingest digital assests, store and index assets for easy searching, retrieve assets for use in many environments, and manage the rights associated with those assets.

Digital Preservation

Digital preservation is the set of processes and activities that ensure continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats, and is an ongoing process. This includes the preservation of materials resulting from digital reformatting (e.g. scanning from print), but particularly information that is born-digital and has no analog counterpart. Digital preservation is defined as: long-term, error-free storage of digital information, with means for retrieval and interpretation, for the entire time span the information is wanted. Long-term is defined as "long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community". Interpretation means that the retrieved digital files can be decoded and transformed into usable representations that are meaningful to the user (human or computer). There is an important distinction between "bit preservation", i.e., preserving the original files intact and unchanged over time) and "functional preservation", or preserving the information in the files by means of reformatting, added documentation, or other processes that will enable users to interpret the information in the future.

e-Research

The term e-Research here refers to the use of information technology to support existing and new forms of scholarly research in all academic disciplines, including the humanities and social sciences. E-research encompasses computational and e-science, cyberinfrastructure and data curation. E-Research projects often make use of grid computing or other advanced technologies, and are usually data intensive, but the concept also includes research performed digitally at any scale. E-research is useful here as a way to bridge the concept of e-science to other fields such as social science and the humanities. Just as e-science applies large-scale computing to processing vast amounts of scientific research data, e-research could include studies of large linguistic corpuses in the humanities, or integrated social policy analyses in the social sciences.

e-Science

E-Science is computationally intensive science carried out in highly distributed network environments, such as science that uses immense data sets requiring grid computing or High Performance Computing to process. The term sometimes includes technologies that enable distributed collaboration, such as the Access Grid, and is sometimes used as an alternative term for Cyberinfrastructure (e.g. e-Science is the preferred term in the UK). Examples of e-Science research include data mining, and statistical exploration of genome and other –omic structures.

Institutional Repository

An institutional repository is an online locus for collecting, preserving, and disseminating - in digital form - the intellectual output of an institution, particularly a research institution.

Metadata

Metadata is data that provides information about other data.

Metadata Standards

Metadata standards are requirements which are intended to establish a common understanding of the meaning or semantics of the data, to ensure correct and proper use and interpretation of the data by its owners and users. To achieve this common understanding, a number of characteristics, or attributes of the data have to be defined, also known as metadata.

MSpace

MSpace is one component of the University of Manitoba's institutional repository. MSpace is built with the open access software DSpace, and is dedicated to collecting, storing, preserving and making accessible digital versions of the intellectual output of UM community members. An institutional repository can include items such as faculty publications; theses; learning objects; research, conference and working papers; technical reports; images; audio and video files; datasets and computer programs.

Open Access

"Open access" to the literature means its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

Primo

Primo® is a one-stop solution for the discovery and delivery of local and remote resources, such as books, journal articles, and digital objects.

Research Data

Research data are defined as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated. This term does not cover the following: laboratory notebooks, preliminary analyses, and drafts of scientific papers, plans for future research, peer reviews, or personal communications with colleagues or physical objects (*e.g.* laboratory samples, strains of bacteria and test animals such as mice).

Research Life Cycle

The life cycle approach observes each stages of a process, to understand the overall process better. The research life cycle begins with the conception of a research project (hypothesis), continues through its methodology design and data collection, analysis, and finally publication and archiving of research outputs (e.g. articles, datasets, software, models, etc.). Understanding the research life cycle helps libraries identify who in involved and what information is produced or transformed during each phase of the project. For a more detailed explanation of this key concept, see e-Science and the Life Cycle of Research by Charles Humphrey.

Strategic Agenda

A strategic agenda is a set of coherent and aligned strategies that an organization identifies and which serve as an organization's platform for a strategic plan. The strategic agenda a library develops will guide implementation over the course of a few years as it is translated into the organization's strategic plan.

For further information, see the Association of Research Libraries, "Chapter 4: Applying the Scenarios to Create Strategy" in *The ARL 2030 Scenarios: A User's Guide for Research Libraries*, (2010) pp. 41-46.

Virtual Research Environment

A Virtual Research Environment (VRE) is a set of online tools, systems and processes interoperating to facilitate or enhance the research process within and without institutional boundaries. A VRE can aid with collaboration and communication amongst members of a research group, whether they share an office or work on different sides of the world.