Creating a Knowledge Map for the Research Lifecycle

Sai Deng [1], and Xiao Hu [2]

[1] Sai Deng, University of Central Florida Libraries, Orlando, FL
`sai.deng@ucf.edu`
[2] Xiao Hu, University of Hong Kong, Hong Kong S.A.R.
`xiaoxhu@hku.hk`

Abstract. In this study, a Knowledge Map (KM) was created based on the Research Lifecycle at the University of Central Florida to provide campus-wide services and resources to researchers. The KM aims to meet the needs of researchers and delivers guided searching and assistance in all aspects of research, including literature review, citation management, research data management, grant management, research work publication and dissemination. It elaborates the research processes and their associated services as presented in the Research Lifecycle, and links these points to various campus resources including those provided by the University Libraries, the Office of Research and Commercialization, the Institute for Simulation and Training and the Faculty Center for Teaching and Learning. It gives unified support to the researchers during their entire research lifecycle and it will keep evolving and developing.

Keywords: Knowledge map, information needs, information seeking, research services, research resources, unified research support.

1   In a Wider Context

The National Science Foundation (NSF) in 2007 published Cyberinfrastructure Vision for the 21st Century Discovery in which they define the various areas forming the cyberinfrastructure as: high performance computing, data management, data analysis and visualization, virtual organizations for distributed communities, and learning and workforce development. According to NSF, a "virtual organization" is created by a group of researchers from different geographic places to share access to resources and services often in real-time through the use of cyberinfrastructure systems [1]. Borgman et. al. studied the emerging virtual organizations and their data practices, trying to capture the best practices to inform the design of scalable virtual organizations [2]. The virtual organization is built on the cyberinfrastructure and functions as a large knowledge environment for scientists and researchers.

In a smaller scale, such as for a domain field or for an institution, a similar knowledge environment can be built to facilitate research and collaboration. To help creating this type of environment, research communities and universities have started examining the research process from the researchers' point of view. One approach is to illustrate the research process, stages and their related resources in graphic forms. The Research Lifecycle by OpenWetWare [3] is created for researchers in biology and biological engineering; the lifecycle itself only shows the research process, but in its site, it mentions meta-level and specific services and resources organized by the community. Some universities have developed research lifecycles at an institutional level such as the University of Virginia Library Research Life Cycle and the University of Bath's Research 360 Institutional Research Lifecycle [4, 5]. The former focuses more on working with

research data, and the latter has included stake holder names. In either research lifecycle, however, the research processes and related services are not integrated together. There are also data models illustrating the various stages needed for successful curation and preservation of research data. The Digital Curation Lifecycle Model created by the Digital Curation Centre at the United Kingdom [6] is a widely adopted data model which aims to help researchers, curators and librarians to work with and manage research data.

While these research lifecycles and data models have been created to help researchers, they are more of "mental models" at an abstract level. How to deliver the resources and services to researchers and to facilitate building an amiable and effective knowledge environment remains an important question. In particular, how to make graphic research lifecycles "alive" and present relevant information and knowledge to researchers in their information seeking process is still a critical issue. Early graphical search engines such as Kartoo [7] present search results in visual display interfaces. Kartoo gathers results from multiple search engines and delivers these results in an interactive visual interface. The TouchGraph GoogleBrowser [8] displays its search results as a graph, showing relationships between different hit sites. These search engines create visualizations of search results on the fly, based on automated hyperlink extraction.

This paper describes a case of developing knowledge maps for supporting research lifecycles at an academic and research institution. Through knowledge maps of different cycles in the research process, we strive to develop a model that can be integrated into cyberinfrastructure supporting research in various domains.

2 The Research Lifecycle at UCF

To support scholarly research, promote campus wide cooperation and infrastructure building, The University of Central Florida (UCF) built a Research Lifecycle (Figure 1) at the institutional level by embedding campus-wide scholarly services into a framework of research flow [9]. As for its presentation, it has been delivered as a graphical concept map in PDF format at the University Libraries' scholarly communication website, with resource links connected to the ideas, concepts and services defined in the map.
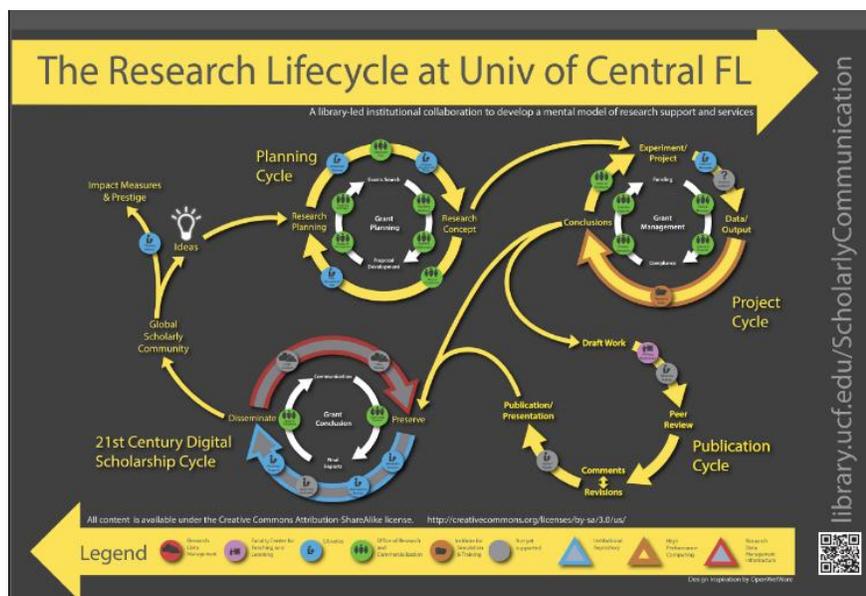

Fig. 1: The Research Lifecycle at UCF. It is available at:
http://library.ucf.edu/ScholarlyCommunication/UCFResearchLifecycle.pdf

The Research Lifecycle at UCF is a more complex model compared to many other research lifecycles. First, it embeds institutional research support services into the research process. Besides a typical research process and its related services, it displays a separate suite of services for funded research. Second, it presents a different perspective and places emphasis on connecting researchers to the embedded service points. It is built at the institutional level with broad purposes: facilitating strategic planning and a campus-wide solution to researchers' needs, and connecting the services to the researchers. It promotes infrastructure building and encourages partnerships within and beyond the campus.

The Research Lifecycle includes four sub-cycles: the planning cycle, the project cycle, the publication cycle and the 21st century digital scholarship cycle. The lifecycle graph (Figure 1) shows how each cycle illustrates the general research process in its outer circle and highlights the processes and services for funded research in its inner circle. The embedded services provided by individual campus units are presented in different colored dots, covering existing services and services under development.

The Research Lifecycle was created by the UCF Libraries Research Lifecycle Committee in collaboration with Office of Research and Commercialization (ORC), the Institute for Simulation and Training (IST) and the Faculty Center for Teaching and Learning (FCTL). Its further development and dissemination has been coordinated by the university library's Scholarly Communication Advisory Group. It has been presented to the Provost and Vice-Provost of Information Technologies & Resources, the Faculty Senate Library Advisory Council and the Faculty Senate. It has sought feedback from faculty members and graduate students. It has also been brought to the researchers' attention at ORC Grants Day and the FCTL Winter Institute.

The Knowledge Map (KM) presented in this study is built upon the Research Lifecycle and it further expands the research resources and illustrates the relationships among the concepts, processes and services.

3   The Researchers' Survey

A Research Data Management Survey was conducted at UCF in fall 2013 at the suggestion of members from the libraries, the IST and Computer Services and Telecommunications (CS&T). According to Beile, 549 invitations were sent to the UCF researchers, and 97 people responded (18.2% response rate), including 84% faculty and 10% administrators. Among those who responded, 90% collaborate with external researchers, 57% work with teams of 1 to 5 people, 84 people identified 120 different funding agencies, and 61% indicated that the funding agency requires them to manage, store, or share data. On data description and sharing, most researchers do not add metadata to their datasets; of the 34% who do most do not use any specific standards. On data storage and preservation, 68% of respondents said they take measures to preserve their data, but mostly only by backing it up [10]. The survey results present great needs in many areas of their research, including collaborator identification, grant management, research data management and preservation. It also calls for a united approach to facilitate information seeking of the researchers and provide meaningful services for their research.

4   The KM Design based on the Research Lifecycle Model

To further meet the Researchers' needs and provide comprehensive information to them, the design of the KM was based on the concepts and ideas defined in the Research Lifecycle; it links all the research activities, services and resources together and provides a path for UCF researchers. It gives details on how and where to gain the different types of service supports and resources, and the covered resources include those in the library's scholarly communication webpages, librarians' research guides, library databases, and information from the ORC, the IST and the FCTL.

Chung et. al. identified several relationship categories for KMs and suggested some descriptors for relationships [11]. The relationships can help the users to better understand the semantics and structure of the KM so as to perform navigation and searching more efficiently. The KM designed for the UCF Research Lifecycle has adopted some of these ideas and its main linking relationships include: Inclusion (contains/ is contained in, includes/ is included in, has part/ is part of, has example/ is example of); Action (provides/ is provided by, requires/ is required by); and Temporal (has stage/ is staged in, precedes/ follows).

In our initial experiment, the KM was developed for each sub-cycle of the Research Lifecycle to facilitate users to locate information relevant to specific concepts, research activities and services. In the KM infographics created using The Institute for Human and Machine Cognition (IHMC) CmapTools (http://cmap.ihmc.us/conceptmap.html), these sub-cycles have been illustrated separately due to the space limit in presentation. In an ideal online interactive environment, lower-level information such as the sub-cycles and services can be presented when a user issues a search query or clicks corresponding concepts in the KM.

The KM follows the convention of color coding in the Research Lifecycle. As shown in Figure 2-5, the blue boxes indicate services and resources provided by the Library, the green ones represent services by the ORC, the orange boxes include those by the IST, and the pinkish purple ones represent services provided by the FCTL. The gray boxes indicate services which are not yet officially supported, but they can still be linked to relevant resources. This tool also allows the KMs to be exported as webpages in which the links are interactive.

4.1   The Planning Cycle

The Planning Cycle (Figure 2) starts from "Research Planning" which is originated from "Ideas," and goes to formulating "Research Concept." "Research Planning" includes "Literature Review," "Citation Management Tools" and "Collaboration Tools." "Literature Review" contains "Literature Research Assistance" which is provided by subject librarians and "Locating Materials" through "Library Catalog" and "Databases." The databases are further divided by disciplines and can be linked to different database pages. "Citation Management Tools" include "EndNote" and "RefWorks;" their instruction sessions are provided by an Information Literacy & Outreach librarian. "Collaboration Tools" contain "Pivot," a database that allows users to search for funding opportunities and instantly view matching faculty from inside or outside of UCF. It maintains a comprehensive database of funds and 3 million pre-populated scholar profiles. After formulating "Research Concept," "Ethics and Compliances" need to be considered. For some funded research, a "Data Management Plan (DMP)" is required. The DMP service is linked to the Digital Initiatives Librarian and the DMP resource site. For the ORC provided services, the "Grant Planning" stage is further divided into "Funding Options," "Identify Collaborators," "Proposal Manager Guidance" and "Proposal Development Writing."
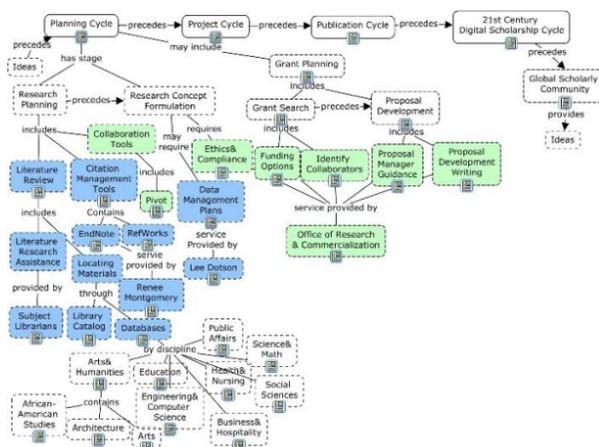

Fig. 2: Knowledge Map Prototype for the Research Lifecycle: Planning Cycle. See full images of KMs for all Research Lifecycle sub-cycles at: http://guides.ucf.edu/KMexpri

4.2   The Project Cycle

The Project Cycle (Figure 3) takes researchers through "Experiment/project," "Data/Output" and "Conclusions." The "Experiment/Project" stage provides services such as "Data Set Metadata," and is divided into "Data Documentation," "Metadata Standards," and "Controlled Vocabularies and Thesauri." The KM is linked to webpages of all these services and the Metadata Librarian who's providing them. The "Data/Output" stage mainly contains "Research Computing" services provided by Advanced Research Computing Center at the IST. The Project Cycle also involves "Ethics & Compliance" which can be assisted by the ORC. If the project is a funded one, the Project Cycle includes "Grant Management," which is further divided into "Award Manager (service)," "Internal Reports," "Project Accounts" and "External Reports."
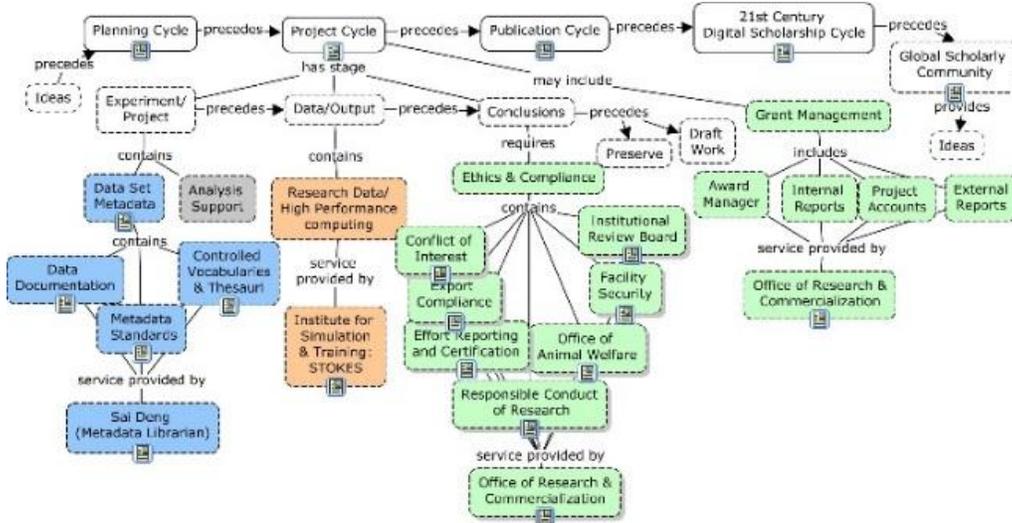
Fig. 3: Knowledge Map Prototype for the Research Lifecycle: Project Cycle.

4.3   The Publication Cycle

The Publication Cycle (Figure 4) has different stages including "Draft Work," "Peer Review," "Comments/Revisions," and "Publication/Presentation." In the "Draft Work" stage, for example, the FCTL provides services in "Writing Workshops." "Where to Publish" and "Author Rights" are not yet officially supported by any university unit, but their related information has been linked to the KM.
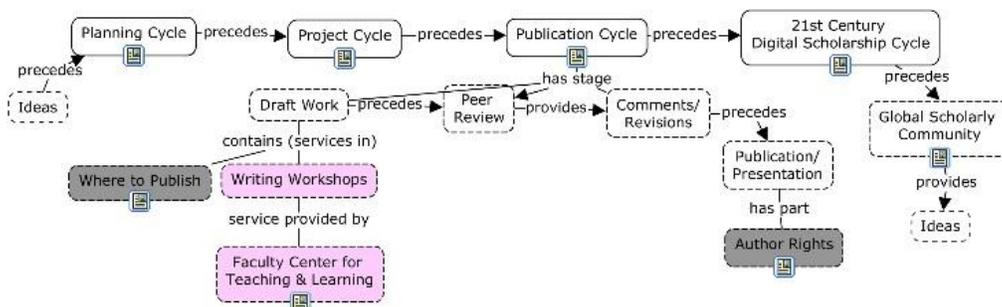

Fig. 4: Knowledge Map Prototype for the Research Lifecycle: Publication Cycle.

4.4   The Digital Scholarship Cycle

The Digital Scholarship Cycle (Figure 5) involves "Preserve" and "Disseminate" stages. The "Preserve" stage includes services like "Metadata," "Open Access Hosting" and "Discovery Support." The librarians are responsible for these services and their information pages are linked to the services. "Long Term Preservation" is not yet fully supported but relevant information is available. Currently it is undecided that which unit will be mainly responsible for providing "Data Sharing" and "Data Curation" services, but information in these areas have been provided and linked. Following preservation and dissemination, the research reaches the "Global Scholarly Community" in which the work's "Impact Measures & Prestige" can be evaluated by using "Citation Metrics." The Research & Information Services Librarian who's responsible for providing the citation service is linked under it.
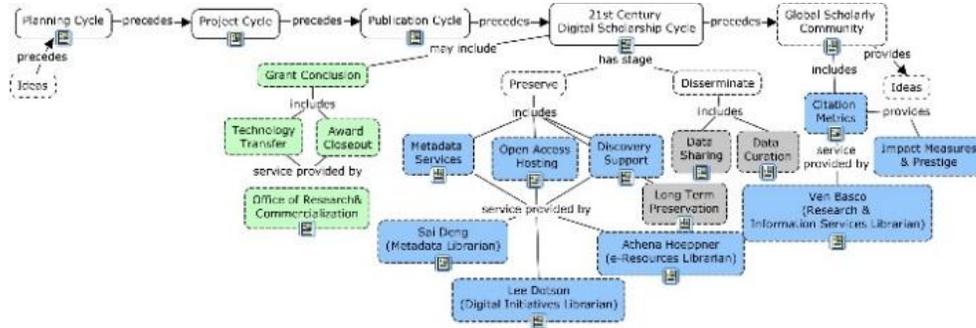

Fig. 5: Knowledge Map Prototype for the Research Lifecycle: Digital Scholarship Cycle.
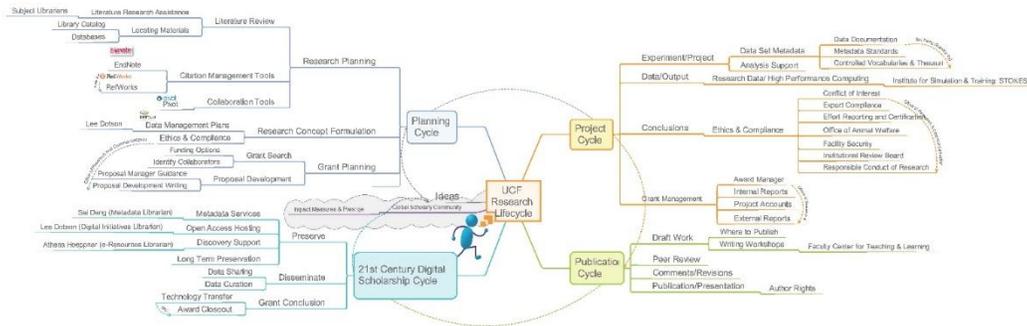

Fig. 6 Knowledge Map Prototype for the Research Lifecycle. It is available at:
http://imm.to/EfkYt

A KM for the full Research Lifecycle has further been experimented using iMindMap (Figure 6). This KM allows a researcher to follow the KM path, collapse and expand the branches, click the resource links, zoom in and out and perform concept searches.

5   Conclusion

The KM prototypes in this study are built to provide a roadmap and unified support to the UCF researchers during their entire research lifecycle. While the Research Lifecycle has sought feedback and has been officially adopted by the university, the KM prototypes created based on the Lifecycle are only proposed designs. Their viability need to be evaluated, the included concepts and the relationships defined between the concepts need to be assessed, and the resources need to be enriched. The next step would be to get feedback for the proposed KMs and make edits and amendments to the KMs based on that feedback. The KM development will be iterative in that prototypes will be presented to faculty members and graduate students for evaluation and their feedback will be used to adjust the content and design of the system. Overall, the KMs will keep evolving and new services and resources will continue being added

to the system. Since the research lifecycle is largely universal across different countries and fields of study, the KMs can be applied to other research institutions. To build a KM for the research lifecycle in the 21st century, a collaborative approach is desirable and even necessary. The library can act as the coordinator, seeking support from a wide range of resources, including campus units (e.g., faculty support center, grants and funding office, computing and systems services), the researchers themselves, as well as other parties providing technical implementation of interactive KMs.

6   Acknowledgements

References

1. National Science Foundation Cyberinfrastructure Council. Cyberinfrastructure vision for 21st century discovery. March 2007. Retrieved July 28 from http://www.nsf.gov/pubs/2007/nsf0728/
2. Borgman, Christine L., Bowker, Geoffrey C., Finholt, Thomas A. and Wallis, Jillian C. Towards a virtual organization for data cyberinfrastructure. *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, pages 353-356, 2009. Retrieved July 28 from http://dl.acm.org/citation.cfm?id=1555459
3. OpenWetWare.org. OpenWetWare: The Research Lifecycle. Retrieved June 20, 2014 from http://openwetware.org/wiki/OpenWetWare:Headquarters/Research_Pathway
4. The Data Management Consulting Group. University of Virginia. Making data management easier: Research life cycle. Retrieved June 20, 2014 from http://dmconsult.library.virginia.edu/
5. University of Bath. Research 360 Institutional Research Lifecycle. Retrieved June 20, 2014 from   http://blogs.bath.ac.uk/research360/about/
6. Digital Curation Centre (DCC) Lifecycle Model Retrieved June 20, 2014 from http://www.dcc.ac.uk/resources/curation-lifecycle-model
7. Dürsteler, Juan C. KartOO. August 16, 2002, *Inf@Vis!: The digital magazine of InfoVis.net*. Retrieved July 28 from http://www.infovis.net/printMag.php?num=97&lang=2
8. Smarty, Ann. TouchGraph Google Browser: Visualize relations between relevant sites. *SEJ: Search Engine Journal*. June 16, 2009. Retrieved July 28 from http://www.searchenginejournal.com/touchgraph-google-browser-visualize-relations-between-relevant-sites/11043/
9. University of Central Florida Libraries Research Lifecycle Committee. 2012. The research lifecycle at UCF [Online Graphic]. Retrieved June 20, 2014 from http://library.ucf.edu/ScholarlyCommunication/ResearchLifecycleUCF.php
10. Beile, Penny. 2014. UCF research data management survey: A report of faculty practices and needs. Research Data Computing Day at the Institute of Simulation and Training, the University of Central Florida (Orlando, FL, February 11, 2014).
11. Chung, Gregory K. W. K.; Cheak, Alicia M.; Lee, John J. & Baker, Eva L. 2012. Development model for knowledge maps. Resource paper, No. 14. National Center for Research on Evaluation, Standards, & Student Testing (CRESST).