The Evolving Scholarly Record and The Utility of OCLC Research Models

Keith Webster
Dean of University Libraries and Director of Emerging and Integrative Media Initiatives
Road Map
The future of libraries and the role of the evolving scholarly record
Wondering about and planning for the future are core to being human.
Where are we going?

- Future of higher education
- Future of scholarly communication
- Research priorities and funding
- Future of technology
- Future of library and information science
Foresight
Draft a forecast
Envision futures
Headline the future

Insight
Ride two curves
Map cross-impacts
Prioritize actions

Action
Build a roadmap
Rally a network
Identify assets and gaps

Prepare
Orient to the future
Identify drivers of change
Catalog signals of change
Conduct foundational research get a sense of the landscape, interview experts and stakeholders to find the right frame on the future you need to explore. Develop alternative futures, Work with decision makers and stakeholders to uncover opportunities and threats in each of the scenarios and prioritize actions that appear in multiple scenarios.
They serve the communities and organizations of which they are a part—they serve the interests of their parent universities and colleges.

The most important long-term influence on the library is the requirement placed on it by changing patterns of research and learning.

The library can no longer define its quality in terms of collection size; services are the new differentiating factor.

Academic library services may benefit from—and some indeed are already benefiting from—undergoing a disruption.

A notable finding is a shared desire among academic libraries to reduce allocations to facilitate information access in favor of other areas.

As universities diversify their educational offer, libraries adapt services based on institutional fit.
Institutional typologies

- **Research**: primary focus on doctoral research and scholarship
- **Career**: primary focus on preparation for career professions
- **Liberal Education**: primary focus on interdisciplinary baccalaureate education
Library services framework
Systematise library planning and assessment

- Academic success
- Off-campus
- Information access
- Unique collections
- Support scholarship
- Transform publishing
- Showcase expertise
LIBRARIES
Past, present, future
Researcher Workflows and the role of libraries

- Discovering & Planning
  - Researching scientific literature and patents
  - Designing Studies
  - Procuring funding
- Experimenting & Managing Data
  - Searching for collaborations
  - Experimenting
  - Managing and analysing data
- Driving & Establishing Impact
  - Managing reputations
  - Measuring impact
- Authoring, Publishing & Disseminating Findings
  - Writing up findings
  - Publishing article(s)
  - Disseminating outputs and outcomes
Evolution of libraries
The journey since the 1980s

Collections-centric
Library as a place to house and access predominantly print-format collection

Client-focused
First widespread availability of computer-based information assets. Quality management and customer service focus

Experiential
New pedagogies drive students to library in a best for primary non-teaching academic space. Different learning activities need different spaces.

Connected
Library becomes a node in a global network of learning environments and connected collections

Interdisciplinary
Library seen as convening space on campus. Creation of digital and physical maker spaces and collaborative studios.
Changes in media delivery
From bricks and mortar to digital convenience
The P-index
Penn, Penn State, Pitt, Princeton, Purdue
COLLECTIONS
From print to digital, local to network
Information formats
The journey since the 1980s
The ‘external’ collection:
Pointing researchers at Google Scholar;
Including freely available ebooks in the catalog;
Creating resource guides for web resources.

The ‘borrowed’ collection

The ‘shared print’ collection

The ‘shared digital’ collection

The evolving scholarly record

The ‘licensed’ collection

The ‘demand-driven’ collection

The ‘facilitated’ collection

Note: Libraries have variable investments across the entire spectrum

A collections spectrum

Purchased and physically stored

Meet research and learning needs in best way
EVOLVING SCHOLARLY RECORD
The ways and means of conducting scholarly inquiry are experiencing fundamental change, with consequences for scholarly communication and ultimately, the scholarly record.

A shift from a print-centric scholarly record to one that is in digital form and resides on the network.

Pathways by which materials comprising the scholarly record are created, managed, and consumed are changing in a variety of ways.

Four roles exist: create (authors), fix (publishers), collect (libraries), use (scholars) - these remain relevant in evolving record model.

In the digital environment, versioning can be a much more complex issue.
Springer Nature publishes its first machine-generated book

Innovative book prototype provides a compelling machine-generated overview about the latest research on lithium-ion batteries, automatically compiled by an algorithm developed in collaboration with the Applied Computational Linguistics lab of Goethe University Frankfurt/Main (Germany)

London | Heidelberg, 02 April 2019

Springer Nature published its first machine-generated book in chemistry. The book prototype provides an overview of the latest research in the rapidly growing field of lithium-ion batteries. The content is a cross-corpus auto summarization of a large number of current research articles in this discipline. Serving as a structured excerpt from a huge set of papers, the innovative pipeline architecture aims at helping researchers to manage the information overload in this discipline efficiently.

In close collaboration between Springer Nature and researchers from Goethe University Frankfurt/Main, a state-of-the-art algorithm, the so-called Beta Writer, was developed to select, consume and process relevant publications in this field from Springer Nature’s content platform SpringerLink. Based on this peer-reviewed and published content, the Beta Writer uses a similarity-based clustering routine to arrange the source documents into coherent chapters and sections. It then creates succinct summaries of the articles. The extracted quotes are referenced by hyperlinks which allow readers to further explore the original source documents. Automatically created introductions, table of contents and references facilitate the orientation within the book.
AIDR 2019

ARTIFICIAL INTELLIGENCE FOR DATA DISCOVERY & REUSE

An NSF - Supported Conference
May 13 - 15, 2019
Carnegie Mellon University

https://events.library.cmu.edu/aidr2019/
Conscious co-ordination or the invisible hand?

- The invisible hand has worked quite well as a means of distributing and coordinating stewardship responsibilities attached to a largely print scholarly record.
- Conscious coordination is likely to replace the invisible hand as the key principle underpinning stewardship models for the scholarly record.
- The features of the evolving scholarly record suggest that autonomous, institution-scale stewardship of the scholarly record is becoming less and less feasible.
- A strategy of deliberate engagement with—and growing dependence on—cooperative agreements, characterized by increased reliance on network intelligence.
- A university’s allocation of resources to stewardship of the scholarly record will increasingly be for the “common good”, rather than primarily for local benefit.

Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination

Brian Lavoie and Constance Malpas
Conscious coordination

System-wide awareness
Align local action with collective effort

Explicit commitments
Move commitments above the institution

Division of labor
Collect more of less

Reciprocal access
Curate locally, share globally
Research Data Management

Managing throughout the research lifecycle

PART ONE
A Tour of the Research Data Management (RDM) Service Space
Rebecca Bryant, Brian Lavele, and Constance Melonas

PART TWO
Scoping the University RDM Service Bundle
Rebecca Bryant, Brian Lavele, and Constance Melonas

PART THREE
Incentives for Building University RDM Services
Rebecca Bryant, Brian Lavele, and Constance Melonas

PART FOUR
Sourcing and Scaling University RDM Services
Rebecca Bryant, Brian Lavele, and Constance Melonas
RDM service categories

**EDUCATION**
- Raise awareness of RDM's importance, encourage RDM skill-building, and disclose RDM tools and resources

**EXPERTISE**
- Decision support for, and customized solutions to, specific research data management problems

**CURATION**
- Technical infrastructure and related services that support data management throughout the research cycle
Compliance with external policies is a strong driver in some locales

Evolving scholarly norms are not uniform, disciplinary diversity is also a challenge

Institutions built services in anticipation of rather than response to researcher demands

Developing RDM expertise may be part of an institutional strategy to increase library support for research workflows

“Incentives to Acquire RDM Capacity” by OCLC Research, from The Realities of Research Data Management Part Three: Incentives for Building University RDM Services (https://doi.org/10.25333/C3S62F), CC BY 4.0
RESEARCH

INFORMATION

MANAGEMENT
Research Information Management
Aggregation and curation of research information

Convenience and Compliance: Case Studies on Persistent Identifiers in European Research Information Management
Rebecca Bryen, Annette Dormund, and Constance Melges

Practices and Patterns in Research Information Management
Findings from a Global Survey
Rebecca Bryen, Anna Clements, Pablo de Castro, Joanne Cottrell, Annette Dormund, Jan Franzen, Peggy Gallagher, Nicole Merrin

Research Information Management: Defining RIM and the Library's Role
Rebecca Bryen
OCLC Research (USA)
Anne Clements
University of St. Andrews (UK)
Carol Felten
Rockefeller University (USA)
David Greenspan
Rutgers University (USA)
Bérenger Hugard
La Trobe University (Australia)
Holly Mercer
University of Western Australia (Australia)
Renate Moxham
Australian National University (Australia)
Malcolm Osmers
University of Arizona (USA)
Anna Rauh
Syracuse University (USA)
John Wright
University of Calgary (Canada)
"RIM Uses" by OCLC Research, from Research Information Management: Defining RIM and the Library’s Role (doi.org/10.25333/C3NK88), CC BY 4.0
Registering researchers in some type of authority file or identifier (ID) system has become more compelling

- manage scholarly output
- correct data for assessment and rankings
- funder mandate tracking

Registration files and authority files

Researchers need to have persistent IDs to link them to their scholarly output

Libraries should assign persistent IDs to authors if they don’t already have them

Challenges:

- A scholar may be published under many name variations
- Multiple people can share the same name
- Some researchers already have multiple unlinked profiles or IDs
- Lack of interoperability
STRATEGIES
<table>
<thead>
<tr>
<th>Individual Experience</th>
<th>University Community</th>
<th>Societal Impact</th>
<th>Strategic Recommendations</th>
<th>About</th>
</tr>
</thead>
</table>

Our Strategy

Strategic Plan 2025
Carnegie Mellon University
Creating a 21st Century Library

**Learning Spaces**
Create 21st-century library spaces for 21st-century learners

**Global Collections**
Collaborate with peer institutions to provide coordinated access to a global collection of information resources.

**Scholarly Record Curation**
Steward the evolving scholarly record, and champion new forms of scholarly communication.

**Information Specialists**
Develop information specialists as partners in research, teaching, and learning
Learning Spaces
Create 21st century library spaces for 21st century learners

Student journey
The library as the primary non-teaching space on campus

Digital visitors & residents
How do we blend physical and digital relationships?
Global Collections

Collaborate with peer institutions to provide coordinated access to a global collection of information resources
Global Collections
Collaborate with peer institutions to provide coordinated access to a global collection of information resources

Press and Journal Management Software
From the Birkbeck Centre for Technology and Publishing

August 28, 2017
CMU, Pitt Receive Andrew W. Mellon Foundation Grant To Plan Platform for Digital Scholarship
Scholarly Record Curation
Steward the evolving scholarly record and champion new forms of scholarly communication

Welcome, Curtis Meyer
Mellon College of Science (MCS)
cmeyer@andrew.cmu.edu

My Summary

370 publications

- 64 Scopus
- 63 Web of Science

49 professional activities
- 39 committee memberships
- 3 consulting / advisories
- 2 distinctions
- 1 fellowship
- 4 offices held

13 teaching activities
- 13 courses taught

3 grants
- 3 grants
Open Science Framework

Open Science Framework is a free and open source tool that can be used for managing projects and collaborations in any discipline. OSF is a great way to keep track of all of the different files that are part of a complex research project. You can store files directly on OSF cloud storage (unlimited number of individual files that are under 5 GB each) or sync popular third-party applications such as Google Drive, Box, Dropbox, Amazon S3, GitHub, figshare, Mendeley, and Zotero to the project. You can add collaborators from CMU or anywhere else to a project and set their access so that they can view or edit parts of a project. You can also use OSF to share your work publicly so that it can be cited. Overall, it’s a powerful tool for research management, documentation, and collaboration that can help you see the sum of your work across many different platforms.
protocols.io

About protocols.io

The protocols.io platform is a free and open access repository for recording and sharing detailed up-to-date research methods and protocols. It allows researchers to easily create, edit, share, and get credit for their protocols, and provides an open access hub for scientists to communicate improvements and corrections to scientific methods. The platform is useful for researchers in any discipline that uses a step-by-step methodology, such as biology, chemistry, neuroscience, engineering, data science, and other computational or life science fields.

protocols.io's editor provides a structured way to create and modify a protocol in minutes. They can be made from scratch or uploaded and converted from an existing Word or PDF document quickly and easily. Protocols and changes to them can be published for everyone to see, or shared with private groups or individuals. Publicly published protocols can be assigned DOIs and made easily discoverable and citable. Mobile apps make it easy to follow along with the "recipe" for a method while doing your research.
This research output has an **Altmetric Attention Score** of 749. This is our high-level measure of the quality and quantity of online attention that it has received. This Attention Score, as well as the ranking and number of research outputs shown below, was calculated when the research output was last mentioned on 18 May 2017.

Altmetric has tracked 7,862,537 research outputs across all sources so far. Compared to these this one has done particularly well and is in the 99th percentile: it's in the top 5% of all research outputs ever tracked by Altmetric.
Information Specialists
Develop information specialists as partners in research, teaching and learning

Ana Van Gulick
Librarian & Program Director, Open Science
Research & Academic Services
anavangulick@cmu.edu
+1 412 268 7311
Hunt 1090, Baker Hall 352
Brain Sciences, Psychology
Research Guides

As the research liaison to the psychology department and the BrainHub initiative at the Carnegie Mellon University Libraries I collaborate with faculty and students to support the entire research and teaching process at the university. I also work with researchers at the Center for the Neural Basis of Cognition together with colleagues in Biology at CMU and at Pitt. I am embedded in the departments where I bring together my experience as a psychologist and as a data curation specialist in the library. I am particularly involved in working with faculty and students on research data management from grant writing to establishing lab protocols to sharing data publicly. I am an advocate for open science and reproducibility initiatives and work to support researchers' needs at the university level. I am particularly interested in promoting the ever-growing number of open source tools being developed for psychology and neuroscience. I completed an ScB in cognitive science at Brown University in 2009 and in 2014 I received a PhD in cognitive psychology from Vanderbilt University where I studied visual object perception and categorization. From 2014-2016 I was a CLIR-DLF postdoctoral fellow for data curation at the CMU libraries where I worked on campus-wide data management and open access initiatives. I hold a faculty position in the University Libraries as well as a courtesy appointment in Psychology and an affiliation with the Center for the Neural Basis of Cognition. My own research program continues to span human cognition and perception as well as studying open science practices and their impact in psychology and neuroscience.

Huajin Wang
Librarian
Research & Academic Services
huajinw@cmu.edu
+1 412 268 3172
Mellon Library 429
Bioinformatics, Computational Biology, Language Technology Institute, Machine Learning
Research Guides

I am a cell biologist by training, with more than 10 years of research experience in molecular and cellular biology, lipid metabolism, bioinformatics, and computational analysis of large biomedical datasets. I have deep understanding of research data. During my PhD studies at University of Alberta and postdoctoral research at Yale/Harvard, I have collaborated with biologists, clinicians, and data scientists on many successful research projects. As a librarian, I am passionate about helping the research community with their data needs, facilitating scientific data reuse and reproducibility, and educating the next generation of biomedical scientists. My current research focus is to engage various stake holders to build incentives and infrastructure for scientific rigor and data sharing, as well as developing data-driven approaches for academic services.
<table>
<thead>
<tr>
<th>Service Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discover Your Research Topic</td>
<td>Discover and refine your research topic with resources that connect you with specialists and databases.</td>
</tr>
<tr>
<td>Find Funding &amp; Collaborators</td>
<td>Uncover potential collaborators and the funding to support your project.</td>
</tr>
<tr>
<td>Manage Information &amp; Data</td>
<td>Work with our specialists to evaluate, select, and navigate the tools to organize your data and keep your project on track.</td>
</tr>
<tr>
<td>Publish &amp; Share Your Research</td>
<td>Navigate the many options for publishing your research to optimize its accessibility and impact.</td>
</tr>
<tr>
<td>Measure Impact &amp; Manage Your Digital Identity</td>
<td>Identify when and where your research has been cited, as well as the scope of its reach.</td>
</tr>
<tr>
<td>Enhance Your Teaching</td>
<td>Advise your students about research methods and resources that enhance their learning inside and outside the classroom, unlocking your teaching potential.</td>
</tr>
</tbody>
</table>

**Our Service Model**

We use this model as a structure for our engagement with the CMU community.

It serves as a structure for information and data skills sessions, for service innovation, and as a checklist for faculty/researcher interactions.

We also use the model as a framework for interactions with the university for planning and budgeting purposes.
Prepare
Orient to the future
Identify drivers of change
Catalog signals of change

Foresight
Draft a forecast
Envision futures
Headline the future

Insight
Ride two curves
Mao cross-impacts
Prioritize actions

Action
Build a roadmap
Rally a network
Identify assets and gaps
EXPLORE FOUR FUTURE FEELINGS

Over the past decade, scientists have searched for tools to help individuals develop insights for the present by imagining the future more vividly. One of the most powerful techniques they’ve devised is to guide thinkers through a series of four specific and distinct emotions. In other words: Sometimes, we can’t think our way to the future—we have to feel our way.

WHY THIS TOOL?

- Humanize future scenarios, using emotions to drive new insights into positive and negative implications.
- Build insights through empathy for how the future will affect people differently.
- Gain cognitive flexibility by powering up the posterior cingular cortex, the region of the brain that allows us to be more open to new information.
Carnegie Mellon University

kgw@cmu.edu

cmkeithw

Keith Webster

cmkeithw