

Think Like an Economist... Then What?

Ronald L. Larsen

July 17, 2014

Roadmap

- The emergence of *iSchools*
- Unsettling trends in enrollment & employment
- Imbalance in education supply and demand
- Reinterpreting career prospects
- Repositioning to reinforce digital scholarship
- Reinforcing the infrastructure for evidence
- Reframing scholarly communication
- A broader mandate for “archivists”

The Emergence of iSchools – 2003

The Founding Core of the iCaucus



Growth of the iSchools – 2008

21 Members



Growth of the iSchools – 2014

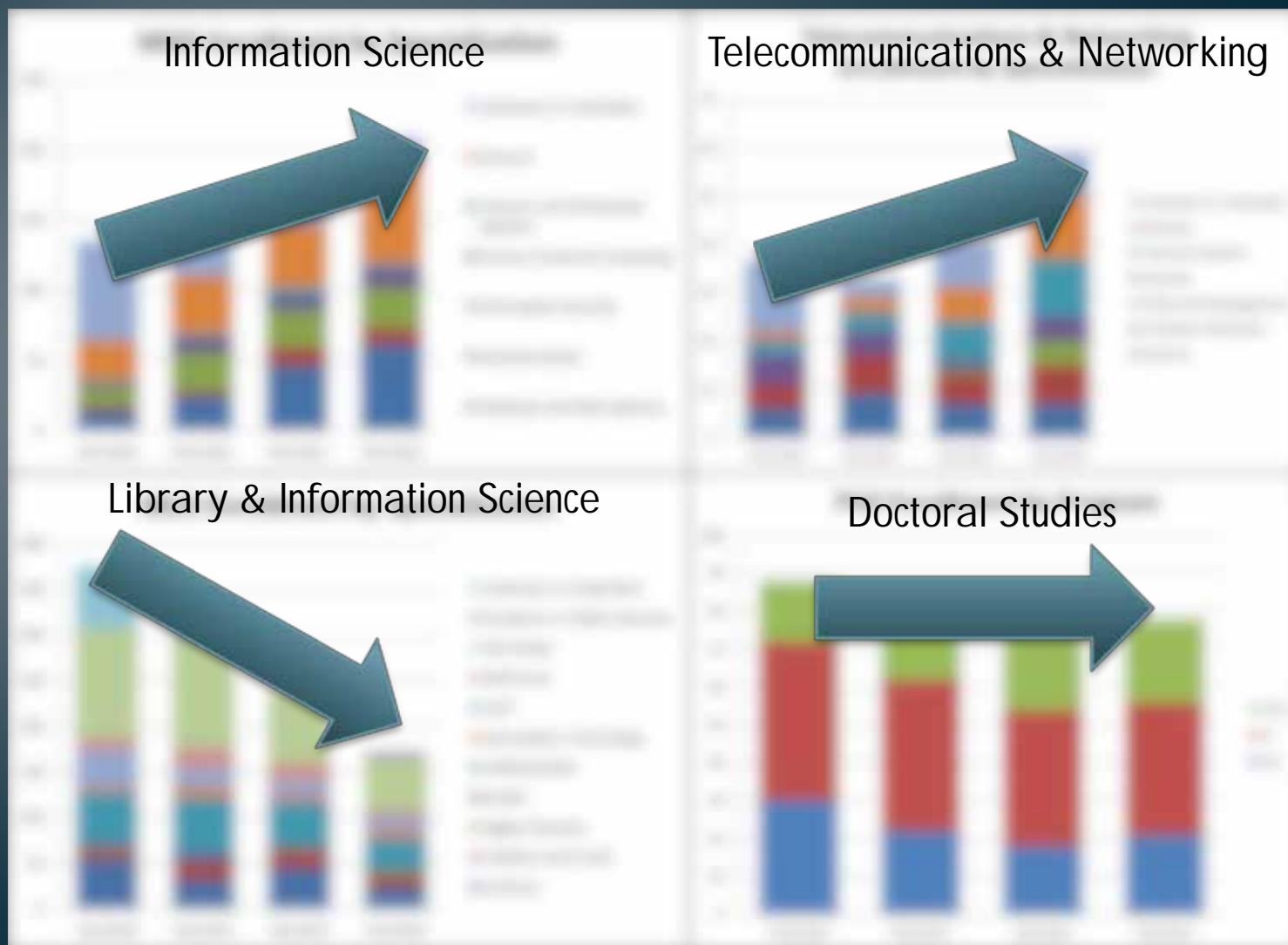
55 Members



iSchools with Archives Programs – 2014

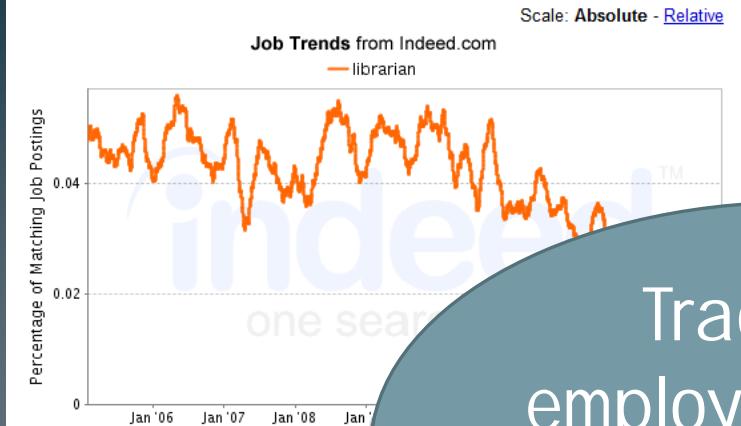


Graduate Enrollment Trends

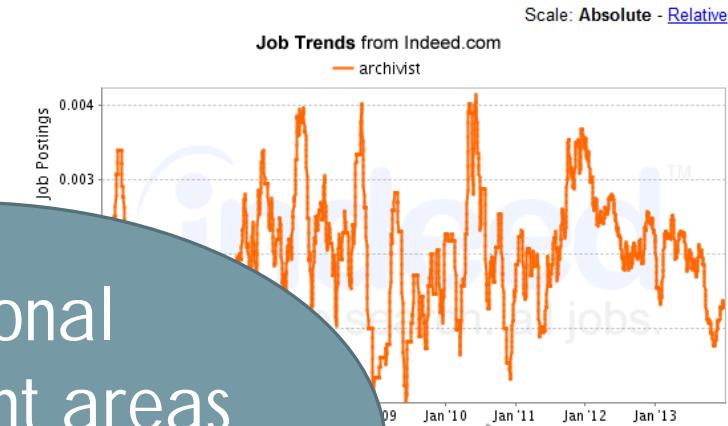


Eight years of job posting data (Indeed.com) ...

librarian Job Trends

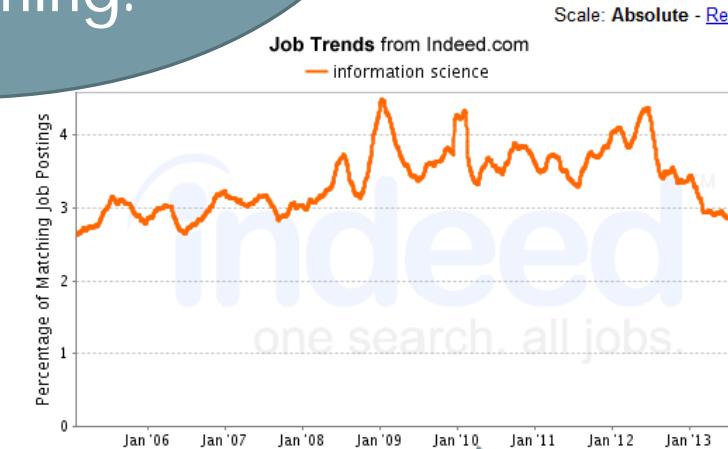


archivist Job Trends



Traditional
employment areas
appear to be flat
or declining.

informatics Job Trends



Eight years of job posting data (Indeed.com) ...

information steward Job Trends

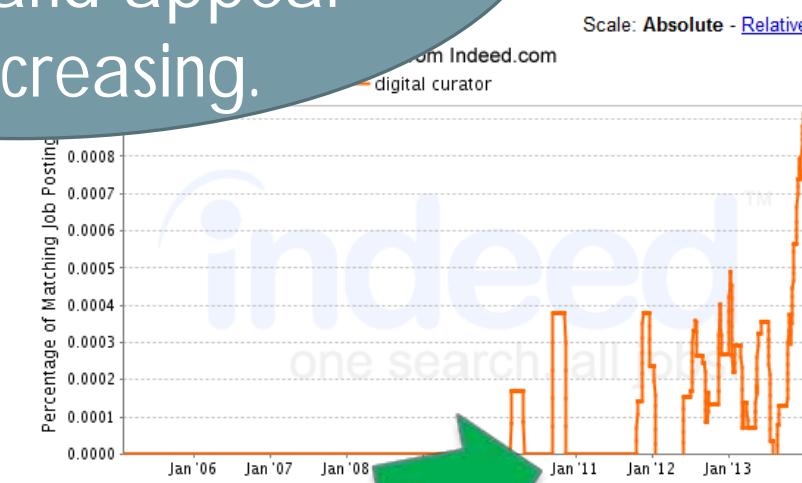


data steward Job Trends



Non-traditional
employment
opportunities are
emerging and appear
to be increasing.

data governance Job Trends



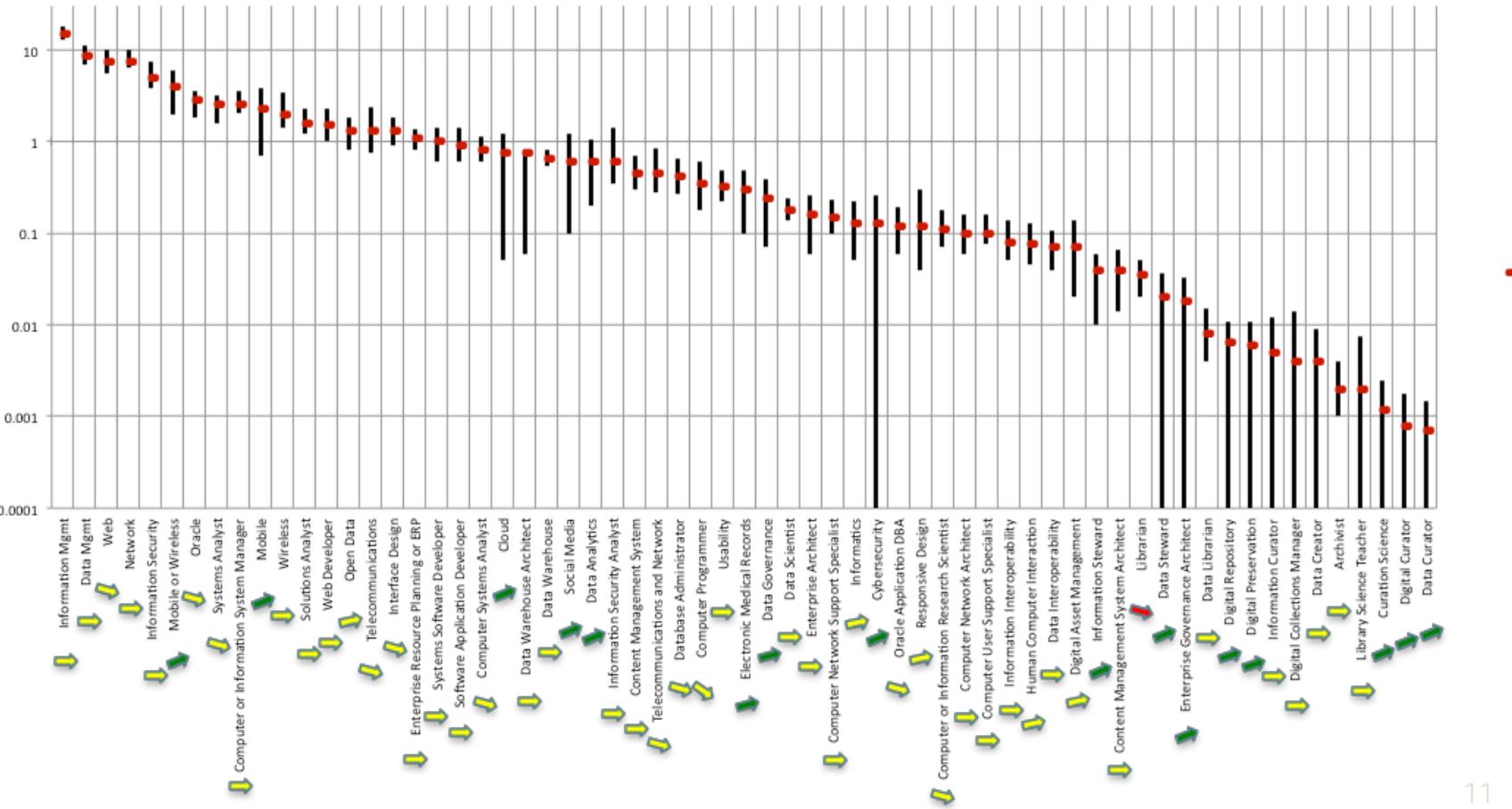
Summarizing Job Posting Trends...



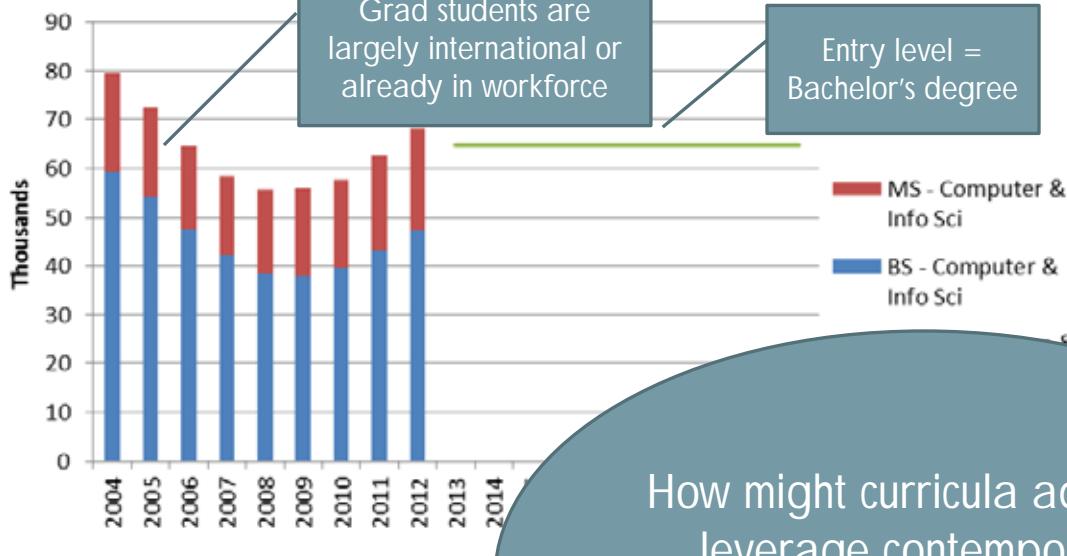
Job Posting Trends

Range for Job Trends (Indeed.com, 2006-13)

Percentage of Matching Job Postings



Comp & IS Degrees vs. Projected Demand*



Grad students are largely international or already in workforce

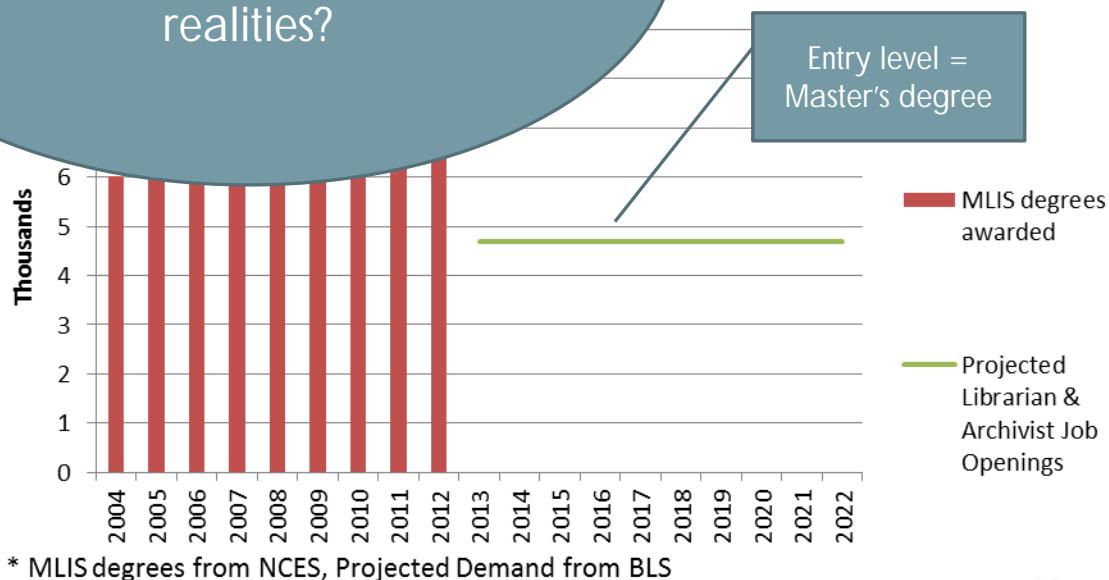
Entry level = Bachelor's degree

Workforce Demand

How might curricula adapt to leverage contemporary needs, opportunities and realities?

Workforce Surplus

MLIS Degrees vs. Projected Demand*



Entry level = Master's degree

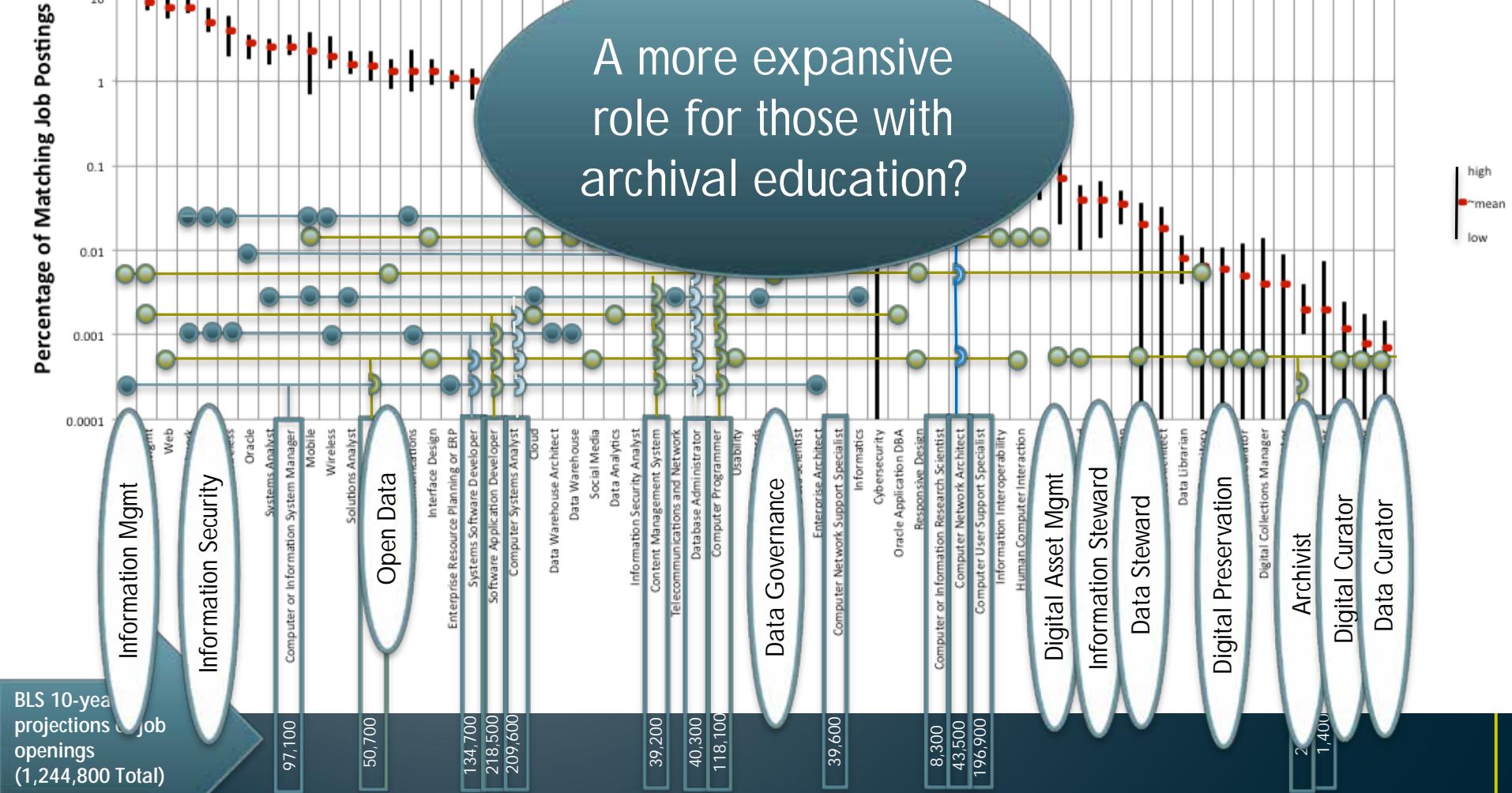
MLIS degrees awarded

Projected Librarian & Archivist Job Openings

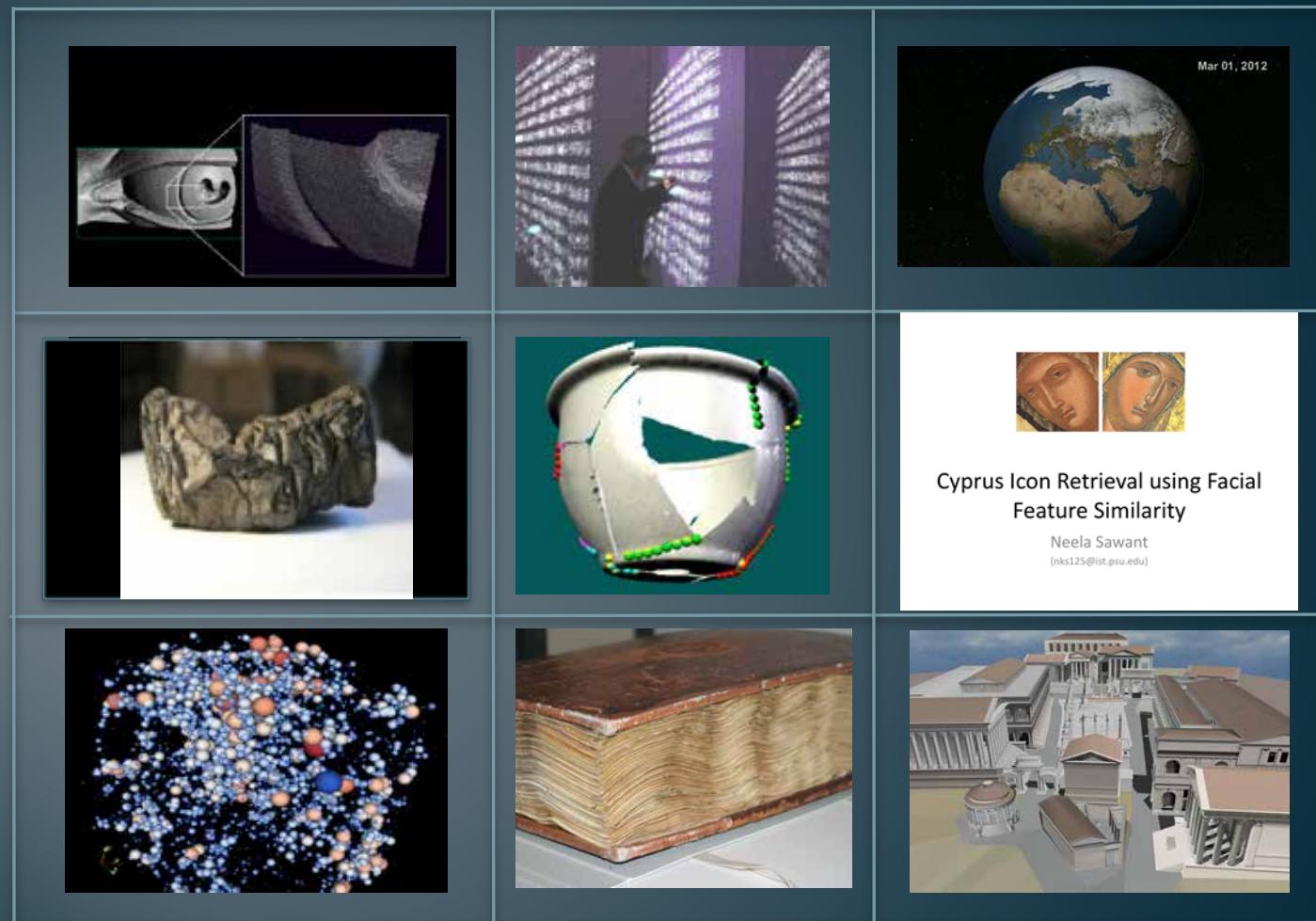
Bureau of Labor Statistics Projections

Range for Job Trends (Indeed.com, 2006-13)

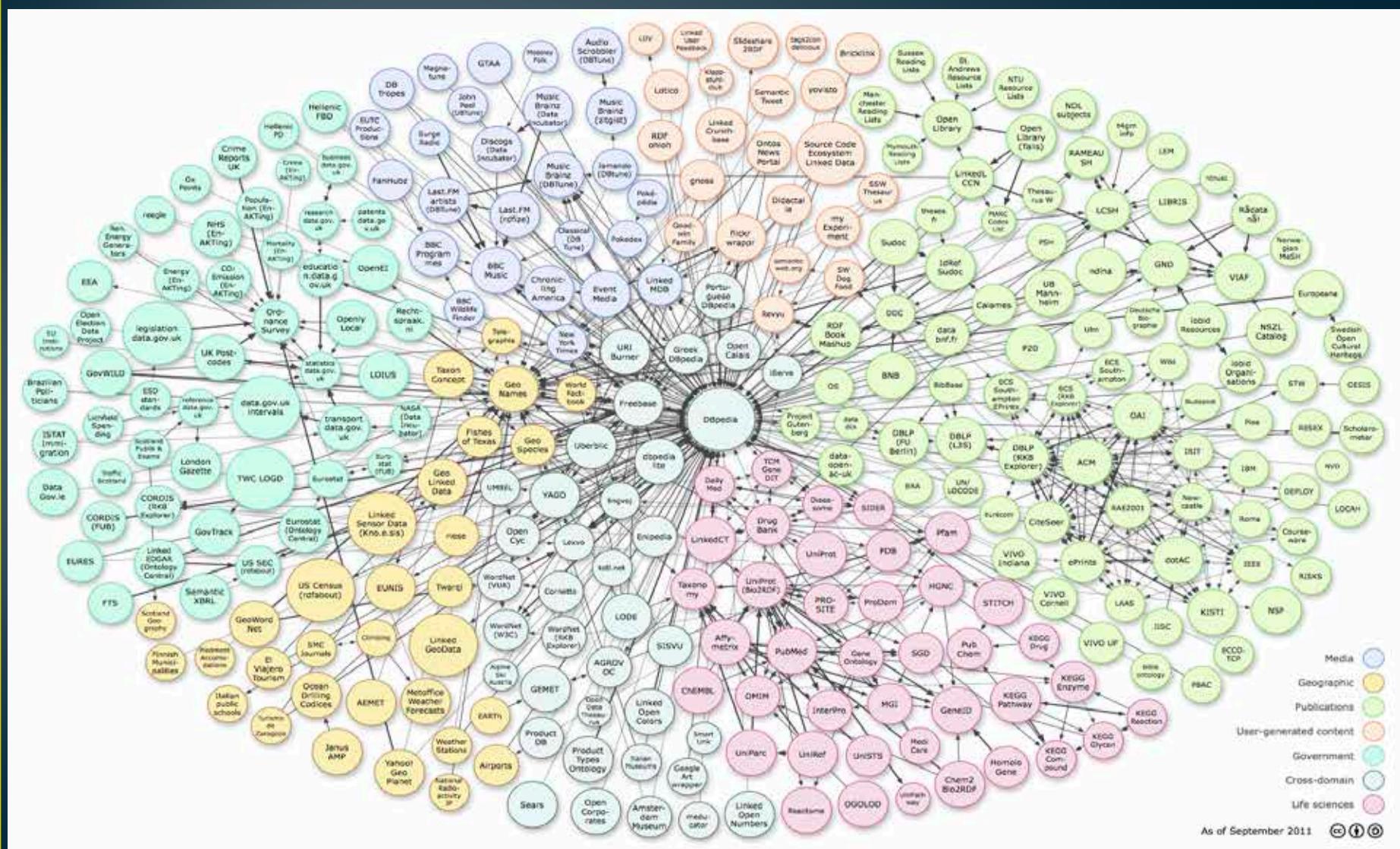
A more expansive role for those with archival education?



Digital Scholarship



Linked Open Data Cloud



Richard Cyganiak & Anja Jentzsch (<http://lod-cloud.net>)

The Burden of Evidence

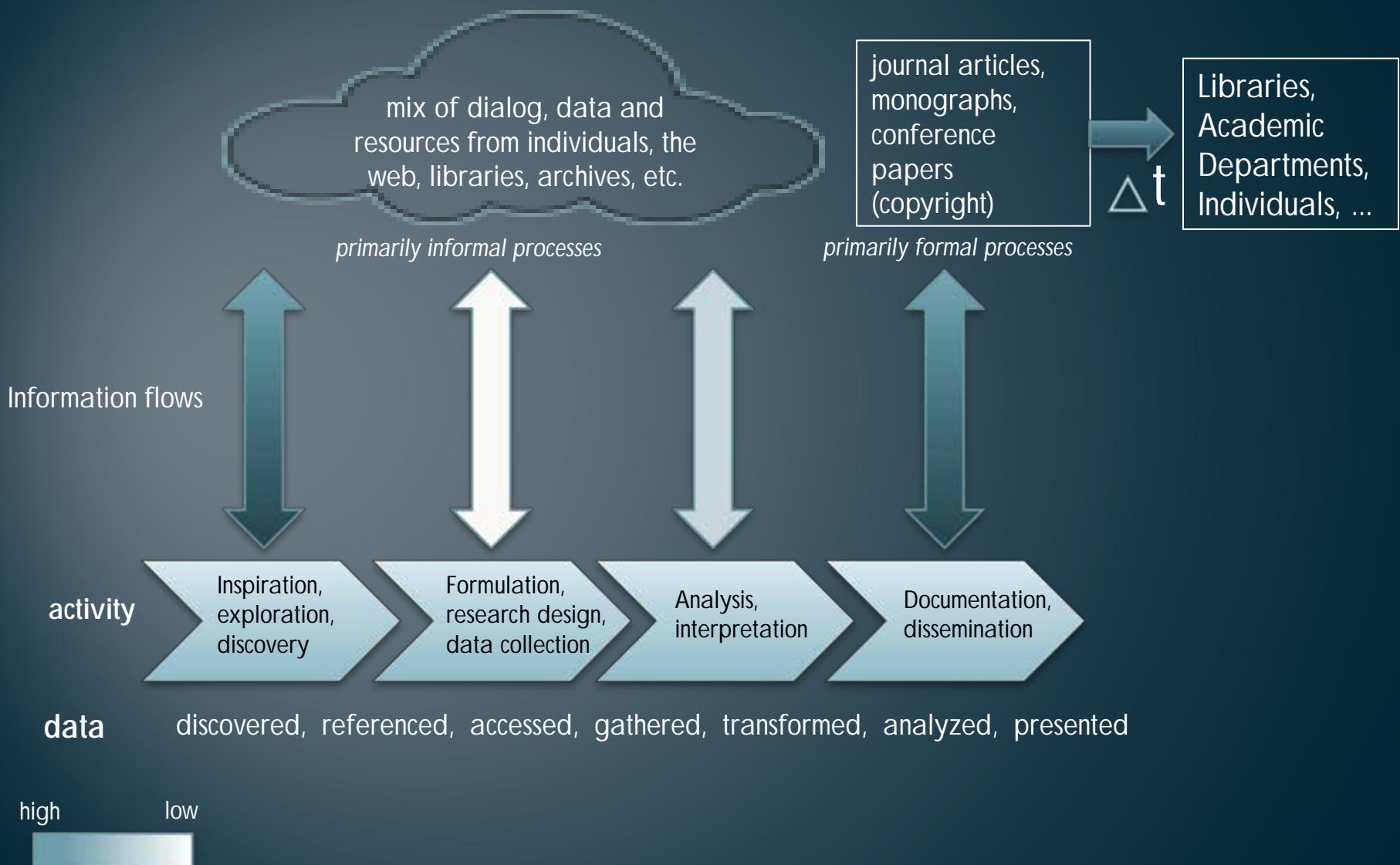
- *Repeatability* - the ability to duplicate an experiment under the same conditions and obtain the same result.
- *Reproducibility* - the ability for others to replicate the work in different environments and obtain the same result.

Necessary Conditions for Reproducibility

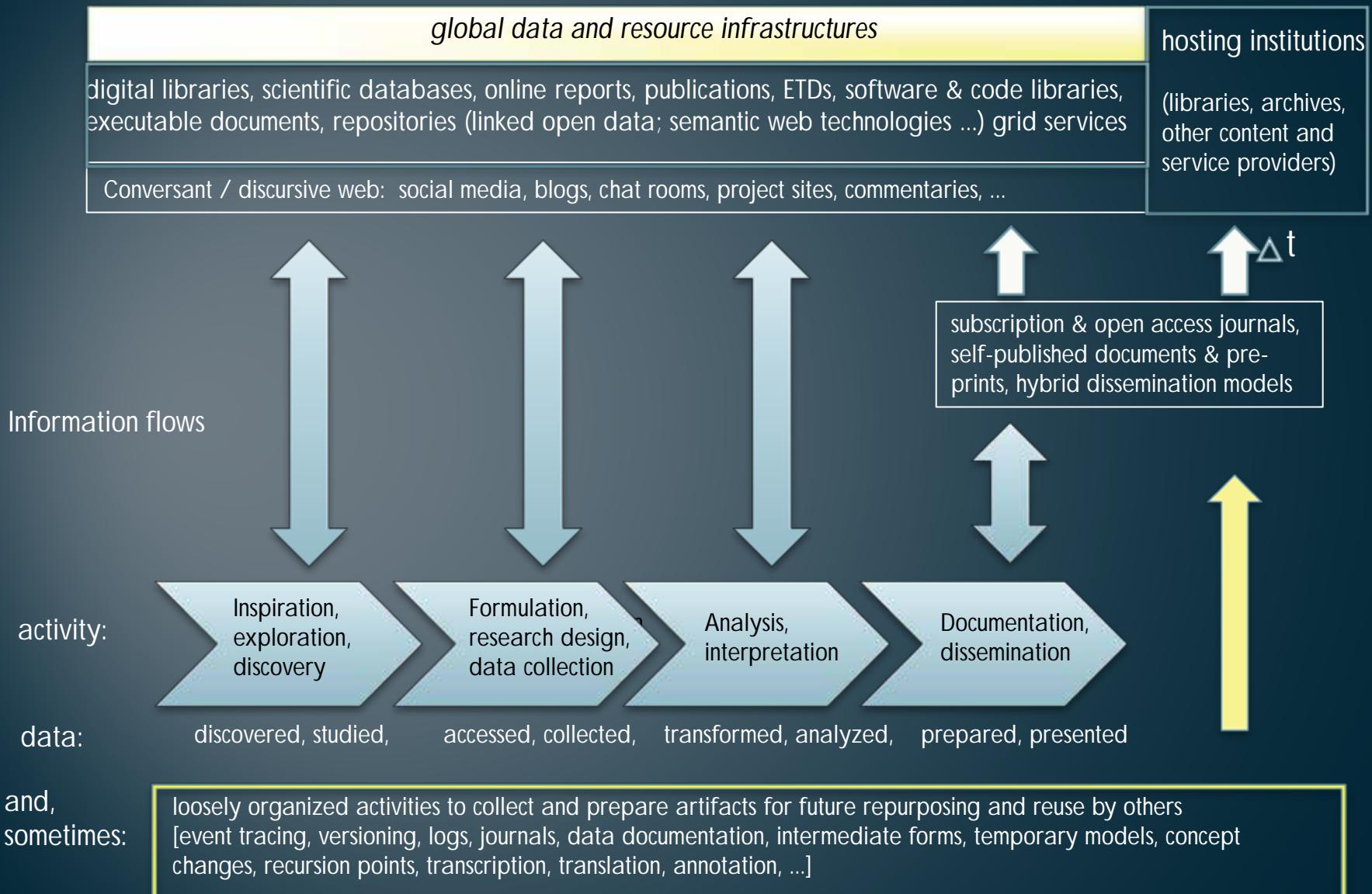
A comprehensive record of the research process and scholarly workflow including:

- Øprocess records: algorithms, software pipelines and versioning, datasets and transformations, storage formats and protocols, event tracing, ...
- Øresource descriptions: journals, logs, tools, methods, dialog, collaborative activities and external contributions, ...
- Øintermediate forms: temporary models, concept changes, recursion points, software versions, external dialogs and contributions ...
- Øworkflow artifacts: transcriptions, translations, annotations, steps taken to acknowledge distribution of effort, attribution and credit, ...

Traditional Scholarly Research Workflow



Evolving Workflow and Communication

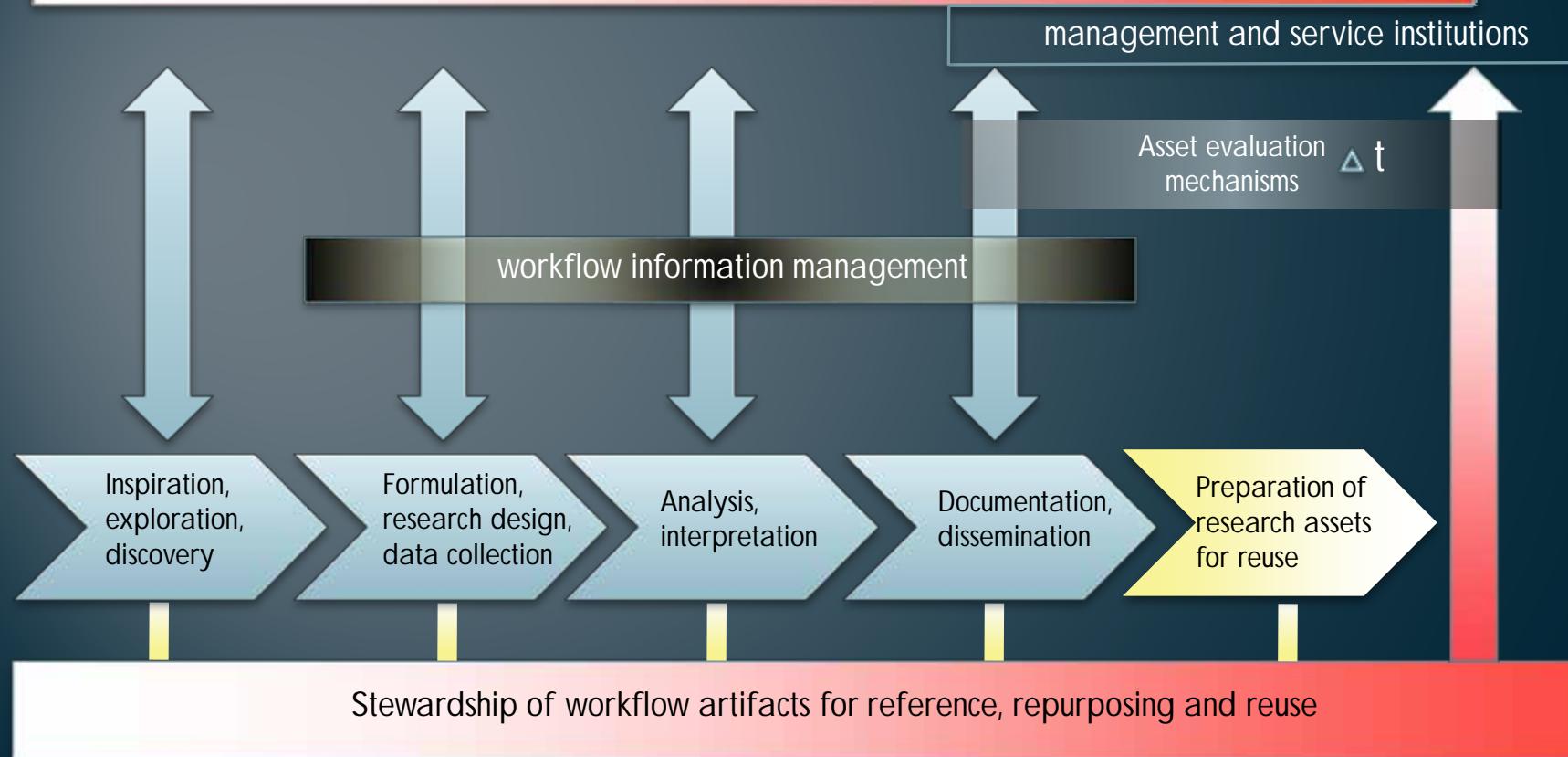


An Emergent Model for Scholarly Communication

scholarly communications layer: dynamic research reports with detailed descriptive information of workflow, methods and concepts as well as access to software, data and other experimental assets, provenance and citation linkages, etc. meeting community-endorsed practices for presentation, access, preservation and archiving

global data and research cyberinfrastructure: research data infrastructures, digital libraries, scientific databases, reports, publications, ETDs, software & code libraries, executable documents, repositories (linked open data; semantic web technologies ...), processing, storage, cloud and grid services

Conversant / discursive web: social media, blogs, chat rooms, project sites, commentaries, ...

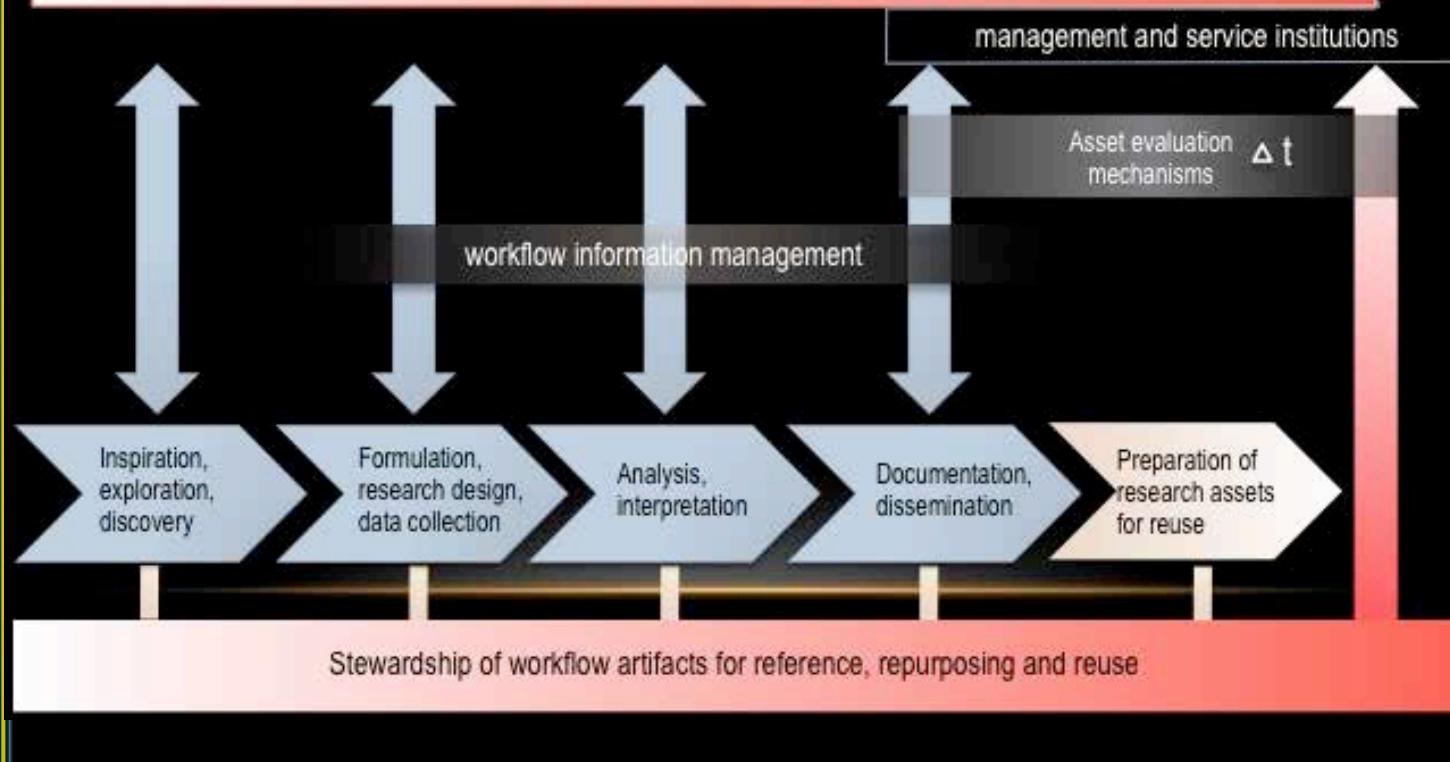


Libraries, Archives and Service Providers

scholarly communications layer: dynamic research reports with detailed descriptive information of workflow, methods and concepts as well as access to software, data and other experimental assets, provenance and citation linkages, etc. meeting community-endorsed practices for presentation, access, preservation and archiving

global data and research cyberinfrastructure: research data infrastructures, digital libraries, scientific databases, reports, publications, ETDs, software & code libraries, executable documents, repositories (linked open data; semantic web technologies ...), processing, storage, cloud and grid services

Conversant / discursive web: social media, blogs, chat rooms, project sites, commentaries, ...



In this model the role of Libraries evolves from one of holders and providers of knowledge resources to one of being an active partner in the research process. Libraries and librarians provide tools and expertise that expedite research and scholarship. Libraries have the institutional structure and many of the resources needed to advance and sustain scholarly workflows.

New Opportunities to Fill Emerging “Skill Gaps”*

1. Approaches to preserving research outputs
2. Knowledge of data management and curation, including ingest, discovery, access, dissemination, preservation, and portability
3. Knowledge to comply with the mandates of funders, including open access requirements
4. Knowledge of data manipulation tools used in the discipline/subject
5. Knowledge of data mining
6. Knowledge on the use of metadata and skills to develop metadata schema appropriate to discipline / subject standards and practices
7. Ability to preserve relevant project records, e.g. correspondence
8. Knowledge of sources of research funding to identify potential funders

* Anne Kenney, 2012 Symposium on Digital Curation, citing Mary Auckland, “Re-skilling for Research,” RLUK, January 2012

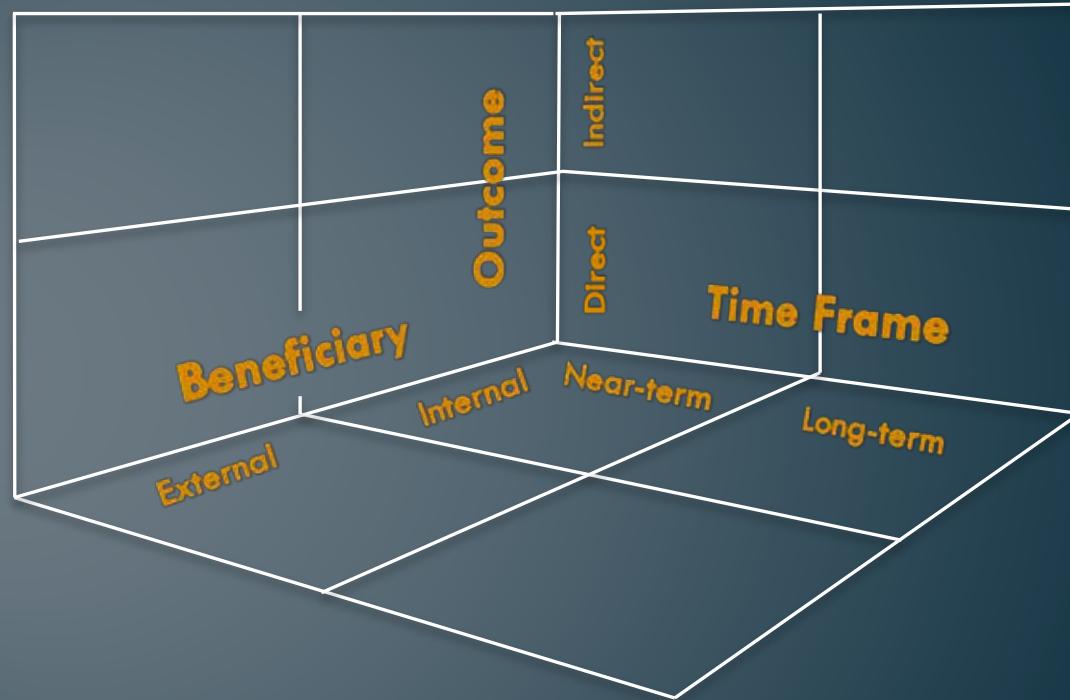
Coming to terms with Terms

- **Curation:** “managing and promoting the use of data from its point of creation, to ensure it is fit for contemporary purpose, and available for discovery and re-use”
- **Archiving:** “ensur[ing] that data is properly selected, stored, can be accessed and that its logical and physical integrity is maintained over time, including security and authenticity”
- **Preservation:** “maintain[ing] specific items over time so that they can still be accessed and understood through changes in technology”

Aspects of data quality

- **Discoverability and Accessibility** – Users can find and access data in straightforward manner.
- **Completeness** – All of the requisite information is available.
- **Interpretation** – The meaning of the data is unambiguous
- **Accuracy** – The data correctly represents the values it models
- **Consistency** – The data does not contradict itself, and values are uniform
- **Provenance & Reputation** – The data can be reliably tracked back to its original source, and the source of the data is legitimate.
- **Timeliness** – The data is up-to-date with regard to the task at hand

Value proposition for data quality



Getting it right matters

- Liability - Deepwater Horizon oil spill
 - BP, USCG, NOAA, DOE, USGS, ...
- Policy setting - Healthcare reform
 - Agency for Healthcare Research and Quality (AHRQ)
- Historical insight – 200K JPL mission tapes
 - Ozone hole verification
- Long-term, synoptic integration- International Comprehensive Ocean-Atmosphere Data Set (ICOADS)
 - Surface marine data over 3 centuries

Educating for digital curation

- Information professionals - specialists in digital or data curation
 - post-graduate education in schools of Library and Information Science (LIS) or Schools of Information (iSchools)
- Disciplinary specialists - scientific, industry, and government employees with data curation responsibilities
 - hybrid STEM programs, such as bioinformatics, geospatial research, and environmental informatics
- Mid-career employees - preparing for curation roles
 - on-the-job training, short courses, workshops, and conferences

“The future ain’t what it used to be.”

(Yogi Berra)

Come gather 'round people
Wherever you roam
And admit that the waters
Around you have grown
And accept it that soon
You'll be drenched to the bone.
If your time to you
Is worth savin'
Then you better start swimmin'
Or you'll sink like a stone
For the times they are a-changin'.



So, what's an archival educator to do?



Is that all there is?
Is that all there is?
If that's all there is, my friends,
Then let's keep dancing.*

*Peggy Lee, "Is That All There Is?" 1969