Big data in Library and Information Science

Koraljka Golub, koraljka.golub@lnu.se
& Joacim Hansson, joacim.hansson@lnu.se

*Big data: från hype till handling*

Linnaeus University

4 December 2015
Background

• Moore’s law: processing power per chip doubles every 18 months
  • So does data storage power
  • Also, exponentially growing data

• Many uses of big data
  • Gaining new knowledge, e.g., to provide new services, improve existing…
  • Problem: also, the complexity of the data has also accelerated

• USA: big data and data scientists adds
  • A shortage of 140,000 to 190,000 people with analytical expertise and 1.5 million managers and analysts who can understand and make decisions (McKinsey)
The essential task of the data science professional is to **transform raw, messy data into actionable knowledge** that can be used by decision makers.

The mission of librarianship is **to facilitate knowledge creation in communities** (source).

Challenges both similar and different to managing collections of documents:
- To make it re-usable, retrievable, preserved…
- Challenges: diversity of metadata and taxonomies, different standards, user needs… how to achieve interoperability?

Library skills occupy **the central ground between understanding information user needs on one end and data curation on the other** (source).

http://www.dcc.ac.uk/
Welcome to Research Data

We apply the principles of library and archival sciences to meet the challenges and needs of researchers and students in finding, using, preparing, managing, curating and preserving research data. Working with our liaison librarians, we can help you plan, describe, disseminate, steward and archive your datasets.

Have a question? Contact us at: researchdata@lib.purdue.edu

Our Services

- Purdue University Research Repository
- Data Management Planning
- Metadata & Data Documentation
- Distributed Data Curation Center
- Data Management & Curation Education
- Big Data @ Purdue Libraries
- Data Consultation
- Digital Object Identifiers (DOIs)
An addition to searching: visualisation

• Common search algorithms used by libraries may not suffice for big data
• Also in addition to existing search options for documents, e.g.,
  • The Visible Archive Project
  • 9,000,000 archival items in 65,000 series

• every box is a series
• chronological
• size of the box maps the size of the collection
• interactive
• linked to others

• source
Value and Problems of Big Data

• Value?
  • We do not really know the value of raw data

• Problems
  • Potential miscalculations and bad decisions
    • Popular science examples
    • Atheoretical, like bibliometrics

• Legal and policy issues

• The role of libraries:
  • Information literacy; quality selection
  • Data and Data analysis literacy; quality selection
Data Science Specialization in Master of Library Science: example

Academic libraries are hungry for librarians who can work with and manage big data projects. With a specialization in data science, you can work on the forefront of this new science and support the work of academic data scientists.

The specialization will appear on your transcript.

Requirements (Minimum of 36 Credit Hours)

M.L.S. Requirements (18 cr.)
- M.L.S. Foundation Requirements (18 cr.)

You must also complete the digital literacy requirement.

Specialization Requirements (15 cr.)

More than half are basic LIS courses
Library and Information Science at LNU: A selection

Topics

• Research data management and preservation
• Metadata
• Semantic web
• Information retrieval
• Evaluation of information
• User services and tools
• Bibliometrics and scientometrics

Projects

• With the University Library
  • Terminology registry
  • RDM with SND
• Interoperability of archives, libraries, museums in Sweden, Finland, Croatia
  • Metadata, resources, data
• A master course on research data management
• A digitisation analysis project
• Bibliometrics and scientometrics
BIG DATA ++
1. Big Data

• Cross-university
  • Library and Information Science, Economics, Biology, Social Sciences, Media Technology, Mathematics, Computer Science, Languages

• Cross-sector

• Towards an international Master at LNU
• Towards an LNU Centre of Excellence
2. Digital Humanities (*ideas incubator*)

- As a bridge between humanities and technology
- LNU recently joined the Scandinavian community on digital humanities
- Support for the initiative expressed by:
  - Lund’s Humanities Lab
  - Göteborg’s Digital Humanities Centre
3. Something bigger? *(concept)*

- Data, Information, Resources, Knowledge…
  - Data Science (Big Data)
  - Knowledge Management
  - Information Retrieval
  - Business Informatics
  - Health Informatics
  - Media and Communications
  - …
Specialisations at University of Illinois

Specializations

M.I.S.
- Chemical Information
- Data Science
- Digital Libraries
- Information Architecture

M.L.S.
- African Studies Librarianship
- Archives and Records Management
- Art Librarianship
- Chemical Information
- Children’s and Young Adult Services
- Data Science
- Digital Libraries
- Information Architecture
- Library Technology Management
- Music Librarianship
- Rare Books and Manuscripts Librarianship
A modern, international university in Småland
- with the regional platform as its base and the world as its stage.

Welcome to a university where everything is possible.
Suggestions and ideas?

- Buzz me (smirk)

Thank you!