In this together?

Scholars and Librarians in Digital Transformation

https://www.nps.gov/cany/planyourvisit/cataractcanyon.htm
THOUGHTS ABOUT MECHANISMS OF TRANSFORMATION IN LIBRARIES

ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?
“DIGITAL” BECOMES
“DIGITAL TRANSFORMATION”
The WWW as scholarly communication system

http://www.flickr.com/photos/hoonynoo/3092027765
Google as digital library project

Growth and Differentiation of WWW and Data

http://evolutionofweb.appspot.com/
Growth and Differentiation of WWW and Data, continued

http://evolutionofweb.appspot.com/
Growth and Differentiation of WWW and Data, cont.
Annual global IP traffic will reach 3.3 ZB (ZB; 1000 Exabytes [EB]) by 2021. In 2016, global IP traffic was 1.2 ZB per year or 96 EB (one billion Gigabytes [GB]) per month. By 2021, global IP traffic will reach 3.3 ZB per year, or 278 EB per month. (Cisco Visual Networking Index: Forecast and Methodology, 2016–2021)
25 years later: Digital Transformation

What’s Actually New About Today’s Newfangled Birth Control Apps?

These futuristic-sounding apps are on the rise, but it’s key to separate the data from the hype.
25 years later: An Image of Libraries
ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?
Trajectories of Digital Transformation
Innovations in Scholarly Communication

https://101innovations.wordpress.com
Innovations in Scholarly Communication

Impacts on Digital Research

• Abundance of connectivity
  – Fostering unhindered collaborative research

• Abundance of data
  – Fostering the rise of data-intensive research

• Abundance of tools
  – Fostering open science and reproducible science
Examples of Digital Research

DIGITAL HUMANITIES
From Digital Imaging to Web-Corpora
From Digital Imaging to Web-Corpora
Exploiting Web-Corpora in Maps
Example of Projects Using New Tools
Examples of Digital Research

DIGITAL SCIENCE
Digital Science

One possible definition:

*Digital science means a radical transformation of the nature of science and innovation due to the integration of ICT in the research process and the internet culture of openness and sharing.*

Example: Open Notebook Science

- “Executable Papers” generate papers online
- Reproducible Science is implemented
  - Transparent methods, always open to all
  - Worldwide online collaboration possible

1. **Workflow** R Markdown is a format for writing reproducible, dynamic reports with R. Use it to embed R code and results into slideshows, pdfs, html documents, Word files and more. To make a report:

   i. **Open** - Open a file that uses the .Rmd extension.
   ii. **Write** - Write content with the easy to use R Markdown syntax
   iii. **Embed** - Embed R code that creates output to include in the report
   iv. **Render** - Replace R code with its output and transform the report into a Slideshow, pdf, html or ms Word file.

http://rmarkdown.rstudio.com/ | CC-BY
Demo

• Distribution of Hawks in the USA
  – Data directly from the web
    • BISON, the United States Federal Resource for Biological Occurrence Data (JSON API)
  – Paper can be packed and archived (R Markdown)
Example: Blue Brain Project

• Objective is a massive model simulation
• Method is to test model against evidence
• Evidence is fed from literature mining
• > Combination of text-based / semantic and numeric methods
Large-scale extraction of brain connectivity from the neuroscientific literature.

Richardet R¹, Chappeller JC¹, Telefont M¹, Hill S¹.

Author information

Abstract

MOTIVATION: In neuroscience, as in many other scientific domains, the primary form of knowledge dissemination is through published articles. One challenge for modern neuroinformatics is finding methods to make the knowledge from the tremendous backlog of publications accessible for search, analysis and the integration of such data into computational models. A key example of this is metascale brain connectivity, where results are not reported in a normalized repository. Instead, these experimental results are published in natural language, scattered among individual scientific publications. This lack of normalization and centralization hinders the large-scale integration of brain connectivity results. In this article, we present text-mining models to extract and aggregate brain connectivity results from 13.2 million PubMed abstracts and 630 216 full-text publications related to neuroscience. The brain regions are identified with three different named entity recognizers (NERs) and then normalized against two atlases: the Allen Brain Atlas (ABA) and the atlas from the Brain Architecture Management System (BAMS). We then use three different extractors to assess inter-region connectivity.

RESULTS: NERs and connectivity extractors are evaluated against a manually annotated corpus. The complete in litero extraction models are also evaluated against in vivo connectivity data from ABA with an estimated precision of 78%. The resulting database contains over 4 million brain region mentions and over 100,000 (ABA) and 122,000 (BAMS) potential brain region connections. This database drastically accelerates connectivity literature review, by providing a centralized repository of connectivity data to neuroscientists.

© The Author 2015. Published by Oxford University Press.
DIGITAL TRANSFORMATION IN LIBRARIES
Rediscovering information privileges of libraries

OPEN ACCESS, RESEARCH ANALYTICS AND INTELLIGENCE
1. OPEN ACCESS SERVICES IN LIBRARIES
2. REFERENCE SERVICES
3. INSTITUTIONAL REPOSITORIES
4. SUBJECT REPOSITORIES
5. DISSERTATION SERVICES
6. SOFTWARE-DEVELOPMENT AND OPERATIONS
7. OPEN ACCESS POLICIES
8. REGISTRIES & AGGREGATORS
9. OPEN ACCESS JOURNALS
10. SERVICES FOR OA MONOGRAPHS
11. PUBLICATION FUNDS
12. BIBLIOGRAPHIC & BIBLIOMETRICS
13. LICENSING
Repositories as Bibliographic Tools
Reporting and Information Systems

• Systems
  – Dspace, Eprints, Fedora, Pure, Symplectics, Converis...

• Interoperability
  – Financial System
  – Human Resource Management
  – Facility Management System
  – Campus Management System
  – Bibliographic Databases
    • WoS, Scopus, ArXiV, PMC, IRs/BASE
  – Authoritative Data Resources /Disambiguation
    • Vocabularies, Ontologies, ORCID/AuthorClaim
Novel forms of Information curation

RESEARCH DATA
The Göttingen eResearch Alliance is an initiative of the University of Göttingen to assist all researchers on the Göttingen Campus (GC) with eResearch related questions and data management issues. As a central point of contact for researchers, research associations and faculties the eResearch Alliance represents the University’s joint forces of the central infrastructure providers, the Göttingen State and University Library (SUB) and the Göttingen University’s Computing and IT Competence Centre (GWDG).

Your research project! | Your data! | Our services!

We understand eResearch as enhanced research, which to us means an optimized usage of digital technologies and methods for innovative research. We offer information, personal advice and support for key issues related to digital research through all phases of the research life cycle:

- **Ideas**
  - Project proposal support
  - Data management planning
  - Expert network

- **Research**
  - Workshops & Trainings
  - ICT services
  - Visualisation & Exploration
  - Data strategy implementation

- **Results**
  - Persistent Identification
  - Data publication
  - Long term archiving

News

- 19.03./20.03.2018: Pre-RDA Symposium in Göttingen: The critical role of university RDM infrastructure in transforming data to knowledge
- New project GRACE investigates costs and scalability of RDM
- 12.07.2017: Göttingen Research Bazaar at Data Science Summer School
- 10.07. - 21.07.2017: Data Science Summer School

Guidelines

- Policies on Research Data and Open Access as "Amtliche Mitteilung" (PDF, German only)
- Research data policy of the Georg-August-University Göttingen (incl. UMG) - English version
- Open Access policy of the Georg-August-University Göttingen (incl. UMG) - English version
CRC 990 EFForTS/INF-Project
Role model *embedded data manager*

- Large interdisciplinary research project (CRC) with integrated service project for information & infrastructure (INF)
- 100+ researchers, field work: Indonesian Rainforest
- Lots of different data require management
- Provide an information system for research data, write an Excel transformation script, moderate sharing agreements...
- “Embedded Data Managers” know: IT Basics, data modeling, concepts of metadata… and how to: analyse requirements, moderate discussions, search information, learn new things rapidly…

Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (Sumatra, Indonesia)
Example: CRC 1002
Electronic laboratory notekeeping

Paper lab notebook

Electronic lab notebook
Libraries as a space—same but very different

OPEN SCIENCE
Open Science: Göttingen Meet-up

- Initiated in December 2016
- Focus on Open Science topics: research data, publishing, peer review, research integrity, etc.
- Junior researchers, librarians, research management, HE didactics, from UGOE, Max Planck Institutes, etc.
- Outreach primarily via graduate schools’ mailing lists

Work mode
Quarterly Meet-ups with short presentations & group discussions
Working groups (monthly): Teaching Open Science, HackyHour
Open Science: Replication WIKI

Welcome to Replication in Economics!
8,566 articles

Find one of 2,097 studies to replicate
Browse in 144 journals, 29 Working paper series & blogs and 18 books
Browse by methods, software used, JEL codes, keywords, authors, data type, geographical origin of data, and data source
Find one of 350 replications and 6 Corrections

Enter a new:
replication  replicated study  Study lacking replication

Join the project
You can join the project by entering a new replication or empirical studies that could be replicated. For the latter you can vote which ones you find most relevant to be replicated. Furthermore, you can improve the articles in the wiki and make comments. Under current events you can announce news and upcoming events. The community portal lists open tasks and suggestions for the wiki.

Most recently added replication studies
- 3 October 2017: News Shocks and the Slope of the Term Structure of Interest Rates: Comment (AER 2017)
- 3 October 2017: The Cyclicality of Sales, Regular, and Effective Prices: Business Cycle and Policy Implications: Comment (AER 2017)
- 1 October 2017: Shy of the Character Limit: "Twitter Mood Predicts the Stock Market" Revisited (EJW 2017)
- 1 October 2017: A comment on "Three reasons to use annual payments in contingent valuation" (J Env Econ & Management forthcoming)

News
AIRLEAP @ Conference "An Urgency for Evidence and Transparency in Economic Analysis and Policy" in St Louis, USA, October 13th-14th will have a session on replication in empirical economics

Find studies and replications for your own research. You can search for software and methods used.

7 of 11

About the project
For scientific progress, it is pivotal to review research findings by independently replicating results, thus making the findings more reliable. However, in econometric research, it is not yet common practice to publish replication findings. This wiki serves as a database of empirical studies, the availability of replication material for them, and of replication studies. It can help teaching replication to students. Seminars at several faculties internationally were already taught for which the information of this database was used. So far the focus has been on some leading journals in economics. Replication results can be published as...
Re-thinking information literacy

DIGITAL “ENLIGHTENMENT”
Education Engagement

• Propaedeutika Digitale
  – A minimum standard for the university graduates
  – A segway between different degrees

• Librarians teaching in formal degrees
  – e.g. Digital Humanities MSc, Data Science MSc

• Librarians addressing digital “ethics”
  – e.g. digital research conduct, security
Libraries are transforming, indeed

- Open Access
  - A vehicle for innovation in libraries in an “old” field
- Research Analytics and Intelligence
  - Rediscovering information privileges of libraries
- Research Data
  - Novel forms of information curation
- Open Science
  - Libraries as a space—same but very different
- Digital “Enlightenment”
  - Re-thinking information literacy

> Nota bene: all of these examples build on established services
But libraries do not change fast

Most assessment measures in libraries are still traditional
THOUGHTS ABOUT MECHANISMS OF TRANSFORMATION IN LIBRARIES

ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?
SUB Göttingen
**Fakten und Zahlen in Kürze**

**Medienbestand**

7 932 106

Media holdings

**Facts and numbers in short**

**Medienbestand**

7 932 106

Media holdings

**Kurse zur Informations- und Medienkompetenz**

369

Courses on information and media literacy

**Standorte**

9

Locations (incl. LSG)

**Aktive Benutzerinnen und Benutzer**

43 208

Active users

**Öffnungstage pro Jahr**

354

Opening days per year (Central Library)

**Anteil elektr. Medien an Erwerbungsausgaben**

71.5%

Share of electronic media in acquisition expenditure

**Öffnungsstunden pro Jahr**

6 032

Opening hours per year (Central Library)

**Erwerbungsausgaben (Euro)**

4 699 554

Purchase orders (Euro)

**davon für elektronische Medien**

3 360 091

Thereof for electronic media

**Digitisierte historische Bestände (Seiten)**

20 834 335

Digitized historical holdings (pages)

**Bibliotheksbesuche (mit Lern- und Studiengebäude – LSG)**

2 720 913

Library visits (incl. Learning and Study Building – LSG)

**Alle 1,1 Sekunden wurde 2016 ein Service der SUB Göttingen genutzt – sei es, dass Sie im Katalog recherchiert, ein Dokument heruntergeladen oder unsere Bibliothek besucht haben, um bei uns zu arbeiten und Kurse oder andere Veranstaltungen zu besuchen.**

In 2016, every 1.1 second a service of SUB Göttingen was used – be it your search in the catalogue, the download of a document or your visit in our library to study, take a course or visit an event.

**Kurse zur Informations- und Medienkompetenz (Teilnehmende)**

4 824

Courses on information and media literacy (participants)

**Wissenschaftliche/kulturelle Veranstaltungen**

264

Scientific/cultural events

**Öffnungstage pro Jahr**

354

Opening days per year (Central Library)

**Anteil elektr. Medien an Erwerbungsausgaben**

71.5%

Share of electronic media in acquisition expenditure

**Öffnungsstunden pro Jahr**

6 032

Opening hours per year (Central Library)

578 318

Face-to-Place-Beratung

**Download/Dokumentenzugriffe**

1 543 414

Downloads/document access repositories

**Dokumentenserver**

8 298 411

Scans, copies, print outs

**Wissenschaftliche/kulturelle Veranstaltungen (Teilnehmende)**

25 258

Scientific/cultural events (participants)

**Buchungen von Arbeitsplätzen (mit LSG)**

189 453

Booking of study spaces (incl. LSG)
# Library Statistics 2016

**Facts and numbers in short**

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active users</td>
<td>43,208</td>
</tr>
<tr>
<td>Locations</td>
<td>9 (incl. LSG)</td>
</tr>
<tr>
<td>Opening days per year (Central Library)</td>
<td>354</td>
</tr>
<tr>
<td>Opening hours per year (Central Library)</td>
<td>6,032</td>
</tr>
<tr>
<td>Media holdings</td>
<td>7,932,106</td>
</tr>
<tr>
<td>Digitized historical holdings (pages)</td>
<td>20,834,335</td>
</tr>
<tr>
<td>Götttingen University Press publications</td>
<td>48</td>
</tr>
<tr>
<td>Purchase orders (Euro)</td>
<td>4,699,554</td>
</tr>
<tr>
<td>Thereof for electronical media</td>
<td>3,360,091</td>
</tr>
<tr>
<td>Share of electr. Media in acquisition expenditure</td>
<td>71.5%</td>
</tr>
<tr>
<td>Courses on information and media literacy</td>
<td>369</td>
</tr>
<tr>
<td>Scientific/cultural events</td>
<td>264</td>
</tr>
<tr>
<td>Share of external funding in overall funding</td>
<td>15.2%</td>
</tr>
</tbody>
</table>
Usage of services of SUB Göttingen

In 2016, every 1.1 seconds a service of SUB Göttingen was used—be it your search in the catalogue, the download of a document or your visit in our library to study, take a course or visit an event. We are pleased with your intensive usage of our services!

<table>
<thead>
<tr>
<th>Service</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library visits (incl. Learning and Study Building—LSG)</td>
<td>2,720,913</td>
</tr>
<tr>
<td>Website visits (without catalogue)</td>
<td>1,939,929</td>
</tr>
<tr>
<td>Catalogue search</td>
<td>10,049,724</td>
</tr>
<tr>
<td>Circulation and inter-library loans</td>
<td>798,867</td>
</tr>
<tr>
<td>Downloads e-journals</td>
<td>2,001,557</td>
</tr>
<tr>
<td>Downloads/document access repositories</td>
<td>1,543,414</td>
</tr>
<tr>
<td>e-Book usage sessions</td>
<td>1,017,458</td>
</tr>
<tr>
<td>Online database record views</td>
<td>327,960</td>
</tr>
<tr>
<td>Booking of study spaces (incl. LSG)</td>
<td>189,453</td>
</tr>
<tr>
<td>Scans, copies, print outs</td>
<td>8,298,411</td>
</tr>
<tr>
<td>Face-to-face-advising</td>
<td>578,318</td>
</tr>
<tr>
<td>Courses on information and media literacy (participants)</td>
<td>4,824</td>
</tr>
<tr>
<td>Scientific/cultural events (participants)</td>
<td>25,258</td>
</tr>
</tbody>
</table>
Changed Requirements

User Survey: How important is the development of the following services for you?

- Electronic Media: 80.0% Extend, 19.2% Maintain, 0.7% Reduce
- Printed Media: 45.9% Extend, 46.1% Maintain, 8.0% Reduce
- Discovery: 42.1% Extend, 54.1% Maintain, 3.8% Reduce
- Publication Management: 35.6% Extend, 57.8% Maintain, 6.7% Reduce
- Research Data Management: 36.9% Extend, 52.2% Maintain, 10.9% Reduce
- Communication: 24.7% Extend, 69.6% Maintain, 5.7% Reduce

n=9507
Göttingen Library Services

1. Learning Space & Venue
2. Information Provision
3. Publication Services
4. Research Services
5. Cultural Heritage
CAMPUS GÖTTINGEN
# University Figures

## Students

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>Students in total</th>
<th>thereof: female</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>27,556</td>
<td>13,995</td>
<td></td>
</tr>
</tbody>
</table>

## Graduates

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>thereof: female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4,287</td>
<td>2,325</td>
</tr>
</tbody>
</table>

## Doctorates

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>thereof: female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>751</td>
<td>372</td>
</tr>
</tbody>
</table>

## Section: Staff

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>Staff incl. Assistants</th>
<th>thereof: female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15,245</td>
<td>9,078</td>
<td></td>
</tr>
</tbody>
</table>

## Staff

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>thereof: female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12,199</td>
<td>7,403</td>
</tr>
</tbody>
</table>

## Section: Finances

<table>
<thead>
<tr>
<th>Year</th>
<th>University incl. Medicine</th>
<th>Expenses</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>1,063,141,147</td>
<td>1,085,157,546</td>
<td></td>
</tr>
</tbody>
</table>
Over 5,200 researchers work at the Göttingen Campus. The institutions have an overall budget in excess of 1.3 billion euros (basic funds and third party funding).

26 Directors of a non-university research institution at the Göttingen Campus have simultaneously received a joint appointment as University Professor.
LIBRARY INNOVATION CAPACITY
Innovation Capacity at SUB—History

►►► Shifting c. 50 permanent FTE (25%) in 20 years, c. 5M € /pA Grant Funding

1997—Digitalisation Centre (Imaging): 7,5 FTE ►► 2009 Closing Imaging Studio
2002—Licensing: 4 FTE ►► 2015 e-only
2003—Research & Development Dep.: 5 FTE
2004—Electronic Publishing Dep.: 7 FTE
2009—Digital Library Dep.: 15 FTE
2009—Metadata: 3,5 FTE ►► also, 2003 Schließung Nds. Zentralkatalog
2014—eResearch Alliance: 2,0 FTE (+ Grant funding)
2016—First “wave” of analysts: 5,0 FTE
2017—eResearch Alliance: another 3,0 FTE (+ additional funding)
Innovation Capacity at SUB – Budget overview

SUB—Budget 2015, in Mio. EURO

[Diagram showing budget allocations]

- Recurrent Budget: 16.57 Mio. EURO
- Tuition Share: 4.22 Mio. EURO
- Other Funds: 3.76 Mio. EURO
- Grants: 0.93 Mio. EURO
- Other Incomes: 1.22 Mio. EURO
Innovation Capacity at SUB—Appointment Capacity

Positions becoming available per year
Recurrent Prefilling Program, 25% each year

2017 Appointments

• Legal Analyst—IPR and Licensing
  – Exploitation, open rights

• Data Analyst—Scholarly Information
  – Bibliometrics, Information Market

• Data Analyst—Spatial and Geospatial Data
  – GIS as cross-cutting data type

• Data Analyst—Language and Text Data
  – Corpuslinguistics und ‘Mining’ as a Service

• Process Analyst—ILS, Cataloguing, Search
  – Refactoring the ‘Enterprise’ Systems
Summary—Transformation Capacity at SUB

- Loss of 28 FTE 2003–2013 ‘digested’
- Transformation of c. 50 FTE 1996–2016
- High fixed-term innovation capacity through grant funding
  - Capacity for advance appointments very limited
- First wave of continuous pre-filling program
  - We need more waves and a substantial temporary bridge investment
Considerations of Transformation

- The pace of library transformation \( (T) \) is determined by the ratio of tasks being reduced \( (R) \) and tasks added \( (A) \), dependent on factors:
  
a. Appointment Capacity
    - The ability to appoint new staff for new tasks
      - Can be boosted by additional grants
      - Can be boosted by pre-filling
  
b. Re-Skilling Capacity
    - The ability to train new staff for new tasks
  
c. Task Efficiencies
    - The amount of existing tasks being compressed
  
d. Task Cancellations
    - The amount tasks discontinued

\[
T = \frac{A}{R} = \frac{a+b}{c+d}
\]
Example with 2% change per year

- With 2%, it takes 25 years to transform 50% of tasks.
Example with increased pace

Growth factors of “innovation demand” and “continuity demand” added

- It takes c. 10 years to transform c. 50% of tasks; transforming 5% a year.
CONCLUSION

ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?

THOUGHTS ABOUT MECHANISMS OF TRANSFORMATION IN LIBRARIES
Research & Library: same trajectory?

• No
  – Libraries require conservative function and have (thereby) intrinsically limited innovation capacity
  – Stronger innovation drivers in research

• Yes
  – Performance Indicators / values still based on traditional publishing
  – Both are in transformation
Trajectories of Digital Transformation

Intensity

Transformation of Scholarly Communication

This is what can happen, when keeping the current pace

Transformation of Library Services

Time

1990  2000  2010  2020  2030
Trajectories of Digital Transformation

This is what can happen with some luck and a major transformation program.
THANK YOU