In this together?

Scholars and Librarians in Digital Transformation



ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?

THOUGHTS ABOUT MECHANISMS OF TRANSFORMATION IN LIBRARIES

The WWW as scholarly comunication system



Google as digital library project





WANTER DRIVEN THE REAL OF

The Original Googl

8

In 1996 Larry Page and Serge students in Stanford CSD, Digital Library Project, new amount of diskspace to tes Pagerank ** algorithm on a wide-web data. At that time hard disks were the largest

they assembled 10 of these low-cost cabinet. In Nov 1999, Google Inc. by the one of the primary search e

web, provided replacement capacity to the Digital Lifra that we could move this original assembly into our history d As of September 2000, Google, in Mountain View, operated searching and web crawling LINUX operating system.

25 years later: Digital Transformation

Smithsonian.com

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



Fertility apps promise to help women both get pregnant and avoid pregnancy. But how reliable are they? (Przemyslaw Klos / Alamy)

25 years later: An Image of Libraries



ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?

Trajectories of Digital Transformation



Trajectories of Digital Transformation



Innovations in Scholarly Communication



https://101innovations.wordpress.com

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

TextGridLab			- 0	X
Datei Bearbeiten XML Werkzeug	e Fenster <u>H</u> ilfe			
🟫 🕂 🛨 🔛 🔛 🏠 🕯	🕯 🔍 🔄 🧶 📄 🖳 🤣 🏠 🖻 🔍 🕐 😾 🖉	- 🖓 - 🕴 🖋 🦼 🔛 🗟 📩 🗙 🛼	🖹 🔝 Text-Bild-L	in) »
🐼 Navigator 🛛 🗖 🗖	🔠 * Faksimile 🛛		🍋 🐚 👘 🥓	
	Aleonin Gr Morio La Morio La Intrastian Morio La Intrastian Morio La	Text mit gewählter Markierung verknüpfen Verknüpfung für ausgewählte Markierung(en) aufhebe Aktive Markierung(en) entfernen Alle Markierungen/Verknüpfungen entfernen Zu verknüpftem Text springen Andocklinien	en	
20	International In	Schreibmodus	+ 200	
	🖹 *Transkription.xml 🕅	Textebenen	· · ·	
→	der> ⊖ ⊖y> ⊖ <l n="1" xml:id="l1"> ⊖ <app> ☐ lem wit="#stmatthiasT0048" re <rdg #pv"="" resp="#Baehrens1881 #Scaf
</app></td><td>Bildgröße
sp=">{anchor xml:id="a4 start" />Oram<a fai1982">iram</a </rdg></app></l>	anchor xml:id="a4 end" />	* <mark>=</mark>	
	<pre>pande mihi</pre>	<pre>sp="#pv">pelidae "#Baehrens1881 #Scaffai1982">Pelidae icaffai1982">, icaffai1982"> icaffai1982"></pre>	,	-
0			max.mustermann@textgrid.de	



From Digital Imaging to Web-Corpora



Distant Reading Showcase

200 Jahre deutsche Dramengeschichte auf einen Blick

Rechts im Bild sind der Personennetzwerke von 465 Dasmen aus den sinher 1331 – 1533, die Statik auf der Bernetzeit zu ander Ferne geeignet: Man erkennt zum Belopial, wann die Autoren anfangen, Shakespaare zu lesen: Die Tigenennetzwerke exploideren. Die Darmentlich ind abgelährt fluchdnes Ratespiel für zwischnetwicht, Generien warde das Poster autorantich mit ussenen Superposterstof dramaväk.















Example of Projects Using New Tools

Projekte und ihre Institutionen im Kontext von DARIAH-DE und TextGrid



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.

Digital Science in Horizon 2020 https://ec.europa.eu/digital-single-market/en/news/digital-science-horizon-2020

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



- 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

SNCBI Resources 🕑 How To 🗹	
Publed.gov PubMed	
US National Library of Medicine National Institutes of Health	Advanced

Format: Abstract -

Send to -

Bioinformatics. 2015 May 15;31(10):1640-7. doi: 10.1093/bioinformatics/btv025. Epub 2015 Jan 20.

Large-scale extraction of brain connectivity from the neuroscientific literature.

Richardet R¹, Chappelier 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

Rediscoverig information privileges of libraries

OPEN ACCESS, RESEARCH ANALYTICS AND INTELLIGENCE

12 POSSIBLE SERVICES THROUGH OA

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. BIBLIOGRAPHICS & BIBLIOMETRICS**
- **13. LICENSING**

Praxishandbuch Open Access Hrsg. v. Söllner, Konstanze / Mittermaier, Bernhard DE GRUYTER SAUR ISBN 978-3-11-049406

From collecting to connecting – the role of libraries in Open Access

Wolfram Horstmann, 13.12.2016

A short history of libraries and Open Access

In the beginnings of Open Access, libraries were not directly involved. The Budapest Open Access Initiative had no representative of a major library among them (cf. Chan et al., 2002). The Berlin Declaration had the German library association as an initial signatory but still a systematic involvement of libraries worldwide was missing (Berliner Erklärung, 2003). However, first systematic accounts of the role of libraries were discussed (Suber, 2003). Predominantly, these started from the hope that Open Access would be a solution to the 'serial crisis', i.e. the negative spiral of price increases for academic journals introduced by publishers forcing a reduction of subscriptions by libraries, yielding, in turn, to increased pricing for journals since publishers wanted to keep revenues. Whether or not the serial crisis was (or is) the appropriate motivation for Open Access services in libraries remained controversial, even among the Open Access proponents (Bosc & Harnad, 2005). On the one side, rationales dominated that claimed to keep the publishing system largely untouched by

Repositories as Bibliographic Tools





Novel forms of Information curation

RESEARCH DATA



Search here

About

Q

The Göttingen *e*Research Alliance is an initiative of the University of Göttingen to assist all researchers on the Göttingen Campus (GC) with *e*Research related questions and data management issues. As a central point of contact for researchers, research associations and faculties the *e*Research 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 *e*Research 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:



- 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

more

News

Blog

- 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
- 25.07./26.07.2017: Workshop
 "Next Generation Medicine?"
- 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 modelling, 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)



Datamanager Timo Gnadt in Indonesia, photo taken by Daniel Kurzawe



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

Re-thinking information literacy

DIGTIAL ,ENLIGHTENMENT'

Education Engagement

- Propaedeutika Digitale
 - A mininum 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
 - Rediscoverig information privileges of libraries
- Research Data
 - Novel forms of information curation
- Open Science
 - Libraries as a space same but very different
- Digtial ,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



ARE SCHOLARS AND LIBRARIANS ON THE SAME TRAJECTORY?

THOUGHTS ABOUT MECHANISMS OF TRANSFORMATION IN LIBRARIES

Library

SUB GÖTTINGEN







http://www.reinerbeckerarchitekten.de/cms/wp-content/uploads/WH1522-2718-25.jpg

Fakten und Zahlen in Kürze

Facts and numbers in short

Medienbestand

7932106 Media holdings

Kurse zur Informations- und Medienkompetenz

Courses on information and media literacy

Locations (incl. LSG)

Standorte (mit LSG)

Öffnungstage pro Jahr (Zentralbibliothek)

Wissenschaftliche/ kulturelle Veranstaltungen Opening days per year (Central Library)

Öffnungsstunden pro Jahr Scientific/cultural events

(Zentralbibliothek)

Opening hours per year (Central Library)

Erwerbungsausgaben (Euro)

699 Purchase orders (Euro)

Digitalisierte historische Bestände (Seiten)

2083433 Digitized historical holdings (pages)

Aktive Benutzerinnen und Benutzer

43208 Active user

Anteil elektr. Medien an Erwerbungsausgaben

davon für elektronische Medien

Share of electr media

in acquisition expenditure

Thereof for electronical media

Bibliotheksbesuche (mit Lern- und Studiengebäude – LSG)

STATISTIK 2016

272091 3 Library visits (incl. Learning and Study Building - LSG)

578318

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)

Courses on information and media literacy (participants)

Downloads/Dokumentenzugriffe Dokumentenserver

Face-to-Face-Beratung

543414 Downloads/document access repositories

Wissenschaftliche/kulturelle Veranstaltungen (Teilnehmende)

Scientific/cultural events (participants)

Scans, Kopien, Drucke 8 298 411 Scans, copies, print outs

Face-to-face-advising

Buchungen von Arbeitsplätzen (mit LSG)

Booking of study spaces (incl. LSG)

Changed Requirements

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



Göttingen Library Services

- 1. Learning Space & Venue
- 2. Information Provision
- 3. Publication Services
- 4. Research Services
- 5. Cultural Heritage



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

Innovation Capacity at SUB – Appointment Capacity



Recurrent Prefilling Program, 25% each year

2017 Appointments

- 1. Legal Analyst IPR and Licensing
 - Exploitation, open rights
- 2. Data Analyst Scholarly Information
 - Bibliometrics, Information Market
- 3. Data Analyst Spatial and Geospatial Data
 - GIS as cross-cutting data type
- 4. Data Analyst Language and Text Data
 - Corpuslinguistics und 'Mining' as a Service
- 5. Process Analyst ILS, Cataloguing, Search
 - Refactoring the ,Enterprise'-Systems

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



Trajectories of Digital Transformation



THANK YOU