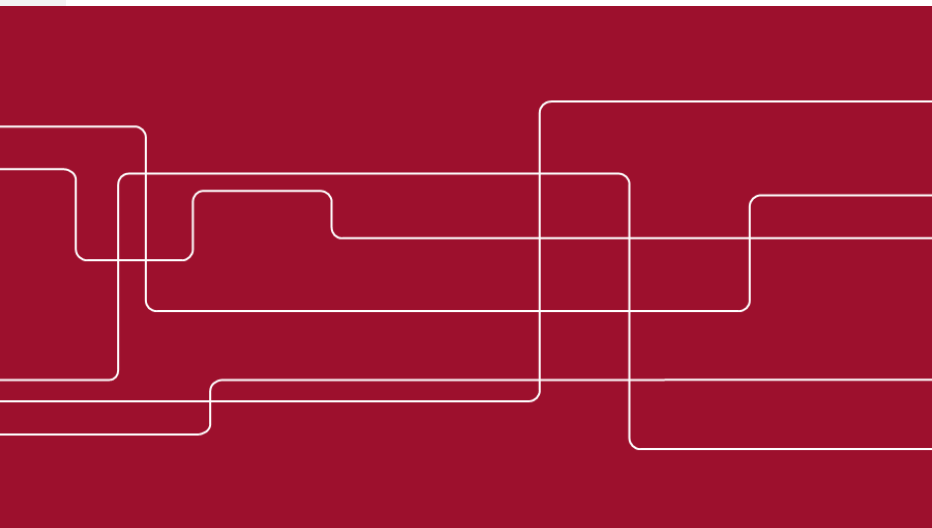




Quadcopters or Linguistic Corpora

Establishing RDM Services for Small-Scale Data Producers
at Big Universities

47th LIBER Annual Conference 2018



Dr Göran Hamrin
Logician Lecturer Librarian
KTH Library Director of Studies

Dr Viola Voß
Linguist Liaison Librarian
ULB Münster Subject Services Humanities

Outline

- KTH + KTHB & WWU + ULB MS
- RD(M) in engineering & in the humanities
- RDM in Sweden & in Germany
- RDM at KTHB & at ULB MS
- Conclusions

Two cultures, two countries, two universities

- RDM questions by researchers are quite similar: information for grant applications, storage possibilities, data protection, DMPs, data reuse, ...
- 1959: C.P. Snow’s Cambridge Lecture “The Two Cultures”
 - › Division between “the scientists” and “the intellectuals”
- 2018: How does the situation look today?
 - › Differences in RD(M) for these two cultures?
 - › Differences in RDM services in our countries?
 - › Differences in RDM services in our libraries?

KTH & KTHB

Kungliga Tekniska högskolan (KTH)

- KTH Royal Institute of Technology
- Sweden's largest technical university
- 13 000 students, 3 700 faculty, 2 000 PhD students

KTH Library (KTHB)

- Founded in 1827, around 50 staff
- KTHB services
 - › Publication database KTH-DiVA
 - › Open Access support
 - › Bibliometrics evaluations

WWU & ULB MS

Westfälische Wilhelms-Universität Münster (WWU)

- Founded in 1780/1902, 5th biggest German university
- ~ 43 000 students, 675 professors, 5 050 faculty
- 15 faculties, ~ 120 subjects in ~ 280 degree courses

Universitäts- und Landesbibliothek Münster (ULB MS)

- University and regional library, founded in 1588
- 248 staff for 182 FTE
- Library system: ~ 100 libraries = 6.25m vols. (print & digital)
- Open Access services since 2002
- WWU Research Data Policy 2017
- WWU “eScience-Center” with RDM services 2017

RD in engineering

- Data is often quantitative and of ordinal type
 - » computations on data are pretty easy
- But data sets can be large! Or sensitive!
 - » RDM is not always easy

Examples of RD in Engineering

- **Empirical-inductive processes**

- › Computer simulation data
- › Fluid mechanics data
- › Geopositioning data

- **Quadcopters!**

- **Some data sensitive!**

- › Biomedical data
- › Traffic data
- › Geopositioning data

RD(M) in the humanities

– Disciplines

- › All non-natural and non-technical sciences
- › *Digital* humanities: use of computer-assisted methods and digital(ised) resources and the reflection of these uses

– Data characteristics and usage

- › Representations of cultural artefacts (texts, images, audio or video recordings, other physical objects)
- › Data often modelled during research (digitising, describing, sorting, annotating, visualising, interpreting)
- › Different perspectives, formats, aggregation levels

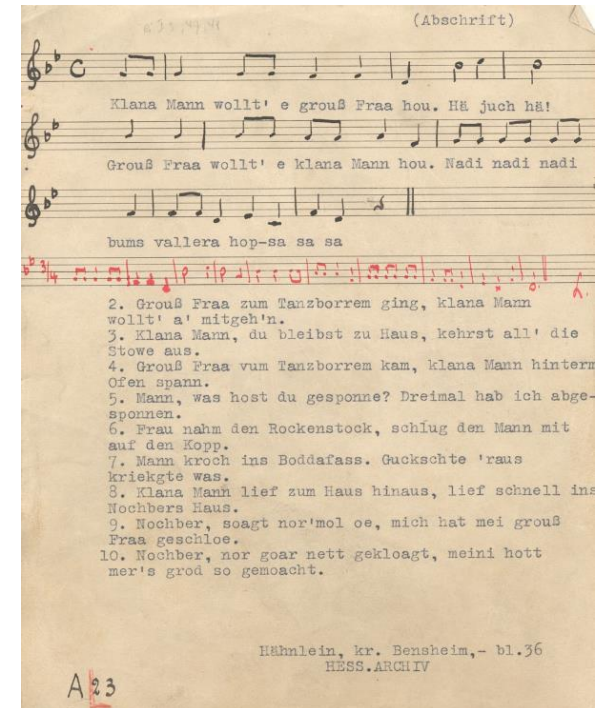
– Complex situation

- › Diverse types of data in different layers
- › Linked to other data types and sources
- › “Corralled” in specific technical settings › “living systems”

Examples of RD in the Humanities

– Music information retrieval for handwritten folksongs

- › Analyze handwritten music scores
- › Transcribe to machine-readable music?
- › Transcription platforms for crowdsourcing?
- › <https://doi.org/10.18452/18952>
- › <http://138.68.106.29>



(Abschrift)

Klana Mann wollt' e grouß Fraa hou. Hä juch hä!

Grouß Fraa wollt' e klana Mann hou. Nadi nadi nadi

bums vallerä hop-sa sa sa

2. Grouß Fraa zum Tanzborrem ging, klana Mann wollt' a' mitgeh'n.
3. Klana Mann, du bleibst zu Haus, kehrst all' die Stowe aus.
4. Grouß Fraa vum Tanzborrem kam, klana Mann hinterm Ofen spann.
5. Mann, was host du gesponne? Dreimal hab ich abgesponnen.
6. Frau nahm den Rockenstock, schlug den Mann mit auf den Kopp.
7. Mann kroch ins Boddafass. Guckschte 'raus kriekgte was.
8. Klana Mann lief zum Haus hinaus, lief schnell ins Nochbers Haus.
9. Nochber, sagt nor'mol oe, mich hat mei grouß Fraa geschloe.
10. Nochber, nor goar nett gekloagt, meini hott mer's grad so gemocht.

Hähnlein, kr. Bensheim, - bl.36
HESS.ARGH IV

A 23

– Coats of arms in practice: heraldry

- › Coats of arms as means of communication?
- › Interdisciplinary semiotics and visual culture
- › images, artefacts, architecture, texts
- › ontology of coats of arms › description, documentation, retrieval, processing
- › <https://go.wwu.de/cdex>



```
<http://digitalheraldry.org/data/>
<http://digitalheraldry.org/ontology/>
<http://w3.org/1999/02/22-rdf-syntax-ns#>
type dho:CoatOfArms.
rdfs:isDefinedBy dho:Ground1.
rdfs:type dho:Plain;
dho:hasTincture dhr:Gules.
dho:hasCharge dhd:Charge1.
dhd:Charge1 rdfs:type dho:CrossCharge;
dho:hasTincture dho:Or;
dho:hasStyle dho:Pommy,
dho:Flory.
dhd:CoA1 dho:hasCharge dhd:Charge2.
dhd:Charge2 rdfs:type dho:Billety;
dho:hasTincture dho:Or;
dho:covers dhd:Charge1.
```

RDM in Sweden

- Small country – small research institutions
- Need for different collaborative approaches
- Individual universities: seldom have repositories
- Large-scale data producers: international / transnational repositories
 - › Example: The Human Atlas Project (www.proteinatlas.org)

Examples of RDM in Sweden

- National initiative: Swedish National Data Service (SND)
 - › <https://snd.gu.se>
 - › “SND 2.0” with university IT provider SUNET
- Ad hoc-solutions: DiVA
 - › <http://www.diva-portal.org>
- Stockholm University Data Repository
 - › Via figshare service: <https://su.figshare.com>
- Swedish University of Agricultural Sciences: Tilda
 - › Climate and environmental data

RDM in Germany 1

- 16 states, 425 universities/colleges, scientific organisations, 6 library consortia, several alliances and initiatives
- **Politics of science: papers and recommendations**
 - › 1998/2009/2013/2018 German Research Foundation (DFG)
 - › 2008/2018 German Initiative for Network Information (DINI)
 - › 2010/2018 Alliance of Science Organisations in Germany
 - › 2012 German Council of Science & Humanities (WR)
 - › 2014/2016 German Universities' Rectors' Conference (HRK)
- **National and regional initiatives 1**
 - › 2014 Council for Scientific Information Infrastructures (RfII)
 - » 2016 proposal: national research data infrastructure (NFDI)
 - › 2017/2018 NFDI initiatives North Rhine-Westphalia & Bavaria

RDM in Germany 2

– National and regional initiatives 2

- › RDA Germany, DARIAH-DE, CLARIN-D, DINI Working Group RD
- › nestor network of expertise in long-term storage of digital resources
- › sites Forschungsdaten.info/org, Forschungslizenzen.de

– Local activities: “role model” universities

- › Bielefeld: Open Access and RDM services
- › Göttingen: eResearch Alliance, Centre for DH (GCDH)
- › Köln: Cologne Center + Data Center for eHumanities (CCeH/DCH)
- › Tübingen: eScience Center, part of the CLARIN-D network

– Conclusions

- › Not easy to keep track of RDM in Germany
- › Scientific landscape › duplicate structures & developments?
- › WWU: reuse, integrate, cooperate, and close gaps where necessary

RDM at KTH(B)

- How can we support RDM at KTH(B)?
- Competent staff in-house
- Combine top-down with bottom-up
- Get KTH President decision to create RDM support function
- Below-the-surface development of services
- Monitoring current state of RDM

Examples of RDM at KTH(B)

- Informal working group with broad focus (KTHB Archive, IT, Research Office, Swedish National Infrastructure for Computing (SNIC))
- Attending selected meetings with researchers
- Web page with information, special mail address, and Q&A
- Improving KTHB staff knowledge
- Recruiting special competence

RDM at WWU & ULB MS 1

– The story so far

- › since 2000: IKM group (uni library, IT services, administration)
- › 2015: IKM survey about research data and RDM
- › 2017: WWU Research Data Policy & Service Point RDM
- › 2017: Center for Digital Humanities (CDH)
- › 2018: RDM and DH in WWU strategic development plan

– The current setting

- › WWU eScience-Center = competence and services center for digital methods and resources for all WWU departments
 - Service Point Research Data Management
 - Service Point Digital Humanities [to be staffed asap]
 - More Service Points to follow (e.g. digitization)

RDM at WWU & ULB MS 2

– Next steps

- › Advice on RDM and on how to “bring the WWU policy to life”
- › Repository for WWU research data
- › Interlink repositories + research information system + ORCID
- › DMP tool (based on RDMO), RDM toolbox (sciebo.RDS)
- › “eScience Cloud” with OpenStack cloud computing (IaaS)
- › Business model for extensive data curation and data storage
- › Training and workshops for students and faculty
- › Coordinating WWU projects and RDM activities › networks

– Some lessons learned

- › Many insights match those reported by other libraries
- › “Be prepared for everything and everyone”

“Same same but different”? – Conclusions 1

– Engineering vs humanities

- › Different: engineers more aware of “handling data” than (traditional) humanists
- › Different: complexity of data types and formats, “keep it safe” vs “keep it alive”
- › Same: funding requirements, questions / reservations / reasons not to publish data

– Sweden vs Germany

- › Different: small central structures may be easier than big decentralized ones
- › Different: smaller financial and staff capacities call for more cooperation, more capacities risk of duplicate structures
- › Same: RDM policies quite abstract, specifications needed

“Same same but different”? – Conclusions 2

– KTH/KTHB vs WWU/ULB MS

- › Different: libraries of smaller universities can concentrate on fewer subjects, libraries of big universities have to be prepared for everything
- › Same: keep track of everything that’s going on in RDM
- › Same: continuous training of staff and cooperation between library and faculty and other libraries important
 - » Cross discipline / country / library type exchanges fruitful for discussing RDM and other library topics

– Snow’s message relevant today

- › “The Rich and the Poor”
- › Start communicating
- › Strive for open science

Literature 1

- **Allianz der deutschen Wissenschaftsorganisationen (2018a):** *Forschungsdatenmanagement. Eine Handreichung.* <http://doi.org/10.2312/allianzoa.029>.
- **Allianz der deutschen Wissenschaftsorganisationen (2018b):** *Research Data Vision 2025 – ein Schritt näher.* <http://doi.org/10.2312/allianzoa.024>.
- **Bohle Carbonell, Katerina (2018):** “The underwear of data science”. In: *Enigmas, Networks, and People at Work*. 28.1.2018. <https://katerinabc.com/make-open-science-successful/>.
- **Curdt, Constanze et al. (2018):** *Zur Rolle der Hochschulen: Positionspapier der Landesinitiative NFDI und Expertengruppe FDM der Digitalen Hochschule NRW zum Aufbau einer Nationalen Forschungsdateninfrastruktur.* <https://doi.org/10.5281/zenodo.1217527>.
- **DHd AG Datenzentren (2018):** *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten.* <http://doi.org/10.5281/zenodo.1134760>.
- **DFG (2018):** *Stärkung des Systems wissenschaftlicher Bibliotheken in Deutschland.* http://www.dfg.de/download/pdf/foerderung/programme/lis/180522_awbi_impulspapier.pdf.
- **DINI (2018):** *Thesen zur Informations- und Kommunikationsinfrastruktur der Zukunft.* <http://dx.doi.org/10.18452/19126>.
- **Dressel, Willow F. (2017):** “Research Data Management Instruction for Digital Humanities”. In: *Journal of eScience Librarianship* 6.2:e1115. <https://doi.org/10.7191/jeslib.2017.1115>.
- **Hansson, Sven Ove (2007):** “What is Technological Science?“. In: *Studies in History and Philosophy of Science* 38:523-527. <https://doi.org/10.1016/j.shpsa.2007.06.003>.
- **HRK (2016):** How university management can guide the development of research data management. Orientation paths, options for action and scenarios. https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-10-Publikationsdatenbank/Beitr-2016-01_Research_Data_Management.pdf.
- **Goble, Carole (2018):** “Building the FAIR Research Commons: A Data Driven Society of Scientists”. Talk given at the Symposium “The Future of a Data-Driven Society” at Maastricht University on the 25th of January 2018. Video available at <https://www.maastrichtuniversity.nl/events/review-symposium-future-data-driven-society> (talk starts at 09:24), slides available at <https://www.slideshare.net/carolegoble/building-the-fair-research-commons-a-data-driven-society-of-scientists>.

Literature 2

- **Kaden, Ben (2018a)**: “Warum Forschungsdaten nicht publiziert werden”. In: *LIBREAS. Library Ideas*. 33. <http://libreas.eu/ausgabe33/kaden-daten/>.
- **Kronenwett, Simone (2017)**: *Dienstleistungsangebote für digital arbeitende GeisteswissenschaftlerInnen. Vorstellung des Cologne Center for eHumanities (CCeH) & Data Center for the Humanities (DCH)*. http://dch.phil-fak.uni-koeln.de/sites/dch/Materialien_Aktivitaeten/2017/Vortrag_Kronenwett-ZBIW-20170328.pdf.
- **Meyer-Doeringhaus, Ulrich / Tröger, Beate (2015)**: “Forschungsdatenmanagement als Herausforderung für Hochschulen und Hochschulbibliotheken”. In: *o-bib* 2.4:65-72. <http://dx.doi.org/10.5282/o-bib/2015H4S65-72>.
- **Peukert, Hagen (2017)**: “Curating Humanities Research Data: Managing Workflows for Adjusting a Repository Framework”. In: *International Journal of Digital Curation* 12.2:234-245. <https://doi.org/10.2218/ijdc.v12i2.571>.
- **Rfii (2016)**: *Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes, and financing for research data management in Germany*. <http://www.rfii.de/?p=2075>.
- **Rfii (2017)**: *Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI)*. <http://www.rfii.de/?p=2269>.
- **Rfii (2018)**: *Zusammenarbeit als Chance. Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland*. <http://www.rfii.de/?p=2529>.
- **Schirnbacher, Peter (2017)**: „Dimensionen des Forschungsdatenmanagements im digitalen Zeitalter“. In: Hauke, Petra / Kaufmann, Andrea / Petras, Vivien (eds.): *Bibliothek – Forschung für die Praxis*. De Gruyter. 389-405. <https://doi.org/10.1515/9783110522334-034>.
- **Snow, C. P. [1959]**: *The Two Cultures*. Cambridge University Press 2012. (= Canto Classics.)
- **Töwe, Matthias (2017)**: “Wie Forschungsdaten die Bibliothek verändern. Erfahrungen aus der ETH-Bibliothek”. In: *B.I.T.online* 20.5:361-370. <http://www.b-i-t-online.de/heft/2017-05/fachbeitrag-toewe.pdf>.
- **Vandegrift, Micah (2018)**: “Designing Digital Scholarship Ecologies” (preprint paper & slides). In: LIS Scholarship Archive. <http://doi.org/10.17605/OSF.IO/93ZVB>.
- **Wissenschaftsrat (2012)**: *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*. <https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>.

Viola Voß

*M.A. Linguistics and French & German Studies
MA Library and Information Science
PhD Linguistics*

ULB Münster | Germany

voss.viola@uni-muenster.de
<https://orcid.org/0000-0003-3056-407X>

Göran Hamrin

*BA Philosophy / BSc Mathematics
MSc Library and Information Science
PhD Mathematical Logic*

KTH Library | Sweden

ghamrin@kth.se
<https://orcid.org/0000-0003-4256-2960>

