

# Research Data Management in (Public) History

A field report from the historical research project  
Stadt.Geschichte.Basel

Moritz Mähr

15.06.2022

# Agenda

- ▶ Stadt.Geschichte.Basel in a nutshell
- ▶ Task and challenges of the research data manager
- ▶ How to tackle Research Data Management?
- ▶ How to do public history with research data?
- ▶ Conclusion



[doi.org/10.5281/  
zenodo.6637118](https://doi.org/10.5281/zenodo.6637118)

# Stadt.Geschichte.Basel in a nutshell

# Factsheet

- ▶ History of the Canton of Basel-City from 200 B.C. to the present day
- ▶ 10 printed volumes, including 8 epochal volumes, 1 thematic volume, 1 overview volume
- ▶ Over 70 authors, between 3 and 15 per volume
- ▶ Several teams of historical and archaeological researchers
- ▶ Over 9 million CHF budget for a project duration of over 8 years
- ▶ Approx. 1 million CHF of the budget is to be used for “digital”

# Project timeline 1/4

- ▶ In 2011, a civil society association and several parliamentary motions called for a new history for Basel-City.
- ▶ In 2016, the cantonal government decided to support the project with CHF 6 million. In its meeting on October 19, the cantonal parliament granted the project a credit of 4.4 million out of a total budget of 9.36 million with a large majority. 1.6 million were contributed from the Lottery Fund. Foundations and private individuals are to provide the remaining funds, of which more than 2.3 million have already been pledged.

## Project timeline 2/4

- ▶ In early 2017, the civil society association handed over the project to a newly established foundation. The foundation's board of directors set up an editorial board to steer the content.
- ▶ In the summer of 2017, the project management began its work, and in the fall of 2018, the researchers started their work at Hirschgässlein 21 (Department of History) and Petersgraben 51 (Department of Ancient Cultures).
- ▶ The digital strategy was developed in 2020 by the digital committee of the editorial board and the digital advisory board and adopted in January 2021.

## Project timeline 3/4

- ▶ In October 2021, I was hired as digital project manager for research data management and digital education.
- ▶ Surveys were conducted among researchers and outreach partners beginning in December 2021.
- ▶ April 2022, the concept for implementing the digital strategy was approved.

# Project timeline 4/4

- ▶ A concept for digital and analog education will be developed by September 2022.
- ▶ In October 2022, the first Data Stories based on research data will be published.
- ▶ In October 2023, educational work with outreach partners will begin.
- ▶ In October 2023, a prototype of the online portal will go online and replace the existing homepage.
- ▶ Starting in January 2024, the volumes will gradually go on sale.



# Task and challenges of the research data manager

# Mission according to the agreement with the canton

- ▶ An online portal. It is designed as a virtual repository providing public access to a wide range of data on Basel's history (...). The portal will be designed and organizationally integrated in such a way that it can be continued and continuously updated after the end of the eight-year project period.
- ▶ The continuous, public visibility and the dialogue with future users of a city history are important concerns of the project. Interested members of the public are already given the opportunity to take a look at the project or to actively participate during the ongoing work on the new city history.

# Specific task according to digital strategy

- ▶ Collection, curation and preservation of research data
- ▶ Development of showcases for the presentation and distribution of relevant data and media

# A challenging organizational structure 1/2

- ▶ The foundation carries the project and is liable to the donors
- ▶ The board of the foundation is responsible for the strategy and large budget items.
- ▶ The editorial board ensures the scientific quality, is responsible for the content and makes strategic decisions together with the project management.

## A challenging organizational structure 2/2

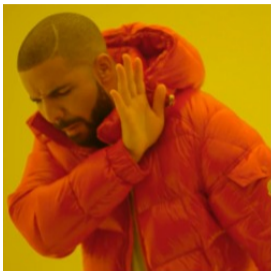
- ▶ In August 2017, the board of the foundation and the editorial board set up a two-person overall project management team.
- ▶ I joined the project management team in October 2021 and am responsible for research data management and digital education.
- ▶ I am supported by a team of data stewards and a digital advisory board.
- ▶ The entire project is administratively handled by the University of Basel.

# Many cooperation partners (and data sources)

- ▶ University Library Basel
- ▶ State Archives Basel-City (and Basel-Country)
- ▶ Historical Museum Basel
- ▶ Art Museum Basel
- ▶ Land registry and survey office
- ▶ Outreach and educational partners (Women's City Tour, etc.)

# How to tackle Research Data Management?

When I ask historians, do you have research data?



**Research  
Data**



**Digitized  
archival  
sources**

imgflip.com



# What's the problem with historians and research data?

- ▶ What is missing, however, is the awareness that we are working with digital or digitized texts, images, object descriptions, etc. **not only work with texts, images and object descriptions, but also with data.** (Hiltmann 2018 translation by MM, emphasis in original)

# What is research data in history? A definition

- ▶ While the narrower term [of research data] actually included only data generated in the concrete research process of the individual projects, such as the text annotations or metadata, the broader term also included the sources underlying the historical research process as a whole - as they are scientifically processed by libraries, archives, and museums and thus made available for potential use by researchers. (Hiltmann 2018 translation by MM)

## One possible solution: standards

- ▶ As long as the direct processing of the source is done by the historian himself, by his reading, hearing or seeing, the requirements for such standards are relatively low and concern primarily technical aspects in the background. If, however, we want to process the sources as data by machine, things look quite different. Especially if the machine is to be able to execute our queries and analyses simultaneously across as many different data sources as possible. (Hiltmann 2018 translation by MM)

## Another possible solution: software

- ▶ FuD is a software system that supports the entire scientific work process in its various research phases and enables collaboration independent of location and time. It is designed for the qualitatively researching humanities and social sciences and consists of three subcomponents: the working environment for data acquisition, processing, analysis and evaluation; the publication environment for the online presentation of digital holdings; the data repository ViDa5 for the creation of data collections and their enrichment with archiving metadata for at least ten years of storage or their long-term provision. (Lemaire 2018 translation by MM)

# Great ideas, but



- ▶ It is hard to enforce standards when the research is already in progress (and paid for). The fact that Stadt.Geschichte.Basel is a large project with several teams, different disciplines and cultures makes it even more difficult.
- ▶ It is expensive to develop digital research environments and time-consuming to train researchers. There is not enough time and budget for this.

# What is research data in Stadt.Geschichte.Basel?

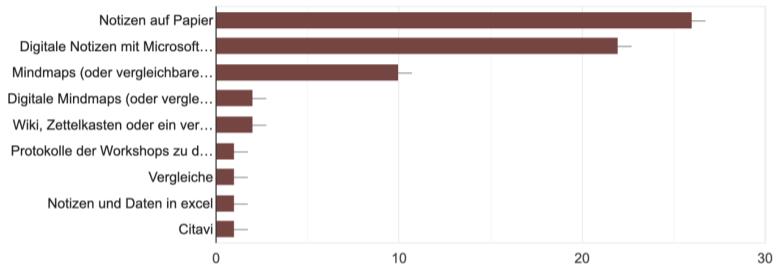
# Understand how research and your researchers work

- ▶ Research data is generated wherever research is carried out. Therefore, it is central to understand researchers and their practices.
- ▶ The question of what exactly is data needs to be addressed for each research project (and sometimes for each group within a project).
- ▶ Therefore, we asked researchers about what data they collect and what tools they use to do so.
- ▶ The survey was very well received by the authors. The high response rate of 40%, 30 responses out of 72 authors, as well as the high interest in the results of the survey (87%) confirm this.

# What tools do you use to develop a research question?

Welche Werkzeuge setzen Sie ein, um eine Forschungsfrage zu entwickeln?

30 Antworten



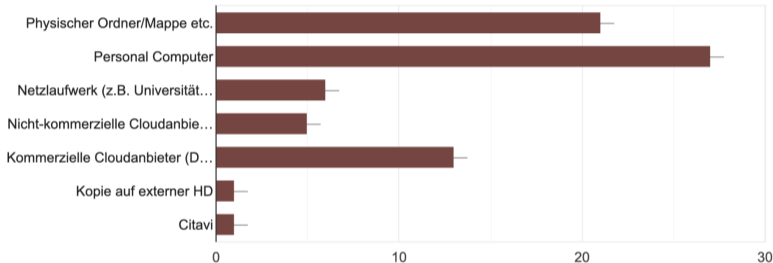
- ▶ Research questions are written on paper (87%) or with word processing programs (73%). 40% use mindmaps or comparable clustering methods, mainly on paper.



# How do you store this data?

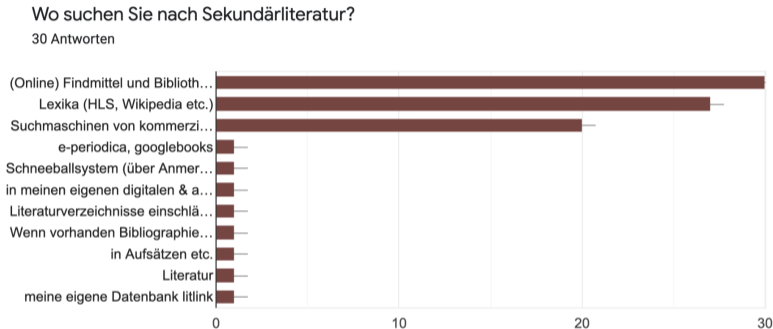
Wie halten Sie diese Daten fest?

30 Antworten



- ▶ The main working tools for data storage are personal computers (90%), physical folders (70%), and commercial cloud providers (43%).

# Where do you search for secondary literature?

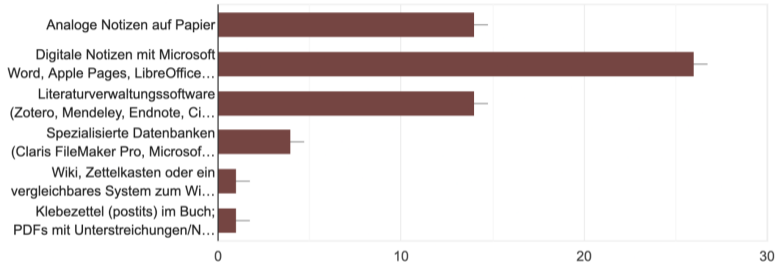


- ▶ Secondary literature is searched using (online) finding aids from archives and libraries (100%), in encyclopedias (90%), and on search engines from commercial providers (67%).

# How do you keep track of secondary literature and excerpts?

Wie halten Sie die Sekundärliteratur und Exzerpte fest?

30 Antworten

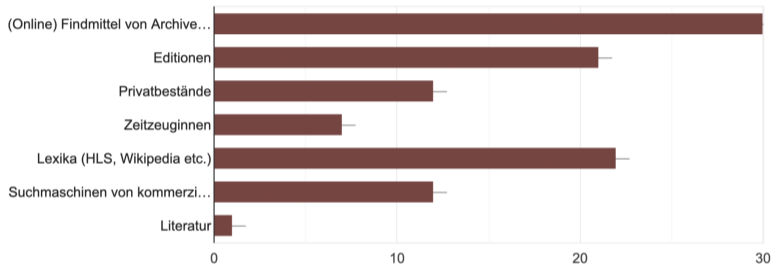


- ▶ Secondary literature and excerpts are recorded using word processing programs (87%), literature management software (47%) and paper (47%).

# Where do you search for sources?

Wo suchen Sie nach Quellen?

30 Antworten

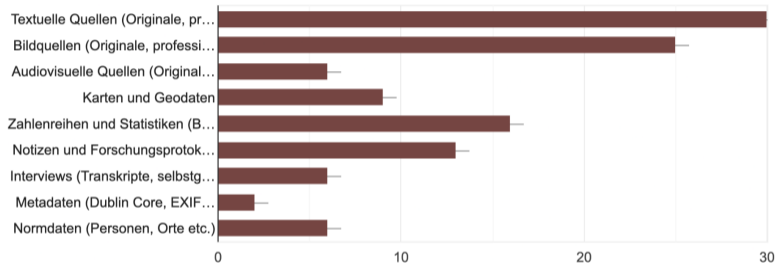


- Sources are searched by means of (online) finding aids of archives (100%), encyclopedias (73%), editions (70%), private collections (40%), commercial search engines (40%) and Zeitzeug:innen (23%).

# What sources (or data) do you collect or use?

Welche Quellen (bzw. Daten) erheben oder verwenden Sie?

30 Antworten

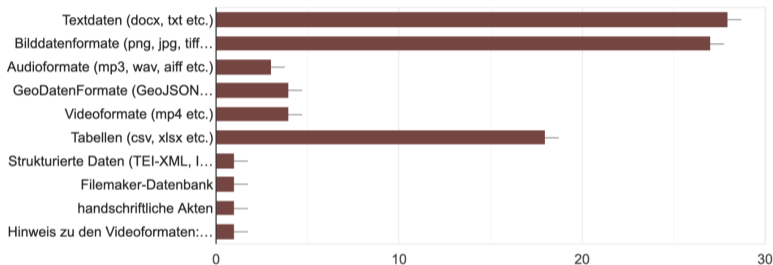


- ▶ Text sources (100%), image sources (83%), time series and statistics (53%), and notes and research protocols (43%).

# What are the formats of these data?

Welche Formate haben diese Daten?

30 Antworten

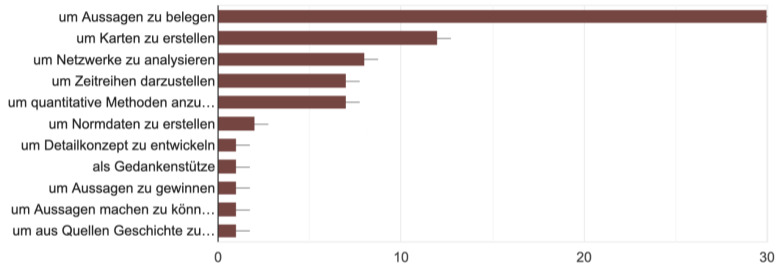


- ▶ These are text data (93%), image data (90%) and tables (60%). Video, audio, and geospatial data are used by less than 15%.

# For what purpose do you collect this data?

Zu welchem Zweck erheben Sie diese Daten?

30 Antworten

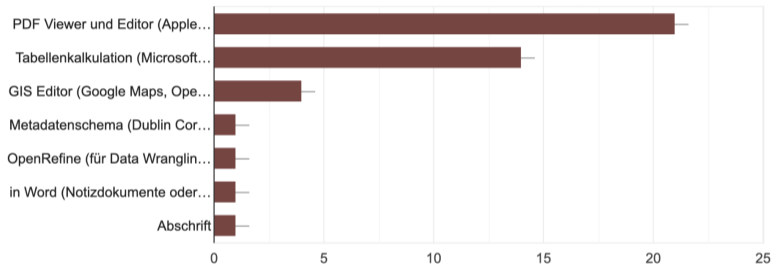


- ▶ Data is used to substantiate statements (100%), create maps (40%), analyze networks (27%), and create time series and quantitative analysis (23%).

# Which tools do you use to annotate/process these sources?

Mit welchen Werkzeugen annotieren Sie diese Quellen bzw. wie extrahieren Sie Daten?

24 Antworten



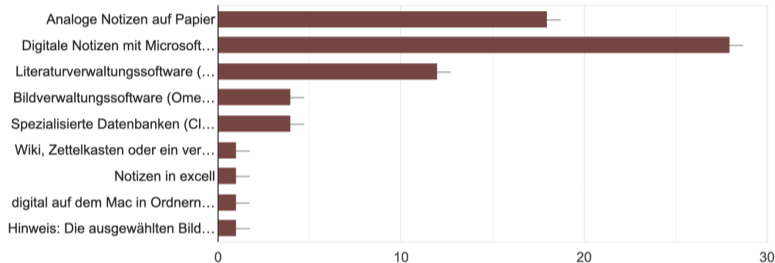
- ▶ PDF software (88%), spreadsheet programs (58%), and GIS programs (17%) are used to annotate sources.



# How do you keep track of the sources?

Wie halten Sie die Quellen fest?

30 Antworten

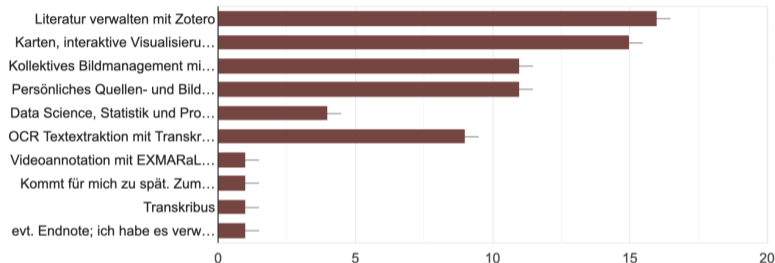


- Sources are recorded with word processing programs (93%), on paper (60%), and with literature management software (40%).

# Would you attend a training session?

Würden Sie an einer Schulung teilnehmen für:

26 Antworten



- There is a high level of interest in training for managing literature with Zotero (61%), creating maps (58%), collaborative and personal image and source management (42%), and text recognition with Transkribus (35%)

# How to solve the problem?

- ▶ Data stewards for the four key data types have been hired:
  - ▶ Images and Figures,
  - ▶ Maps and Georeferences,
  - ▶ Tables and Statistics, and
  - ▶ Bibliographic References.
- ▶ Project-wide solutions have been defined for common data types, and specific solutions are provided on request for little-used data types.
- ▶ The data is collected together with the manuscripts for the printed volume. This minimizes the workload for authors and centralizes data quality control.
- ▶ Free training in the use of digital work tools and research practices according to user needs is provided.

# How to do public history with research data?

# How do research data and public history go together?

- ▶ The simple answer is: they don't fit well together.
- ▶ Research data must (actually always) be produced for a very specific audience.
- ▶ Research data is part of a data ecosystem. Redundancy must be avoided. Existing data pools should be tapped or populated.
- ▶ The public cannot be educated directly with research data. They need intermediaries.
- ▶ That's why we work with professional outreach and education partners.
- ▶ To find out how our partners inform themselves about the history of Basel-City, we conducted a survey.

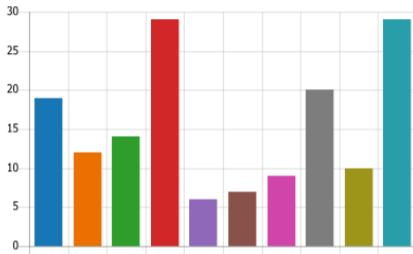
# Know your audience

- ▶ We surveyed selected people from the field of galleries, libraries, archives and museums (GLAM), as well as operators and schools with whom we would like to collaborate.
- ▶ Over 90 people took part in the survey.

# In which field do you work?

## 1. Tätigkeitsbereich

[More Details](#)



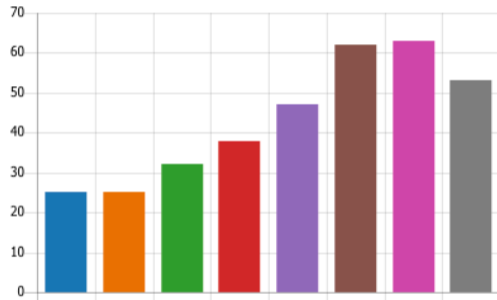
- ▶ The participants indicate research (33%), other (33%), outreach (22%), archives (21%) and education (16%) as their fields of activity. Their average age is 57 years. Two-thirds would like to be informed about the results of the survey.

# What is your interest in the history of Basel?

## 6. Welche Zeiträume der Basler Geschichte sind für Sie von besonderem Interesse?

[More Details](#)

● 200 v. Chr. bis 800 n. Chr.	25
● 800-1270	25
● 1250-1530	32
● 1510-1805	38
● 1760-1860	47
● 1856-1914	62
● 1912-1966	63
● 1960-heute	53



- The interest in the history of the last 200 years is close to 70% and decreases towards antiquity to 28%. Consequently, the highest interest is in volumes 6 and 7.



# What is your interest in personalities, institutions, places and events?

- ▶ Personalities: Christoph Merian, Erasmus of Rotterdam, Fritz Hauser, Hans-Peter Tschudi, Isaak Iselin, Jacob Burckhardt, Johann Rudolf Wettstein, Munatius Plancus, Peter Ochs, Willhelm Klein.
- ▶ Institutions: Basel Mission, FC Basel, GGG, Geigy AG, industry, art museum, museums, Roche, hospitals, University of Basel, guilds.
- ▶ Locations: Old town, train stations, Birsig, Claraplatz, Kleinbasel, Landhof, Lange Erlen, Marktplatz, Münsterplatz, Petersplatz, Rhine, St. Alban, St. Jakob, St. Johannquartier
- ▶ Events: Construction of the Rhine bridges, earthquake, separation of cantons, industrialization, Middle Ages, Reformation, world war

# Which topics in Basel's history are of particular interest to you?

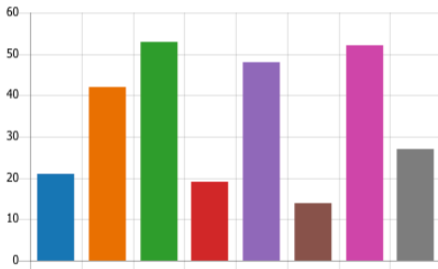
- ▶ Among the topics, urban life (58%), culture (52%), social affairs (49%), education (48%), politics (48%), poverty (45%), migration (43%) and colonialism (42%) stand out.

# Which online information services do you use to learn about the history of Basel?

12. Welche online Informationsangebote nutzen Sie, um sich über die Geschichte von Basel zu informieren?

[More Details](#)

historische Karten der UB http...	21
<a href="https://de.wikipedia.org/wiki/">https://de.wikipedia.org/wiki/...</a>	42
<a href="https://hls-dhs-dss.ch/de/artic...">https://hls-dhs-dss.ch/de/artic...</a>	53
<a href="https://merian.bs.ch/">https://merian.bs.ch/</a>	19
<a href="https://ub.unibas.ch/">https://ub.unibas.ch/</a> bzw. htt...	48
<a href="https://www.basel.com/">https://www.basel.com/</a>	14
<a href="https://www.baslerstadtbuch.c...">https://www.baslerstadtbuch.c...</a>	52
<a href="https://www.bs.ch/">https://www.bs.ch/</a>	27

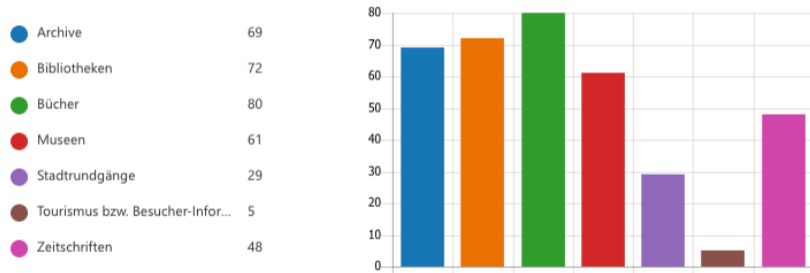


► The most important online information offers are the HLS (60%), the Basel City Book (58%), the UB (54%) and Wikipedia (47%).

# Which offline information services do you use to learn about the history of Basel?

14. Welche offline Informationsangebote nutzen Sie, um sich über die Geschichte von Basel zu informieren?

[More Details](#)



- Offline, people search for information on Basel's history in books (91%), libraries (82%), archives (79%), museums (70%) and magazines (55%).

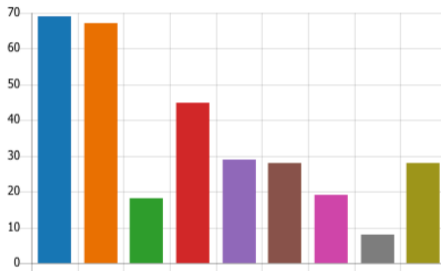
How do you process and/or convey historical content?  
Do you use sources or data for preparation, implementation  
and so on?

# What types of sources do you have in use?

## 15. Welche Quellentypen haben Sie im Einsatz?

[More Details](#)

- Textuelle Quellen (Originale, p... 69
- Bildquellen (Originale, professi... 67
- Audiovisuelle Quellen (Origina... 18
- Karten und Geodaten 45
- Zahlenreihen und Statistiken (... 29
- Notizen und Forschungsproto... 28
- Interviews (Transkripte, selbst... 19
- Metadaten (Dublin Core, EXIF ... 8
- Normdaten (Personen, Orte et... 28

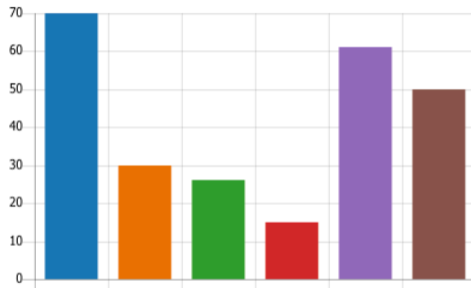


► Text (79%), images (76%) and maps (52%) are predominant.

# Where do you look for sources (or data)?

16. Wo suchen Sie nach Quellen (bzw. Daten)?

[More Details](#)



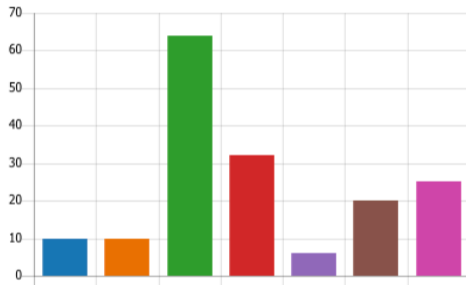
- Sources are searched using the finding aids of archives and libraries (79%), in encyclopedias (70%) and with commercial search engines (57%).

# What forms of presentation do you use?

## 17. Welche Darstellungsformen setzen Sie ein?

[More Details](#)

Audio	10
Interaktive Visualisierungen (O...	10
Texte (Print)	64
Texte, Blogbeiträge (Online)	32
Video	6
Visualisierungen (Online)	20
Visualisierungen (Print)	25



- ▶ Texts (offline 72%, online 37%) and visualizations (offline 28%, online 25%) dominate. Audio, interactive visualizations and video are used by less than 11% of respondents.



# How to solve the problem?

- ▶ There is a need for target group-oriented digital and analog services and products for the mediation of research data
- ▶ Collaborate with existing providers, integrate existing services and data
- ▶ Identify and address gaps (data and services) within the ecosystem
- ▶ Small projects, rapid prototyping, and ongoing monitoring of success (e.g., with surveys)
- ▶ Low technical and legal barriers to entry increase the chance of reuse of data

# Conclusion

- ▶ Research data management is not an end in itself; strategic and operational priorities must be set.
- ▶ Research data management is costly and demands a lot from all parties involved. Therefore, it should be simple, practical, and accessible.
- ▶ Research practices are subject to change. Research data management can influence them.
- ▶ Research data is produced for a specific audience. Researchers are not the only audience.

# Questions?



- ▶ Questions (and/or criticism) on my take on Research Data Management?

# Bibliography

- Hiltmann, Torsten. 2018. "Forschungsdaten in der (digitalen) Geschichtswissenschaft. Warum sie wichtig sind und wir gemeinsame Standards brauchen." Billet. *Digitale Geschichtswissenschaft* (blog). September 17, 2018. <https://digigw.hypotheses.org/2622>.
- Lemaire, Marina. 2018. "Vereinbarkeit von Forschungsprozess und Datenmanagement in den Geisteswissenschaften." *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, December, 237–247 Seiten. <https://doi.org/10.5282/O-BIB/2018H4S237-247>.