LIBER Webinar:
Finding & Reusing Research Data

WEBINAR HOSTS

Dr Birgit Schmidt
Göttingen State & University Library

Co-Chair of LIBER's *Research Data Management Working Group*; member of the *Digital Skills for Library Staff & Researchers Working Group*.

bschmidt@sub.uni-goettingen.de

Rob Grim
Economics (Data) Librarian University Library, Erasmus University

Co-Chair of LIBER's *Research Data Management Working Group*.

rob.grim@eur.nl

SPEAKER

Kathleen Gregory

Researcher and PhD candidate at the Data Archiving and Networked Services (DANS), Netherlands

kathleen.gregory@dans.knaw.nl

NOTES

- **The webinar is being recorded.** A recording will be shared with all attendees, and posted on YouTube.
- **Slides can be downloaded on Zenodo.** Check the chat box for the link.
- **Questions?** Put them in the chat box. We'll put them to the speakers, following the presentations.

Finding and Reusing Research Data

Kathleen Gregory

M.A., MSLIS, PhD candidate

LIBER RDM Webinar Series

April 11, 2019

Motivation

If data are out there, will anyone use them?

Discovery is a pre-condition to reuse.

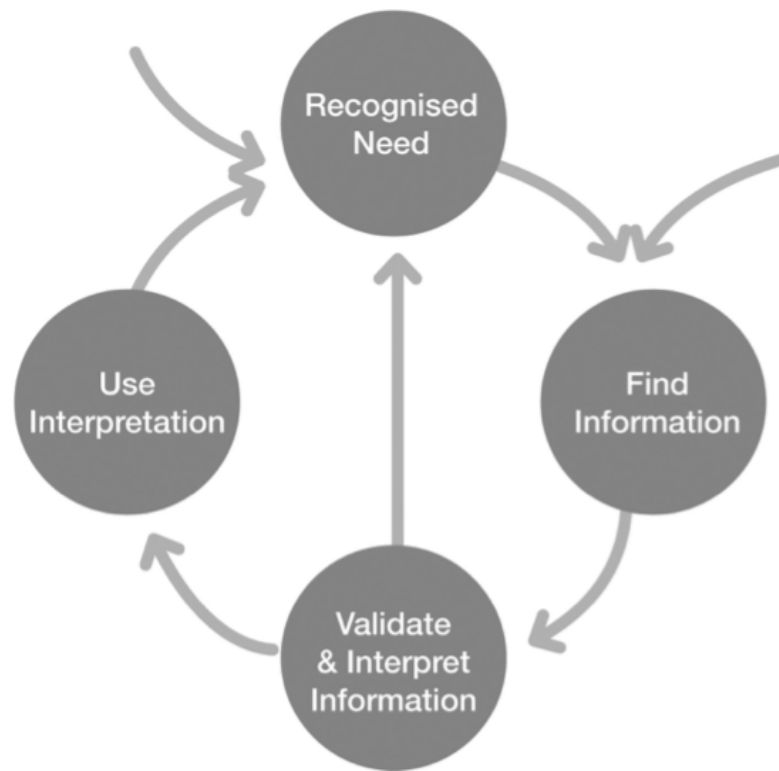How do people discover, make sense of, and
 (re)use research data?

Where are we going?

- Overview of recent research

- Resources and tips for data discovery

- What does this mean for you?

Framework for research
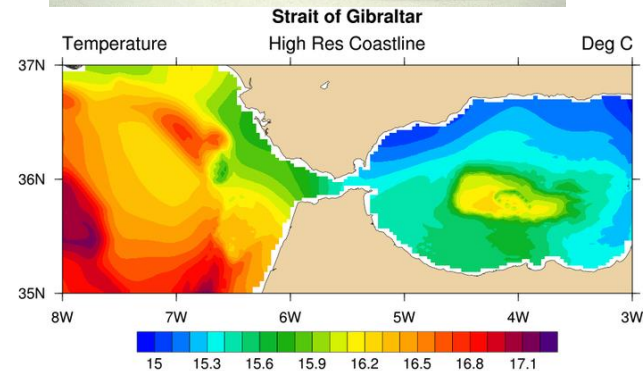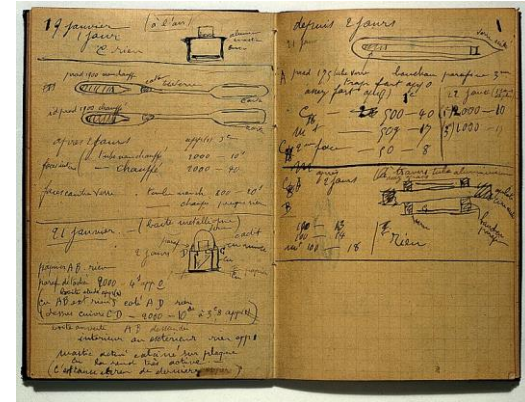
Models of interactive information retrieval

Blandford, A., & Attfield, S. (2010). *Interacting with Information. Synthesis Lectures on Human-Centered Informatics.* San Rafael, CA: Morgan & Claypool Publishers.

Data are complex objects

Data are diverse.

Data do not stand alone.

Data are not always stable and do not travel easily.

Strait of Gibraltar

Borgman, C.L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press.

Leonelli, S., Rappert, B., & Davies, G. (2017). Data shadows: Knowledge, openness, and absence. *Science, Technology, & Human Values, 42(2)*, p.191-202.

Methodology

- Semi-structured interviews with data seekers across disciplines (n=22)

- Multidisciplinary survey (n=1677, still in analysis phase)

Users and data needs - interviews

Who are the "users"?

- Heterogeneity within disciplines

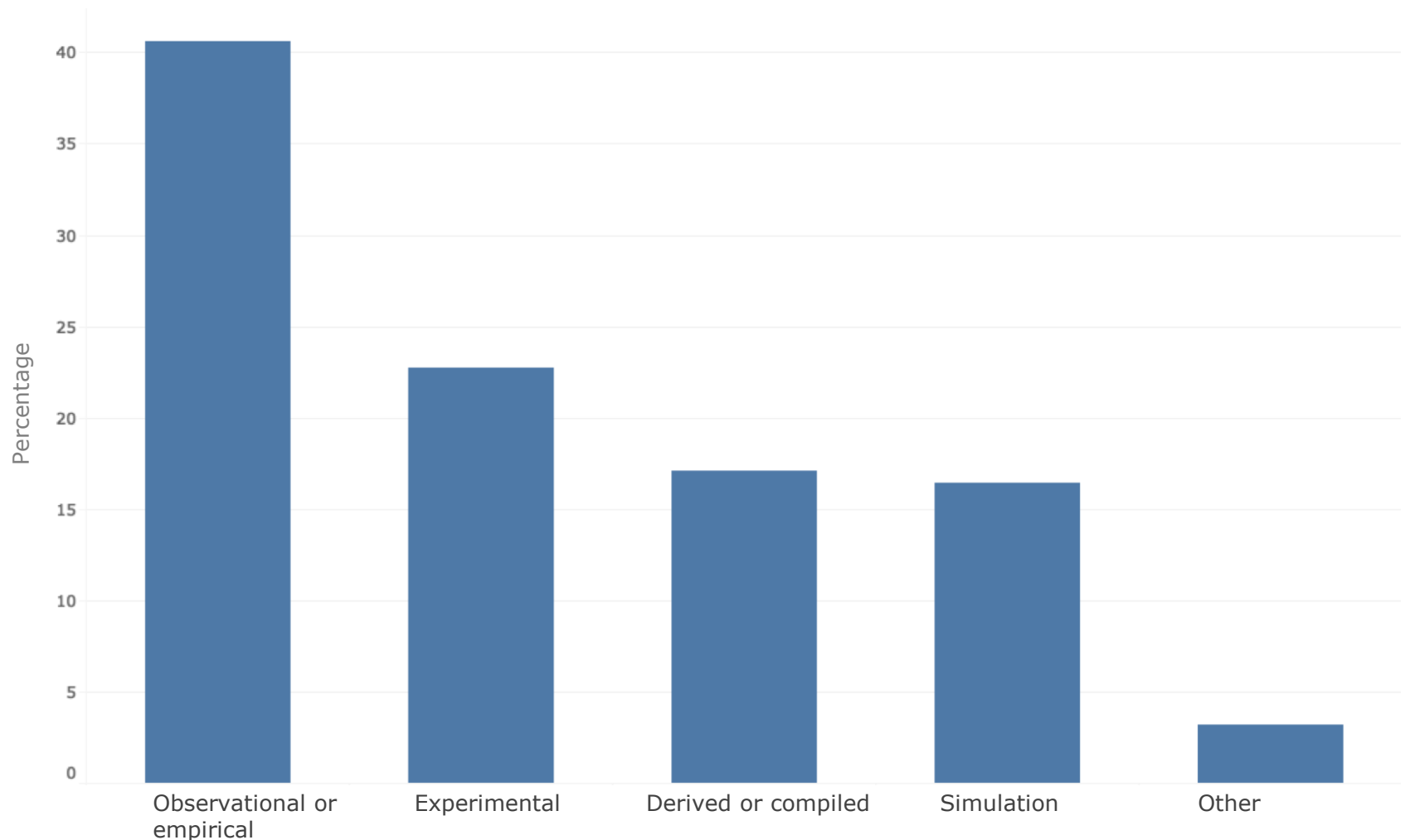- The person seeking data is not always the person using the data

Users and data needs – interviews

What data are needed?

- Data needed for research are not always research data

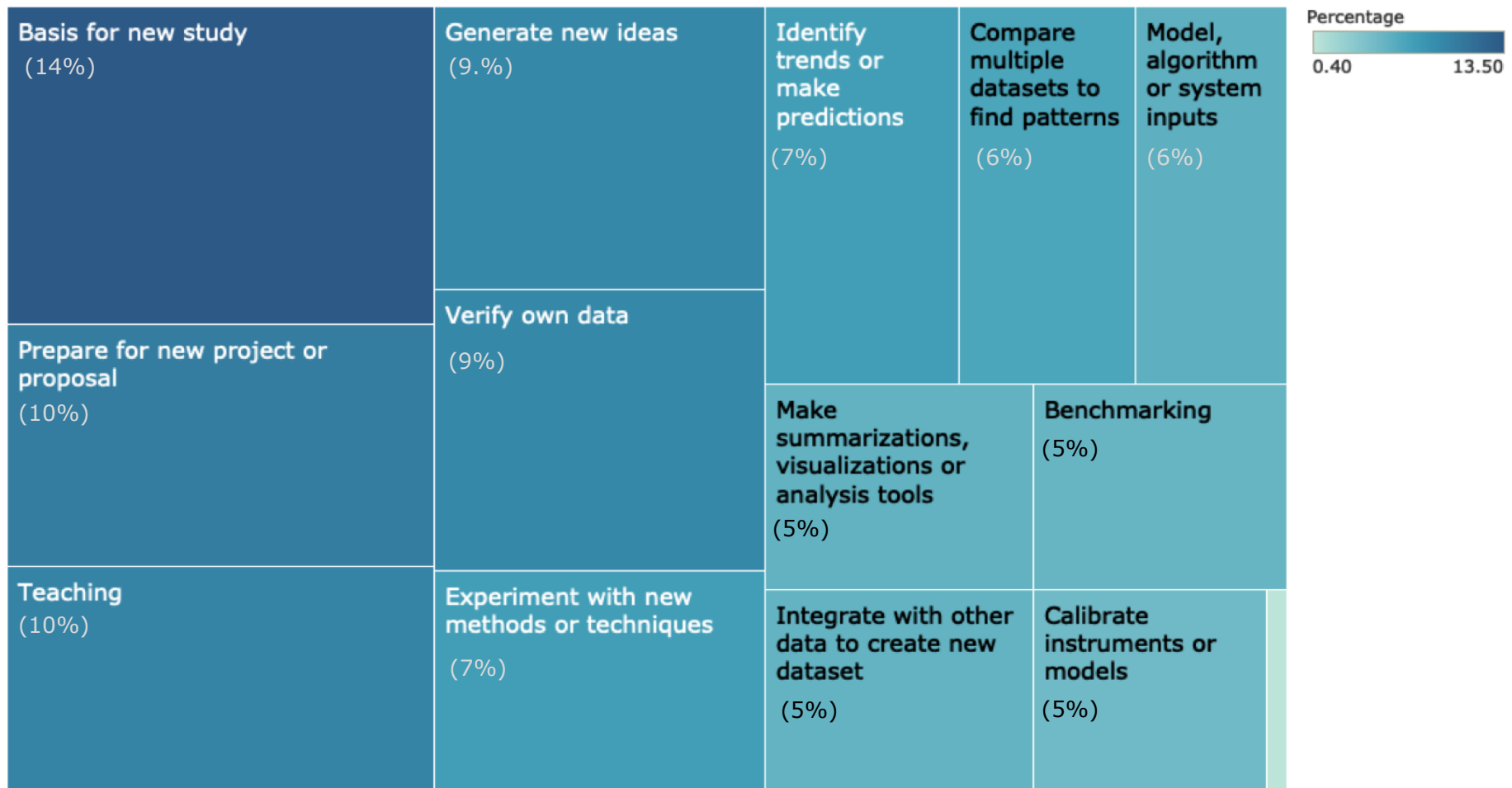- Numerous roles - data as hubs for collaboration and creativity

Users and data needs – survey

Which best describes the type of secondary data that you need?

Users and data needs – survey

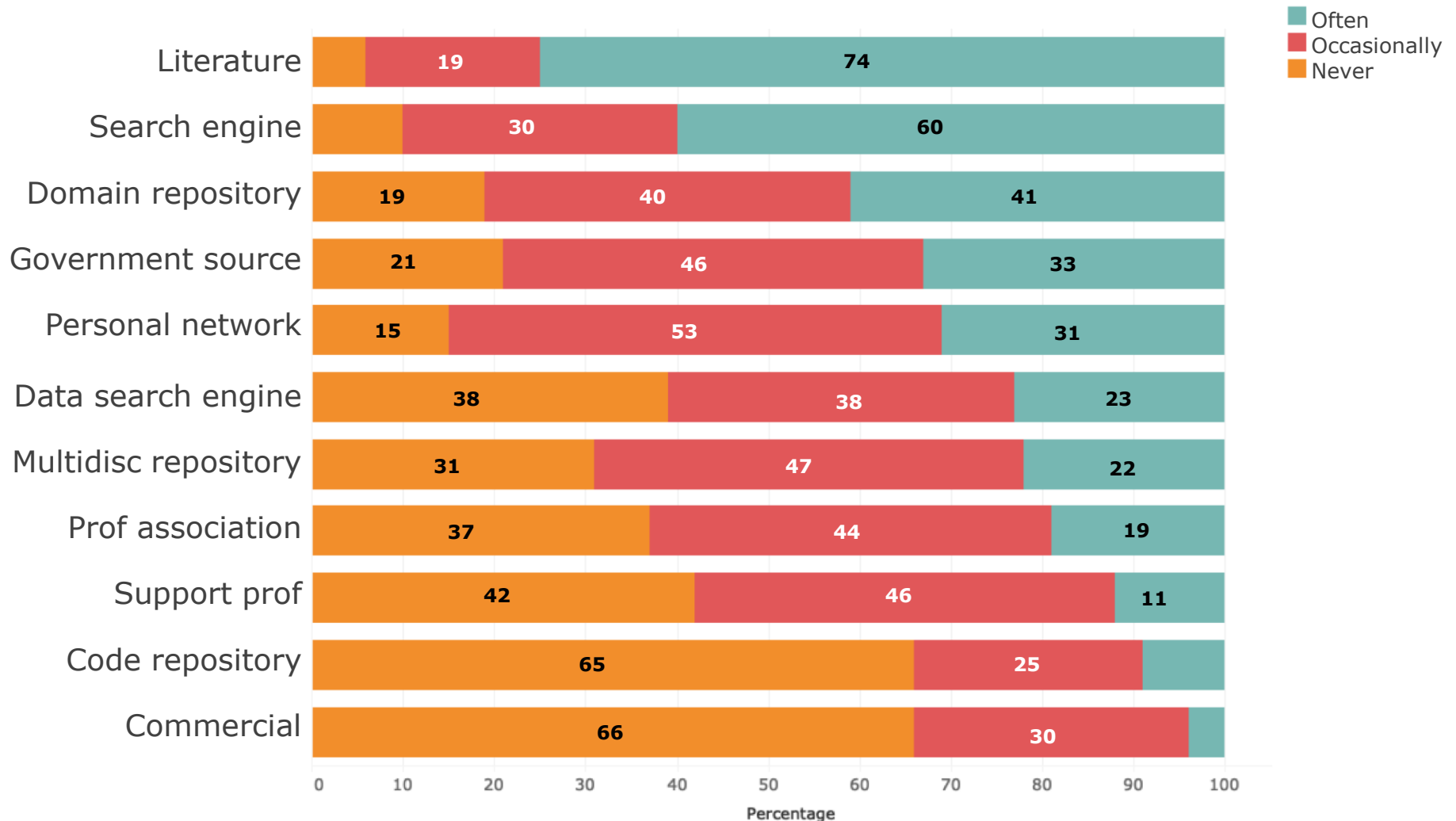## Why do you use or need secondary data?
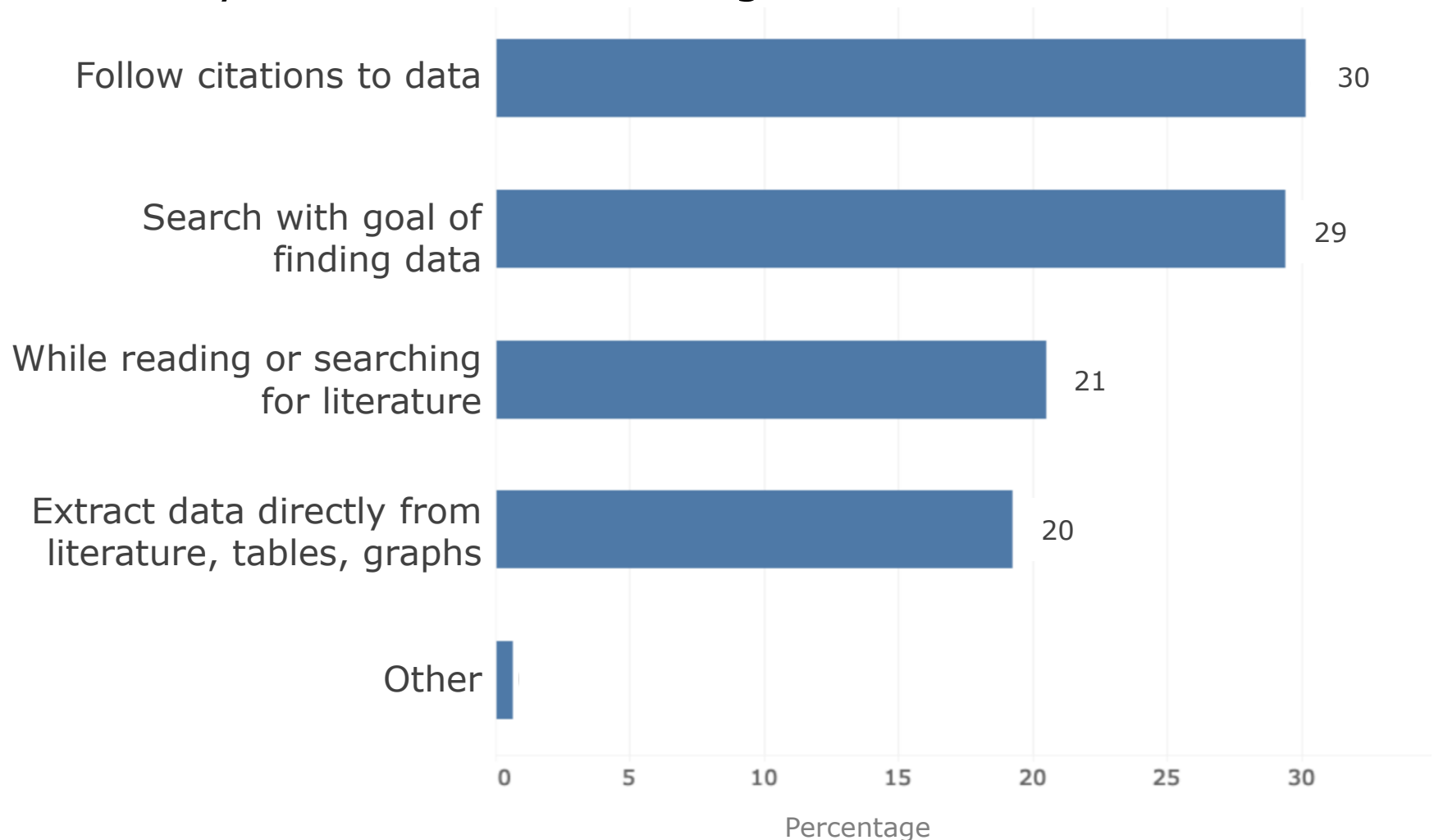
Search and discovery strategies

Multiplicity of strategies

Search and discovery strategies – survey

What sources do researchers use to find data?

Legend: Often, Occasionally, Never

Source	Never	Occasionally	Often
Literature		19	74
Search engine		30	60
Domain repository	19	40	41
Government source	21	46	33
Personal network	15	53	31
Data search engine	38	38	23
Multidisc repository	31	47	22
Prof association	37	44	19
Support prof	42	46	11
Code repository	65	25	
Commercial	66	30	

Percentage

Search and discovery strategies – survey

How do you discover data using the academic literature?

Search and discovery strategies - interviews

Key role of social interactions

Actually, most of the times that I have looked for external data, it has been through (personal) connections (11).

The human network of contacts is still the best way to find the information you want, especially if it is a small group...that is the most powerful and accurate source of information that I use at this point. (17)

Evaluation and sense-making - interviews

Role of social interactions continues

I think if there was a good search engine, then I could get the dataset directly. I would still get in touch with the data author anyway, both for social reasons - developing the network and eventual collaboration - and also **because most of the times the metadata are not enough to really understand the biology behind the species** *(4).*
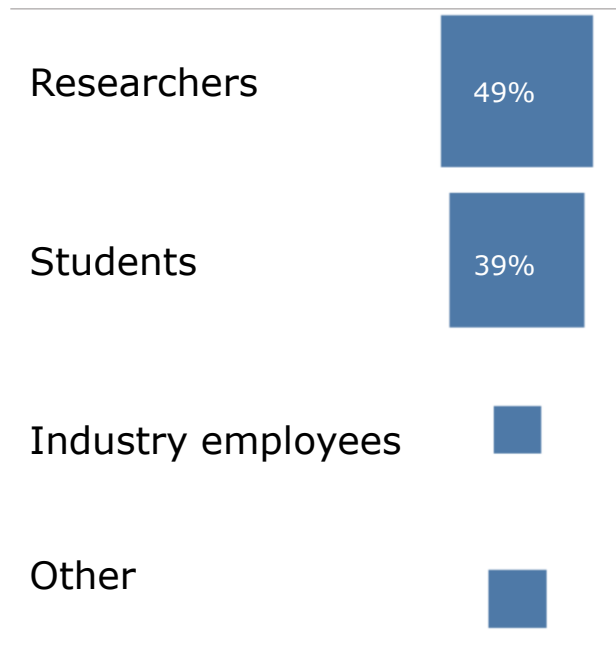
Evaluation and sense-making - survey
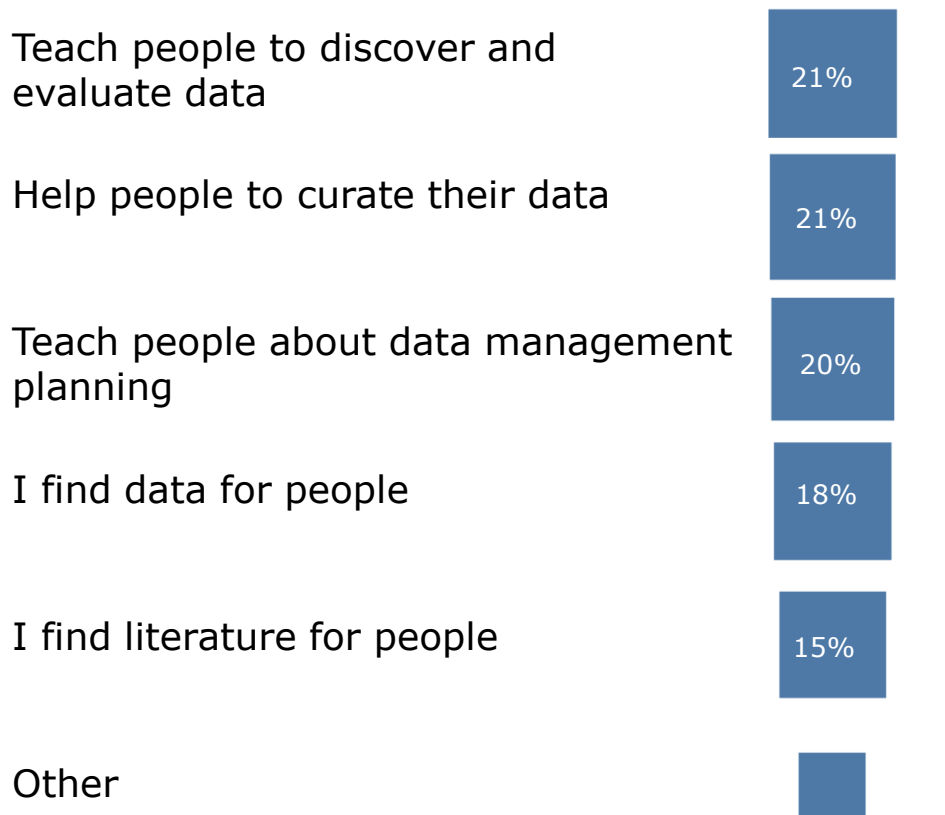
Other sense-making strategies
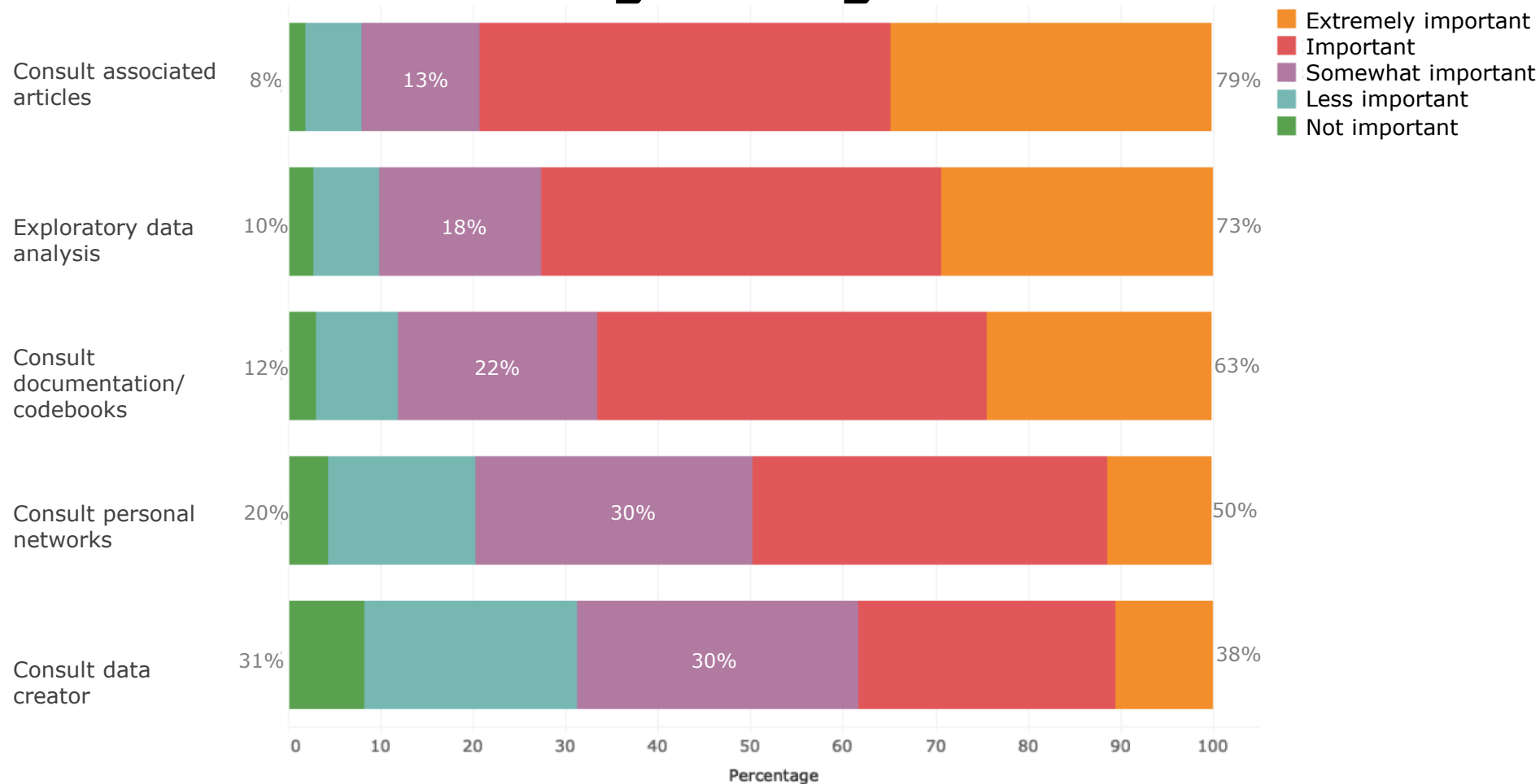
What do librarians/data professionals do?

Who do you support?

Researchers	49%
Students	39%
Industry employees	
Other	

How do you support people with their data needs?

Teach people to discover and evaluate data	21%
Help people to curate their data	21%
Teach people about data management planning	20%
I find data for people	18%
I find literature for people	15%
Other	

Discovery tips

Eleven Quick Tips for Finding Research Data

Tip 1: Think about the data you need and why you need them.

Tip 2: Select the most appropriate resource.

Tip 3: Construct your query strategically.

Tip 4: Make the repository work for you.

Tip 5: Refine your search.

Tip 6: Assess data relevance and fitness-for-use.

Tip 7: Save your search and data source details.

Tip 8: Look for data services, not just data.

Tip 9: Monitor the latest data.

Tip 10: Treat sensitive data responsibly.

Tip 11: Give back (cite and share data).

Discovery tips

CESSDA Data Management Expert Guide

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide

🏠 **EVENT CALENDAR   TRAINING RESOURCES   ABOUT**

Training / **Training Resources** / Data Management Expert Guide

Data Management Expert Guide

This guide is designed by European experts to help social science researchers make their research data Findable, Accessible, Interoperable and Reusable (FAIR).

You will be guided by different European experts who are - on a daily basis - busy ensuring long-term access to valuable social science datasets, available for discovery and reuse at one of the 18 CESSDA social science data archives.

CESSDA Data Management Expert Guide: Data Discovery

Data Management Expert Guide ∨

Case study: Data for a replication study

Kristyna Bašná works at the Institute of Sociology of the Czech Academy of Sciences (n.d.). She needed data for a replication study. How did she discover, access and use such data?

⊕ *What kind of data were you looking for?*

⊕ *How did you locate suitable data resources?*

⊕ *What challenges did you encounter before you could use the data?*

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide

To remember…

Data search and reuse are shaped by community norms, infrastructures and data supplies.

Understanding data needs and practices deeply needs to take these influences into account.

Questions – for you (and for me)

- What do you do now to support researchers with data discovery and reuse?

- What other possibilities are there?

Kathleen Gregory

kathleen.gregory@dans.knaw.nl

References

- Blandford, A., & Attfield, S. (2010). *Interacting with Information. Synthesis Lectures on Human-Centered Informatics*. San Rafael, CA: Morgan & Claypool Publishers.

- Borgman, C.L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press.

- Leonelli, S., Rappert, B., & Davies, G. (2017). Data shadows: Knowledge, openness, and absence. *Science, Technology, & Human Values, 42(2)*, p.191-202.

- Gregory K, Khalsa, Michener WK, Psomopoulos FE, de Waard A, Wu M (2018) Eleven quick tips for finding research data. PLoSJS Comput Biol 14(4): e1006038. https://doi.org/10.1371/journal.pcbi.1006038

- Gregory, K.; Cousijn, H.; Groth, P.; Scharnhorst, A.; Wyatt, S. (2019). Understanding Data Search as a Socio-technical Practice. *Journal of Information Science.* https://doi.org/10.1177/0165551519837182

- Gregory, K.; Cousijn, H.; Groth, P.; Scharnhorst, A.; Wyatt, S. (2019). Searching data: A review of observational data retrieval practices in selected disciplines. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24165

Image credits (in order of appearance)

- Wellcome Images. [Marie Curie: Holograph Notebook](#). Under [CC BY 4.0](#). (on Slide 5: Data are complex objects).

- National Center for Atmospheric Research. [Example of Coastlines Graphics: Straits of Gibraltar](#). (on Slide 5: Data are complex objects.)

- All other images available from [pixabay.com](#) under the [Pixabay License](#) and are free for commercial use, no attribution required.

Funding

This work is part of the project *Re-SEARCH: Contextual Search for Research Data* and was funded by the NWO Grant 652.001.002.

THANKS!

Questions?

Please put them in the chat box.