Daten gemeinsam nutzen
Anreize und Hindernisse beim Data Sharing
Ergebnisse aus zwei internationalen Studien

Prof. Dr. Ulrich Mansmann (DIFUTURE),
ulrich.mansmann@lmu.de

Knowledge Exchange

MEDIZIN INFORMATIK INITIATIVE

DFG Jisc CSC DEff SURF

Ergebnisse stammen aus 22 Interviews von Forschern aus 5 Projekten:

Chemie – Großbritannien
Biologie – Deutschland
Bioinformatik – Niederlande
Sozialwissenschaften – Finnland
Kulturgeschichte – Dänemark

Modes of *Data Sharing*

- ***Private management***: sharing data with colleagues within a research group
- ***Collaborative sharing***: using data within a consortium
- ***Peer exchange***: sharing data with trusted peers in informal networks
- ***Transparent governance***: sharing data with external parties such as funders and institutions for accountability, research assessment, scrutiny or inspection
- ***Community sharing***: with members of a research community
- ***Public sharing***: making data available to any member of the public

Six degrees of openness: Whyte, A. and Pryor, G (2011), Open Science in Practice: Researcher Perspectives and Participation. International Journal of Digital Curation 1(6): 199-213. http://ijdc.net/index.php/ijdc/article/viewFile/173/241

Motivations for *Data Sharing*

- Direct benefits
  - for the research itself (more robust)
  - for the career of the researcher (recognition)
  - for discipline (get wiser)
  - for science (better science)
- Norms of the project, research group, and/or discipline
- External drivers: policies and expectations from research funders and publishers

Data sharing training embedded in research methods training is crucial for data sharing to become standard research practice.

Future incentives for data sharing

An individual researcher sharing data may be at a disadvantage, in spending time doing things that are not rewarded, or taking the risk of being outcompeted when sharing ideas freely.
The field or discipline as a whole could benefit if everyone would share data freely.

- Solving this contradiction needs input from scientific societies, publisher, and funders. They have to promote data sharing;
- Training of students in essential data sharing and management practices is seen as an important factor in increasing the incidence of data sharing;
- Data infrastructure and standards are needed at various levels;
- The sharing of failed experiments is mentioned as being of paramount importance in different research fields.

Influence of policies and support services on data sharing

Researchers who recognize the strong influence norms can have, advocate the need to create research environments where it is routine practice to share data.

Recommendations

- Research funders
- Scientific Societies
- Research Institutions
- Publishers
- Data Centers and Repositories
- Knowledge Exchange

Summary (1-1)

- Rather than finding disciplinary or group patterns in data sharing practices, the interviews show how individual researchers across the case studies have similar data sharing practices and recognize similar incentives and motivations;
- In all but a very few individuals, data sharing occurs in response to extrinsic motivation. It is not something done for its own sake, but to achieve other ends.
- Researchers are motivated by three main types of incentives to share research data: direct benefits; strong influence of sharing norms; and external drivers (funder, publisher, support services);

Summary (1-2)

- Researchers' ***competence*** and ***autonomy*** are clearly promoted when data sharing enhances the research itself and the researchers' career.
- ***Relatedness***, the sense of being a respected and valued member of a community, is supported when that community has data sharing as one of its norms.

Ryan, R. and Deci, E. (2000), Intrinsic and extrinsic motivations: classic definitions and new directions. Contemporary Educational Psychology 25:54-67. doi:10.1006/ceps.1999.1020

MEDIZIN INFORMATIK INITIATIVE

White Paper: PRACTICAL CHALLENGES FOR RESEARCHERS IN DATA SHARING

Population of Survey: N=7656

This data was collected between April and June 2017 by contacting registrants to nature.com, biomedcentral.com and springer.com. It went to ~249,000 recipients of which ~15,000 clicked through to the survey, resulting in 7,719 respondents from 126 different countries.

Figure 18: Number of respondents by region of the world

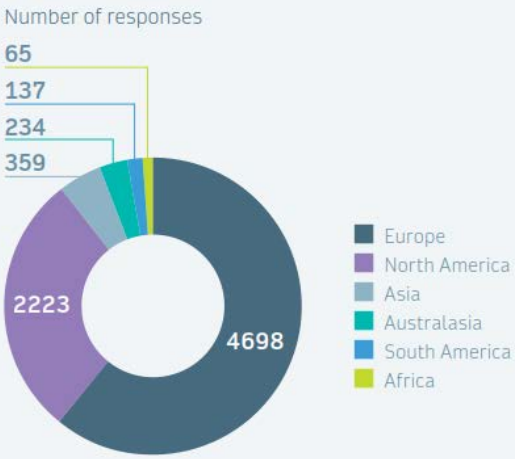
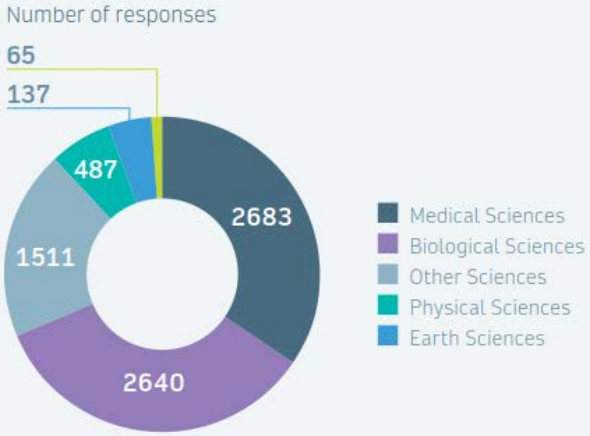


Figure 19: Number of respondents by subject



MEDIZIN INFORMATIK INITIATIVE

Figure 1: Q - How important is it to you that your data are discoverable? (1 is the least important) (n=7,656)

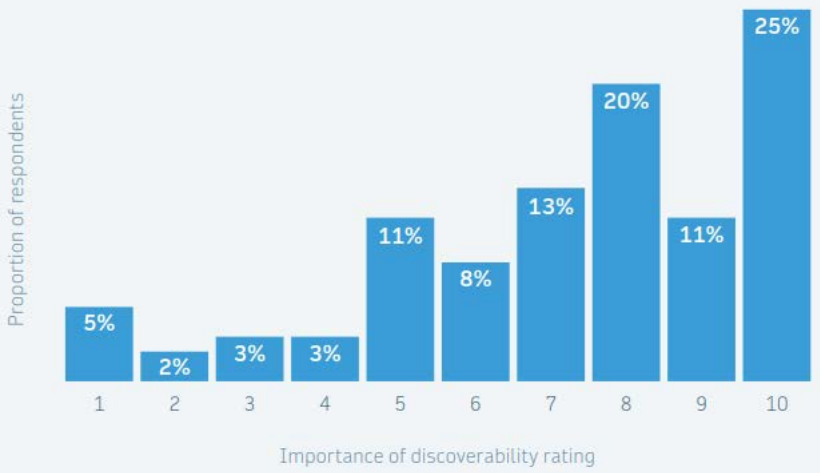
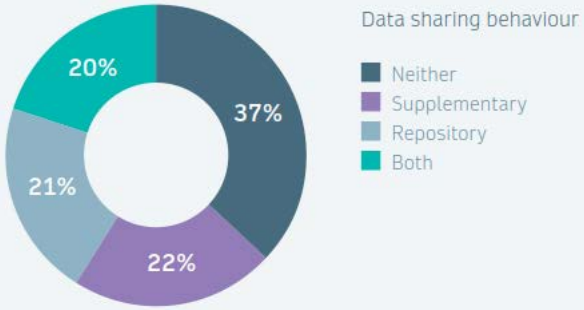


Figure 2: Q - Generally, when submitting a manuscript to a journal what do you do with the data files generated by your research? (n=7,697)



MEDIZIN INFORMATIK INITIATIVE

igure 3: Q - What problems do you have in sharing datasets? (separated by seniority) (n=4,263)

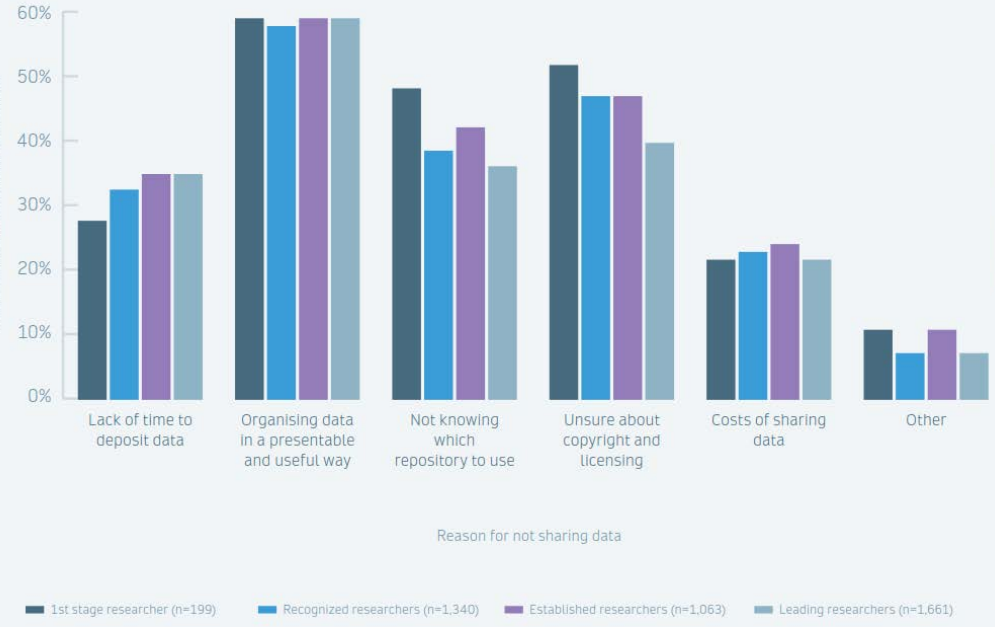


Figure 8: Problems in sharing datasets in different subject areas (n=7,719)

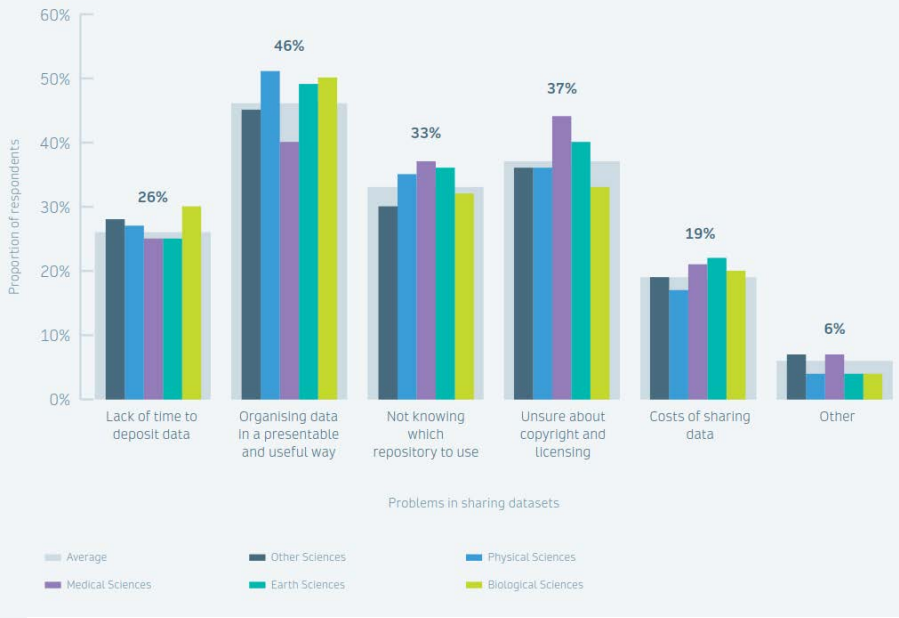


Figure 4: 'Other' problems mentioned in sharing datasets mentioned (n=385)

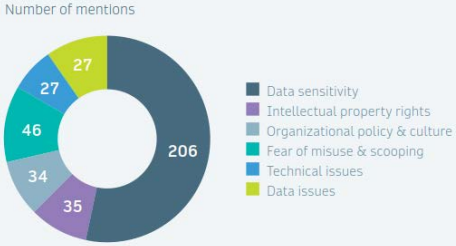


Figure 6: The depositing of data in different subject areas (n=7,664)

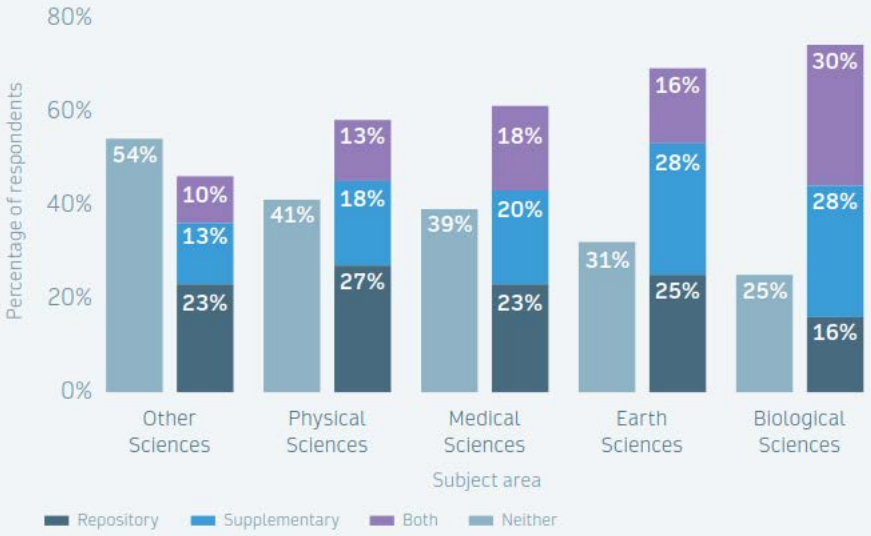


Figure 7: The importance that data are discoverable in different subject areas (1 is the least important) (n=7,626)

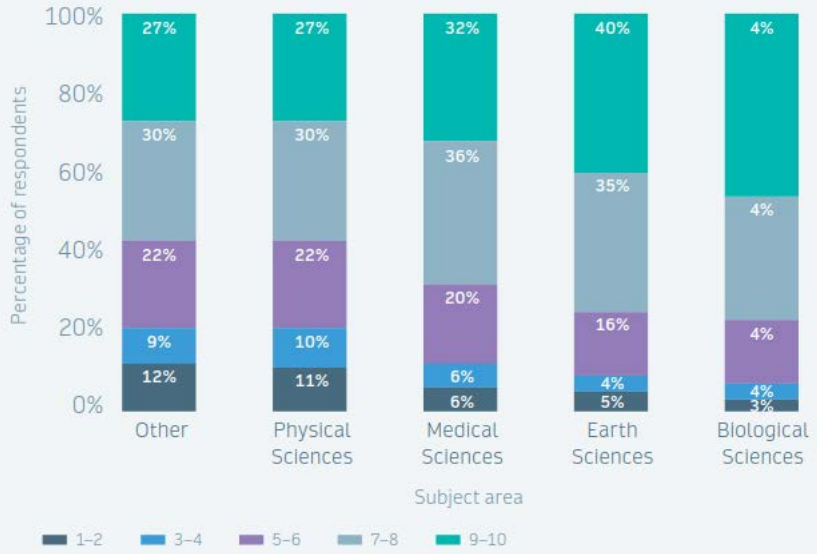
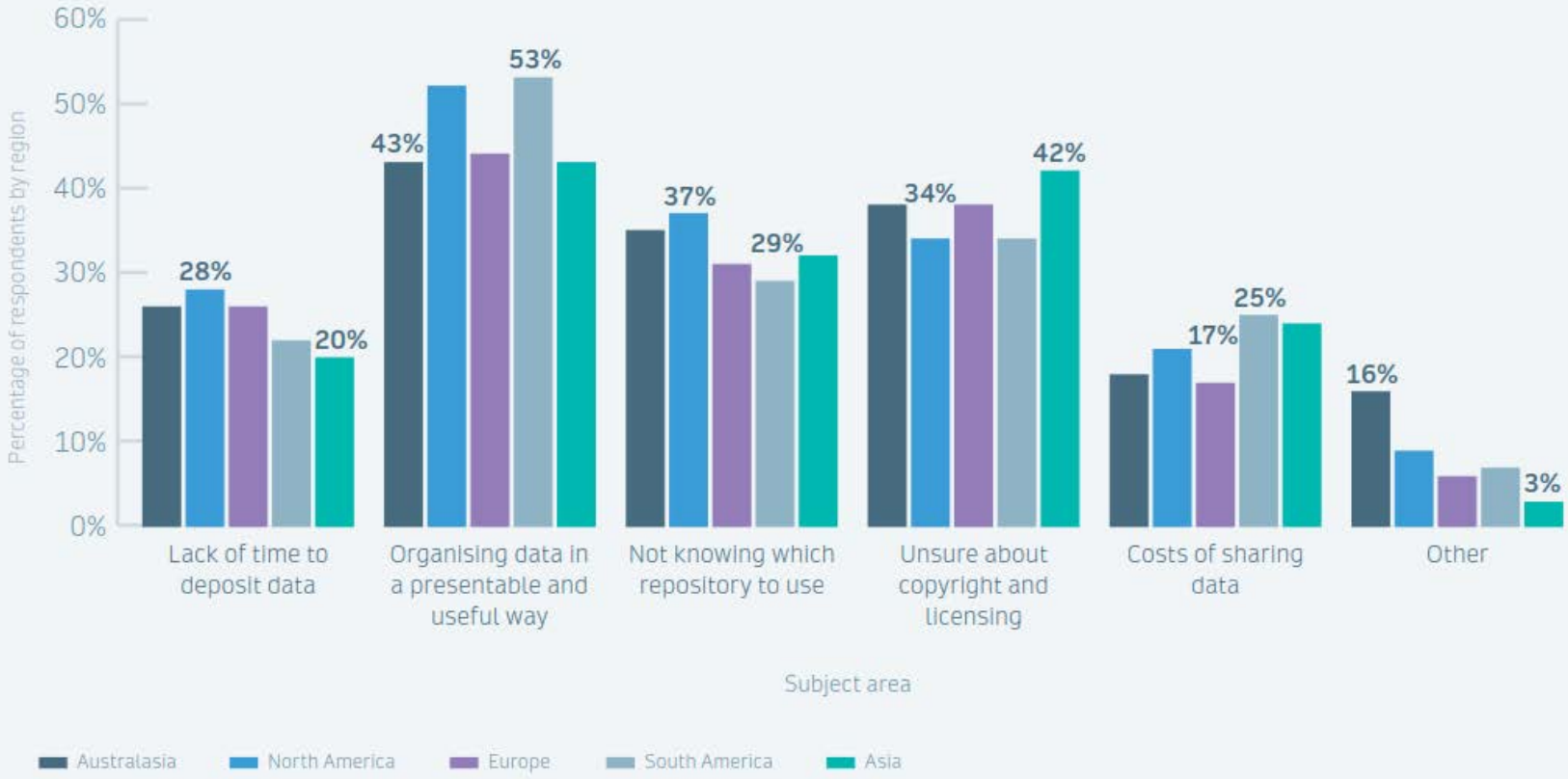


Figure 17: Reasons for not sharing data by region (n=7,654) – including % of highest and lowest responding regions



Summary (2-1)

- There are differences in data sharing for individual researchers, depending on the stage of their career and specifics of their research;
- Relevance of norms and policies;
- There is a need for standardization and harmonization of policy both from funders and journals, where policies are also being more widely adopted;
- A recent analysis of data sharing in The BMJ found that rates of sharing were low despite a strong data sharing policy;
- Data sharing policies are not enough, however. To increase the amount of data that are shared, there is a need for clearer routes to help researchers through the increasingly complex scholarly ecosystem;
- Increased data management, support and education;
- Faster, easier routes to optimal ways of sharing data.

What can be learned for Data Sharing within the MI-I?

- Degrees of Openness
- Implementing uniform data sharing policies
- Teaching data sharing as basic research strategy
- Providing appropriate technical infrastructure for data sharing
- Implement instruments which support the production of machine readable data management plans.