



# 10 Simple Rules for F.A.I.R. data sharing (well...maybe a little more than 10)

**Tina B Lonsdorf** 





Besançon et al. BMC Medical Research Methodology https://doi.org/10.1186/s12874-021-01304-y (2021) 21:117

BMC Medical Research Methodology

#### **COMMENTARY**

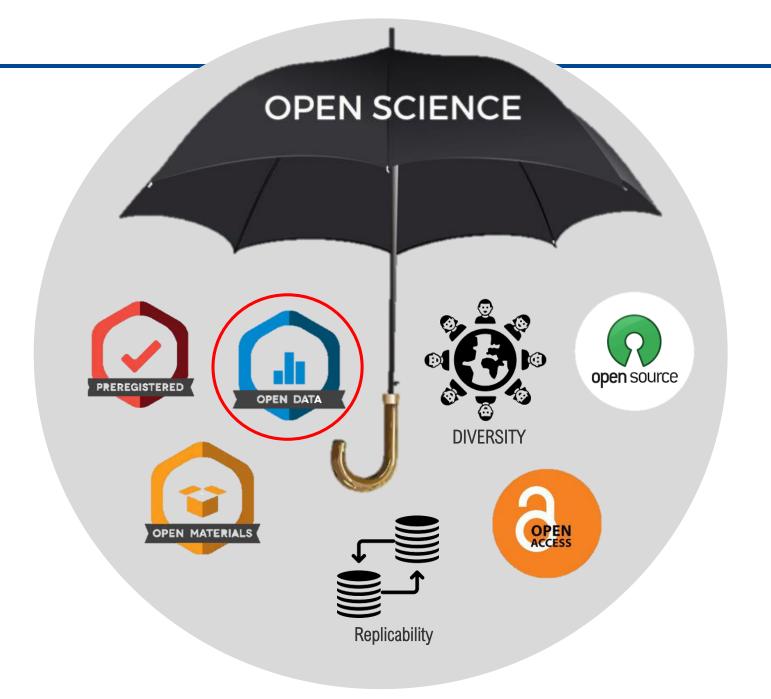
**Open Access** 

# Open science saves lives: lessons from the COVID-19 pandemic



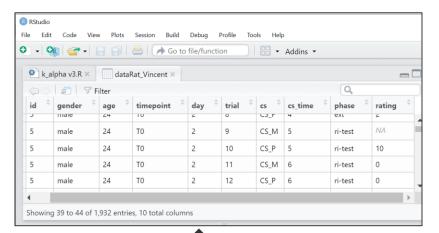
Lonni Besançon<sup>1,2\*</sup>, Nathan Peiffer-Smadja<sup>3,4</sup>, Corentin Segalas<sup>5</sup>, Haiting Jiang<sup>6</sup>, Paola Masuzzo<sup>7</sup>, Cooper Smout<sup>7</sup>, Eric Billy<sup>8</sup>, Maxime Deforet<sup>9</sup> and Clémence Leyrat<sup>5,10</sup>







# What is data sharing?

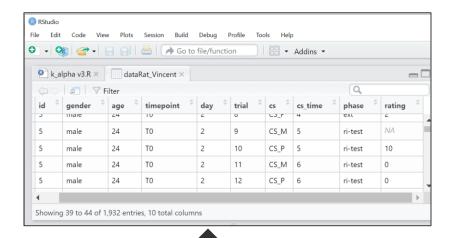


anything that your results are built on

"the activity of making **primary** research data available in an online repository upon publication of the article associated with the data."



# What is data sharing?

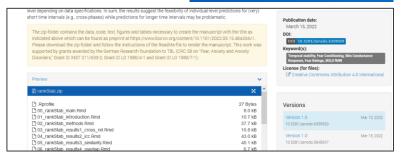


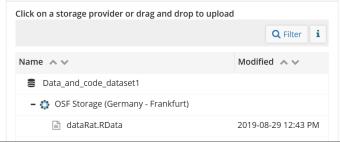
"the activity of making **primary** research data available in an **online** repository

upon publication of the article associated with the data."



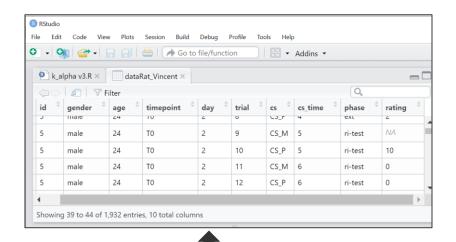








# What is data sharing?

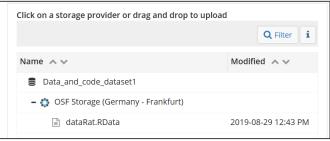


"the activity of making primary research data available in an online repository

upon publication of the article associated with the data."









# Why should we share data publicly?











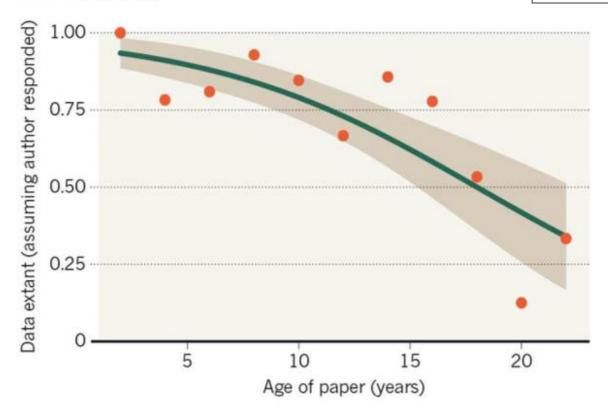


Counteracting data loss over time

### MISSING DATA

As research articles age, the odds of their raw data being extant drop dramatically.

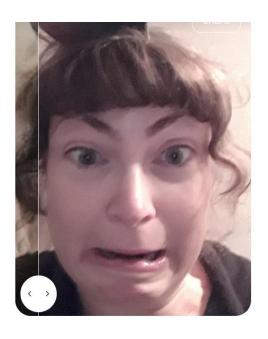
- Decrease of 17% per year.
- Availability as little as 20% for research from 20yrs ago.



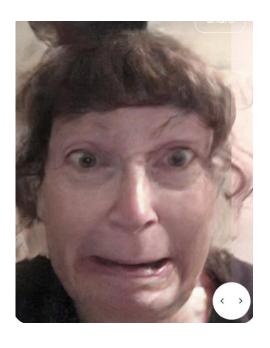


Counteracting data loss over time (also a selfish reason)

Past Me



Future Me





Does unfortunately not answer emails

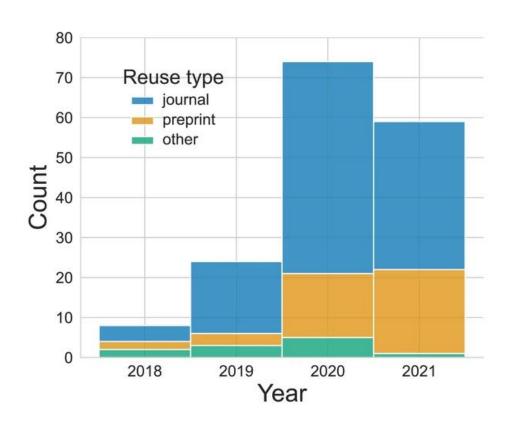


- Counteracting data loss over time (also a selfish reason)
- Maximize the impact of public investments in scientific research
- Use scarce resources (funds, manpower, time) optimally



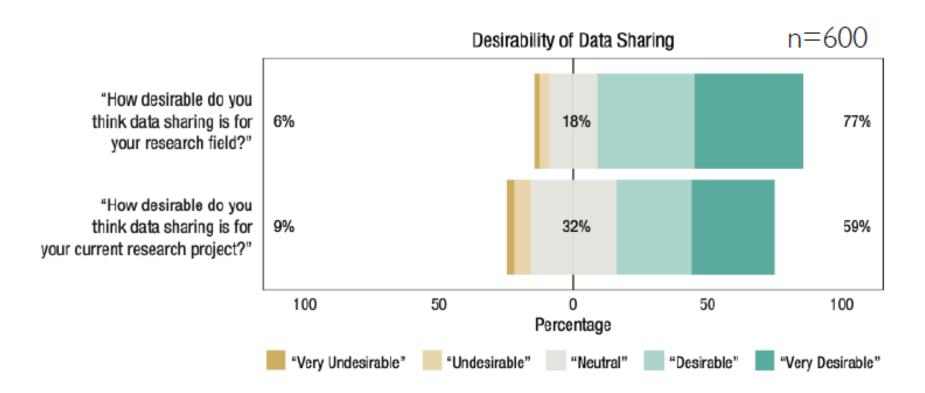
"collecting these data from scratch would have required more than 21.000 individual subjects...and a total reuse value of nearly 21 Mio US dollar"

(cf. Markiewicz et al., 2021, elife)



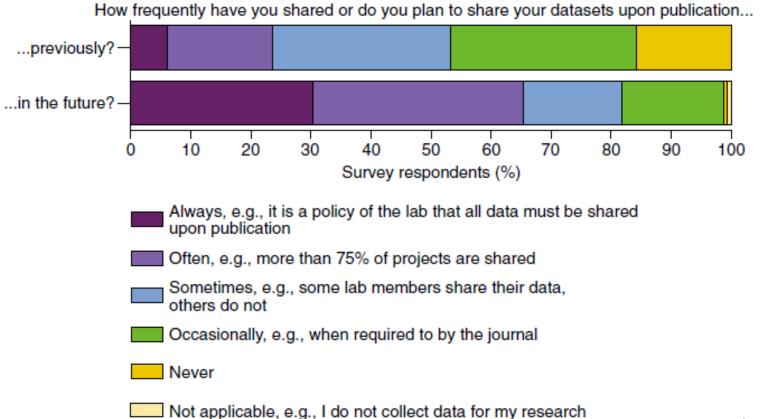


- Counteracting data loss over time (also a selfish reason)
- Maximize the impact of public investments in scientific research
- Use scarce resources (funds, manpower, time) optimally
- It is becoming increasingly valued



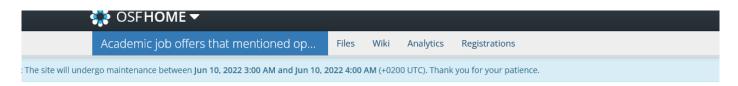


- Counteracting data loss over time (also a selfish reason)
- Maximize the impact of public investments in scientific research
- Use scarce resources (funds, manpower, time) optimally
- It is becoming increasingly valued





- Counteracting data loss over time (also a selfish reason)
- Maximize the impact of public investments in scientific research
- Use scarce resources (funds, manpower, time) optimally
- It is becoming increasingly valued (also for academic positions & grants)



# Academic job offers that mentioned open science

Contributors: Felix Schönbrodt, David Thomas Mellor, Christina Bergmann, Naomi Penfold, Samuel Westwood, Alexandra Lautarescu, Olivia Kowalczyk, Lorenza Dall Leonhard Falk Florentin Schramm, Bianca Weber, Greg Tananbaum, Franka Tabitha Etzel, Amanda Montoya

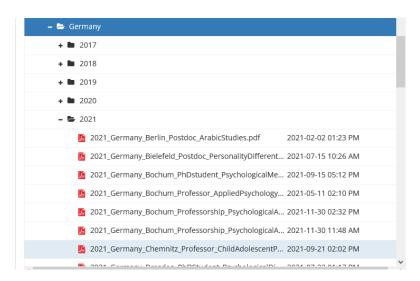
Date created: 2018-01-18 02:25 PM | Last Updated: 2022-05-23 08:19 AM

Identifier: DOI 10.17605/OSF.IO/7JBNT

Category: Project

Description: This project archives job offers (mostly professorship positions) which require or suggest an

License: Other 1



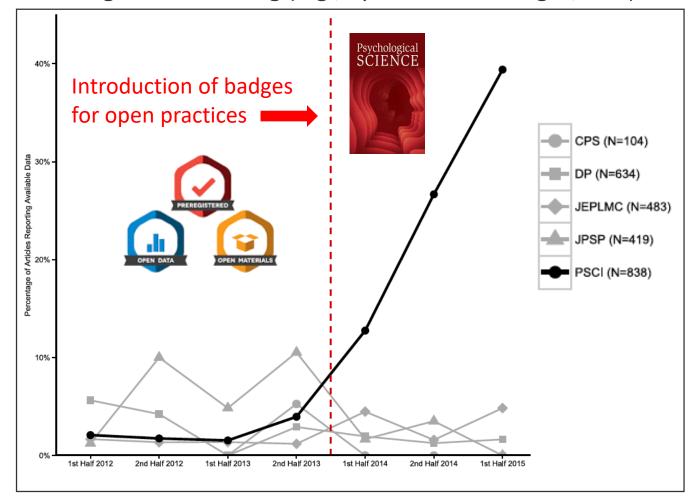




#### **Initiatives from Journals & Research societies**



- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)





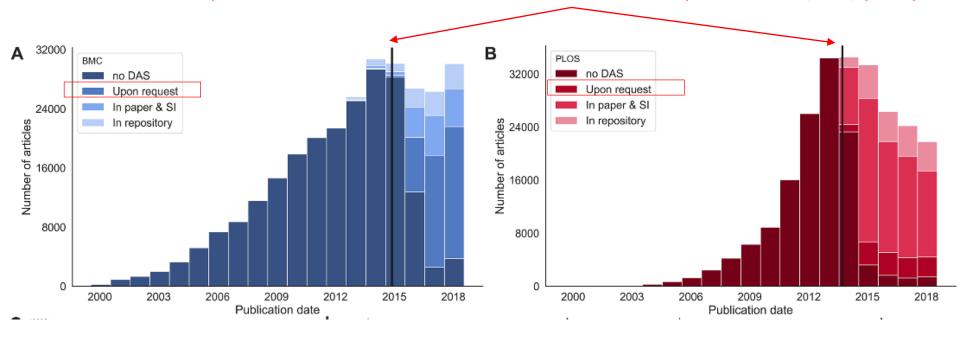


#### **Initiatives from Journals & Research societies**



- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)

publisher introduced a mandated data availability statement (DAS) policy.







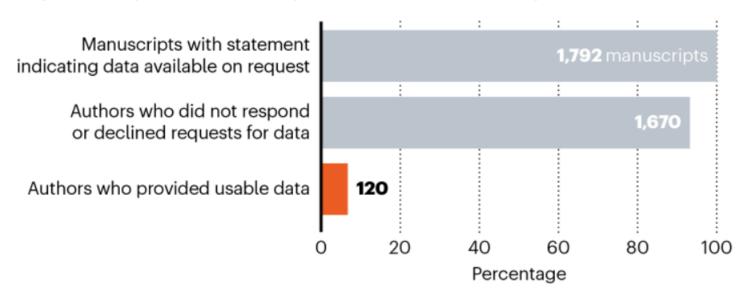
#### **Initiatives from Journals & Research societies**



- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)

# **DATA-SHARING BEHAVIOUR**

Of almost 1,800 manuscripts for which the authors stated they were willing to share their data, more than 90% of corresponding authors either declined or did not respond to requests for data. Only about 7% of authors actually handed over data.



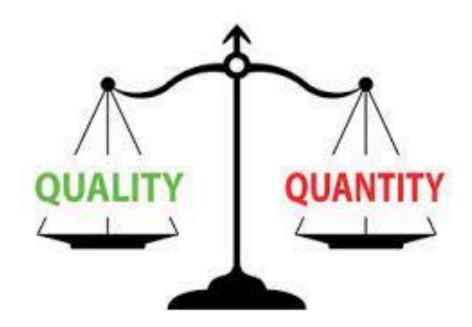




#### **Initiatives from Journals & Research societies**



- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)







#### **Initiatives from Journals & Research societies**

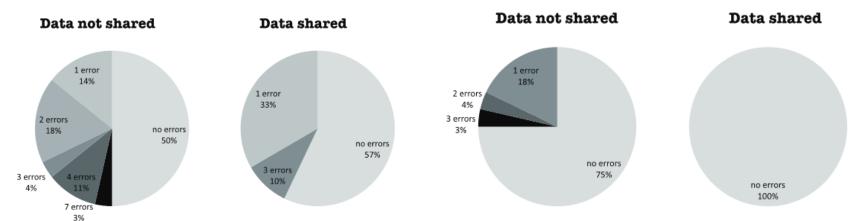


- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)



#### All reporting errors:

# Reporting errors concerned with p < .05:







#### **Initiatives from Journals & Research societies**



EMBARG

- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)



#### **Initiatives from Funders**

- Data sharing policies
- Mandatory data sharing plans



Data that have been collected in a funded research project *should* be made available to the public with a publication (Type I data) or immediately after completion of the research project (Type II data) or within a few month

→ All data (also those not inculded in a paper\*, meta-data, material)

\*applies to Type II data only





#### **Initiatives from Journals & Research societies**



- mandatory data sharing policy
- encouraged data sharing (e.g., open science badges, TOP)



#### **Initiatives from Funders**

- Data sharing policies
- Mandatory data sharing plans



#### **Initiatives from Researchers**

- Registered Reports
- Peer Review Openness Initiative (PRO; Moray, 2016)
- BIDS data structure













# **Data sharing questions**



What is the lawful & ethical basis for sharing?





What are common barriers to sharing?



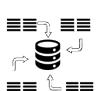
data sharing



When should I share?

What should I share?







With whom may I share?

Where should I share?



"The interest of the scientific community



"The interest of the scientific community



# Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten



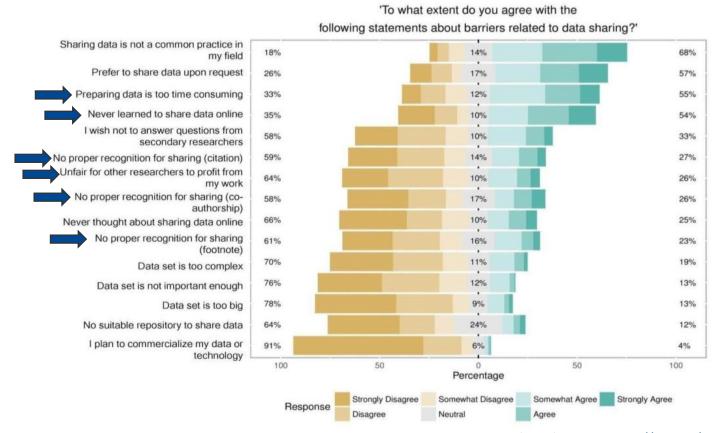
Stand 26.6. 2008

Forschungsprimärdaten bilden einen wertvollen Fundus an Informationen, die mit hohem finanziellem Aufwand erhoben werden. Je nach Fachgebiet und Methode sind sie replizierbar oder basieren auf nicht wiederholbaren Beobachtungen oder Messungen. In jedem Fall sollten die erhobenen Daten nach Abschluss der Forschungen öffentlich zugänglich und frei verfügbar sein. Dieses ist die wesentliche Voraussetzung dafür, dass Daten im Rahmen neuer Fragestellungen wieder genutzt werden können sowie dafür, dass im Falle von Zweifeln an der Publikation die Daten für die Überprüfung der publizierten Ergebnisse herangezogen werden können.

- Reusability
- Reproducibility



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected



Houtkoop et al., (2018) AMPPS <a href="https://doi.org/10.1177/2515245917751886">https://doi.org/10.1177/2515245917751886</a> Schönbrodt, Gollwitzer & Abele-Brehm (2016) Data Management in Psychological Science: Specification of the DFG Guidelines



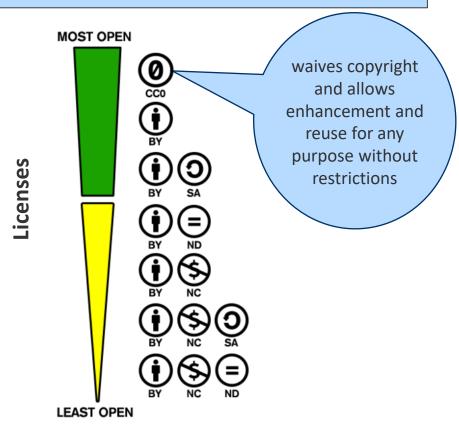
"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected





"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected

- Public sharing of data
- Data under embargo
  - → public after a certain time period
  - → available upon request
- No public sharing of data possible



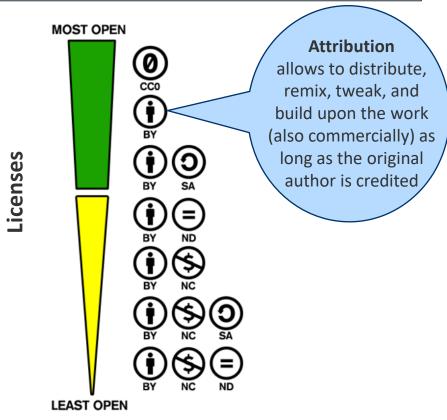
https://choosealicense.com/

https://www.openaire.eu/how-do-i-license-my-research-data



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected

- Public sharing of data
- Data under embargo
  - → public after a certain time period
  - → available upon request
- No public sharing of data possible
- Data sharing agreement forms



https://choosealicense.com/

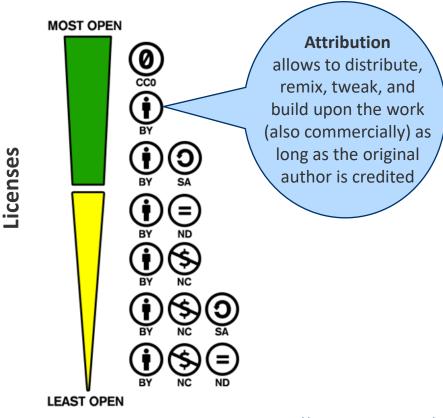
https://www.openaire.eu/how-do-i-license-my-research-data



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected

- Public sharing of data
- Data under embargo
  - → public after a certain time period
  - → available upon request
- No public sharing of data possible
- Data sharing agreement forms

Date created: 2019-03-22 01:52 PM | Last Updated: SFAM
Create DOI
Category: Data
Description: Add a brief description to your component
License: CC-By Attribution 4.0 International



https://choosealicense.com/

https://www.openaire.eu/how-do-i-license-my-research-data



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected as well as the interest of study participants in the ethically responsible handling of their data."



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected as well as the interest of study participants in the ethically responsible handling of their data."



# General Data Protection Regulation (GDPR) Datenschutzgrundverordnung (DSGVO)



Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data

- + Informed and signed consent of the study participant
  - + Data processing contracts if applicable



"The interest of the scientific community must be weighed against the interest of individual researchers to harness the data they collected as well as the interest of study participants in the ethically responsible handling of their data."



# General Data Protection Regulation (GDPR) Datenschutzgrundverordnung (DSGVO)



Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data

- + Informed and signed consent of the study participant
  - + Data processing contracts if applicable



### **Excursion: informed consent under the GDPR**

#### Preprint

Information Provision for Informed Consent Procedures in Psychological Research under the GDPR: A Practical Guide

#### Dara Hallinan

FIZ Karlsruhe — Leibniz Institute for Information Infrastructure

#### Franziska Boehm

FIZ Karlsruhe — Leibniz Institute for Information Infrastructure

#### Annika Külpmann

Faculty of Psychology, Ruhr University Bochum

#### Malte Elson

Faculty of Psychology, Ruhr University Bochum Horst Görtz Institute for IT Security, Ruhr University Bochum

https://psyarxiv.com/znb9m

# N BRAIN MAPPI





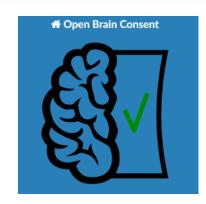
#### The Open Brain Consent: Informing research participants and obtaining consent to share brain imaging data

Elise Bannier, Gareth Barker, Valentina Borghesani, Nils Broeckx, Patricia Clement, Kyrre E. Emblem, Satrajit Ghosh, Enrico Glerean, Krzysztof J. Gorgolewski, Marko Havu, Yaroslav O. Halchenko 🔀 Peer Herholz Anne Hespel, Stephan Heunis X, Yue Hu, Chuan-Peng Hu, Dorien Huijser, María de la Iglesia Vayá, Radim Jancalek, Vasileios K. Katsaros, Marie-Luise Kieseler, Camille Maumet, Clara A. Moreau, Henk-Jan Mutsaerts, Robert Oostenveld, Esin Ozturk-Isik, Nicolas Pascual Leone Espinosa, John Pellman, Cyril R Pernet 🔀 Francesca Benedetta Pizzini, Amira Šerifović Trbalić, Paule-Joanne Toussaint, Matteo Visconti di Oleggio Castello, Fengjuan Wang, Cheng Wang, Hua Zhu ... See fewer authors A

First published: 01 February 2021 | https://doi.org/10.1002/hbm.25351 | Citations: 12

#### Ultimate consent form

The following consent form has been put together, by merging best parts of existing consent forms and consulting with experts in research ethics.





# Lawful basis of data sharing





# Lawful basis of data sharing

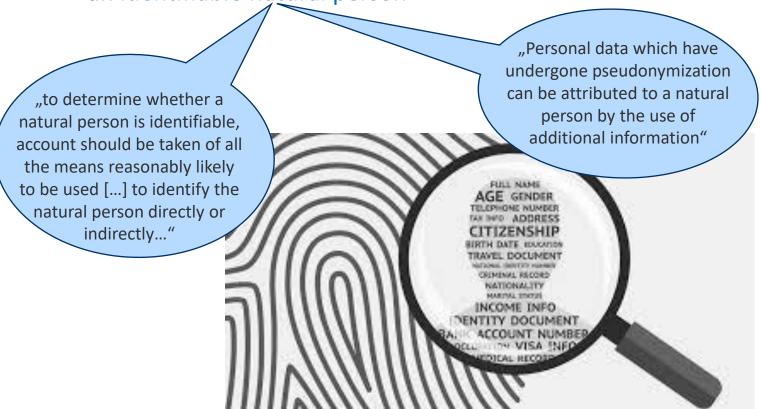
Applies to personal data only





→ any information "relating to an identified person or an identifiable natural person"

effective as of 05/2018





# Lawful basis of data sharing

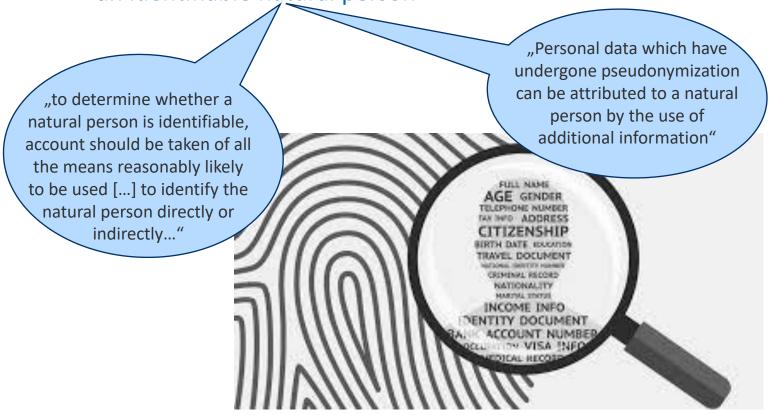
Applies to personal data only





→ any information "relating to an identified person or an identifiable natural person"

effective as of 05/2018





### Lawful basis of data sharing

Applies to personal data only

General
Data
Protection
Regulation



effective as of 05/2018

- → any information "relating to an identified person or an identifiable natural person"
- → does not apply to anonymized data
- → pseudonymized data are personal data





### Lawful basis of data sharing – purpose of data processing

It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognized ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or research projects to the extent allowed by the intended purpose. (GDPR, Recital 33).









Ich willige in die Veröffentlichung meiner pseudonymisierten Daten als "open data" (ohne Weitergabe des Pseudonymisierungscodes) über Forschungsplattformen und deren Nachnutzung die auch über die Ziele dieser Studie hinausgehen können wie von der Deutschen Forschungsgemeinschaft als finanzieller Förderer dieser Studie empfohlen.

O JA

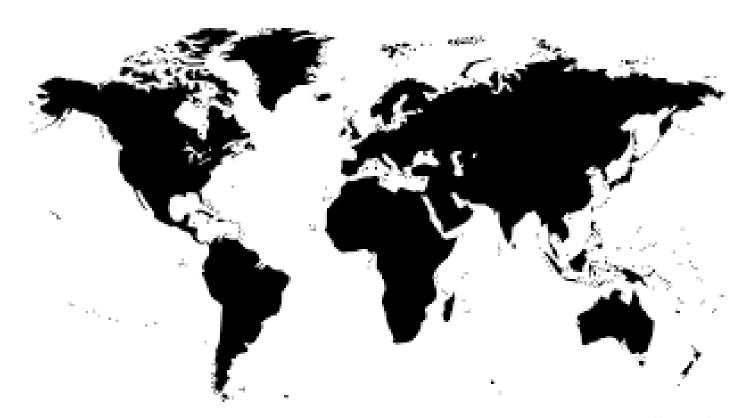
O Nein



# Lawful basis of data sharing – with whom?



- Data transfer only within countries that are deemed to adequately protect data.
- EU recognized countries that adequately protect data: Israel, Japan,
   New Zealand, Switzerland, Uruguay...

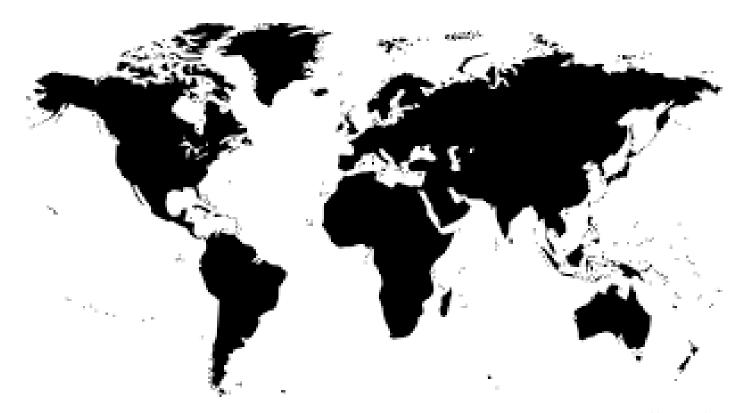




### Lawful basis of data sharing – with whom?



- Data transfer only within countries that are deemed to adequately protect data.
- EU recognized countries that adequately protect data: Israel, Japan,
   New Zealand, Switzerland, Uruguay
- Transfer to other countries requires "lawful basis" (specific and separate consent)
- US is *not* deemed to adequately protect data

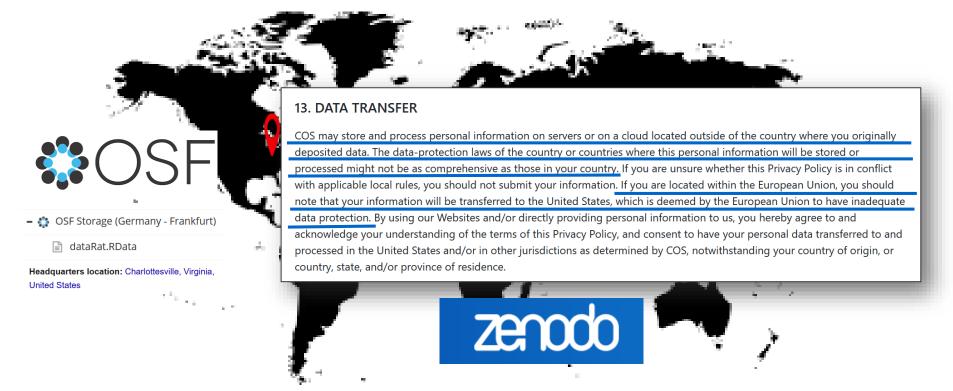




### Lawful basis of data sharing



- Data transfer only within countries that are deemed to adequately protect data.
- EU recognized countries that adequately protect data: Israel, Japan,
   New Zealand, Switzerland, Uruguay
- Transfer to other countries requires "lawful basis" (specific and separate consent)
- US is not deemed to adequately protect data





### Lawful basis of data sharing



- Data transfer only within countries that are deemed to adequately protect data.
- EU recognized countries that adequately protect data: Israel, Japan,
   New Zealand, Switzerland, Uruguay
- Transfer to other countries requires "lawful basis" (specific and separate consent)
- US is *not* deemed to adequately protect data





### **Data sharing questions**



What is the lawful & ethical basis for sharing?





What are common barriers to sharing?



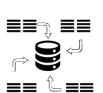
data sharing



When should I share?

What should I share?







With whom may I share?

Where should I share?

Invited Forum: Challenges in Making Data Available *Tutorial* 

### **Practical Tips for Ethical Data Sharing**

Michelle N. Meyer Geisinger Health System, Danville, Pennsylvania



Advances in Methods and Practices in Psychological Science 2018, Vol. 1(1) 131–144 © The Author(s) 2018 Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/2515245917747656 www.psychologicalscience.org/AMPPS

**S**SAGE



- Promise to destroy data
- Promise to limit analyse to certain topics



- Get consent to ratain and share (and include in IRB)
- Consider risk of re-identification
- Use appropriate repositories

Invited Forum: Challenges in Making Data Available Tutorial

### **Practical Tips for Ethical Data Sharing**

Michelle N. Meyer Geisinger Health System, Danville, Pennsylvania



Advances in Methods and Practices in Psychological Science 2018, Vol. 1(1) 131–144 © The Author(s) 2018 Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/2515245917747656 www.psychologicalscience.org/AMPPS





### What about data that were collected without excplicit consent to share?



- Consent was silent about data sharing (vs. Promised not to share)
- Data are not sentitive (re-identification is unlikely to cause harm)
- Re-identification is unlikely (& incentive is low)
- Sharing is limited to secondary research that falls within the scope of the research described in the orginal consent form
- Sharing is limited to research that participants are not known to oppose



ETHICS

Anonymized data do not fall under the GDPR and can be shared

Invited Forum: Challenges in Making Data Available **Tutorial** 

### **Practical Tips for Ethical Data Sharing**

Advances in Methods and Practices in Psychological Science 2018, Vol. 1(1) 131-144 © The Author(s) 2018 Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/2515245917747656 www.psychologicalscience.org/AMPPS

Michelle N. Meyer

Geisinger Health S



Contents lists available at ScienceDirect

(S)SAGE

#### NeuroImage





Review

The ethics of secondary data analysis: Considering the application of Belmont principles to the sharing of neuroimaging data

Beth Brakewood a,\*, Russell A, Poldrack b

- Imaging Research Center, University of Texas, 100 E. 24th
- b Department of Psychology and Neurobiology and Imagin

# **BRAIN MAPPIN**

REVIEW ARTICLE | ① Open Access | ② ①

Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed

Tonya White K. Elisabet Blok, Vince D. Calhoun

First published: 04 July 2020 | https://doi.org/10.1002/hbm.25120 | Citations: 15



### **Data sharing questions**



What is the lawful & ethical basis for sharing?





What are common barriers to sharing?



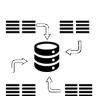
data sharing



When should I share?

What should I share?







With whom may I share?

Where should I share?



# **Data sharing questions**

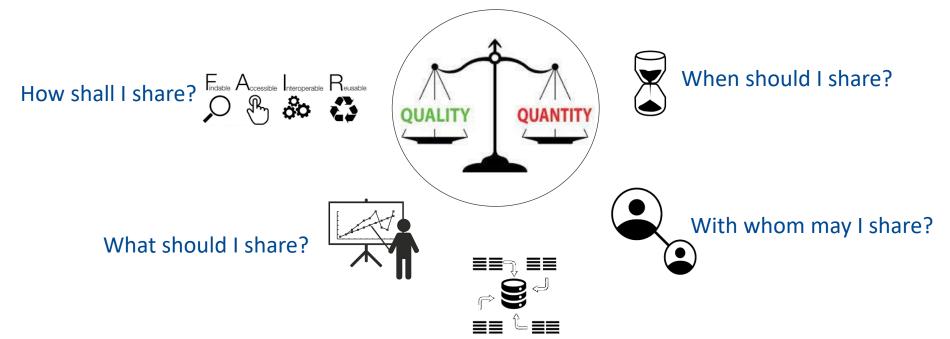


What is the lawful & ethical basis for sharing?





What are common barriers to sharing?



Where should I share?



### Data sharing – what is is good for?

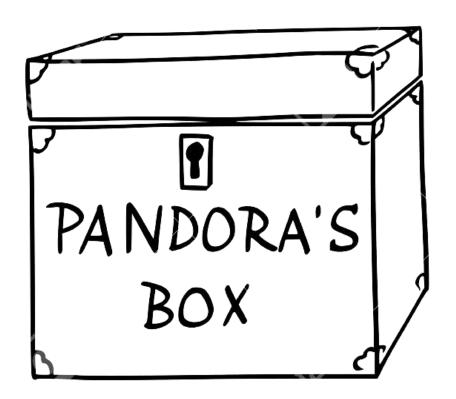
What will happen with all the shared data? Will someone really check/re-use it? Big data or meta-analyses? How do we avoid that it's a lot of effort put into something that no one really cares about?





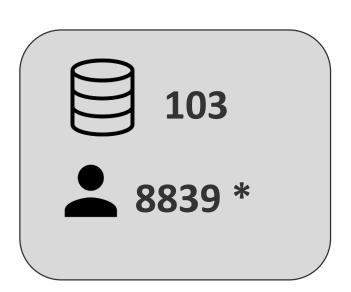
- Google dataset search
- OSF search
- e-Mails among collaborators & colleagues







- Google dataset search
- OSF search
- e-Mails among collaborators & colleagues



\* Trait anxiety: 31 datasets with N=2458 SCR: 65 datasets with N=4908





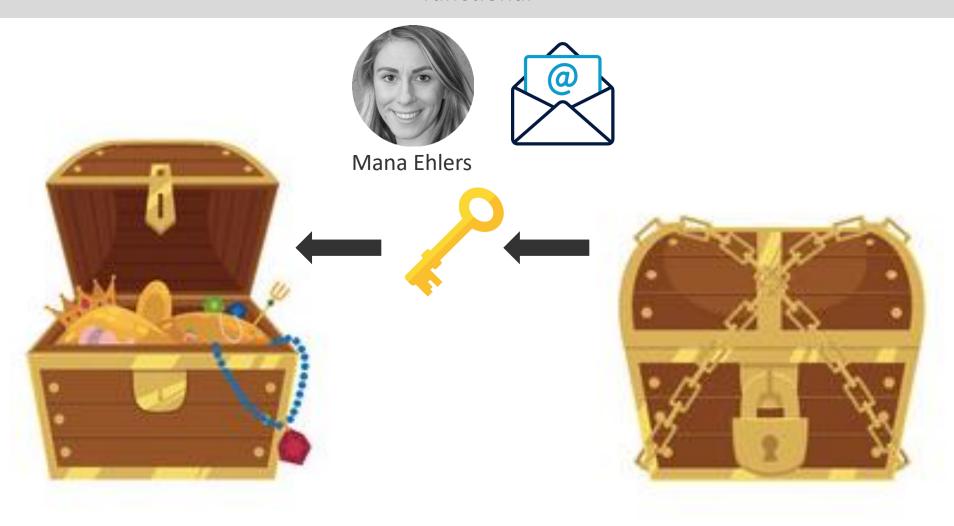


We can only use publicly available data to their full potential if the data are indeed functional



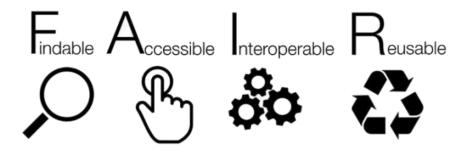


We can only use publicly available data to their full potential if the data are indeed functional





### **Community driven recommendations**



# SCIENTIFIC DATA (11011) ON ATION OF THE PROPERTY OF THE PROPER

Amended: Addendum

SUBJECT CATEGORIES

» Research data » Publication characteristics

# **OPEN** Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson et al.#

Received: 10 December 2015 Accepted: 12 February 2016 Published: 15 March 2016 There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically



### **Community driven recommendations**

### Berichte

Der Umgang mit Forschungsdaten im Fach Psychologie: Konkretisierung der DFG-Leitlinien

Im Auftrag des DGPs Vorstands (17.

Felix Schönbrodt, Mario Gollwitzer und Andrea Abele-

https://econtent.hogrefe.com/doi/pdf/10.1026/00

### Data Management in Psychological Science: Specification of the DFG Guidelines

Felix Schönbrodt, Mario Gollwitzer & Andrea Abele-Brehm\*

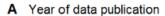
on behalf of the DGPs Executive Board (adopted Sep 17, 2016; current version: Nov 9, 2016)

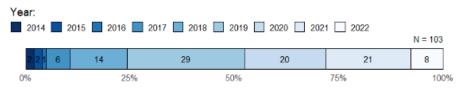
# Data Management and Data Sharing in Psychological Science: Revision of the DGPs Recommendations

German Psychological Society (Deutsche Gesellschaft für Psychologie - DGPs) Commission "Open Science"

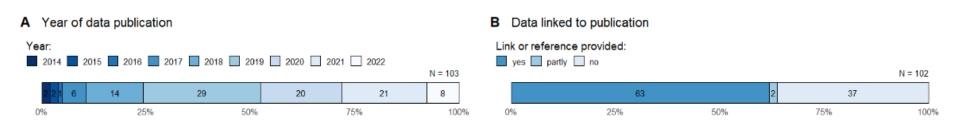
(approved by the DGPs Executive Board on June 26, 2020)



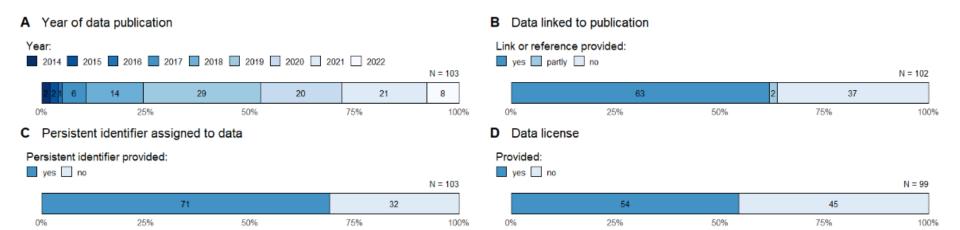










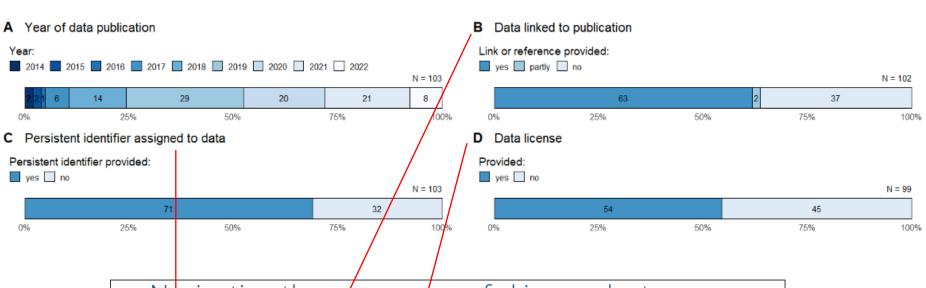




59



### An Inventory of open data in fear conditioning research



Navigating the manyverse of skin conductance response quantification approaches – a direct comparison of Trough-to-Peak, Baseline-correction and model-based approaches in Ledalab and PsPM



Contributors: Manuel Kynn, Tina Lonsdorf, Anna Gerlicher

Date created: 2021-12 02 06:13 PM | Last Updated: 2022-05-30 10:55 PM

Identifier: DOI 10.17605/OSF.IO/FT86V

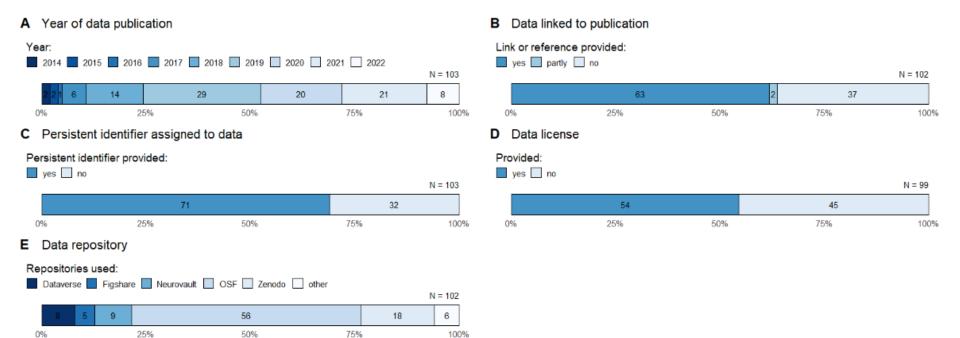
Category: 🛢 Data

Description:

Data and code for https://psyarxiv.com/9hekd (now accepted in Psychophysiology, 2022, DOI:10.1111/psyp.14058)

License: CC-By Attribution 4.0 International







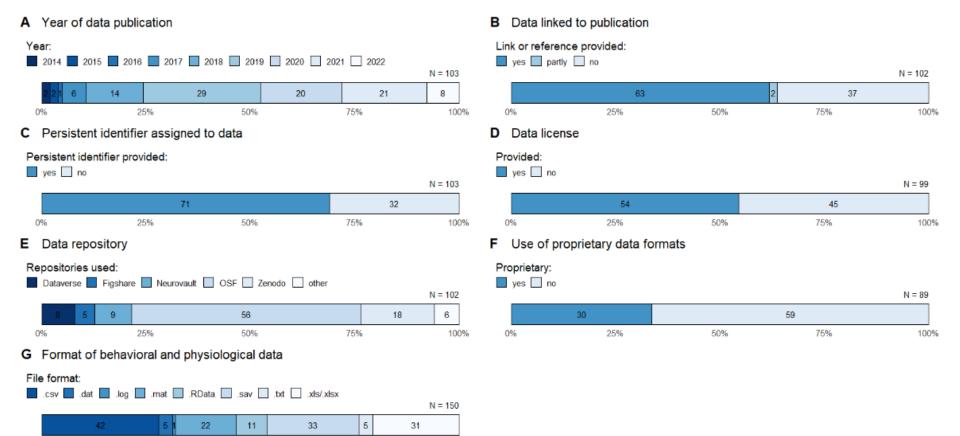
0%

25%

50%

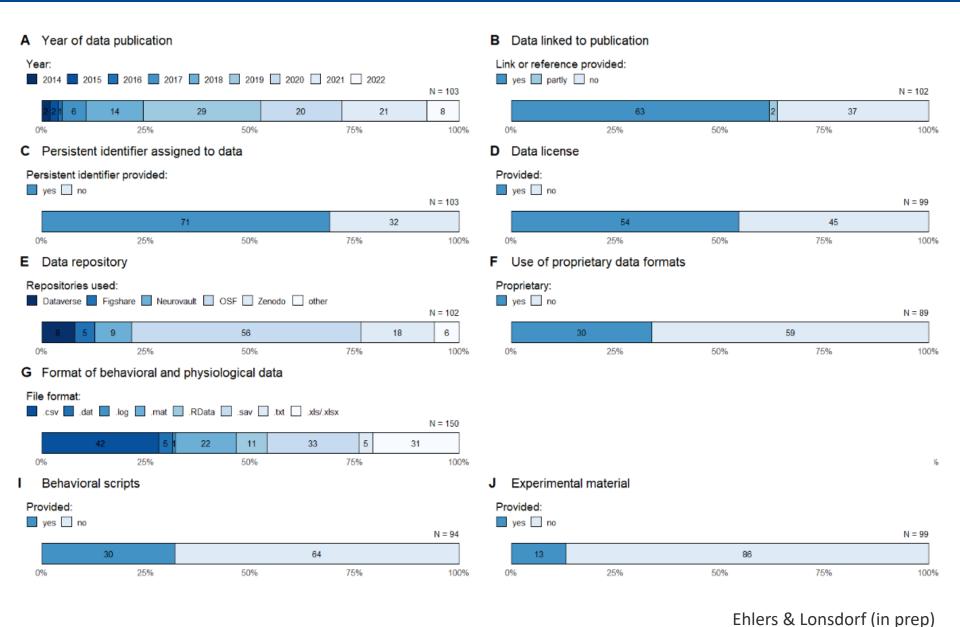
75%

### An Inventory of open data in fear conditioning research

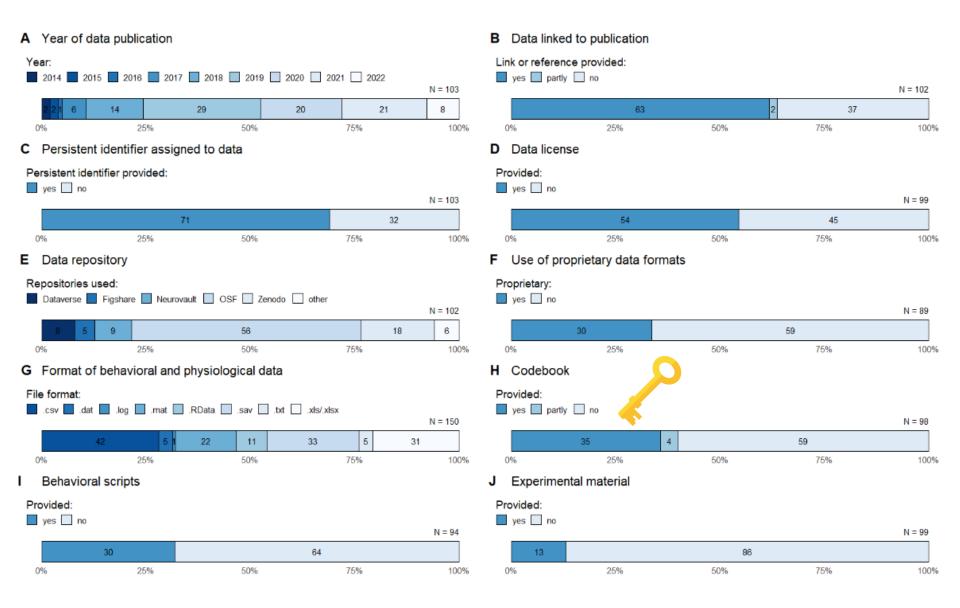


100%











# **Data sharing questions**

What is the lawful and ethical basis for sharing?





What are common barriers to sharing?



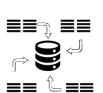
data sharing



When should I share?

What should I share?





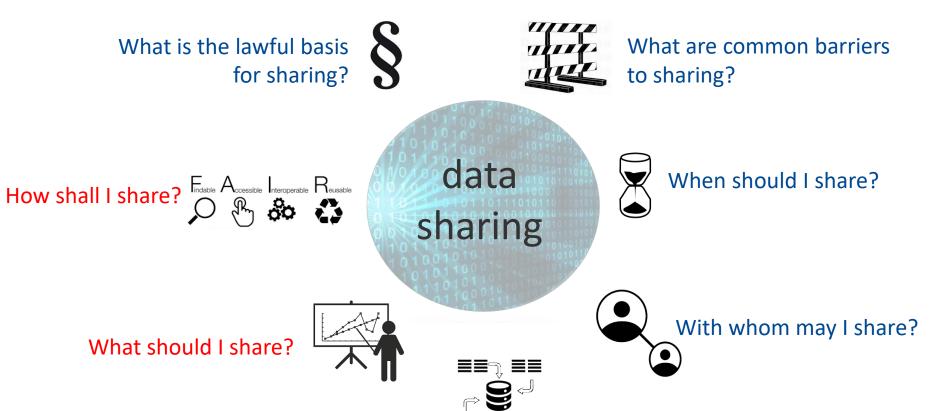


With whom may I share?

Where should I share?



### Take home messages from the field-specific inventory



Where should I share?





# What have we learned from the inventory?



### General



- Include doi
- Include license
- Version control
- Provide link to publication/preprint
- Prefer searchable repositories





#### Versions

Version 1.0 Mar 15, 2022 10.5281/zenodo.6359920

Version 1.0 10.5281/zenodo.5648047 Mar 15, 2022

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.5648046. This DOI represents all versions, and will always resolve to the latest one. Read more.

Data and code for https://psyarxiv.com/9h2kd (now accepted in Psychophysiology, 2022, DOI:10.11111/psyp.14058) License: CC-By Attribution 4.0 International

Mar 15, 2022





### What have we learned from the inventory?



### General

doi



- Include doi
- Include license
- Version control
- Provide link to publication/preprint
- Prefer searchable repositories

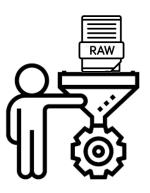
Versions
Version 1.0 10.5281/zen

ersion 1.0 Mar 15, 2022 0 5281/zenodo 6359920

Version 1.0 10 5281/zenodo 5648047

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.5648046. This DOI represents all versions, and will always resolve to the latest one. Read more.

Data and code for https://psyarxiv.com/9h2kd (now accepted in Psychophysiology, 2022, DOI:10.1111/psyp.14058)
License: CC-By Attribution 4.0 International



### Data

- Non-proprietary data format
- Provide codebook (variable names, abbreviations, aggregation, transformations...)
- Provide met-data (folder structure, content)
- Share raw and processed data (!)
- English language



