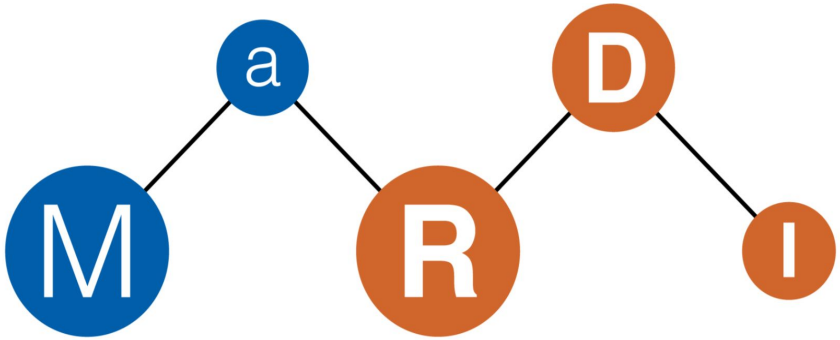
Mathematical research data

& Pizza

Leipzig, 03.08.2022



TABEA BACHER
MPI MATH IN THE SCIENCES

What is research data?

“The recorded factual material commonly accepted in the scientific community as necessary to validate research findings.”

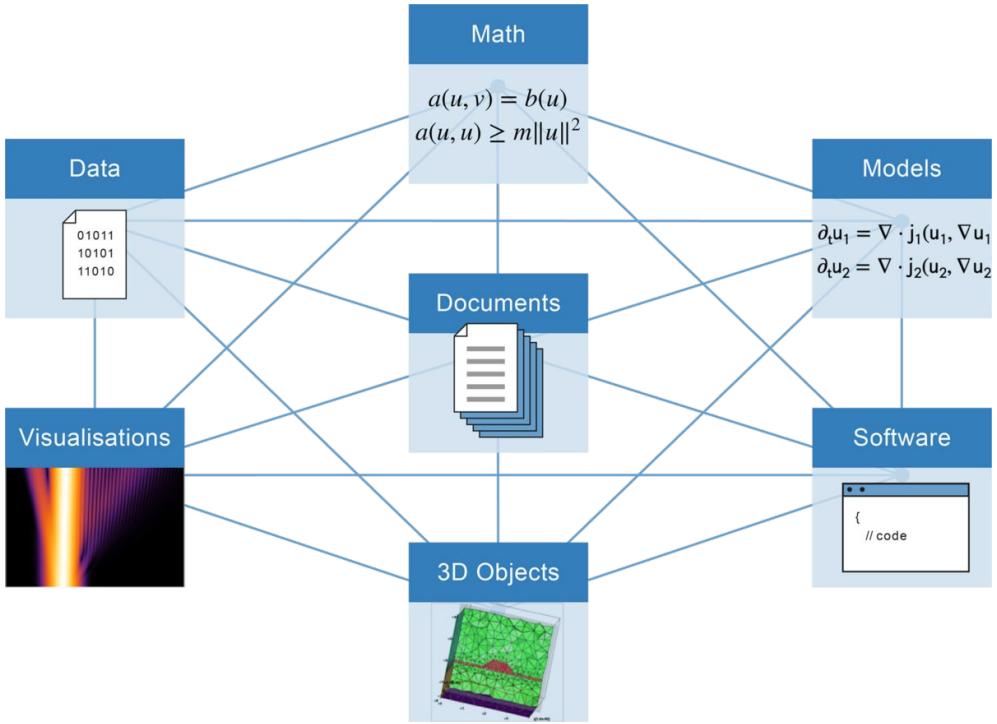
“Alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder ihre Ergebnisse sind.”

This is much broader than data alone!

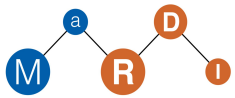
https://www.ukri.org/about-us/epsrc/our-policies-and-standards/policy-framework-on-research-data/scope-and-benefits/

https://www.forschungsdaten.info/themen/informieren-und-planen/was-sind-forschungsdaten/ und Forschungsdaten Definition: Kindling, Maxi und Schirmbacher, Peter: „Die digitale Forschungswelt“ als Gegenstand der Forschung. Information – Wissenschaft – Praxis 64 (2013): S. 130. doi.org/10.1515/iwp-2013-001

What is research data in mathematics?



- mathematical documents: papers, proofs, formulae,...
- notebooks, domain-specific research-software packages and libraries, computer algebra systems, programmes, scripts
- simulation data
- formalised mathematics
- collections of mathematical objects
- mathematical models
- …

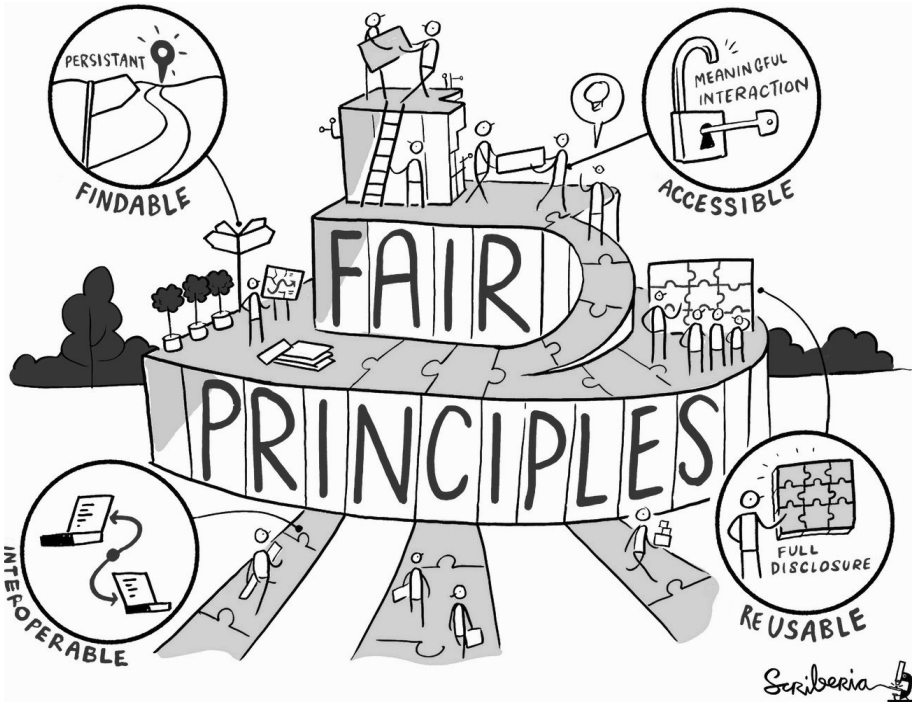


What is the status quo with research data in mathematics?

- research data which was promised in papers and stored on long-gone personal homepages
- research data which would run on one computer but not on another
- rely on colleagues/advisors to know about research data
- research data which would just not provide the promised results

A lot of implicit knowledge and sometimes big hurdles to build on other people's research!

How to handle research data



Ideally, it would be

- Findable
- Accessible
- Interoperable
- Reusable

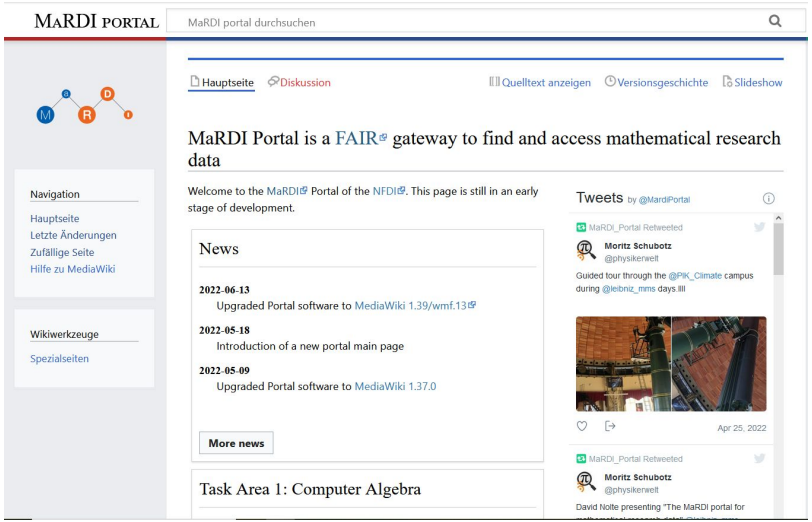
Mark Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gaby Appleton, et al. The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3(160018), 2016.

MaRDI – the Mathematical Research Data Initiative

- kick off in November 2021
- 15 institutions and partners
- 1 out of 30 consortia across all scientific disciplines
- 28 (full-time equivalent) employees
- 10 mio. euro funding over 5 years

Funded by the Deutsche Forschungsgemeinschaft (DFG), Projektnummer 460135501. NFDI 29/1 "MaRDI – Mathematische Forschungsdateninitiative".

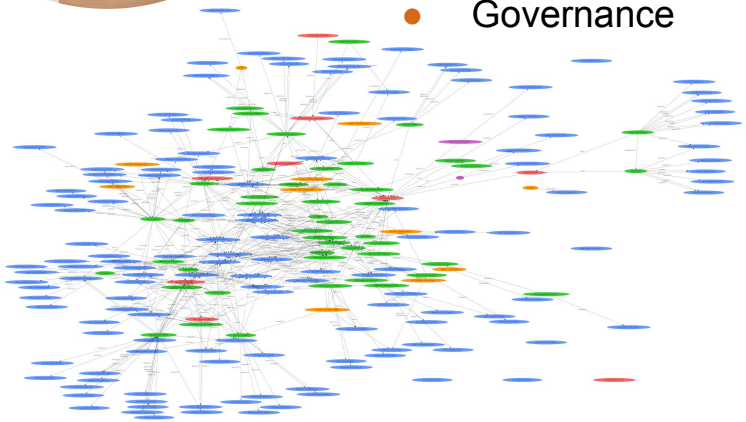
Task Areas and Services



- portal
- knowledge graphs
- workshops and training
- benchmarks
- peer-review standards and processes
- workflows

Task Areas:

- Computer Algebra
- Scientific Computing
- Statistics and Machine Learning
- Interdisciplinary
- MaRDI Portal
- Data Culture and Community Integration
- Governance



MaRDI and YOU

Events:

MaRDI Minisymposium @ DMV Jahrestagung

MaRDI Workshops on

- Digital Humanities and Math
- Scientific Computing
- Computer Algebra (MaRDI, OSCAR and MATHREPO)

https://www.mardi4nfdi.de/news

Newsletter: https://t1p.de/r4c3c

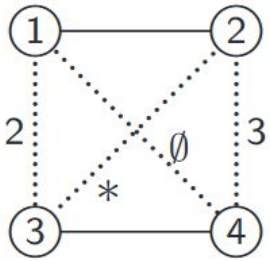
Twitter: https://twitter.com/mardi4nfdi

www.mardi4nfdi.de/community

One example from my research

$$X \perp\!\!\!\perp Y \mid Z$$

$$[X \perp\!\!\!\perp Y] \Leftrightarrow p_{00} \cdot p_{11} = p_{01} \cdot p_{10}$$



Project: Complete classification of conditional independence structures which are realizable by four jointly distributed binary random variables

Field of mathematics: Algebraic Statistics

Types of research data: papers, code, software, database, graphics

Services used throughout:

- google search
- arxiv
- jupyter notebooks
- Internet archive
- deepl
- MS Word

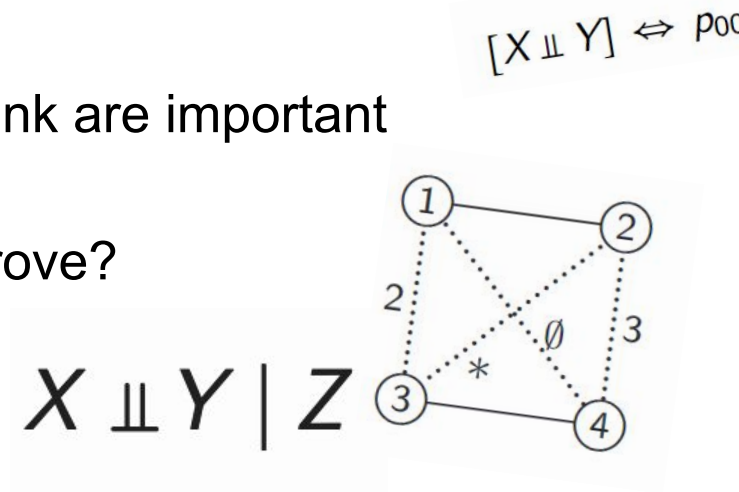
Lněnička, R., & Matúš, F. (2007). On Gaussian conditional independence structures. *Kybernetika*, *43*(3), 327-342

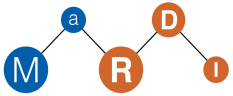
http://web.archive.org/web/20070719141308/http://atrey.karlin.mff.cuni.cz/~simecek/skola/

How FAIR is this?

While you enjoy your Pizza

- Share your own stories!
- What kind of mathematics do you do and what research data do you encounter there?
- Which services would you like to use/ think are important for you?
- Do you have an issue MaRDI could improve?





Design a Postcard

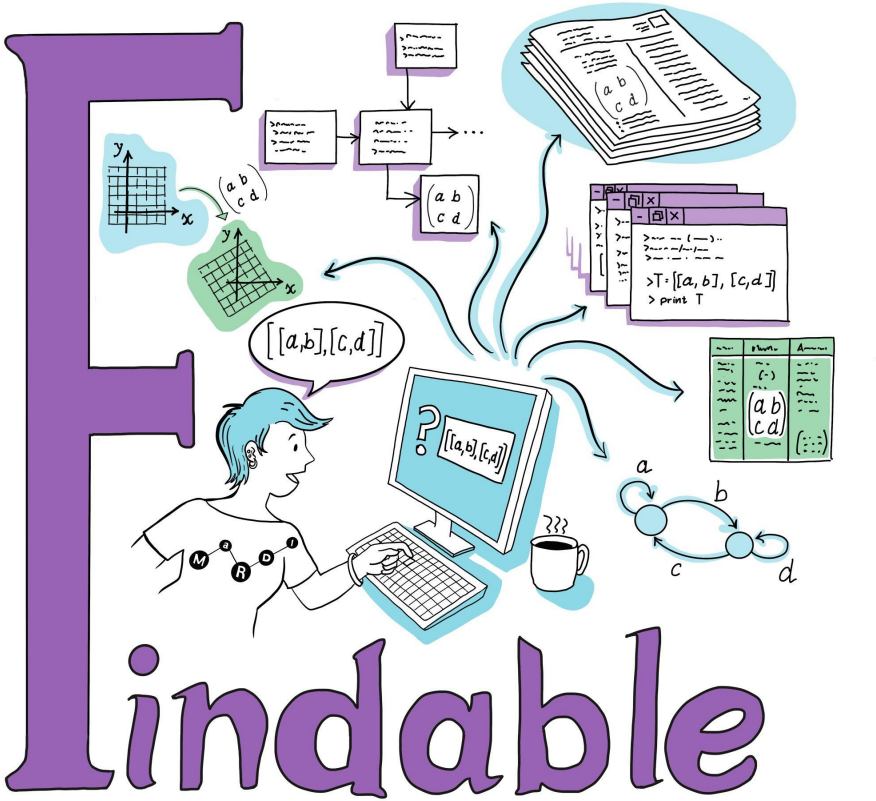
Review your discussion:

- What are the commonalities of your stories?
- What is your conclusion?

Try to capture this essence in a few words and/or pictures on the prepared blanks.

Design a Postcard

- a lesson learned



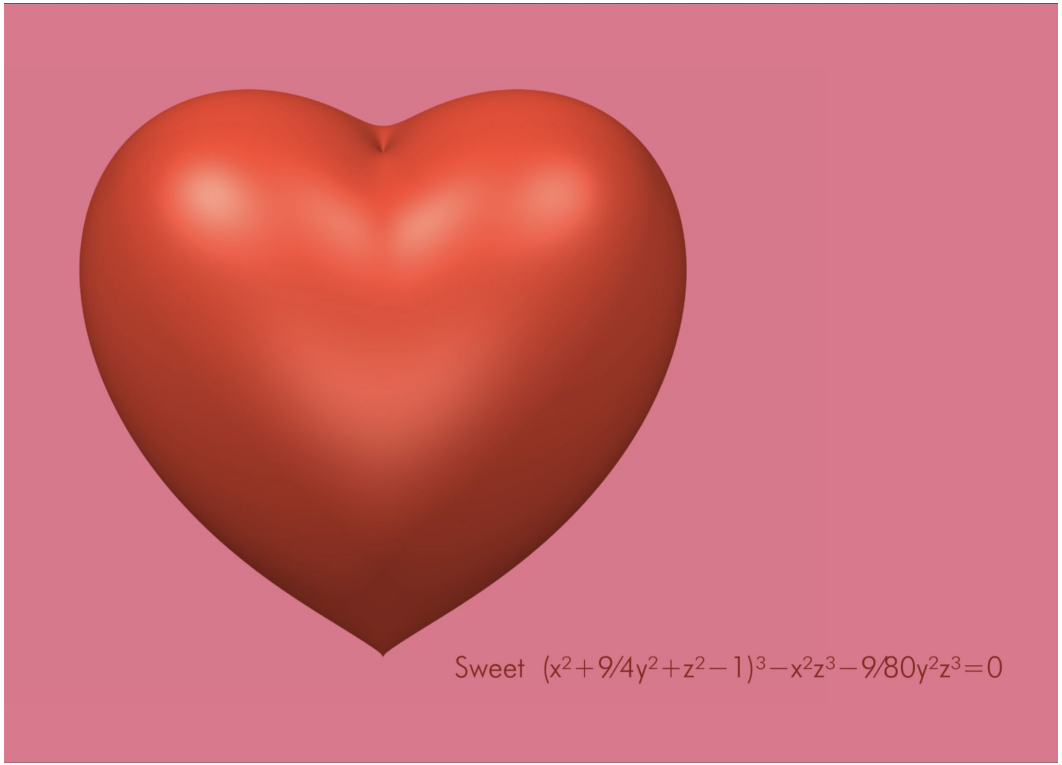
Design a Postcard

- a lesson learned
- something funny

Credit to Halya Pidstrigach

Design a Postcard

- a lesson learned
- something funny
- something beautiful



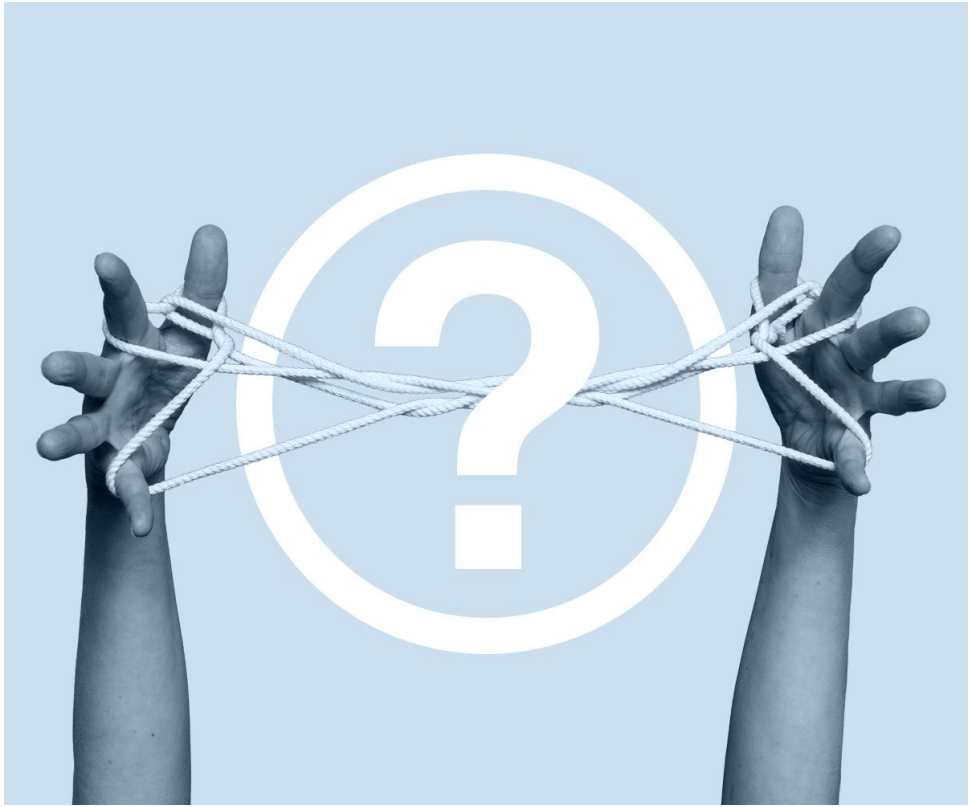
by IMAGINARY

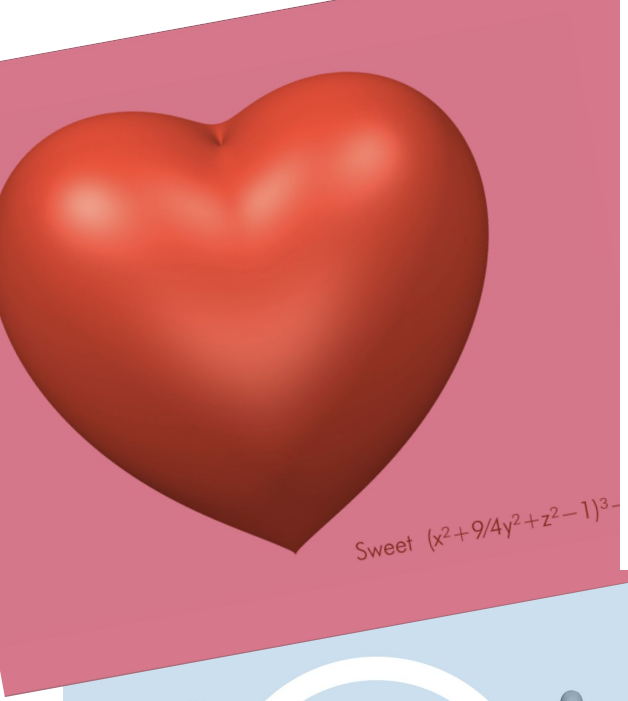
Design a Postcard

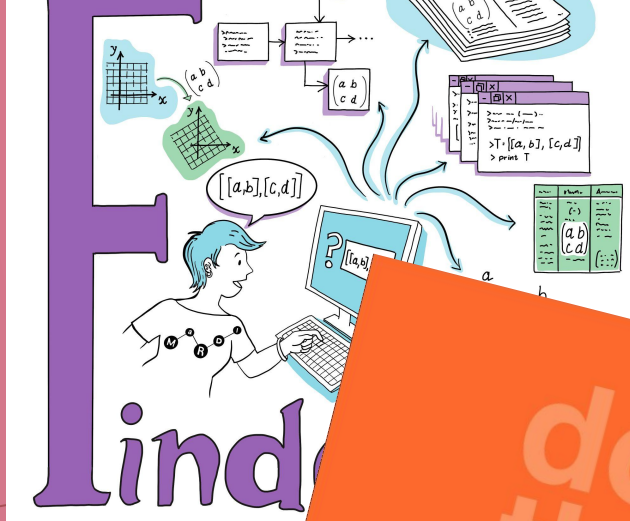
- a lesson learned
- something funny
- something beautiful
- a play on words

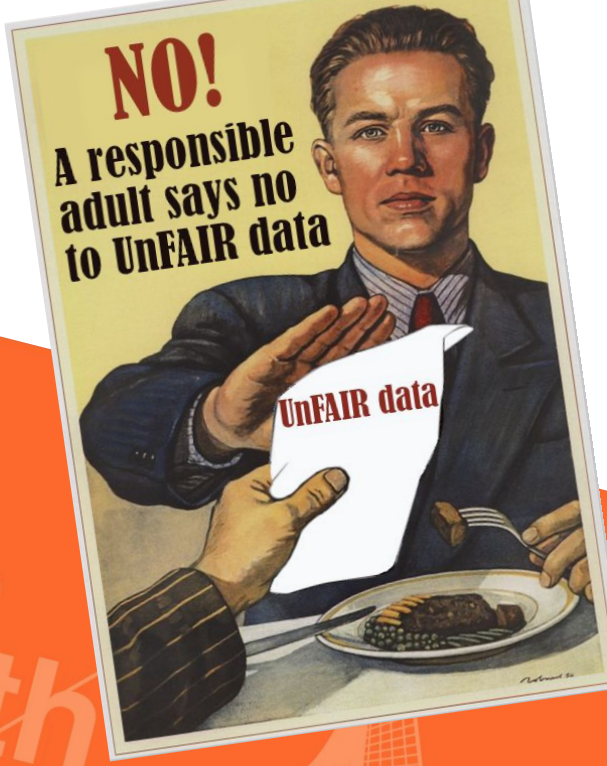
Design a Postcard

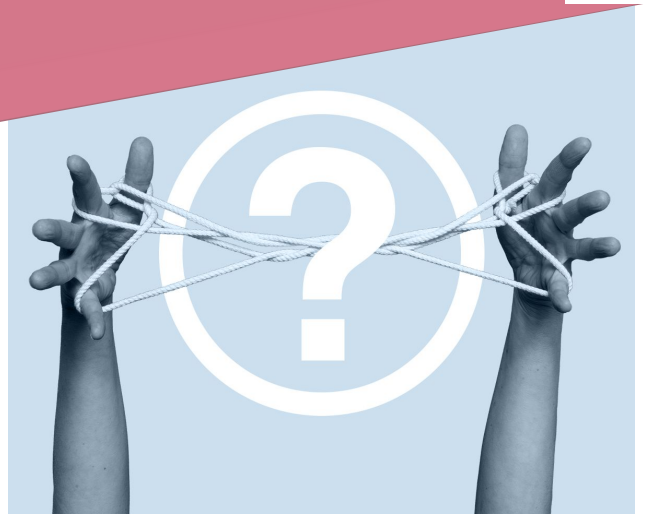
- a lesson learned
- something funny
- something beautiful
- a play on words
- a question that you have











do the math
FAIR your data!