Introduction to Transparent Practices for Quantitative Research

Chatchavan Wacharamanotham
University of Zurich
chat@ifi.uzh.ch

10 June 2025

This tutorial is designed based on the open materials of the courses presented at CHI 2022–23, VIS 2023, and MuC 2023 by Chat Wacharamanotham, Fumeng Yang, Abhraneel Sarma, Xiaoying Pu, and Lace Padilla. https://osf.io/27r5z

Chatchavan Wacharamanotham

Lecturer at the Department of Informatics, University of Zurich, Switzerland

Previously: Assistant professor at University of Zurich; Lecturer at Swansea University (UK); PhD in Human–Computer interaction from RWTH Aachen University

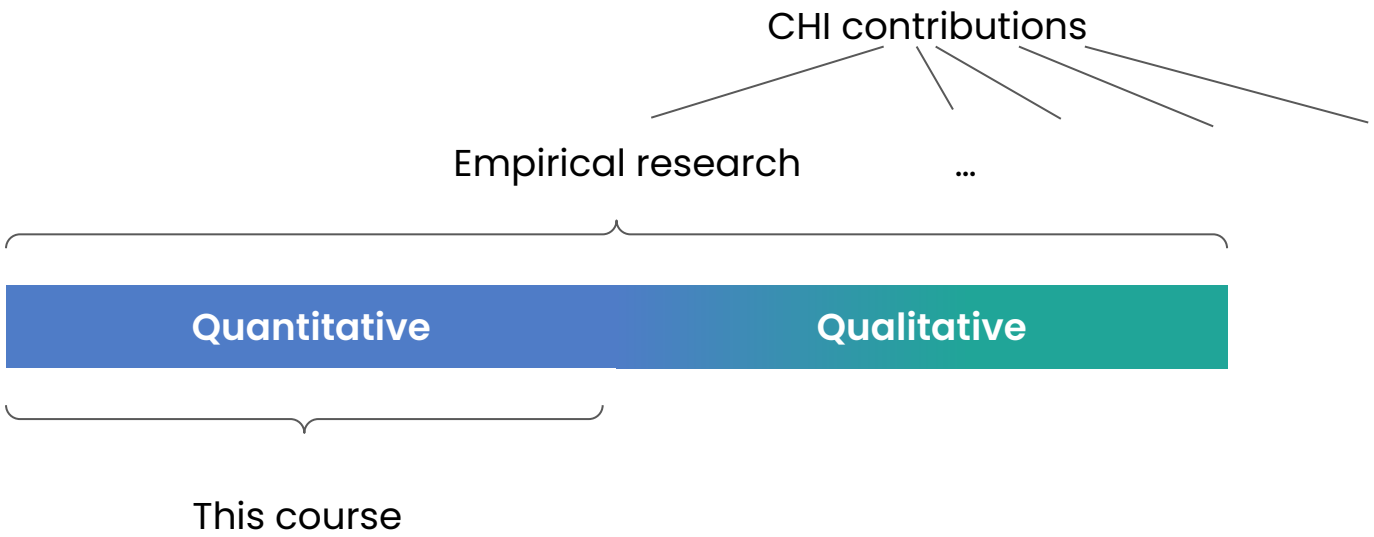
Research: Improving how computer can help people do better and transparent science

Past research: Interaction techniques for touch input on and above screens

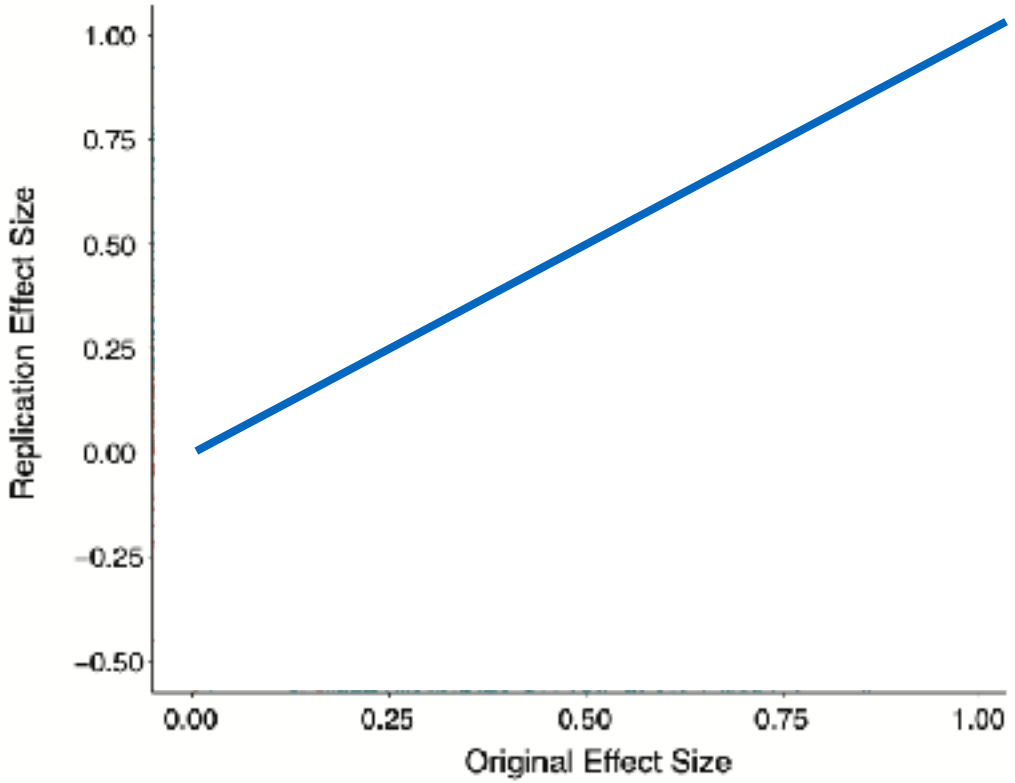
Roles in the CHI conferences: Associate Chair (2022–23) • Best Paper Award Committee (2022) • Student Research Competition Co-chair (2023) • Associate editor of IJHCS (International Journal of Human-Computer Studies) • Organizer of JoVI (The Journal of Visualization and Interaction)

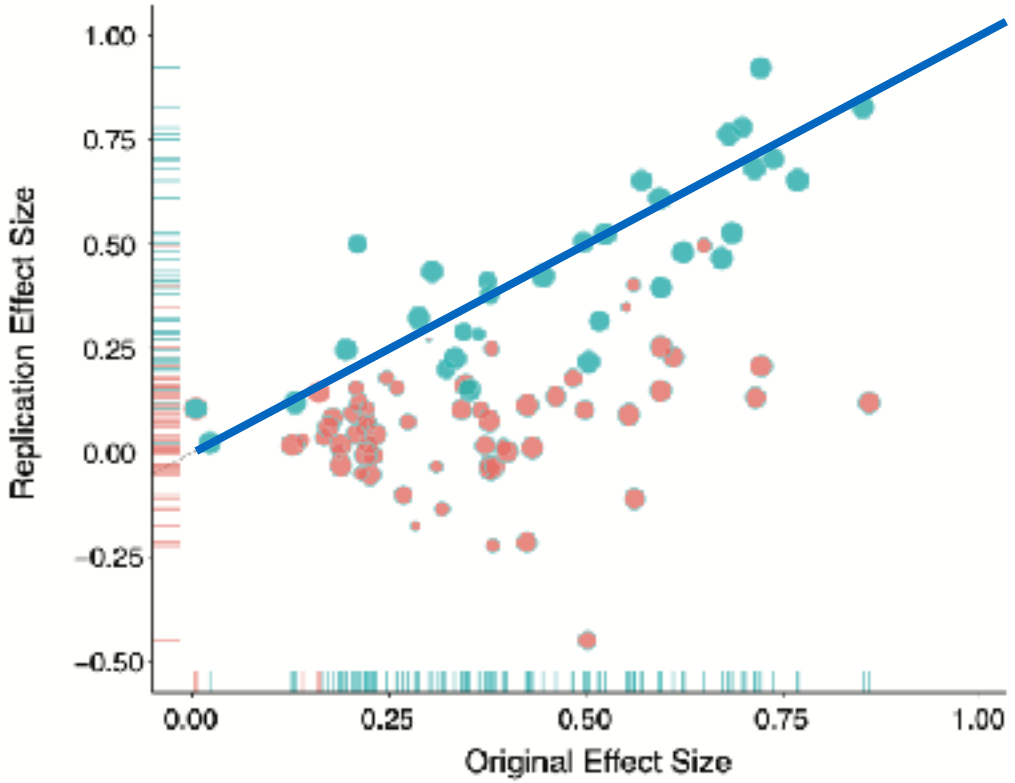
https://chatw.ch

Scope

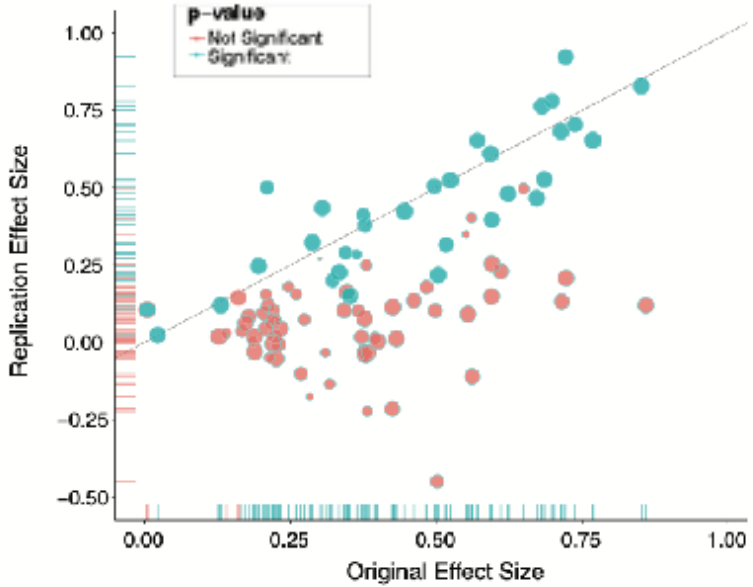

CHI contributions
Empirical research
...
Quantitative
Qualitative
This course

Why should we care about research transparency?


1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
Replication Effect Size
0.00
0.25
0.50
0.75
1.00
Original Effect Size


Replication Effect Size
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
0.00
0.25
0.50
0.75
1.00
Original Effect Size

Over half of psychology studies fail to replicate



100 psychology studies

Smaller effect sizes in 83% of the replication studies

Statistically significant results:

- 97% of original studies
- 36% in the replications

Source: Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science, 349(6251).

Replicability

Closely matched method

+

New data

=

Consistent results

Reproducibility

Same data analysis method

+

Same data

=

Same results

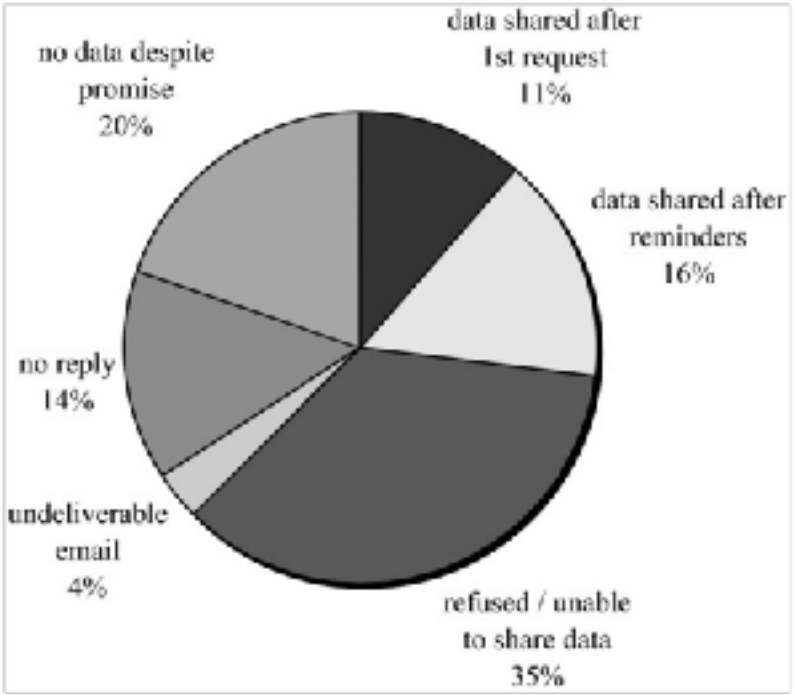
Reproducibility is a lower bar, but still important for evaluating the claims of research results

According to Clearbout terminology, which differs from ACM's terminology. See discussion in Plesser HE (2018) Reproducibility vs. Replicability: A Brief History of a Confused Terminology. Front. Neuroinform. 11:76.

“If researchers want to use data or code from my paper, they can contact me”

A team of psychology researchers requested data from the authors of 141 articles published in prestigious psychology journals in the previous year.

27% success rate

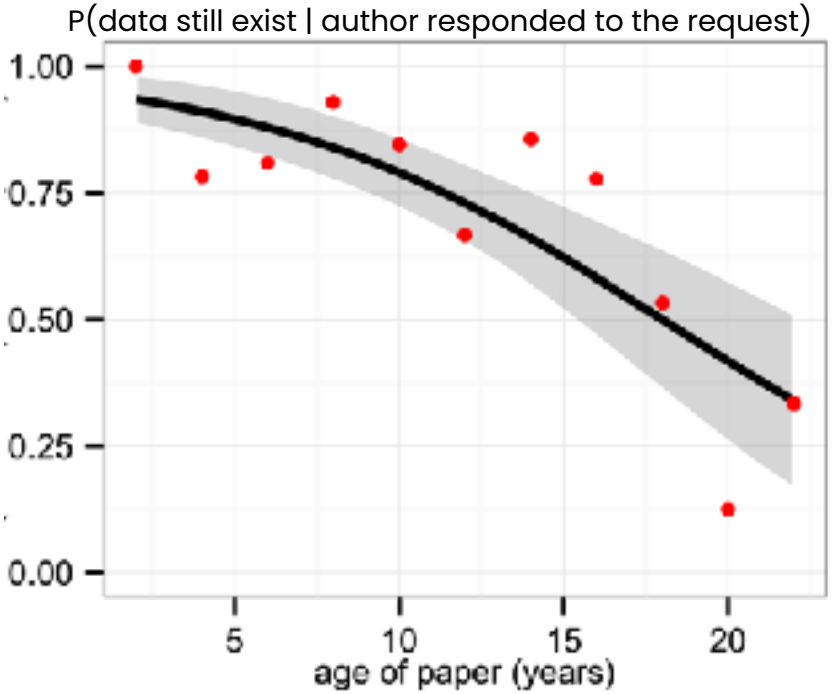


Wicherts et al. The poor availability of psychological research data for reanalysis. American Psychologist 61, 7 (2006), 726 – 728.

“If researchers want to use data or code from my paper, they can contact me”

Researchers from the field of biology requested data from 516 articles published between 2–20 years

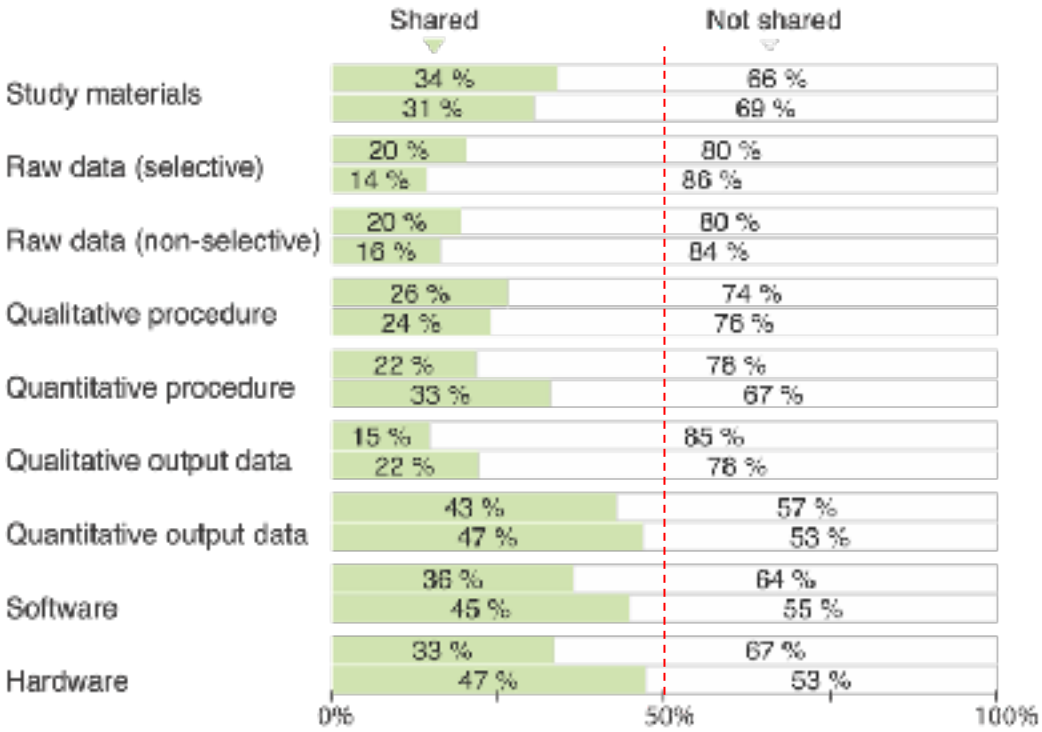
The odds of data still exist fall 17% per year



Vines et al. The Availability of Research Data Declines Rapidly with Article Age. Current Biology 24, 1 (2019), 94–97.

Survey to authors of

CHI 2018, 2019



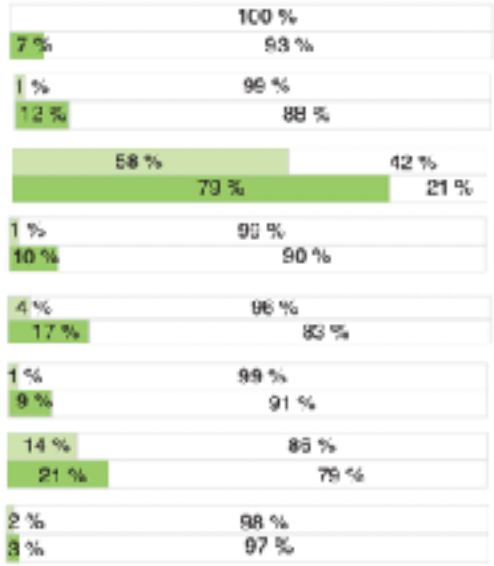
Percents of respondents

Wacharamanotham et al. (2020) Transparency of CHI Research Artifacts: Results of a Self-Reported Survey. In Proc. of CHI 2020.

Content analysis of papers from

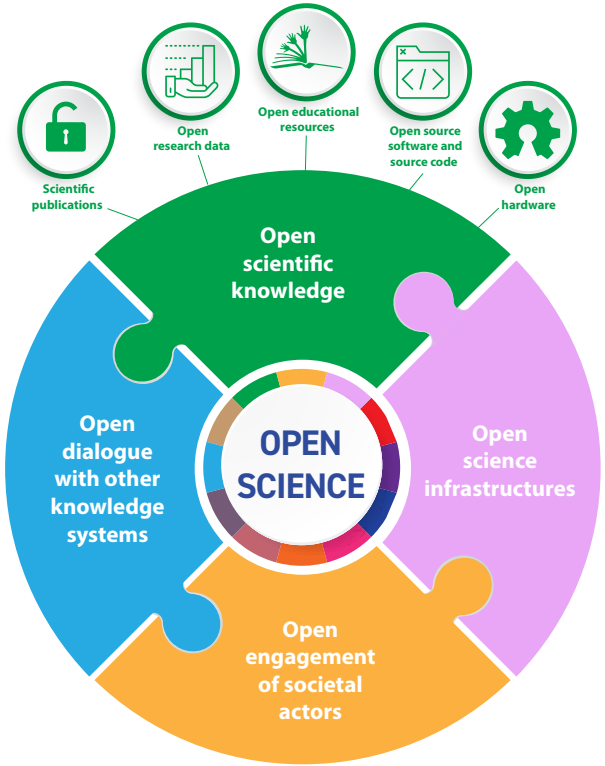
CHI 2017, 2022

Mapped to equivalent categories on the left



Percents of sampled papers

Niksirat et al. (2023) Changes in Research Ethics, Openness, and Transparency in Empirical Studies between CHI 2017 and CHI 2022. In Proc. of CHI 2023.



193 member states of the UNESCO promote Open Science

"[Open scientific knowledge] also refers to the possibility of **opening research methodologies and evaluation processes.**"

Guide to a Successful Submission

Transparency

Research transparency is of utmost importance in a CHI paper. It allows reviewers to understand and assess submitted work thoroughly, and it allows members of the research community to understand, analyze, and build upon the work in published CHI papers. As such transparency is taken into account very seriously in the review process.

CHI papers should strive for research transparency regardless of the contribution type and methodology. Different contribution types, (e.g. technical contributions, quantitative studies, and qualitative studies) use different criteria for assessing transparency.

Contributions that are technology-oriented (e.g., a new technique or algorithm) and **contributions that are quantitative studies** (i.e., experiments with statistically analyzed results) are expected to be verifiable, reproducible (e.g., others should be

Source: CHI 2025 Guide to a Successful Submission

Conceptualizing research transparency for HCI

"**Research transparency** refers to honesty and clarity in all communications about the research processes and outcomes to the extent possible."

- "honesty and clarity" — sometimes have trade-off
- "all communications" — among researchers and beyond
- "process and outcomes" — emphasis may differ across research methods
- "to the extent possible" — weigh transparency with ethics, privacy, intellectual properties, and other values

This preprint gives ideas on how to be transparent in many types of HCI research (also beyond quantitative)

From a working paper by Yvonne Jansen, Jan B. Vornhagen, Olga Iarygina, Kavous Salehzadeh Niksirat, Lonni Besançon, Pierre Dragicevic, Julien Gori, and Chat Wacharamanotham. "The Many Ways of Being Transparent in Human-Computer Interaction Research" https://osf.io/2wze6_v1

Plan

Transparency in **planning** studies

With a focus on experimental studies

Transparency in **data analysis**

General concerns + exercises in preregistration

Transparency in **reporting**

Examples in frequentist statistics + pointers

Transparency in **visualizing research data**

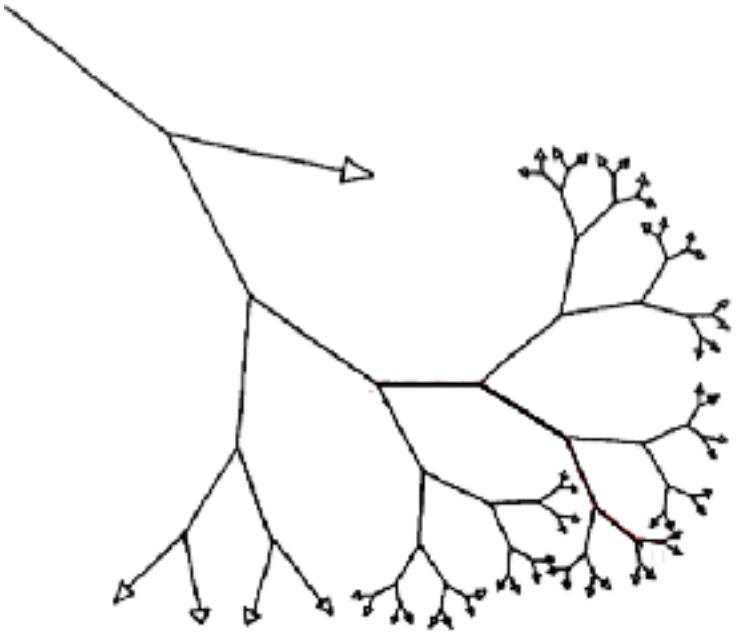
Principles + examples

Transparency in research **materials**

What, how, and where to share

Transparency in planning studies

Choices in research



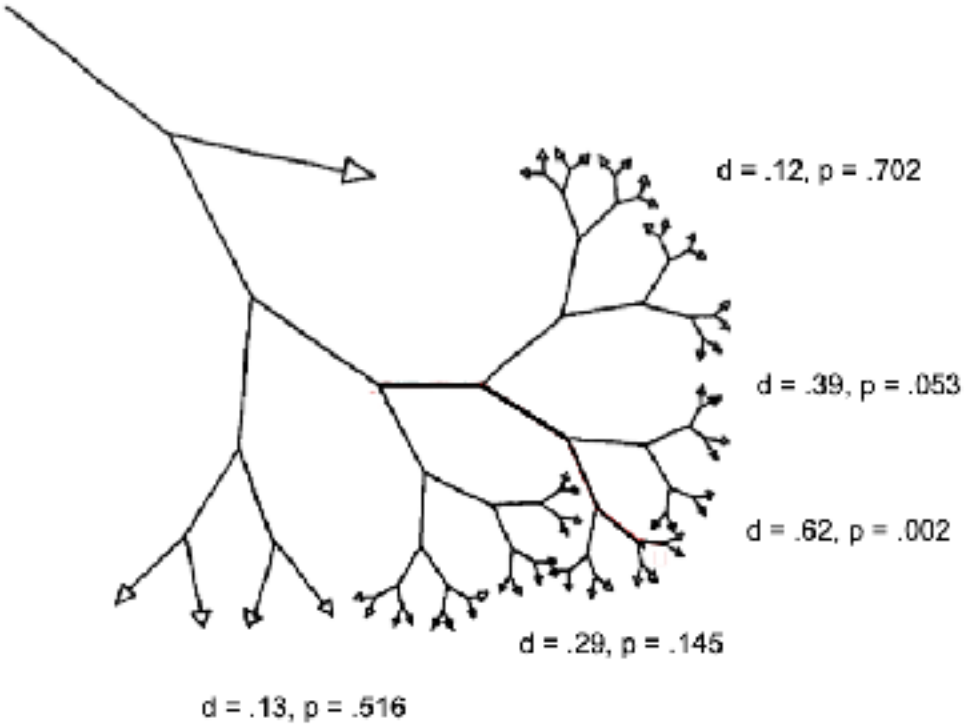
Some choices in data analysis

- Choosing between different options of dealing with incomplete or missing data on ad hoc grounds
- Specifying pre-processing of data (e.g., cleaning, normalization, smoothing, motion correction) in an ad hoc manner
- Deciding how to deal with violations of statistical assumptions in an ad hoc manner

Diagram from Marjan Bakker's slide

Choices from **Degrees of freedom in planning, running, analyzing, and reporting psychological studies:** A checklist to avoid p-hacking (Wicherts et al., 2016)

Choices in research



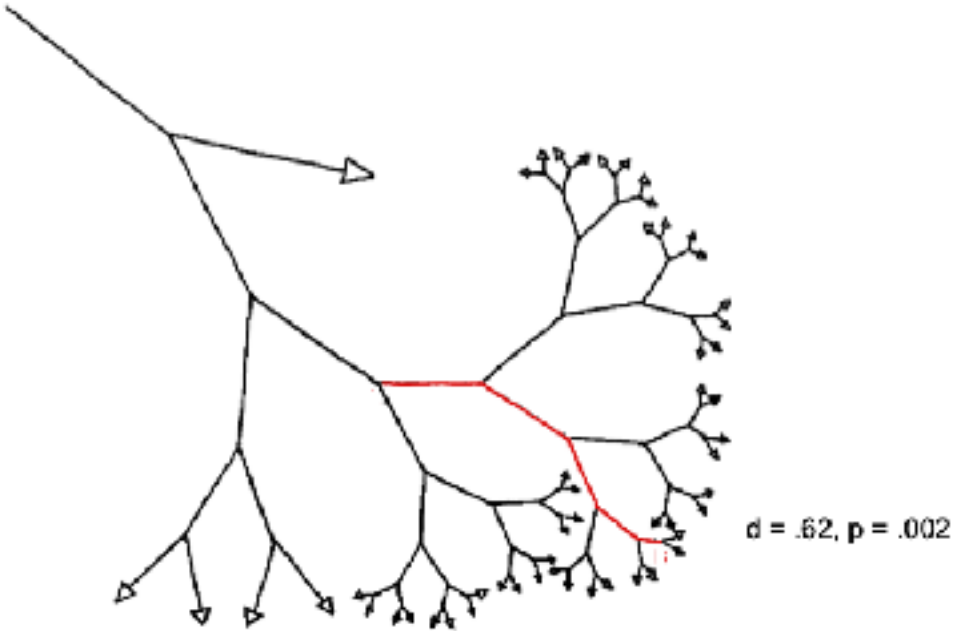
Some choices in data analysis

- Choosing between different options of dealing with incomplete or missing data on ad hoc grounds
- Specifying pre-processing of data (e.g., cleaning, normalization, smoothing, motion correction) in an ad hoc manner
- Deciding how to deal with violations of statistical assumptions in an ad hoc manner

Diagram from Marjan Bakker's slide

Choices from **Degrees of freedom in planning, running, analyzing, and reporting psychological studies:** A checklist to avoid p-hacking (Wicherts et al., 2016)

Choices in the **design phase**



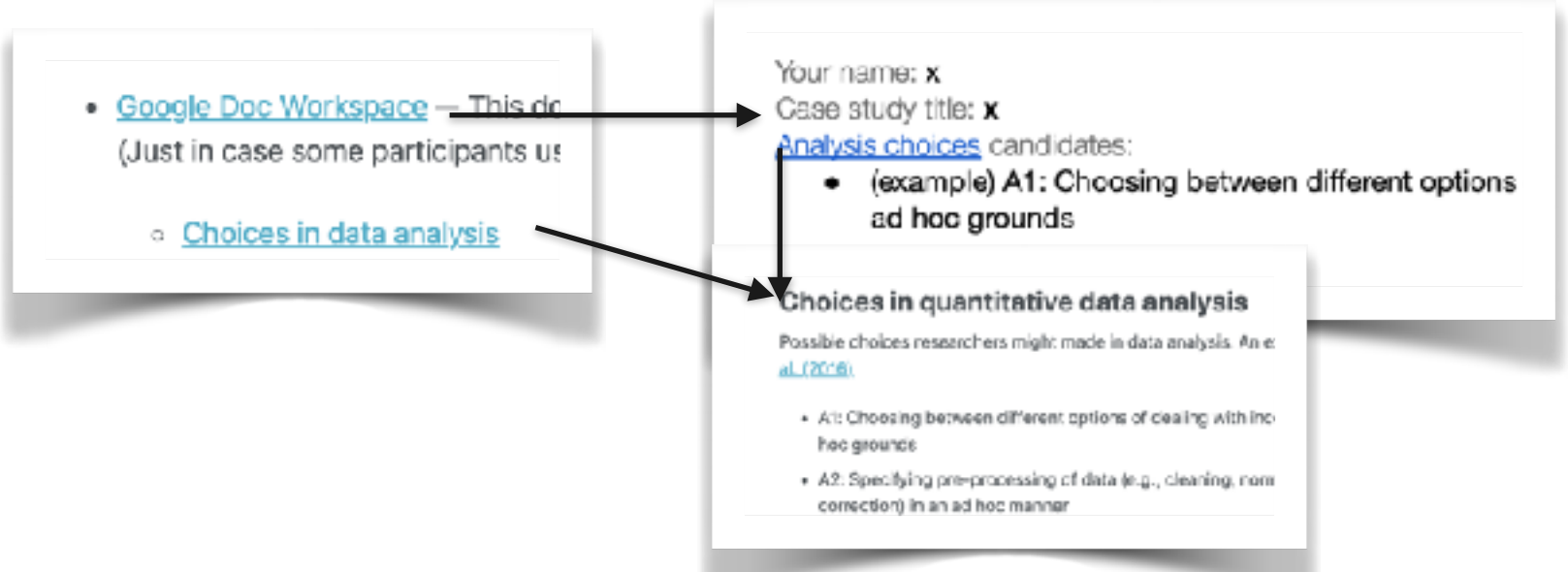
Establishing transparency in decisions made at the research design phase

- Make justifiable choices
- Report the choices made
- Discuss implications of the paths not taken

Diagram from Marjan Bakker's slide

Exercise 1: Choices in data analysis (10 minutes)

1. Go to the talk page https://chatw.ch/transparency-4-quant
2. Click on the link to Google Doc workspace, and grab a space on the template
3. Look through the list of analysis choice and choose 2–3 choices from your case study
4. Briefly describe these choices. (We will use them in a breakout room discussion later.)

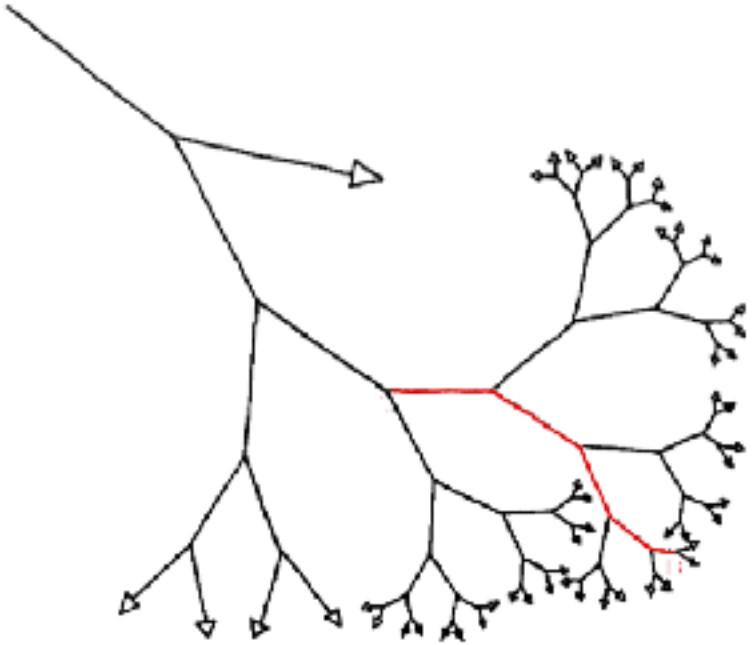


Exercise 2: Discuss choices and changes (20 minutes)

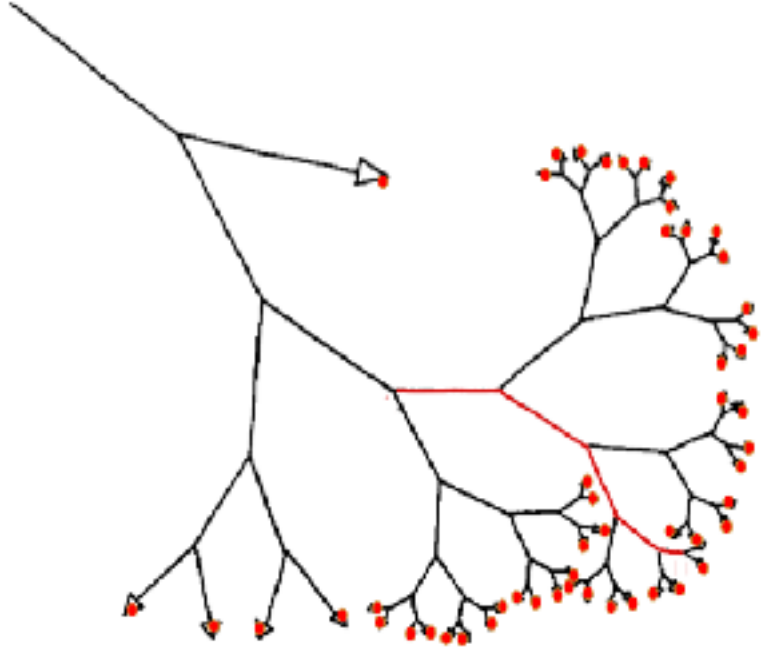
1. In your breakout room, take turn to describe the case study and the analysis choices (max. 3 minutes per person)
2. Choose one analysis choice from a case to work together.
3. Discuss:
    - When might this decision be made?
    - When might the researchers change this decision?
    - Which factor(s) might have led to this change?

Transparency in data analysis

Ways to be accountable for data analysis choices



Declare your choices in advance (preregistration)



Explore how different choices affects the results (sensitivity analysis, multiverse analysis)

Diagrams from Marjan Bakker's slide

Preregistration

A timestamped record of the plan, including information such as

- A brief narrative description of reason to conduct the research
- Explicitly state the intended purpose (exploratory or confirmatory)
- Hypothesis and prediction of the outcome
- Expected analysis method (ideally a script for data analysis)

Evidence of what and when you planned → increase transparency and credibility

Templates and recipes for preregistering various types of studies are listed on [OSF's Registration Forms and Templates page]

AsPredicted **preregistration template**

1. **Have any data been collected** for this study already?
2. What's the **main question** being asked or **hypothesis** being tested in this study?
3. Describe the key **dependent variable(s)** specifying how they will be measured.
4. How many and which **conditions** will participants be assigned to?
5. Specify exactly **which analyses** you will conduct to examine the main question/ hypothesis.
6. Describe exactly how **outliers** will be defined and handled, and your precise rule(s) for excluding observations.
7. How many observations will be collected or what will determine **sample size**? No need to justify decision, but be precise about exactly how the number will be determined.
8. **Anything else** you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Exercise 3: Drafting a preregistration (20 minutes)

Continue with the case you previously chose.

1. Find the relevant preregistration section from the OSF template
2. Pair-write the preregistration text together
    - One person write
    - Another person help thinking and discussing and take notes of findings from this drafting process. Prepare them as input to the plenary

Preregistration

"Preregistration is a plan, not a prison" [1]

Circumstances that unfold after filing a preregistration may necessitate adjustment

- If you haven't seen the data, file a new preregistration with explicit reference to the previous plan
- Explain the reasons for deviation in the paper

You may add further exploratory analyses as long as they are clearly separated from the preregistered analyses in the paper

Use pilot studies to inform your decisions

Preregistration: A Plan, Not a Prison (DeHaven, 2017 on COS blog)

Preregistration

Critique: *"But most studies in HCI are iterative and exploratory"*

- Preregister the exploratory intention and initial hypotheses
- **Benefit:** Reviewers cannot challenge that the exploratory analyses comes from failing to achieve statistical significance from other tests [1]

[1] **HARK No More: On the Preregistration of CHI Experiments** (Cockburn et al., CHI 2018). The arguments from HCI researchers' perspective makes this paper worth reading as a whole.

For CHI double-blind reviewing process, see instruction for sharing anonymized preregistration in section 3 of Open Practices in Visualization Research (Haroz, 2018, BELIV position paper)

Transparency in reporting

Transparent statistics guiding principles

1. **Faithfulness:** Strive to capture and convey the "truth" as accurately as possible, especially concerning the uncertainty within the data.
2. **Robustness:** Prefer data analysis and reporting strategies that are robust to departures from statistical assumptions—or that make few assumptions
3. **Resilience:** Data analysis and reporting strategies should yield similar outcomes across hypothetical replications of the same study.
4. **Process transparency:** Communicate the decisions made during the analysis and report writing as explicitly as possible.
5. **Clarity:** Study reports should be easy to process—even when they target experts.
6. **Simplicity:** Prefer the simplest procedure even if it is slightly inferior in other respects.
7. **Non-contingency:** Outside exploratory analyses, data analysis and reporting strategies should avoid decisions that are contingent on data
8. **Precision and economy:** Plan for data quality, high statistical power, and high statistical precision
9. **Material availability:** Sharing as much study material as possible

Transparent Statistics Guiding Principles. (Transparent Statistics in Human–Computer Interaction Working Group, 2019)

2. Robustness: Prefer data analysis and reporting strategies that are robust to departures from statistical assumptions—or that make few assumptions

Some people tend to avoid extreme answers, the difference between the rating 5 and 6 may be smaller than those of 8 and 9.

A. **Parametric test** (e.g., *t*-test or ANOVA)
B. **Nonparametric tests** (e.g., Wilcoxon tests, or Mann-Whitney U test)

4. Process transparency: Communicate the decisions made during the analysis and report writing as explicitly as possible.

A. The difference is not statistically significant (p = 0.5)
B. The **Wilcoxon test** is not statistically significant (**W = 1762**, p = 0.5)
C. The **Wilcoxon rank sum test** is not statistically significant (W = 1762, p = 0.5)

5.Clarity: Study reports should be easy to process—even when they target experts.

A comparison of a novel physical user interface prototyping system (technique B) to the previous state of the art (A)

A. The feedback time differs by 104 ms (95% CI: [81, 126])
B. Technique B has lower feedback time than A by 104 ms (95% CI: [81, 126])
C. […] Technique B's feedback time tend to be **less than the threshold of human perception (less than about 100ms).**
D. Technique B has lower feedback time with **Cohen's *d* =0.2**

Checklist for reporting statistics: The SAMPL Guidelines

(Lang & Altman, 2016)

General Principles for Reporting Statistical Results

Reporting numbers and descriptive statistics

- Report numbers—especially measurements—with an appropriate degree of precision. For ease of comprehension and simplicity, round as much as is reasonable. For example, mean age can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is 5 points, scores can be reported as whole numbers; decimals are not necessary.

- Report total sample and group sizes for each analysis.

- Report numerators and denominators for all percentages.

- Summarize data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.

- Summarize data that are not normally distributed with medians and interpercentile ranges, ranges, or both. Report the upper and lower boundaries of interpercentile ranges and the minimum and maximum values of ranges, not just the size of the range

- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead.

- Display the data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.[42,43]

Checklist for reporting statistics: The SAMPL Guidelines

(Lang & Altman, 2016)

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one- or two-tailed and for paired or independent samples.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes, such as differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines, report a measure of precision, such as the 95% confidence interval.
- Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. The SE is essentially a 68% confidence coefficient; use the 95% confidence coefficient instead.
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities when possible and to one or two decimal places (e.g., $P = 0.03$ or 0.22 not as inequalities: e.g., $P < 0.05$). Do NOT report "NS"; give the actual P value. The smallest *P* value that need be reported is $P < 0.001$, save in studies of genetic associations.
- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting null-hypothesis significance tests

Choice of the test must match statistical assumptions

Degrees of freedom can rescue your paper

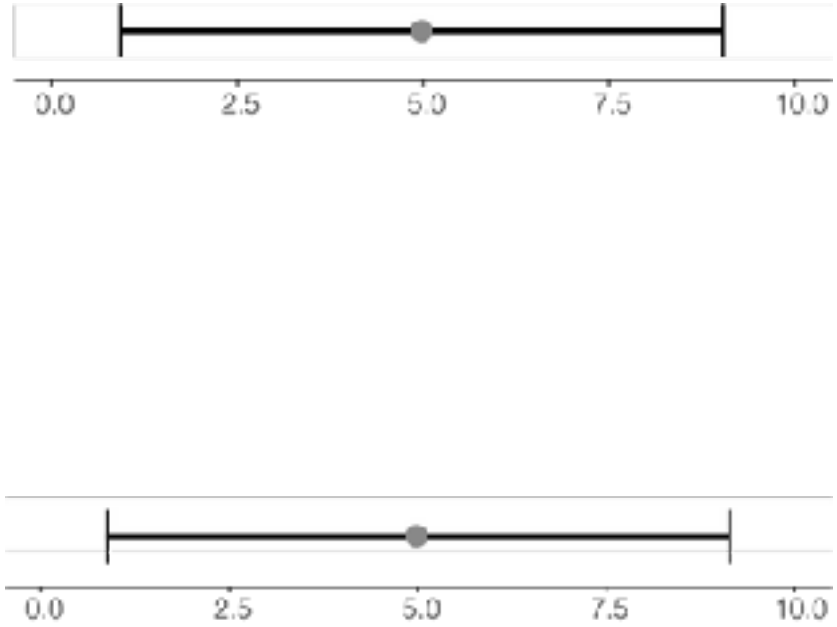
statcheck.io: Check consistency between the *p*-value and parameters (e.g., *t*, *F*, and their degrees of freedom)

The prevalence of statistical reporting errors in psychology (1985–2013) (Nuijten et al., 2016)

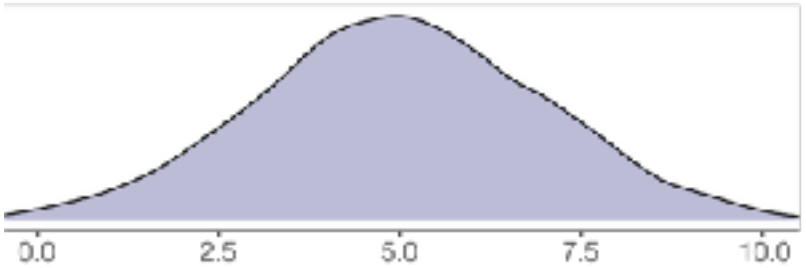
Statcheck tutorial (Nuijten & Polanin, 2020)

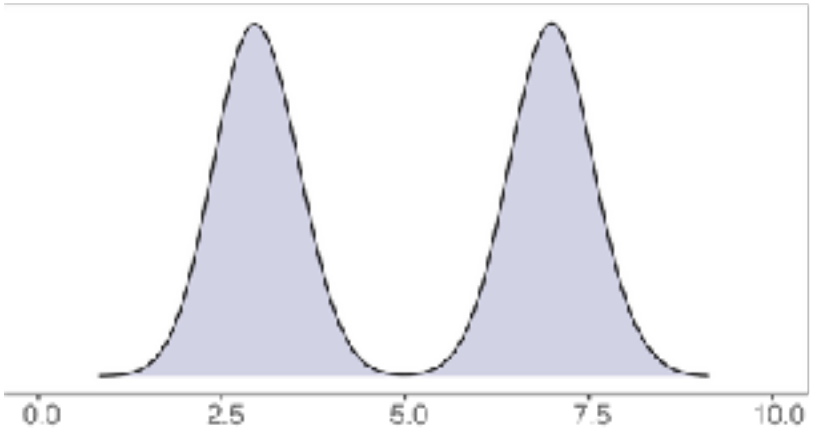
Transparency through visualizing research data

What can you say about these two 95% confidence intervals?



Summaries can **obscure** important relationships in distributional data





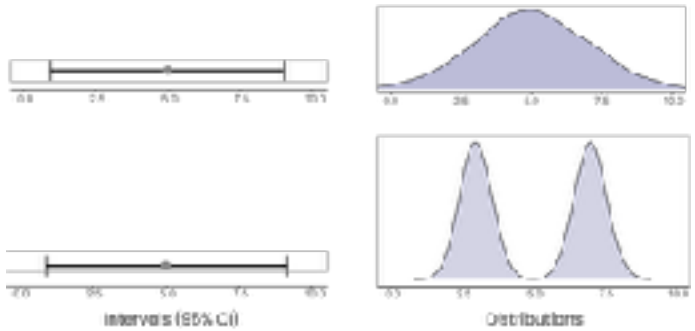
Intervals (95% CI)

Distributions

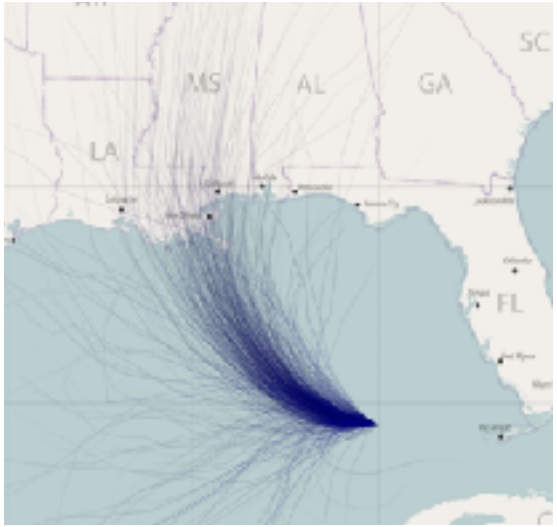
Visualizing uncertainty in the results

Expressiveness principle: the visual representation should represent ***all*** and ***only*** the relationships that exist in the data[1,2]

Expressiveness is a proxy to transparency



[1] Mackinlay, J. (1986). Automating the design of graphical presentations of relational information.
[2] Munzner, T. (2014). Visualization analysis and design. CRC press.



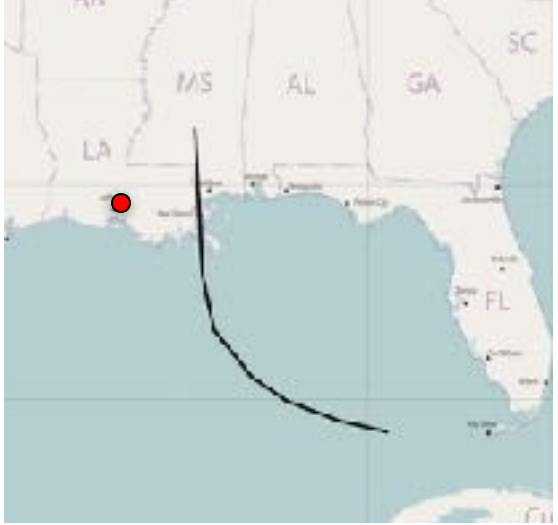
expressiveness →

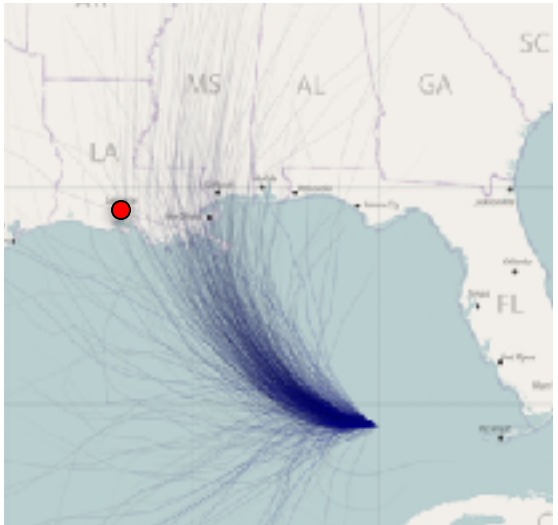
Visualizations of data can vary on a spectrum of expressiveness
Choices of visualization is an aspect of research transparency

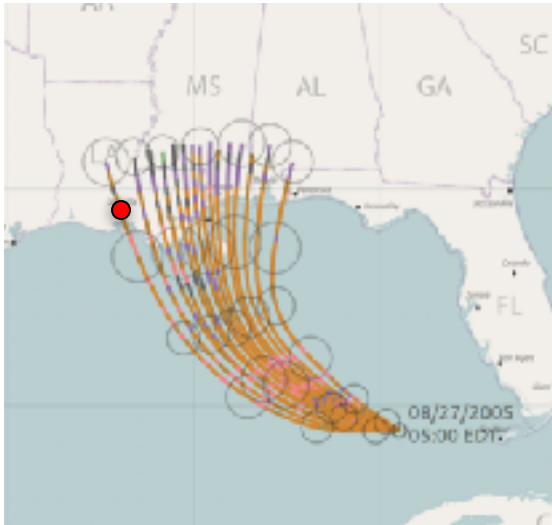
Liu et al., 2018 (IEEE VIS)

Would you stay or evacuate?

Usable visualizations support users in making accurate inferences







Liu et al., 2018 (IEEE VIS)

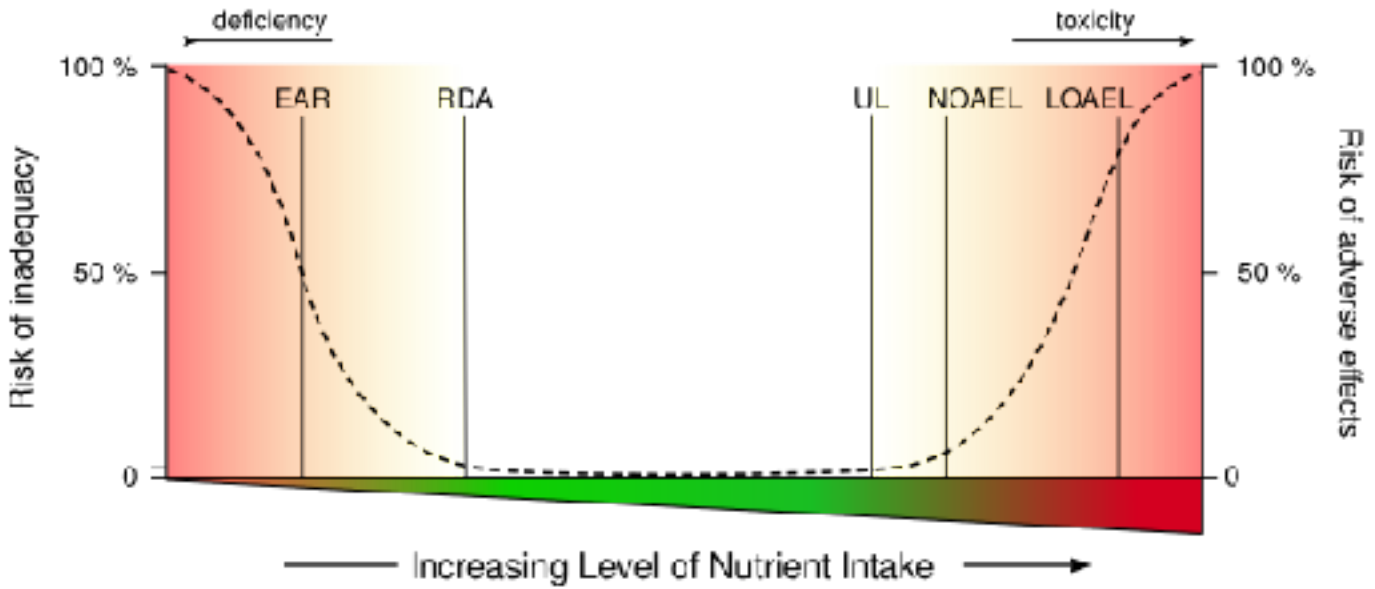
Padilla et al., 2018 (CRPI)

Dietary Reference Intakes (DRIs): Recommended Dietary Allowances and A

Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Vitamin A (µg/d)[a]	Vitamin C (mg/d)	Vitamin D (µg/d)[b,c]	Vitamin E (mg/d)[d]	Vitamin K (µg/d)	Thiamin (mg/d)	Ribo (mg/
Infants							
0–6 mo	400*	40*	10*	4*	2.0*	0.2*	
6–12 mo	500*	50*	10*	5*	2.5*	0.3*	
Children							
1–3 y	**300**	**15**	**15**	**6**	30*	**0.5**	
4–8 y	**400**	**25**	**15**	**7**	55*	**0.6**	
Males							
9–13 y	**600**	**45**	**15**	**11**	60*	**0.9**	
14–18 y	**900**	**75**	**15**	**15**	75*	**1.2**	

Dietary reference intake (Food and Nutrition Board, Institute of Medicine, National Academies)



Dietary reference intake (Julius Senegal)

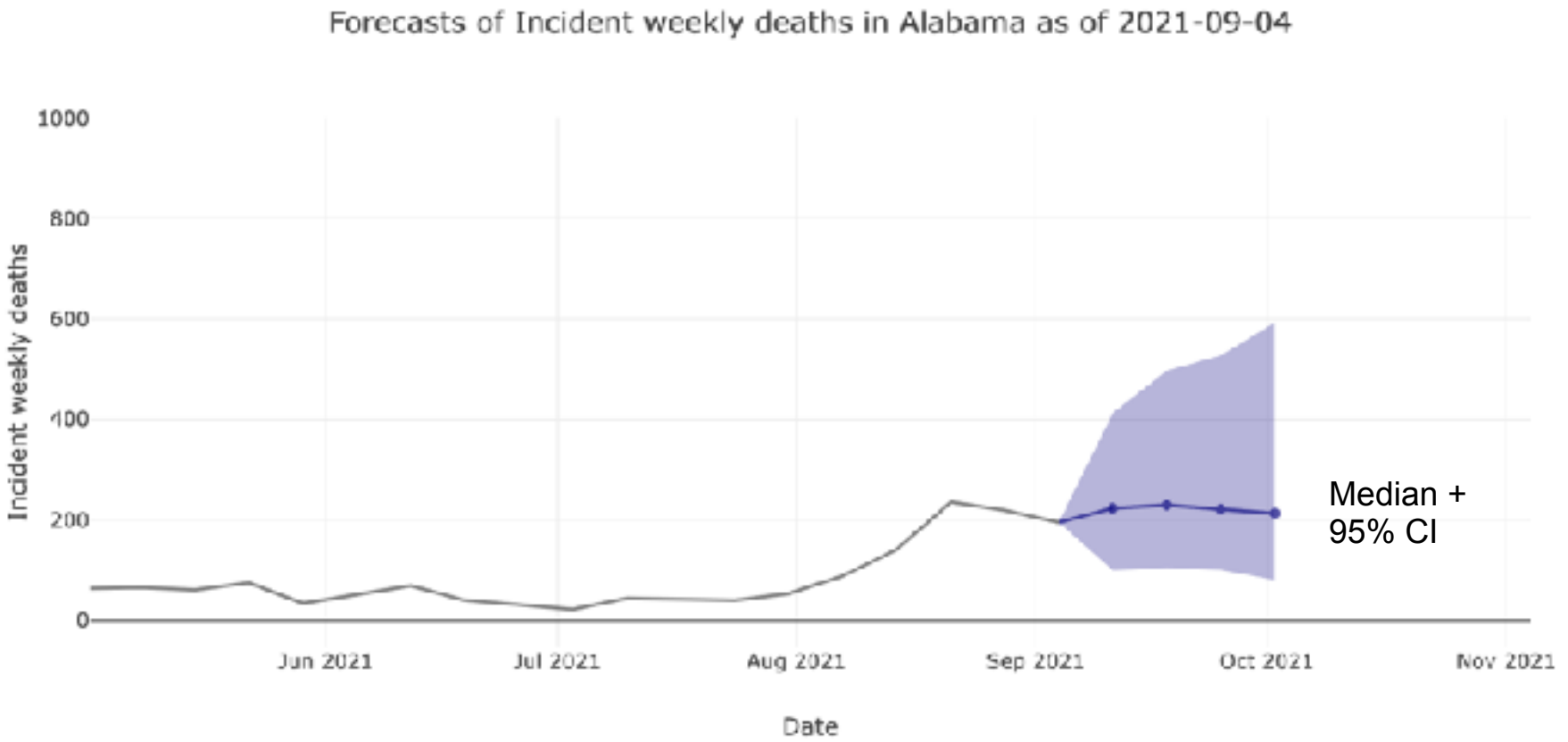
Uncertainty matters

*Without uncertainty, viewers may come to **incorrect** conclusions about the data.*

Showing uncertainty:

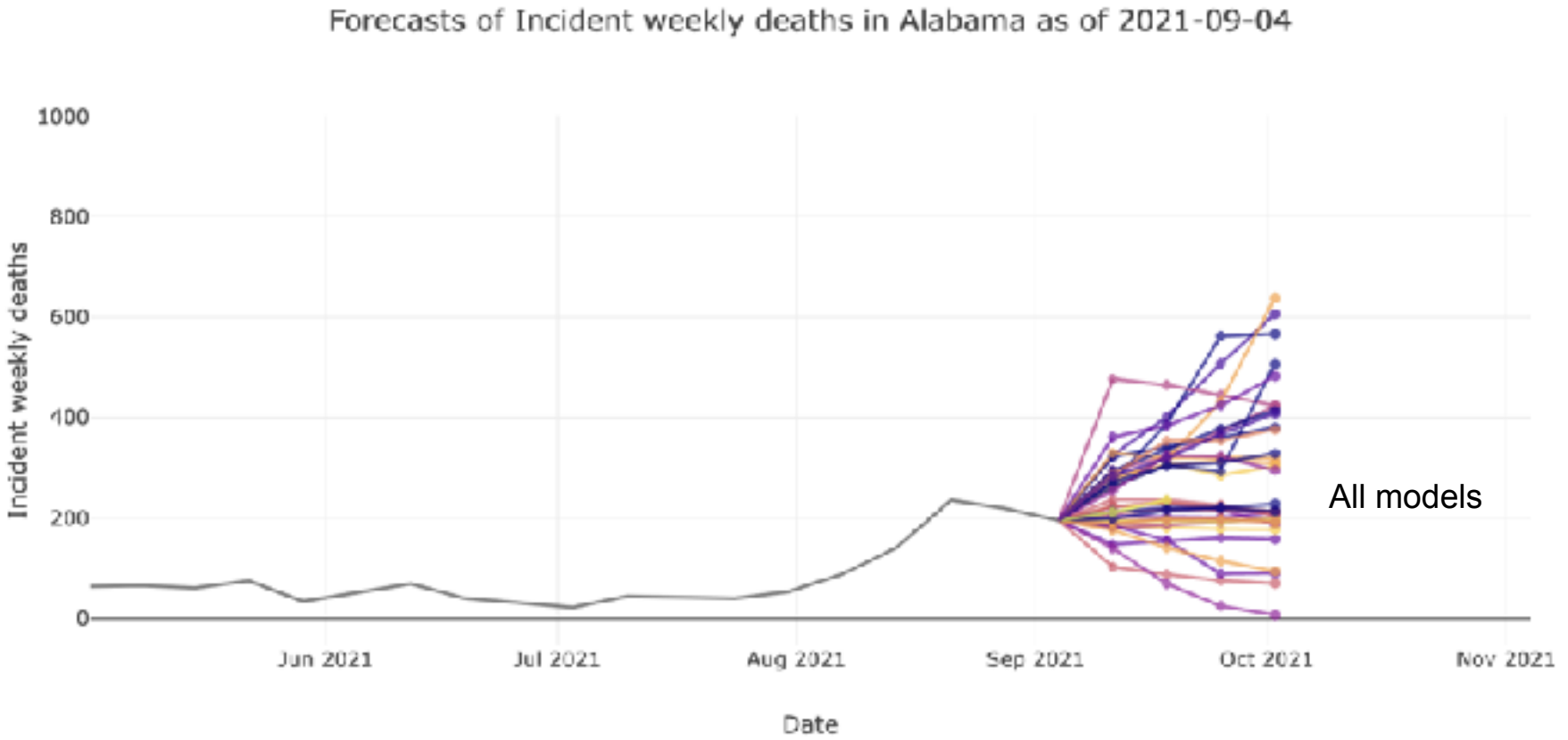
- Increases scientific credibility
- Increases trust
- Let them tune their expectations and assumptions correctly

Usable visualizations support users in making accurate inferences
Showing uncertainty contributes to usability



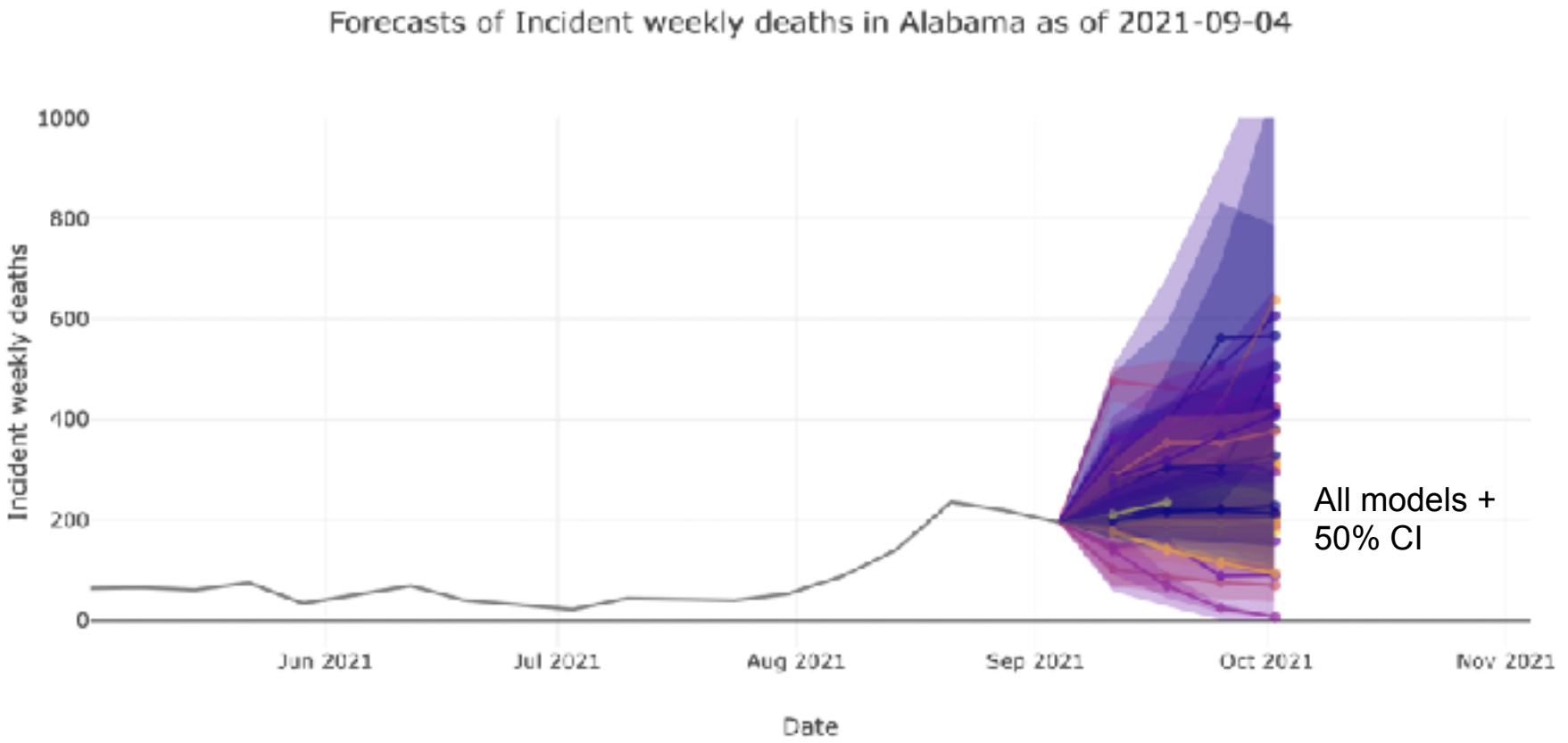
expressiveness

usability



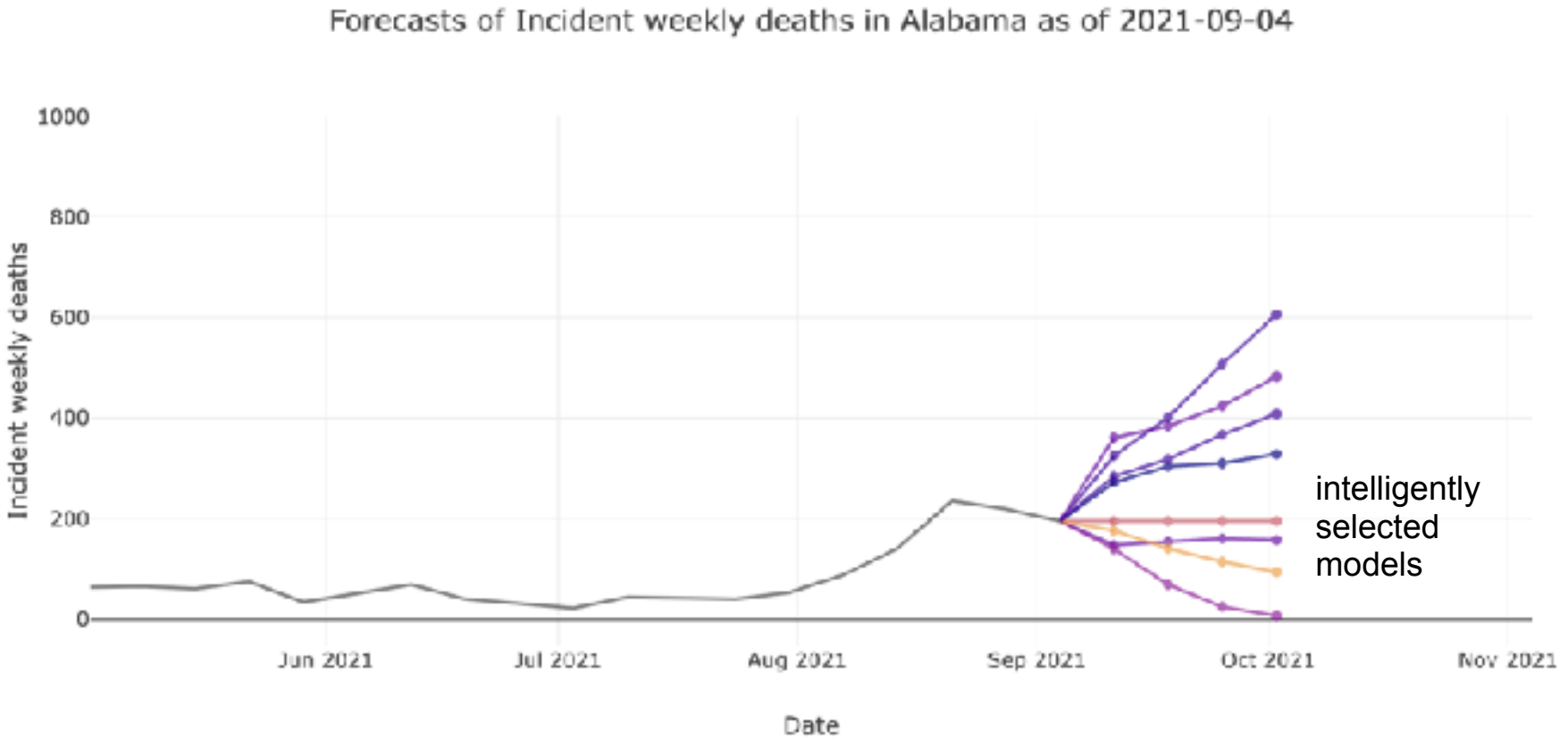
expressiveness

usability



expressiveness

usability



expressiveness

usability

Research transparency through visualization

Balancing tradeoffs between:

Expressiveness: Faithfully represent the data, and

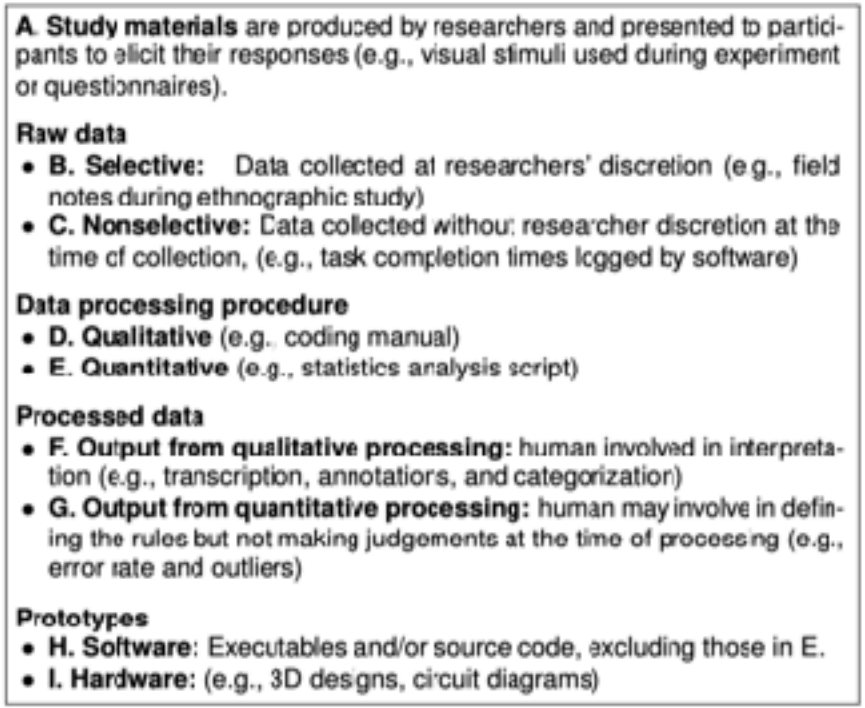
Usability: Support users in making accurate inferences from the data

No single best answer

Consider **context, data set, and audience** when making these decisions

Transparency in research materials

What to share?



A. Study materials are produced by researchers and presented to participants to elicit their responses (e.g., visual stimuli used during experiment or questionnaires).

Raw data

- **B. Selective:** Data collected at researchers' discretion (e.g., field notes during ethnographic study)
- **C. Nonselective:** Data collected without researcher discretion at the time of collection, (e.g., task completion times logged by software)

Data processing procedure

- **D. Qualitative** (e.g., coding manual)
- **E. Quantitative** (e.g., statistics analysis script)

Processed data

- **F. Output from qualitative processing:** human involved in interpretation (e.g., transcription, annotations, and categorization)
- **G. Output from quantitative processing:** human may involve in defining the rules but not making judgements at the time of processing (e.g., error rate and outliers)

Prototypes

- **H. Software**: Executables and/or source code, excluding those in E.
- **I. Hardware:** (e.g., 3D designs, circuit diagrams)

Ethical considerations:

- to study participants
- to taxpayers who fund your research

Consult your IRB.

Simple anonymization (rename participant ID and shuffle the order) sometimes suffice

If cannot share (e.g., research on company confidential data),

- Share aggregated statistics at the as close to raw as possible
- Describe what materials are generated and provide justification in the paper

For extensive discussion on materials, misunderstandings, and how to share, see:
Wacharamanotham et al., (2020). **Transparency of CHI Research Artifacts: Results of a Self-Reported Survey**. In Proc. of CHI 2020.

How to prepare materials for sharing?

Interoperable file formats, e.g., text csv, Excel Open XML (.xlsx)

- 👉 Guide on how to organize data in spreadsheet

A clear entry point: README.txt, README.md, or index.html

- 👉 Github repository template for organizing data

Data dictionary:

- Which file containing what data
- Column: name, readable description, unit of measurement, and range
- 👉 OSF guide on data dictionary

For detailed discussion on the whole research materials management process, see **Good enough practices in scientific computing** (Wilson et al., 2017)

A guide on data organization: A reproducible data analysis workflow with R Markdown, Git, Make, and Docker (Peikert & Brandmaier, 2019)

For ultimate reproducible research compendium based on R, check the rrtools package.

Examples supplemental material organization

Structure of this repository

- analyses
  - exp1.R
  - helper scripts
    - CI.helper.R
    - plotting functions.R
- data
  - exp1.csv
  - exp1_column-description.csv
  - exp2.csv
  - exp2_column-description.csv
    - raw data
      - exp2-complete-column-descriptor.csv
      - exp2-complete.csv
      - exp2-raw-BART-data-column-descriptor.csv
      - exp2-raw-BART-data.csv
- markdown
  - exp1.md - The complete analysis script for experiment 1

Jansen & Hornbæk (2018) https://github.com/yvonnejansen/posture

Examples supplemental material organization

column_id	data_type	range	description	exact_question
Timestamp	time		timestamp when the participant complete the experiment	
in_charge	integer	[1,7]	self-reported feeling in charge measure	To what extend do you feel in charge?
power	integer	[1,7]	self-reported sense of power	How powerful do you feel?
fatigue	integer	[1,7]	self-reported fatigue	Did you find this task fatiguing?
difficult	integer	[1,7]	self-reported task difficulty	Did you find it difficult to hold your body in the required po
p_id	integer	[1-44]	participant id	
painful	integer	[1,7]	self-reported pain	Did you find it painful to hold your body in the required pos
height	integer	[155,195]	participant's height in cm	
gender	string		participant's gender	
condition_nr	integer	[0,1]	numerical condition assignment	
condition_name	string		condition assignment as string	

Jansen & Hornbæk (2018) https://github.com/yvonnejansen/posture

Where to share?

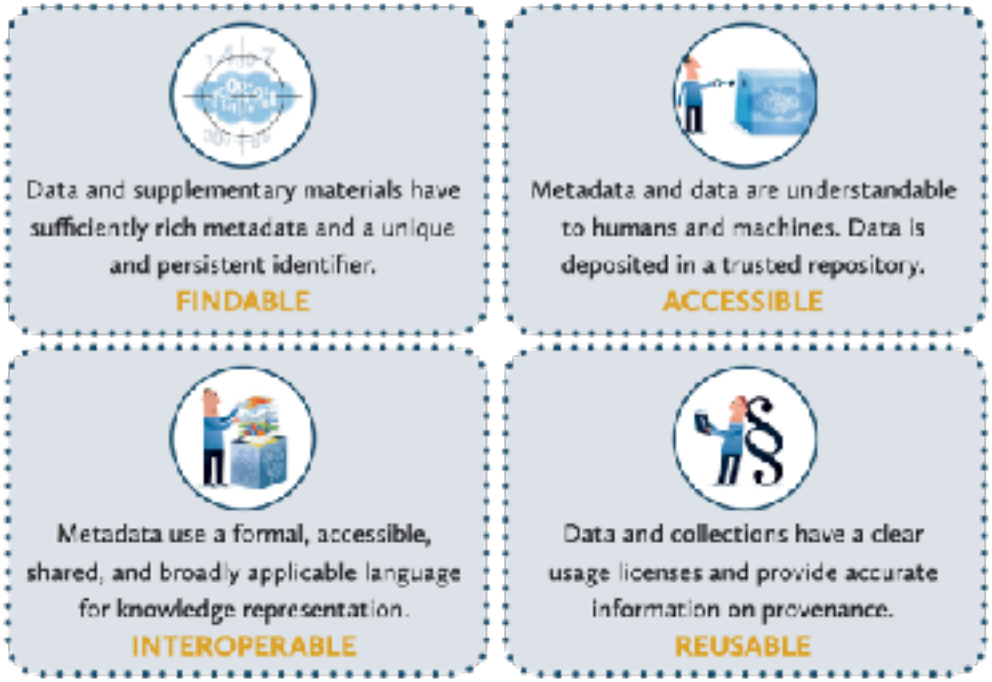


Diagram: LIBER Europe

OSF

One-stop service for whole project life cycle

zenodo

Good for big (>1 GB) files, Has versioned DOIs

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Search engine for specialized data repositories

Where to share?

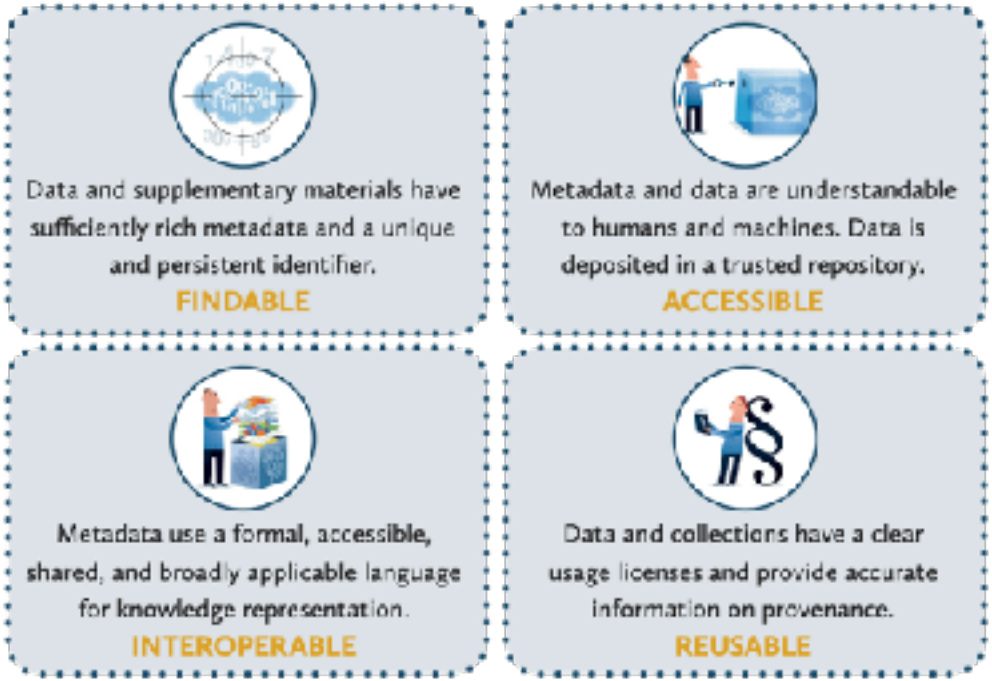


Diagram: LIBER Europe

ACM DL DIGITAL LIBRARY

😐 Findable: Same DOI as the paper, but materials are in single zip file

✅ Accessible: Supplementary materials has no paywall (but not widely known)

Some SIGCHI conferences only allow a video preview as supplementary material

Where to share?

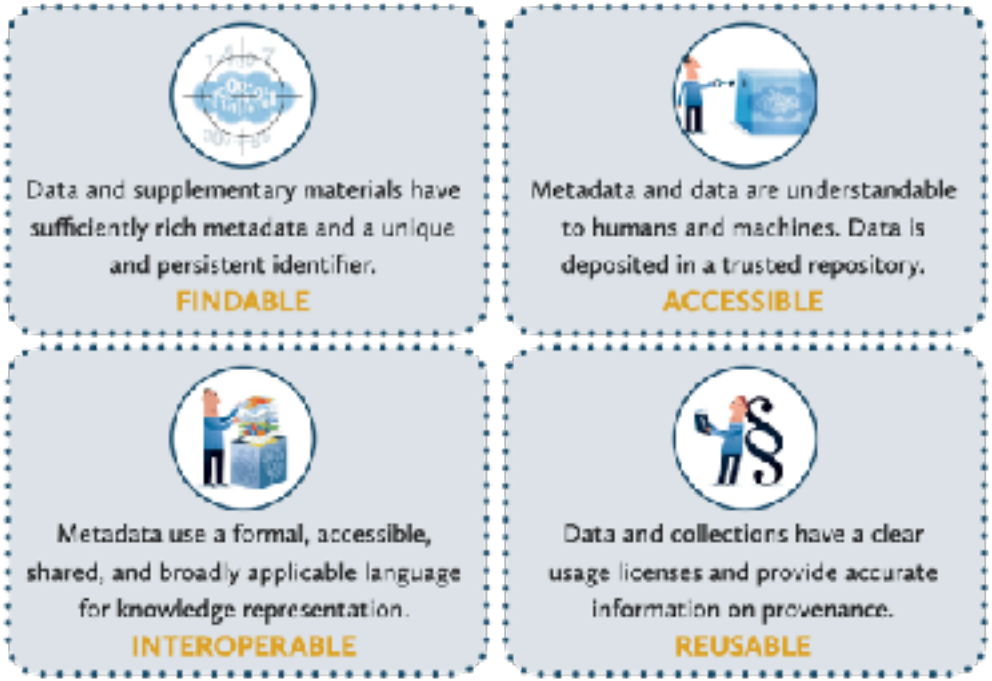


Diagram: LIBER Europe

GitHub

✅ Findable

✅ Interoperable: GitHub forking, Git submodule

❌ ~~Accessible~~: Repositories are deletable → broken link, Whodunit?

Recommendation: Add a snapshot of GitHub to OSF or Zenodo


Be
FAIR
Findable
Accessible
Interoperable
Reusable
and
CARE
Collective Benefit
Authority to Control
Responsibility
Ethics

principles

Originally developed in the context of indigenous data, we think the principles could be applied broadly. Below are our generalized wording; for the original, see: https://www.gida-global.org/care

Collective benefit: Data ecosystems shall be designed and function in ways that enable inclusive development, improved governance, and equitable outcomes

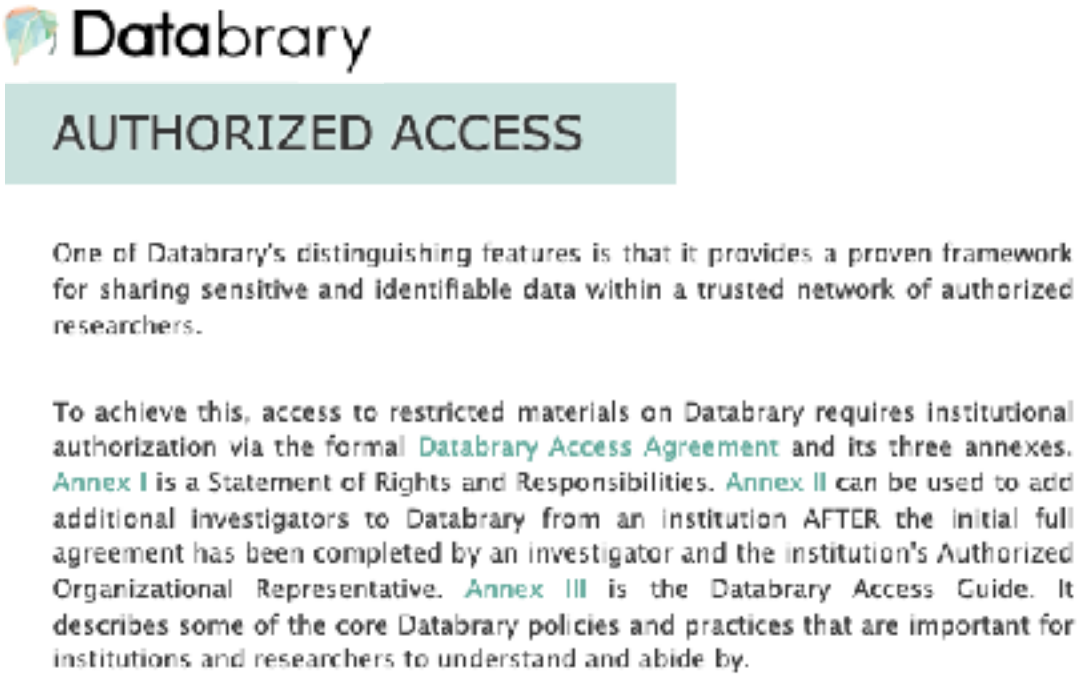
Authority to Control: Recognizing the rights and interests of people involved in generating the data, especially their rights to free, prior, and informed consent in the collection and use of the data

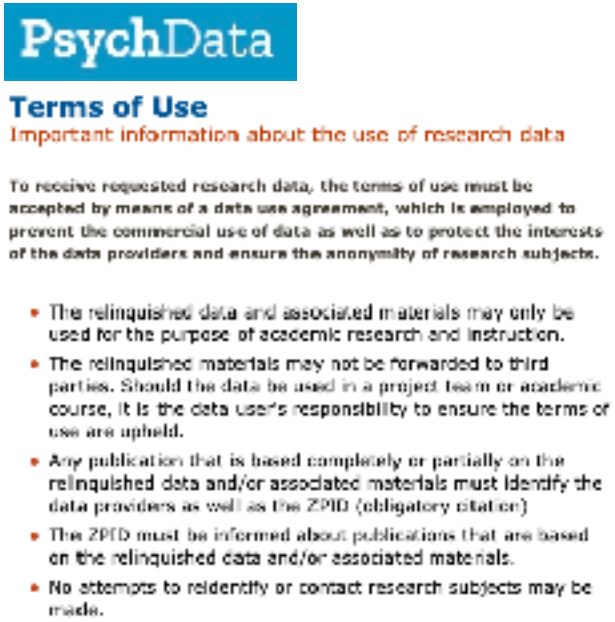
Responsibility: Researchers are responsible for sharing how the data are used to support collective benefits as well as benefits to individuals who involved in generating the data

Ethics: Minimize potential harm and maximize the benefit of people involved

Sharing sensitive data

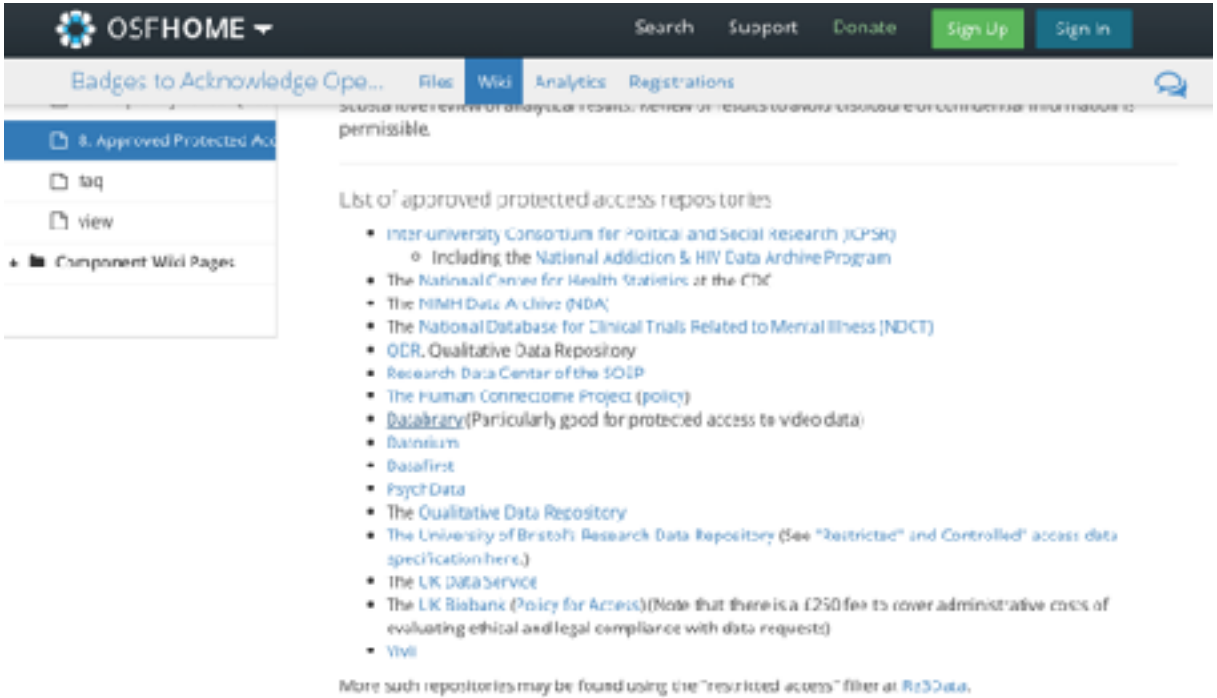
Ethical concerns? Consider using one of the [Protected Access Repositories]





Sharing sensitive data

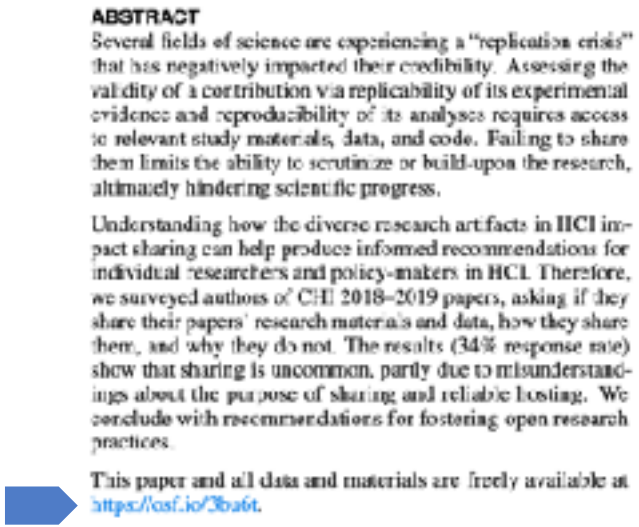
Ethical concerns? Consider using one of the [Protected Access Repositories]



https://osf.io/tvyxz/wiki/8.%20Approved%20Protected%20Access%20Repositories

Pointing the readers to the shared materials

Crossing the paywall: Link to the FAIR repository at the end of the abstract



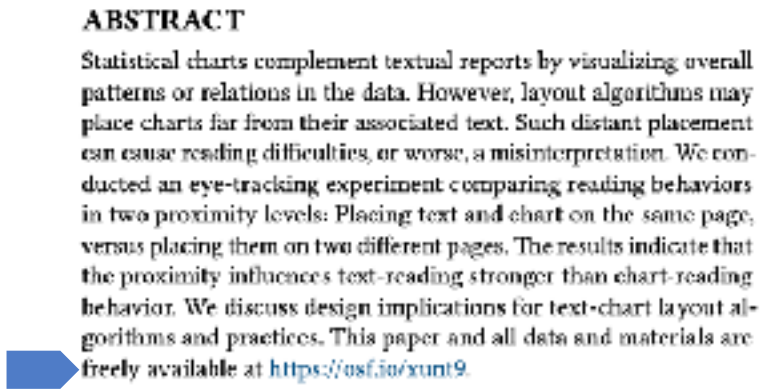
ABSTRACT

Several fields of science are experiencing a "replication crisis" that has negatively impacted their credibility. Assessing the validity of a contribution via replicability of its experimental evidence and reproducibility of its analyses requires access to relevant study materials, data, and code. Failing to share them limits the ability to scrutinize or build-upon the research, ultimately hindering scientific progress.

Understanding how the diverse research artifacts in HCI impact sharing can help produce informed recommendations for individual researchers and policy-makers in HCI. Therefore, we surveyed authors of CHI 2018–2019 papers, asking if they share their papers' research materials and data, how they share them, and why they do not. The results (34% response rate) show that sharing is uncommon, partly due to misunderstandings about the purpose of sharing and reliable hosting. We conclude with recommendations for fostering open research practices.

This paper and all data and materials are freely available at https://osf.io/7bu6t.

Transparency of CHI Research Artifacts: Results of a Self-Reported Survey
(Wacharamanotham et al., 2020)



ABSTRACT

Statistical charts complement textual reports by visualizing overall patterns or relations in the data. However, layout algorithms may place charts far from their associated text. Such distant placement can cause reading difficulties, or worse, a misinterpretation. We conducted an eye-tracking experiment comparing reading behaviors in two proximity levels: Placing text and chart on the same page, versus placing them on two different pages. The results indicate that the proximity influences text-reading stronger than chart-reading behavior. We discuss design implications for text-chart layout algorithms and practices. This paper and all data and materials are freely available at https://osf.io/xunt9.

Asymmetric effect of text-chart proximity on reading behavior
(Yang & Wacharamanotham, 2020)

Sharing vs. the anonymized reviewing process

When you submit the materials to an anonymized reviewing, consider:

- Preregistration: OSF: view-only link AsPredicted: anonymous PDF
- Source code:
  - The absolute paths may contain your name
  - Github URL may contain your name or user ID

Although it is the due-diligence of the authors to anonymize materials, minor oversights is not a reason for rejection

Exercise 4: Brainstorm research materials and sharing concerns

(10 minutes)

Continue with the case you previously chose.

1. Brainstorm possible 2–3 research materials that may be generated
2. Choose one research materials and brainstorm 3 concerns that people may have against sharing
3. Discuss ways to mitigate that concern

Reflection

Reflection on research transparency

More transparent = more work?

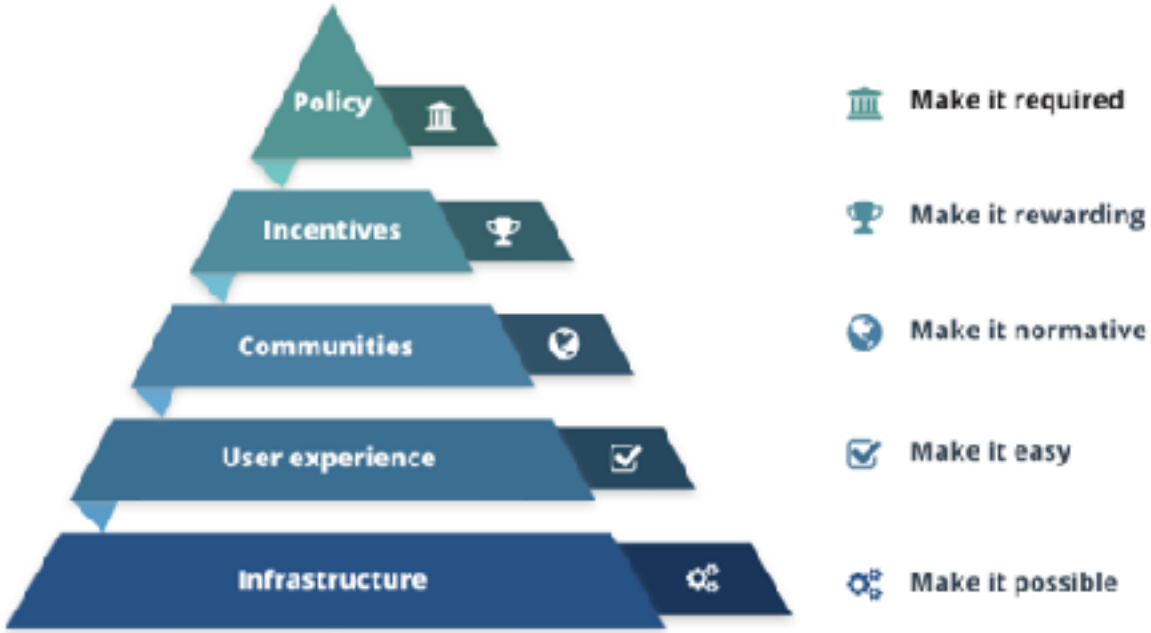
- Some learning needed for the first time, effortless later on
- Pays off: Better methodological rigor to self and to reviewers
- Small step: How can my next paper be more transparent than the last one?

Cultivating research transparency culture

When giving feedback or writing reviews, instead of penalizing the lack of transparency:

- Describe what could be improved
- Describe good consequences of the improvements
- Point to guides and examples

Motivating research transparency in HCI



Source: Center for Open Science strategy for scale sustainable adoption of open behaviors by researchers

Choices in research

Design phase:

D1 Measuring additional variables that can later be selected as covariates, independent variables,...

D2 Measuring the same dependent variable in several alternative ways

Analysis phase:

A2 Specifying pre-processing of data (eg. cleaning, normalization, smoothing, motion correction) in an ad hoc manner

Reporting phase:

R6 Presenting exploratory analyses as confirmatory

Where to share?

OSF zenodo

re3data.org

Motivating research transparency in HCI

Source: Center for Open Science strategy for scale: substantive adoption of open science tools by researchers

63

Chat Wacharamanotham

chat@ifi.uzh.ch

This tutorial is designed based on the open materials of the courses presented at CHI 2022–23 by Chat Wacharamanotham, Fumeng Yang, Abhraneel Sarma, Xiaoying Pu, and Lace Padilla.
https://osf.io/27r5z