

Motivating research transparency in Human-Computer Interaction

Chat Wacharamanotham

chat@acm.org

Twitter: [@chatchavan](https://twitter.com/chatchavan)

Mastodon: [@chat@hci.social](https://mstdn.social/@chat@hci.social)

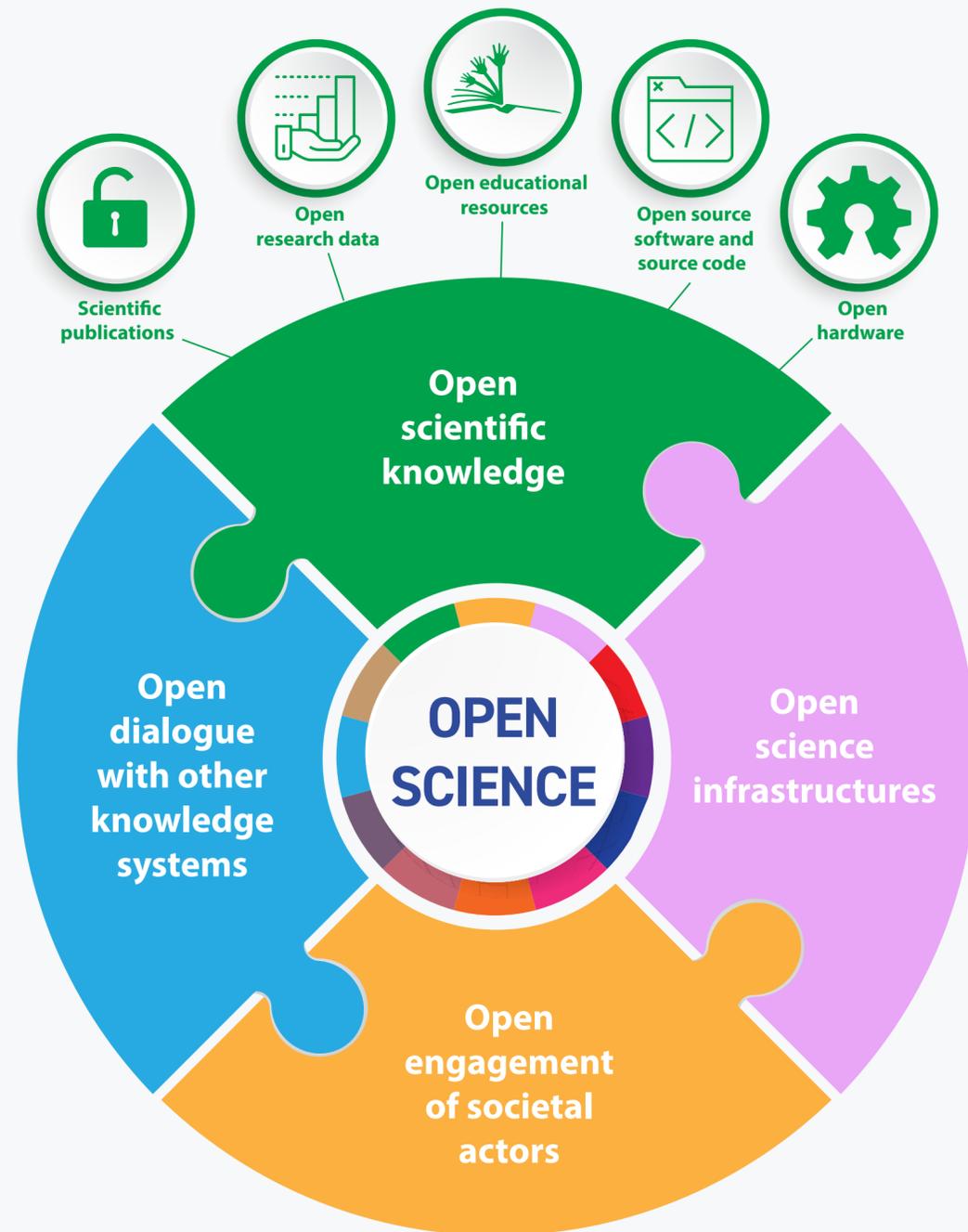
Max Planck Society Open Science Ambassador Meeting 2023

20 September 2023



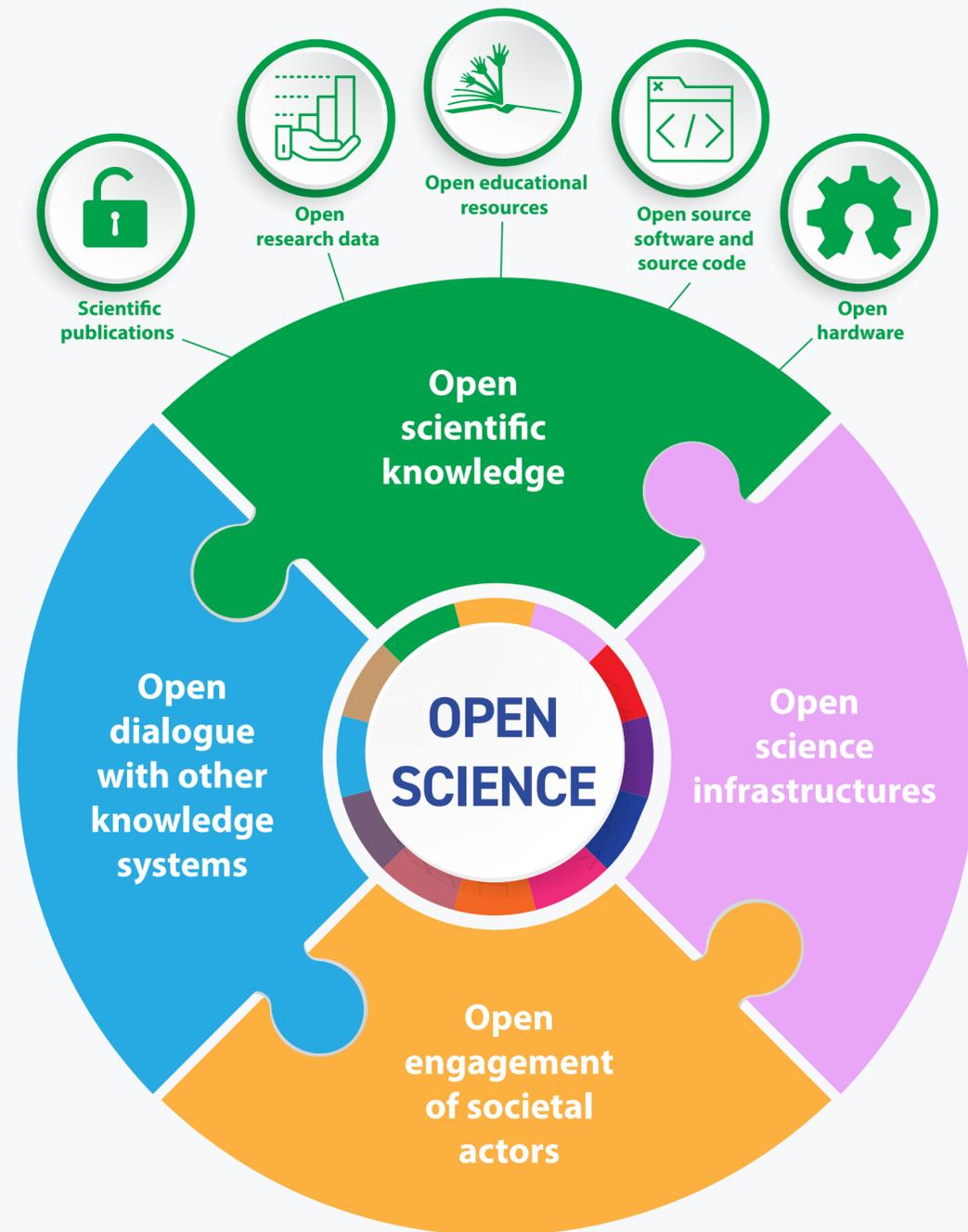
<http://tiny.cc/t-osap-2023>

(slides)



“[Open scientific knowledge] also refers to the possibility of opening research methodologies and evaluation processes.” [1]

[1] UNESCO Recommendation on Open Science



“[Open scientific knowledge] also refers to the possibility of opening research methodologies and evaluation processes.” [1]

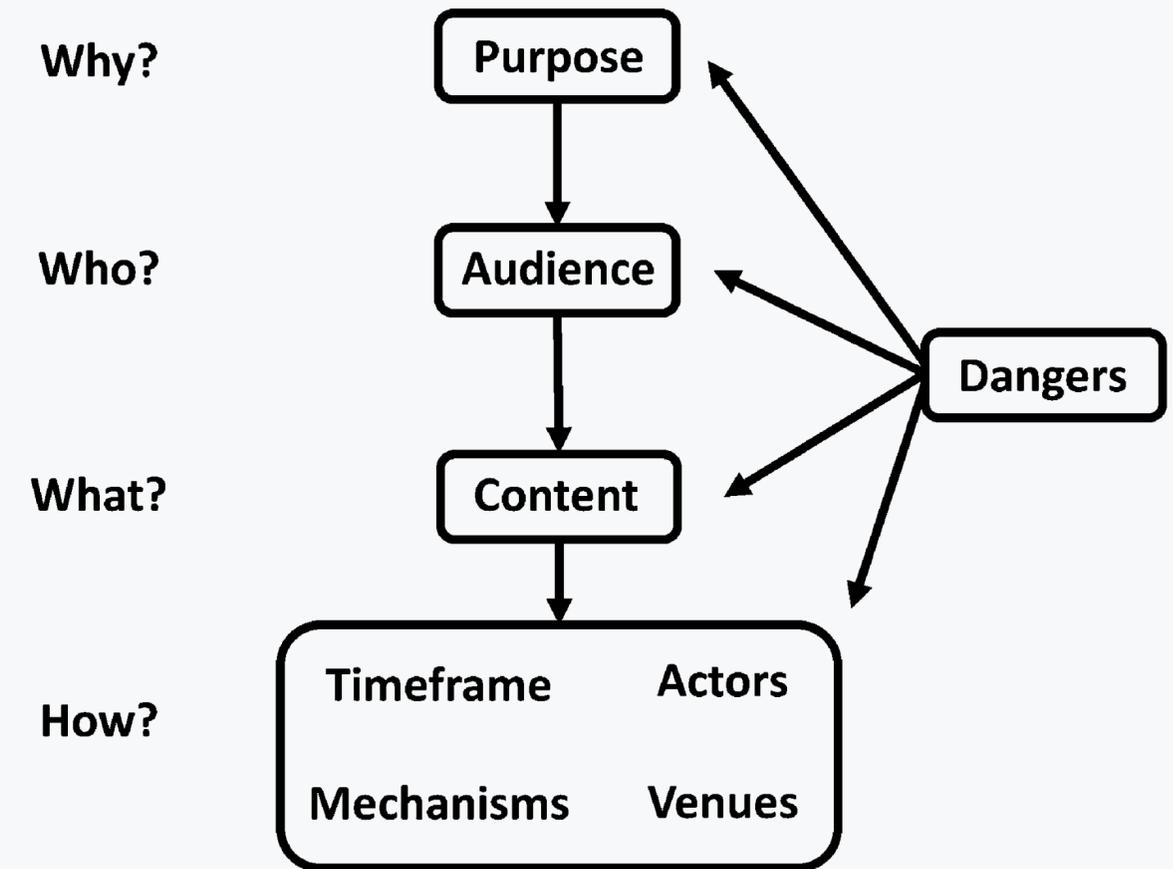
“**Research transparency** refers to honesty and clarity in all communications about the research processes and outcomes to the extent possible.” [2]

[1] [UNESCO Recommendation on Open Science](#)

[2] From a working article by Yvonne Jansen, Jan B. Vornhagen, Olga Iarygina, Kavous Salehzadeh Niksirat, Lonni Besançon, Pierre Dragicevic, Julien Gori, and Chat Wacharamanatham; This forthcoming article is based on “A Manifesto for Transparent Quantitative Research – Definition” by Wacharamanatham et al. in Report from Dagstuhl Seminar 22392: Transparent Quantitative Research as a User Interface Problem. <https://doi.org/10.4230/DagRep.12.9.220>

Research transparency

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

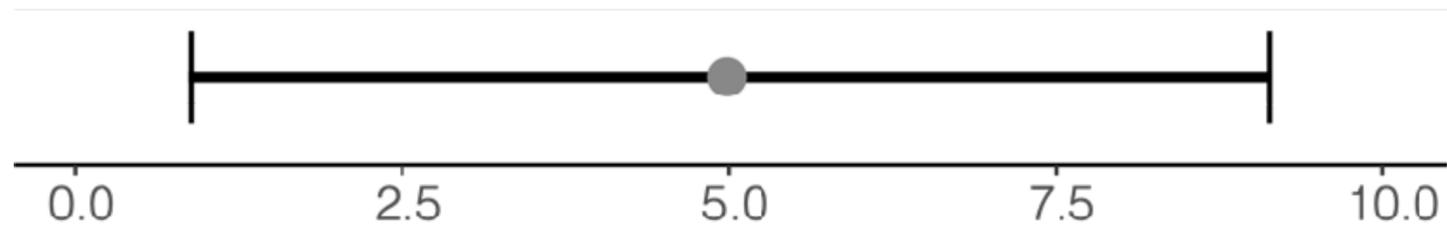
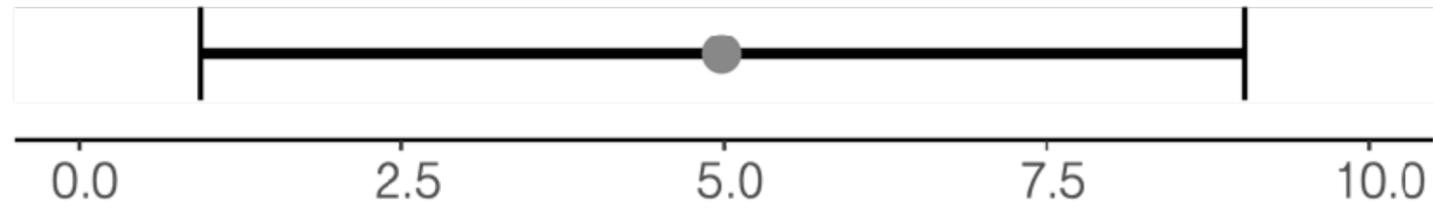


Elliott (2022) A [Taxonomy of Transparency in Science](#). *Canadian Journal of Philosophy*, 52: 3, 342–355

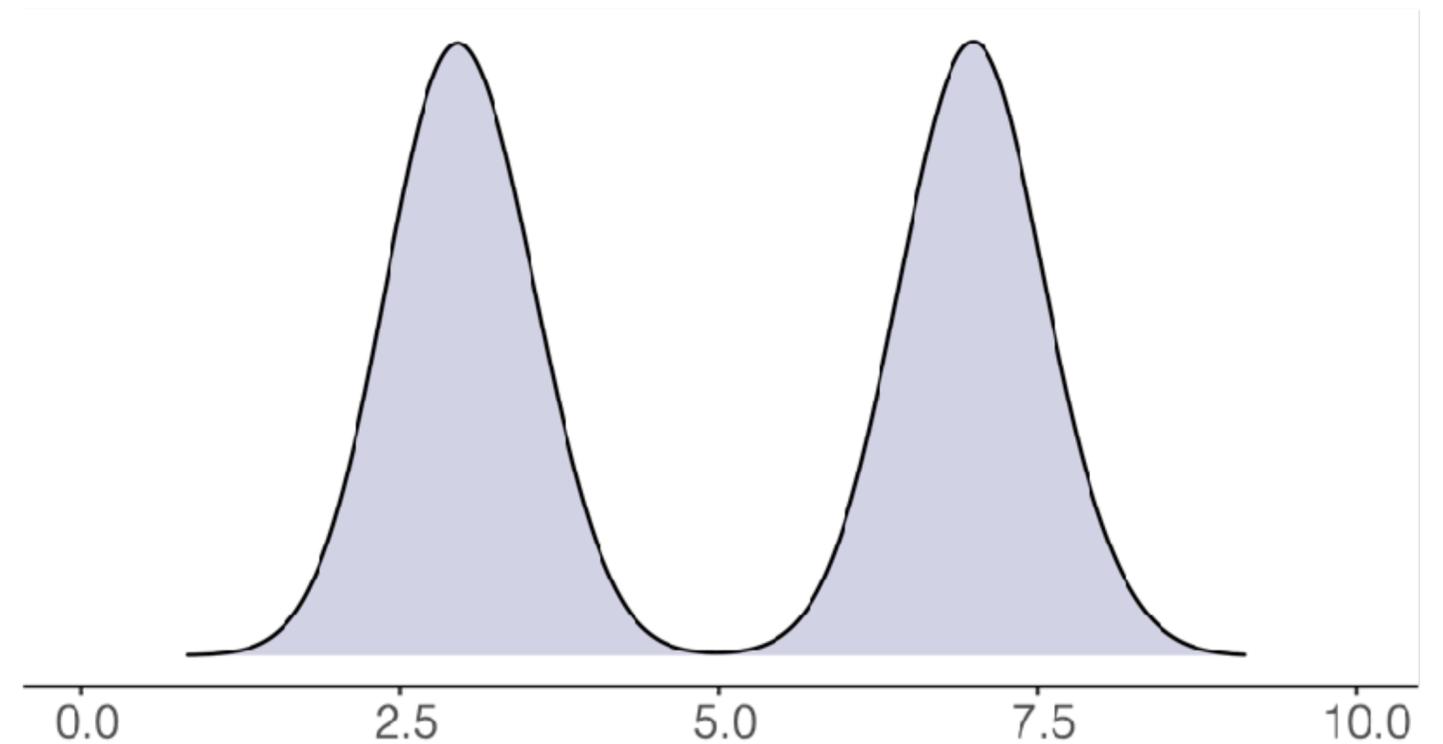
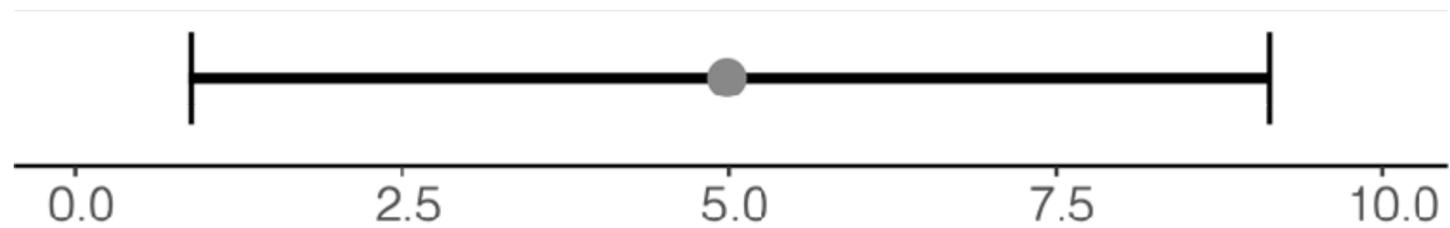
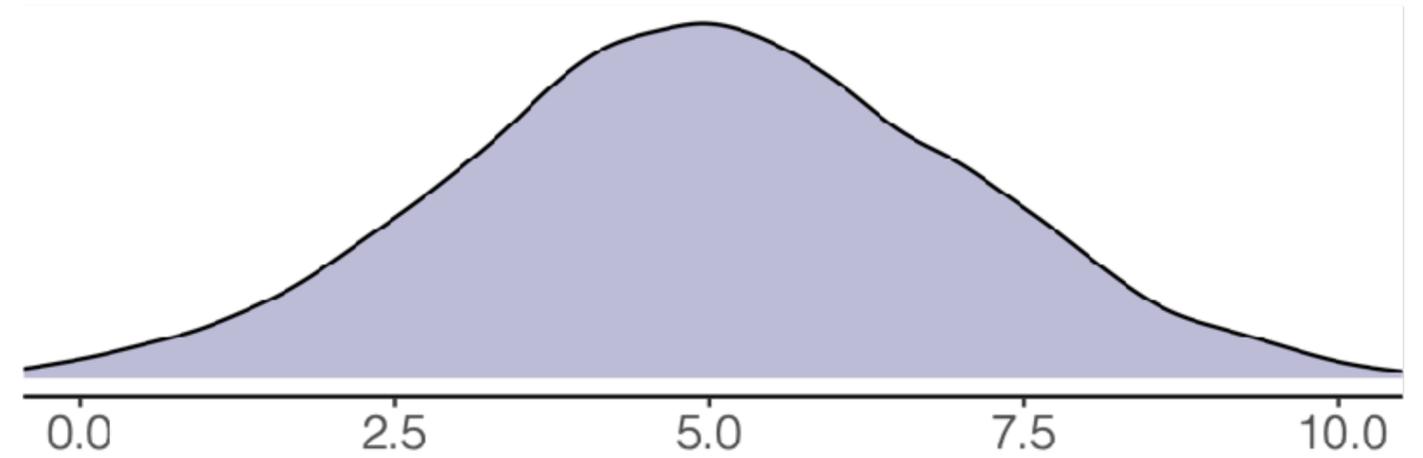
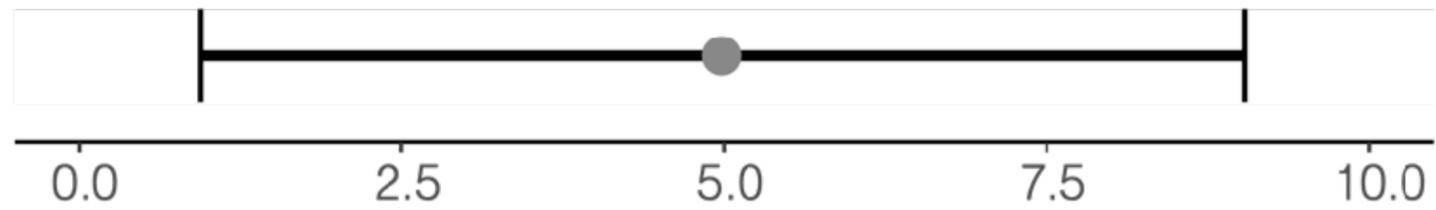
Transparency through visualizing research data



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



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



7

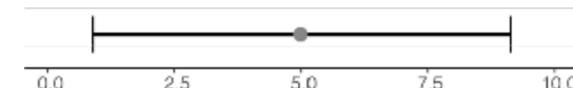
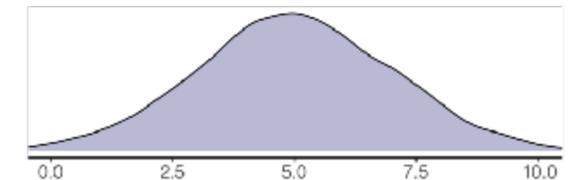
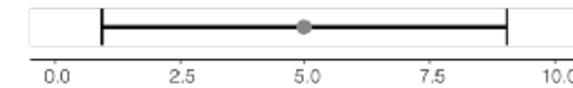
Intervals (95% CI)

Distributions

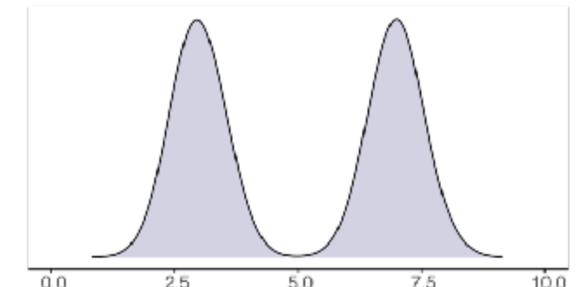
Visualizing uncertainty in the results

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

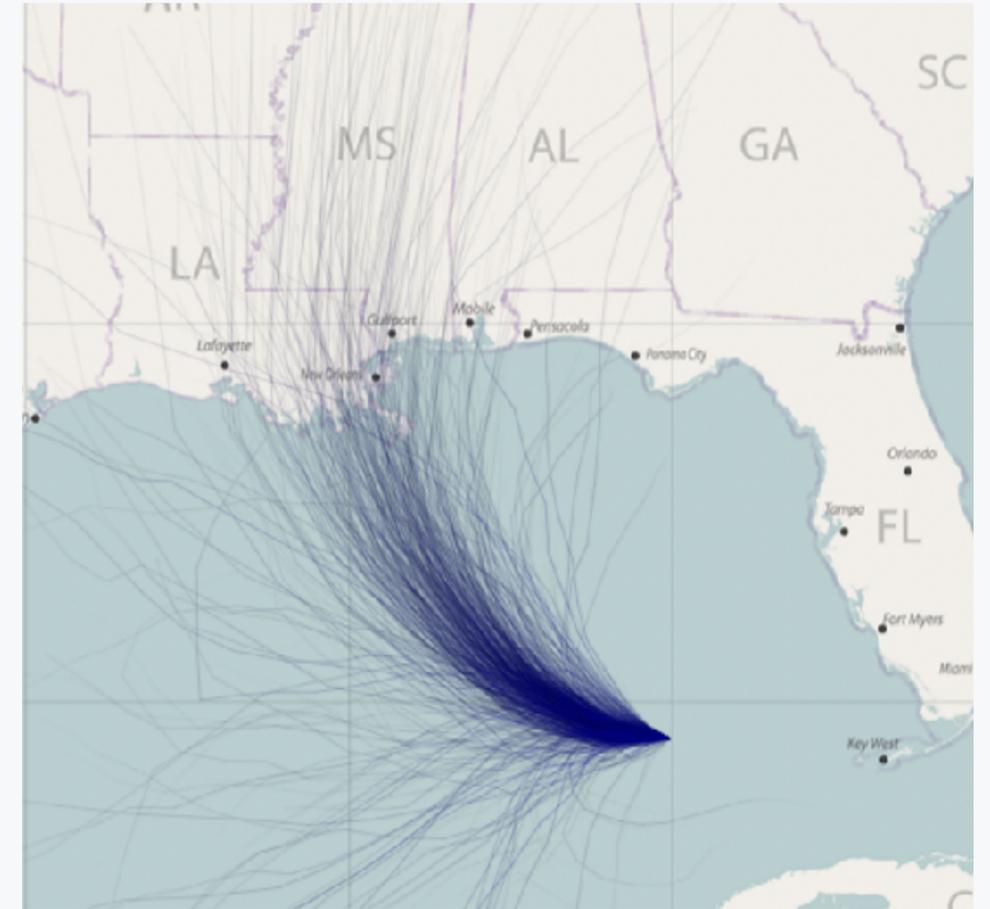
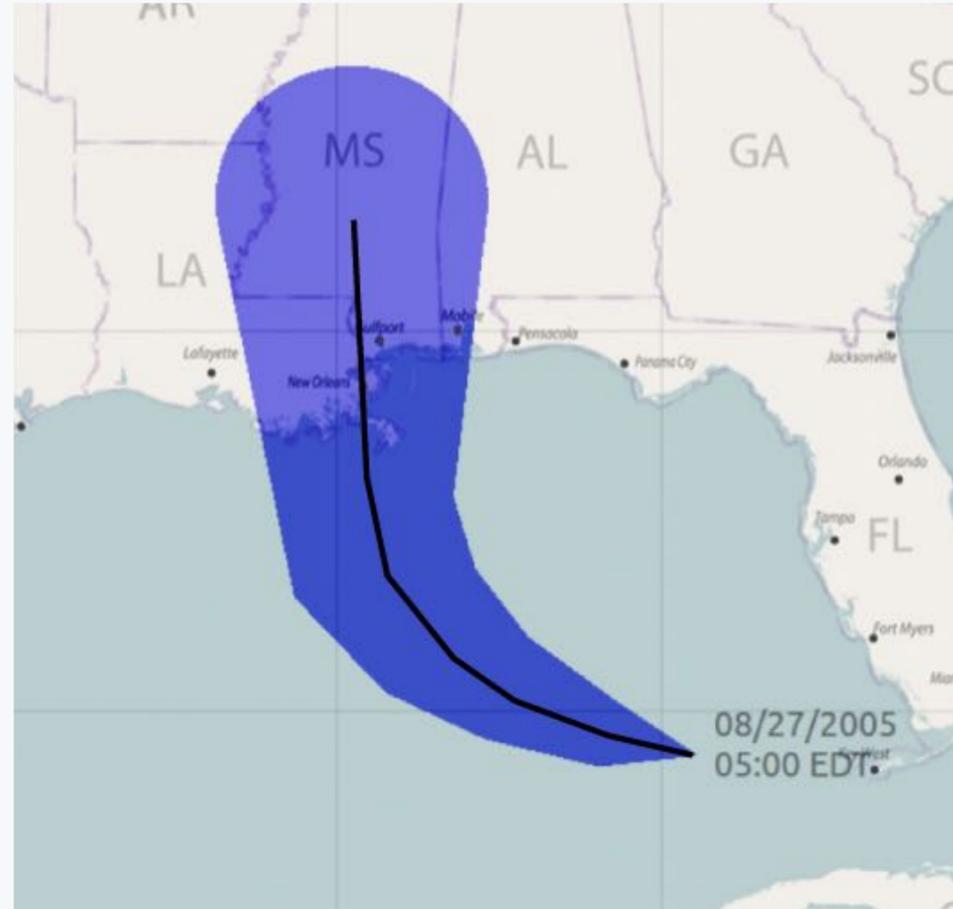
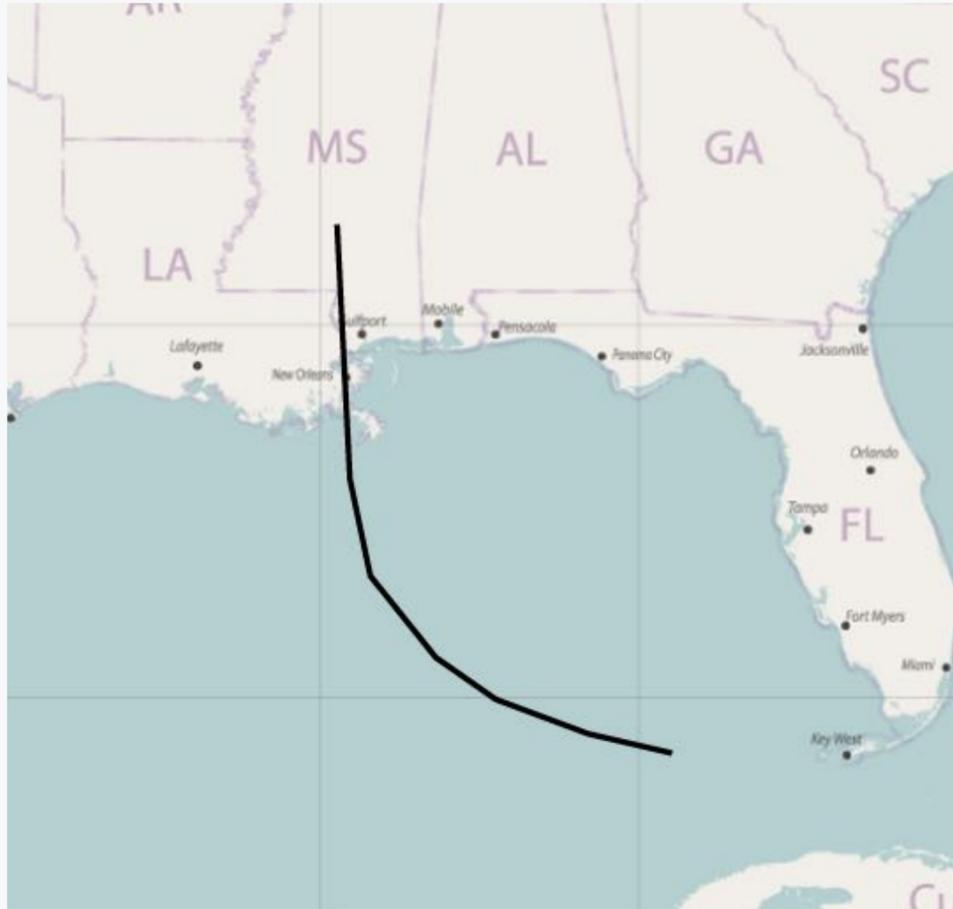


Intervals (95% CI)

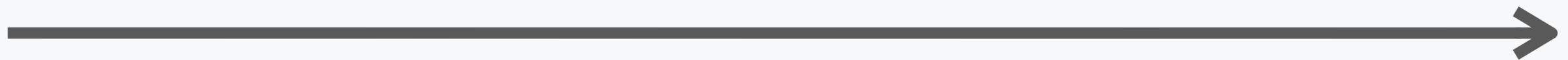


Distributions

8 [1] [Mackinlay, J. \(1986\)](#). Automating the design of graphical presentations of relational information.
[2] [Munzner, T. \(2014\)](#). Visualization analysis and design. CRC press.



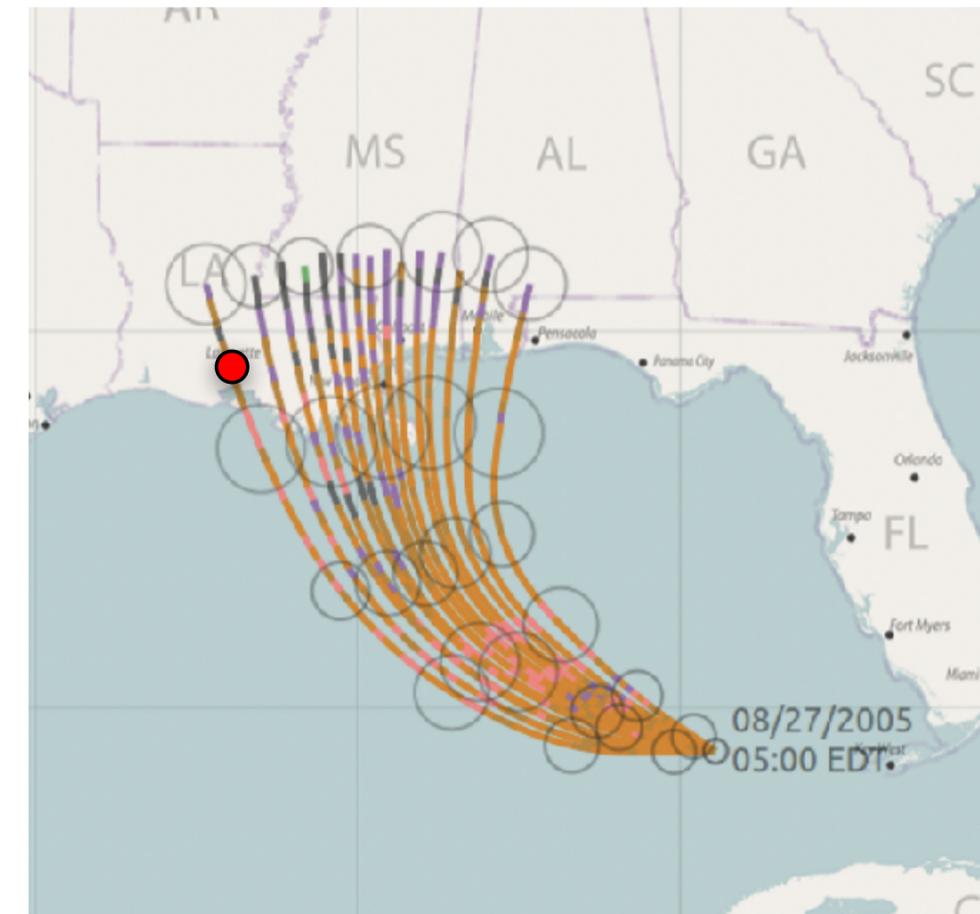
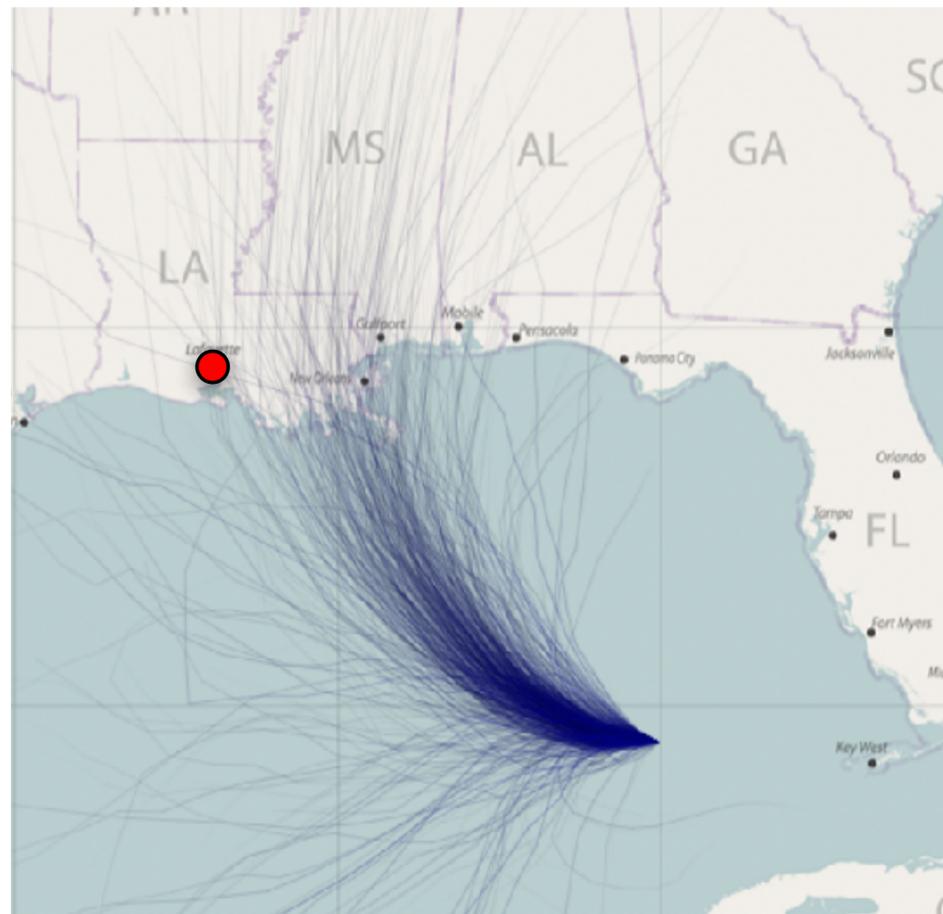
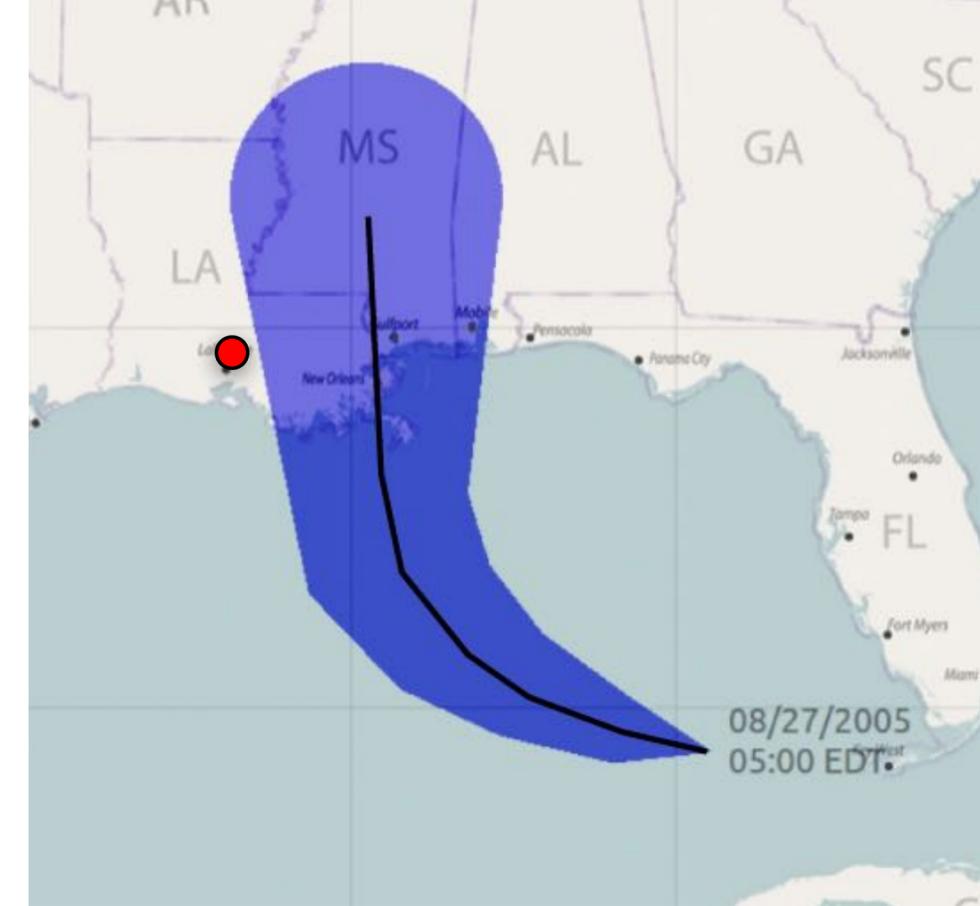
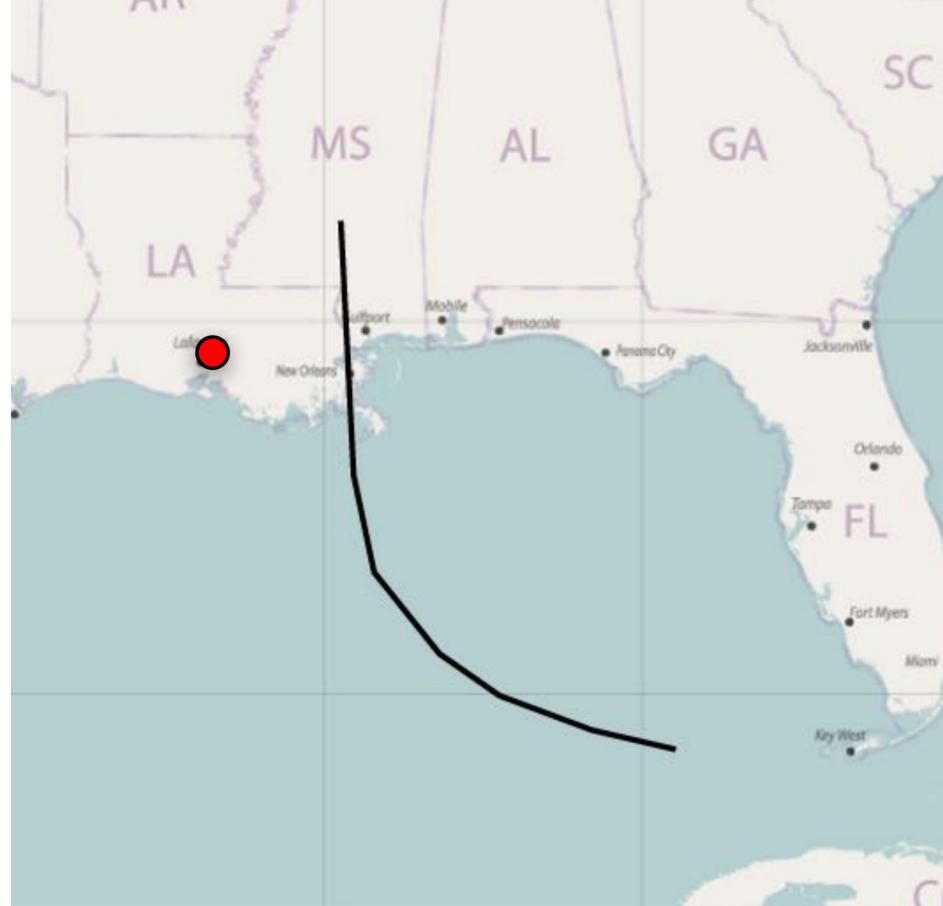
expressiveness



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

Would you stay or evacuate?

Usable visualizations support users in making accurate inferences

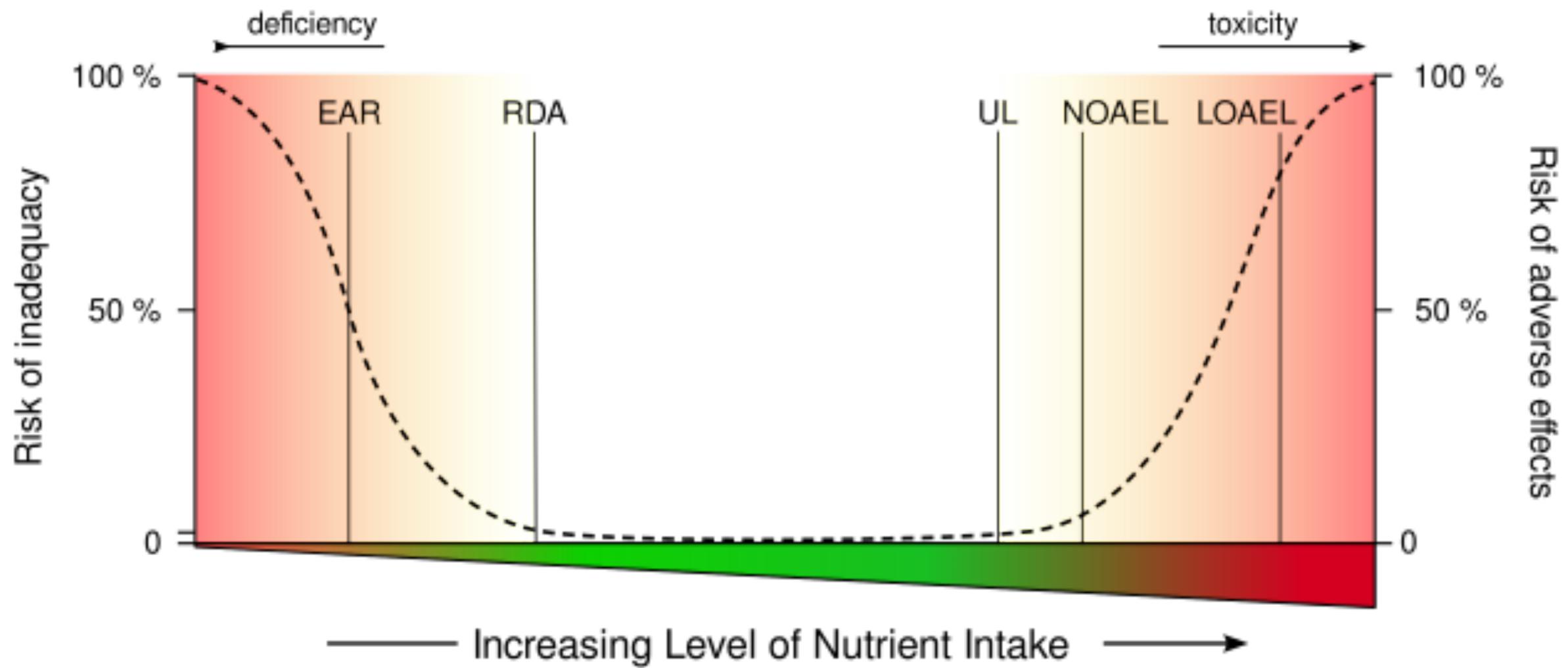


Dietary Reference Intakes (DRIs): Recommended Dietary Allowances and A

Food and Nutrition Board, Institute of Medicine, National Academies

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

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



Uncertainty matters

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

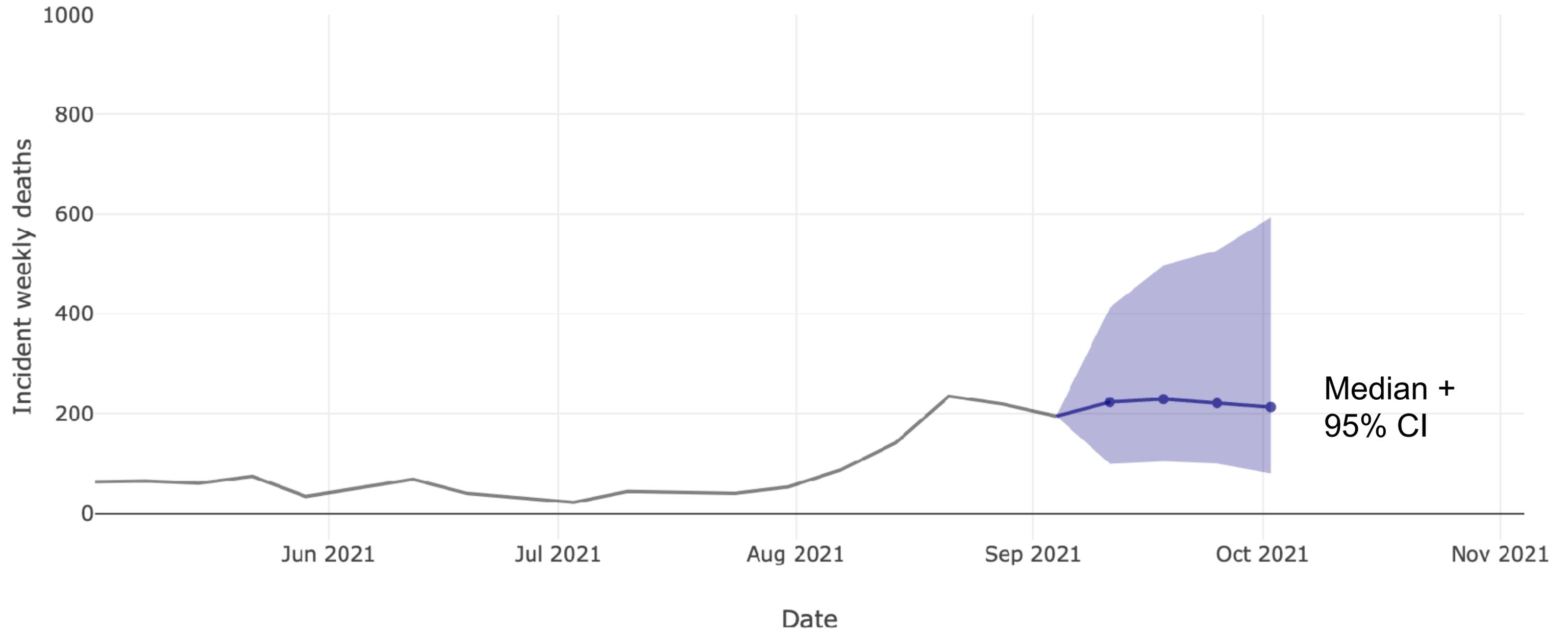
Showing uncertainty:

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

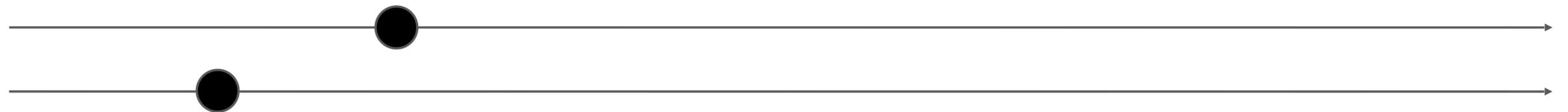
Usable visualizations support users in making accurate inferences

Showing uncertainty contributes to usability

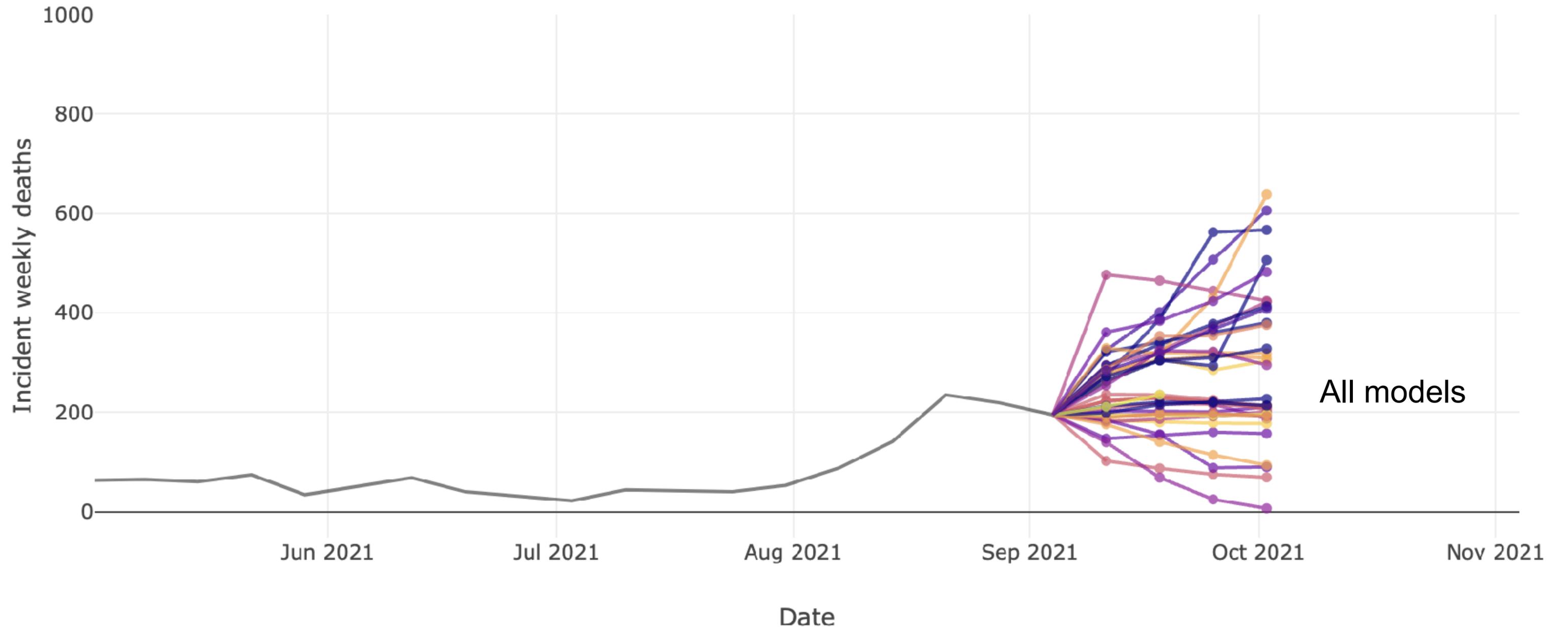
Forecasts of Incident weekly deaths in Alabama as of 2021-09-04



expressiveness
usability



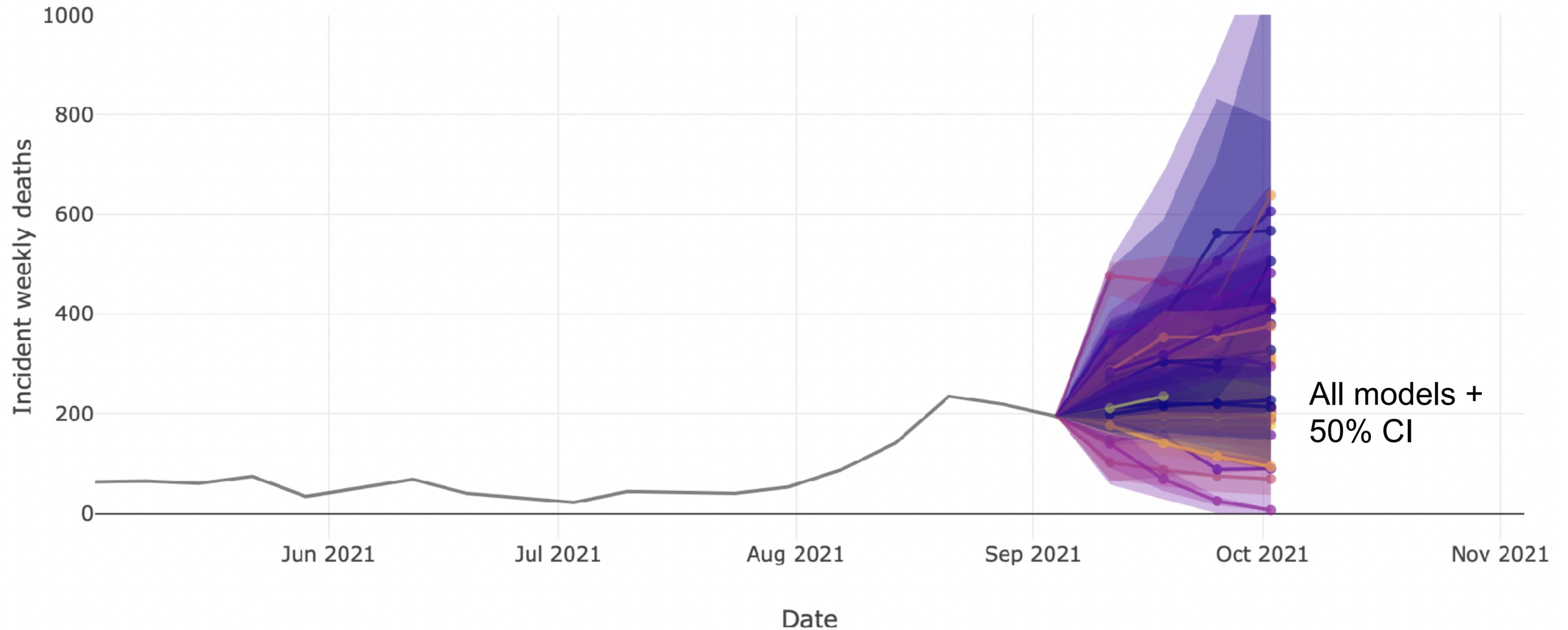
Forecasts of Incident weekly deaths in Alabama as of 2021-09-04



expressiveness
usability



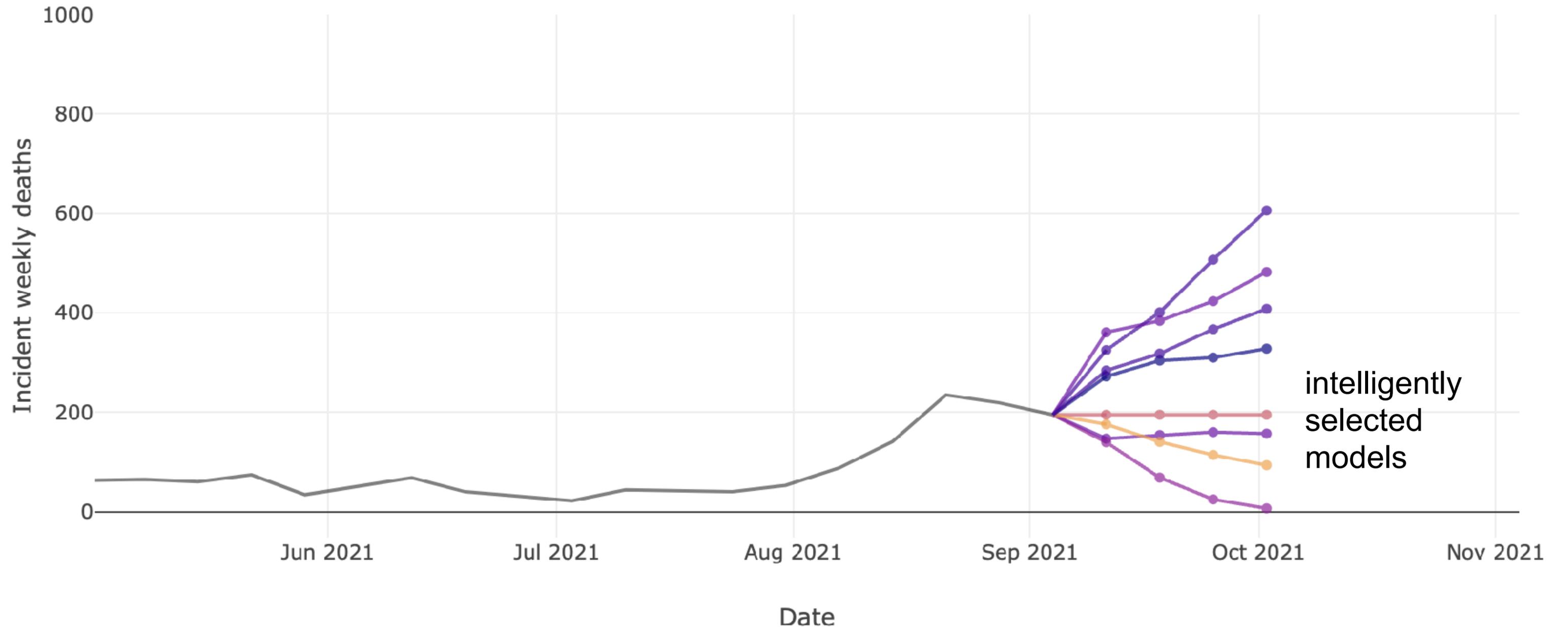
Forecasts of Incident weekly deaths in Alabama as of 2021-09-04



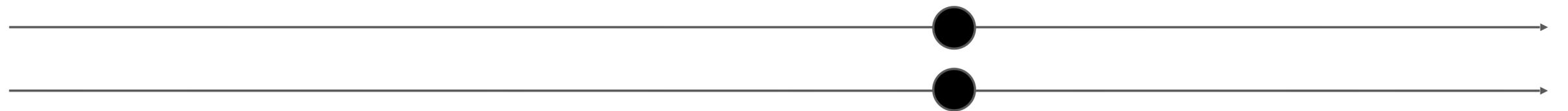
expressiveness
usability



Forecasts of Incident weekly deaths in Alabama as of 2021-09-04



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

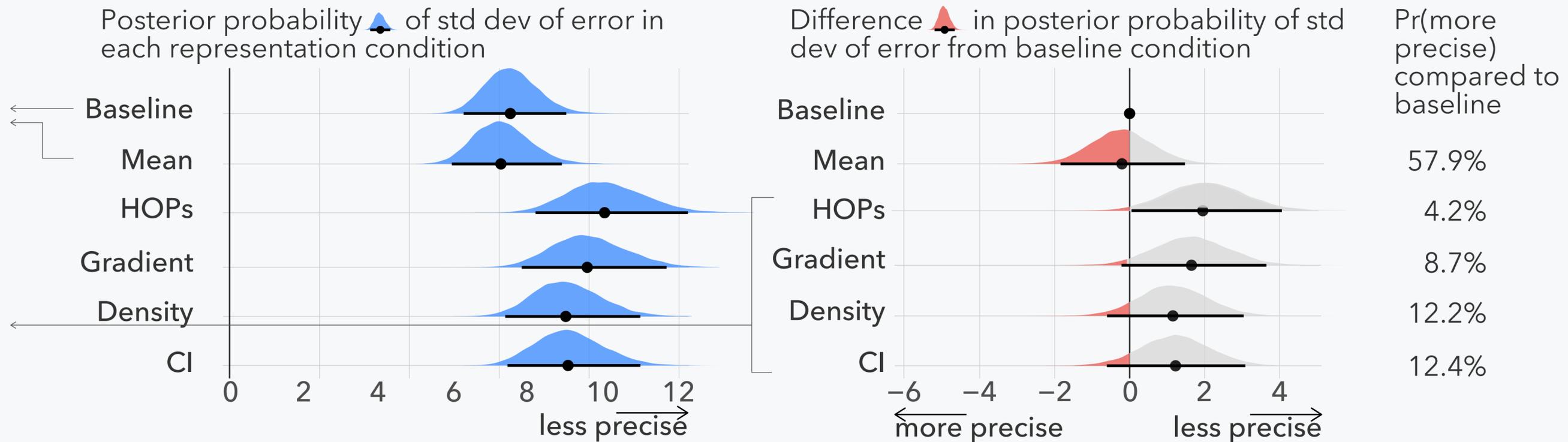
No single best answer

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

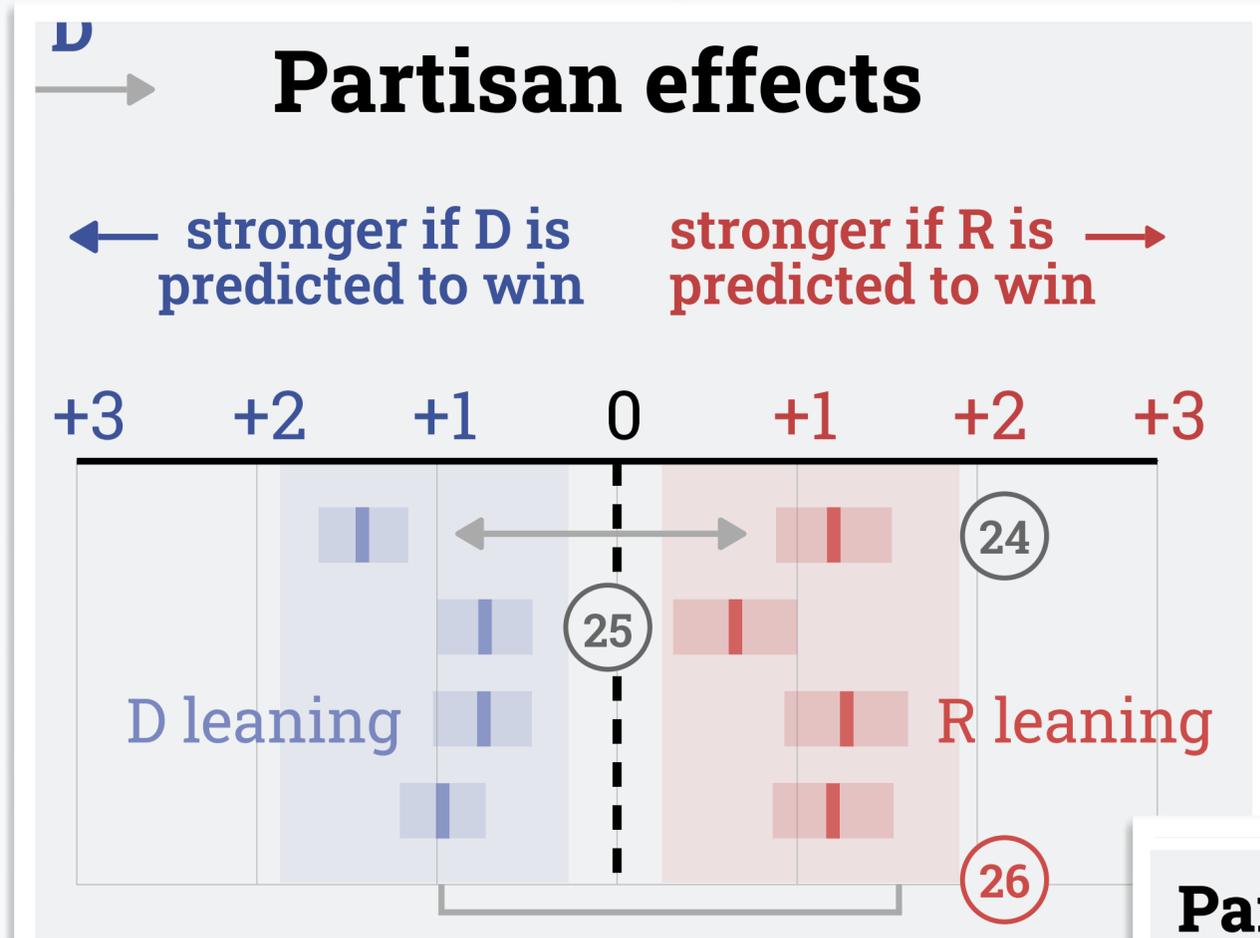
Bridge charts to text with annotations

Baseline and **Mean** will likely lead to the most precise (most consistent or less variable) responses on average

We estimate, with considerable probability (~ 90%), that **uncertainty encodings** will likely lead to less precision i.e. greater variability



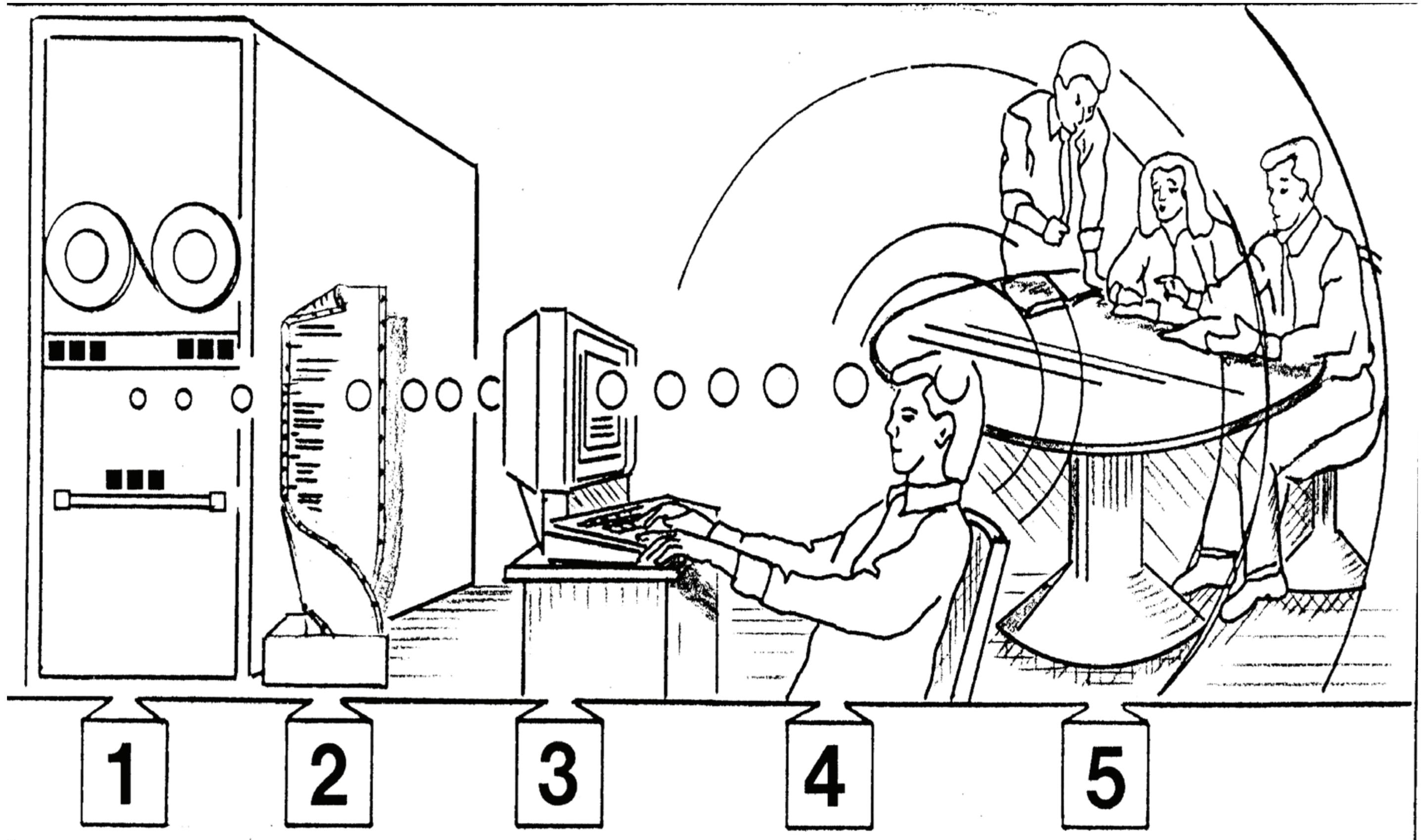
Bridge charts to text with annotations



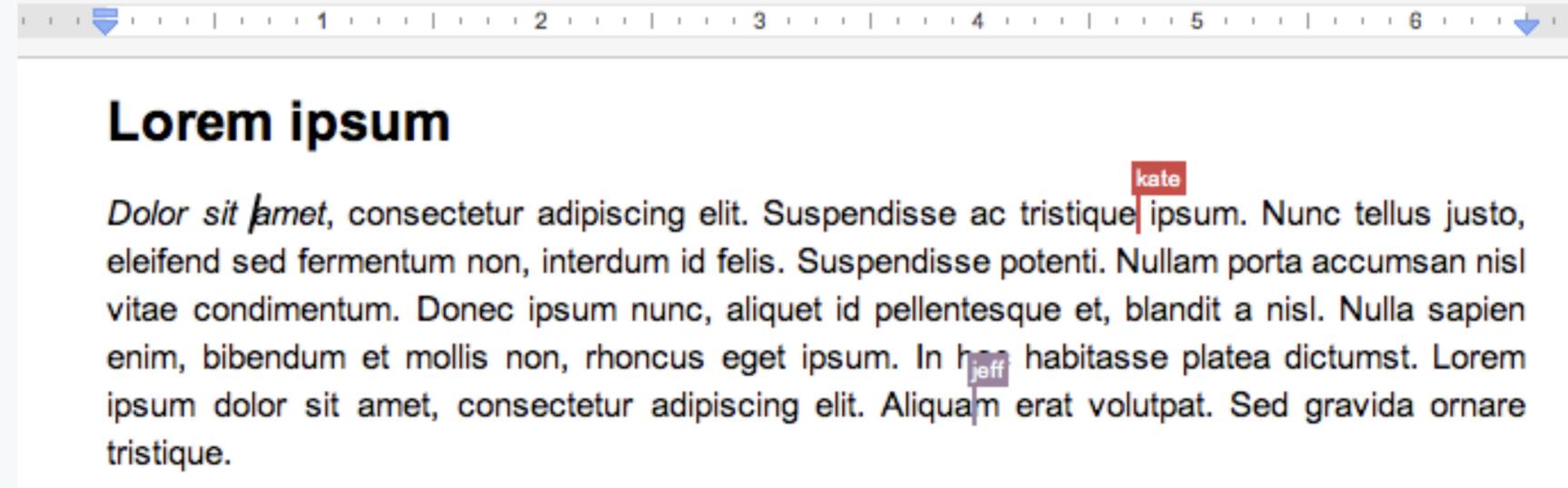
Partisan effects

Republicans ²⁶ feel more positively and ³⁴ ³⁵ trust the forecasts more when their candidate is predicted to win, compared to how Democrats feel when a Democratic candidate is predicted to win. That said, R

The field of Human-Computer Interaction



2010



2006 Google Document beta based on “Writely”

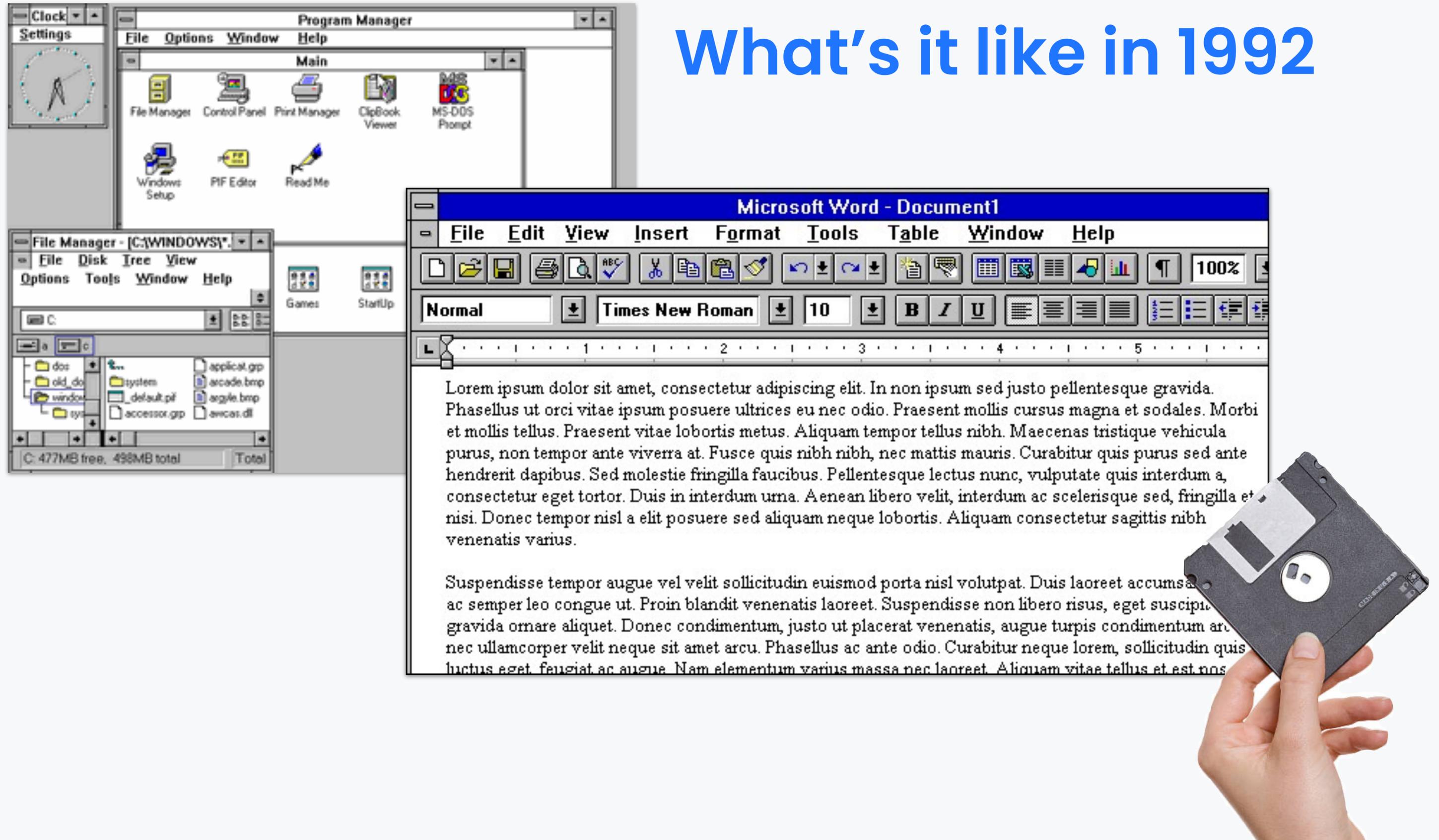
⋮

1993 A study from Judith Olson and colleagues show that brainstorming with ShrEdit as a secondary channel yielded better ideas

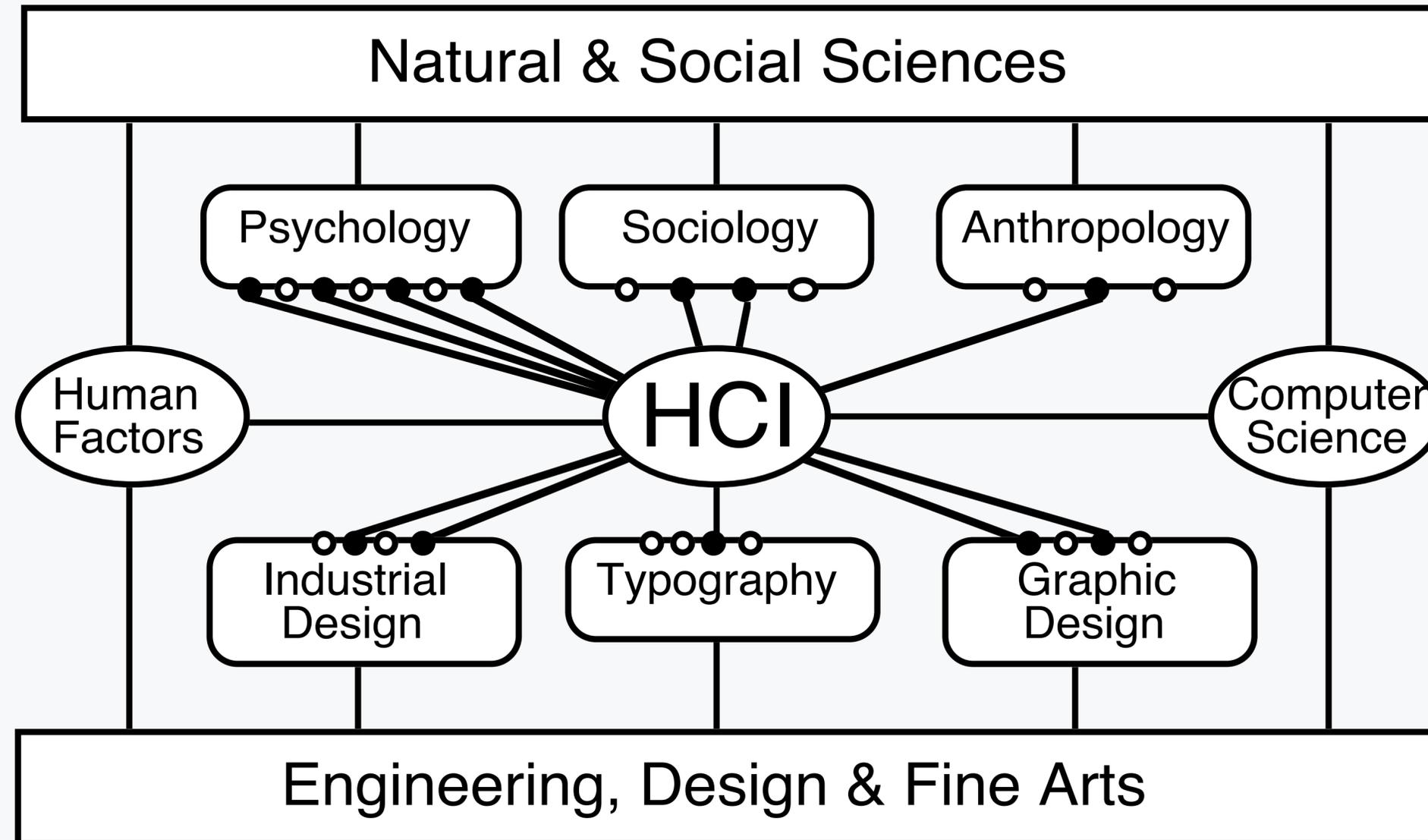
1992 A study from Paul Dourish and Victoria Bellotti found that awareness of others’ cursor position in workspace improved collaboration

1992 Laura McGuffin and Gary Olson developed “ShrEdit”: a real-time synchronous editor running on a network of Apple Macintoshes

What's it like in 1992



The field of Human-Computer Interaction



The situation of research transparency in Human-Computer Interaction

Replicability

Closely matched method

+

New data

=

Consistent results

Reproducibility

Same data analysis method

+

Same data

=

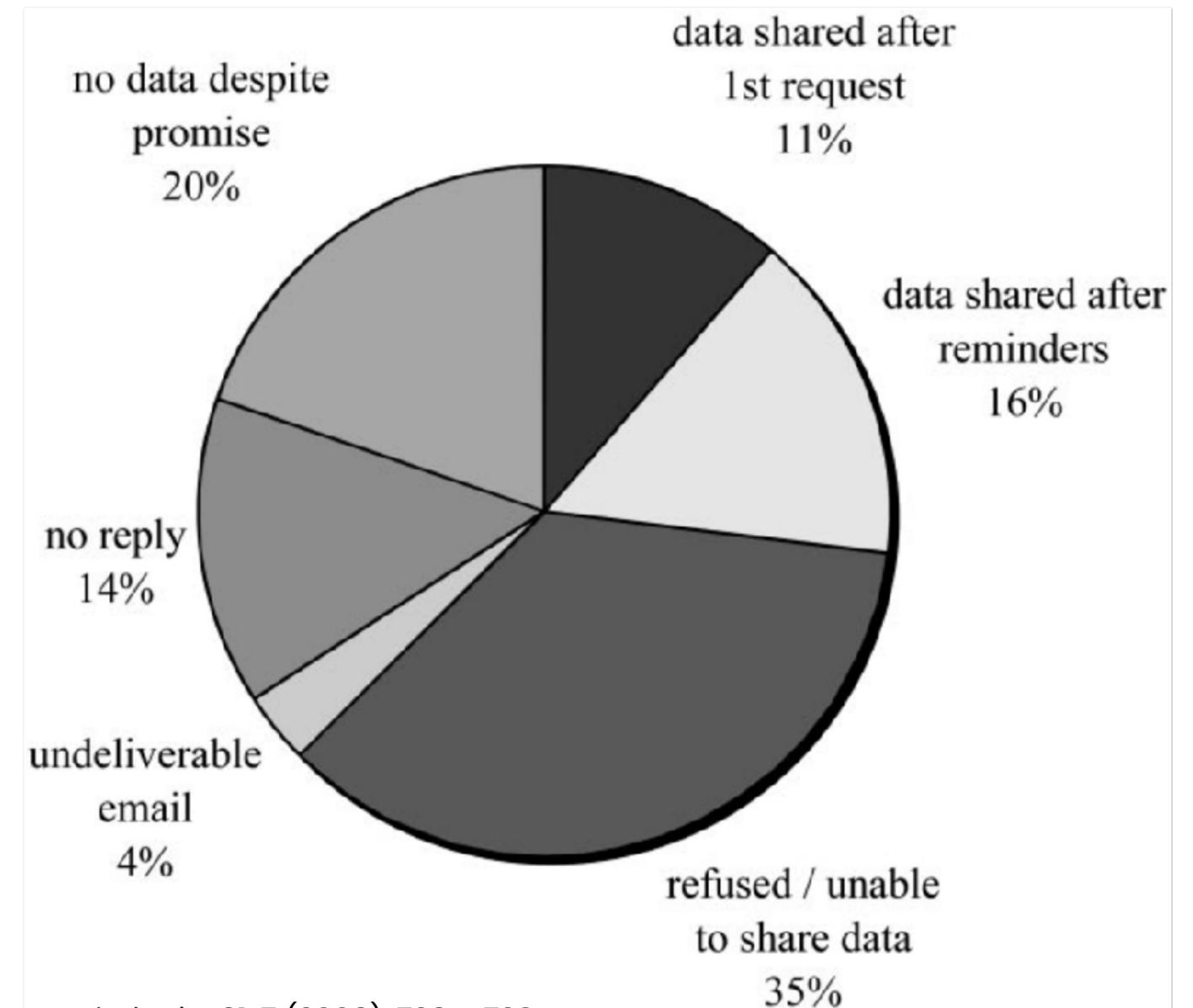
Same results

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

“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.

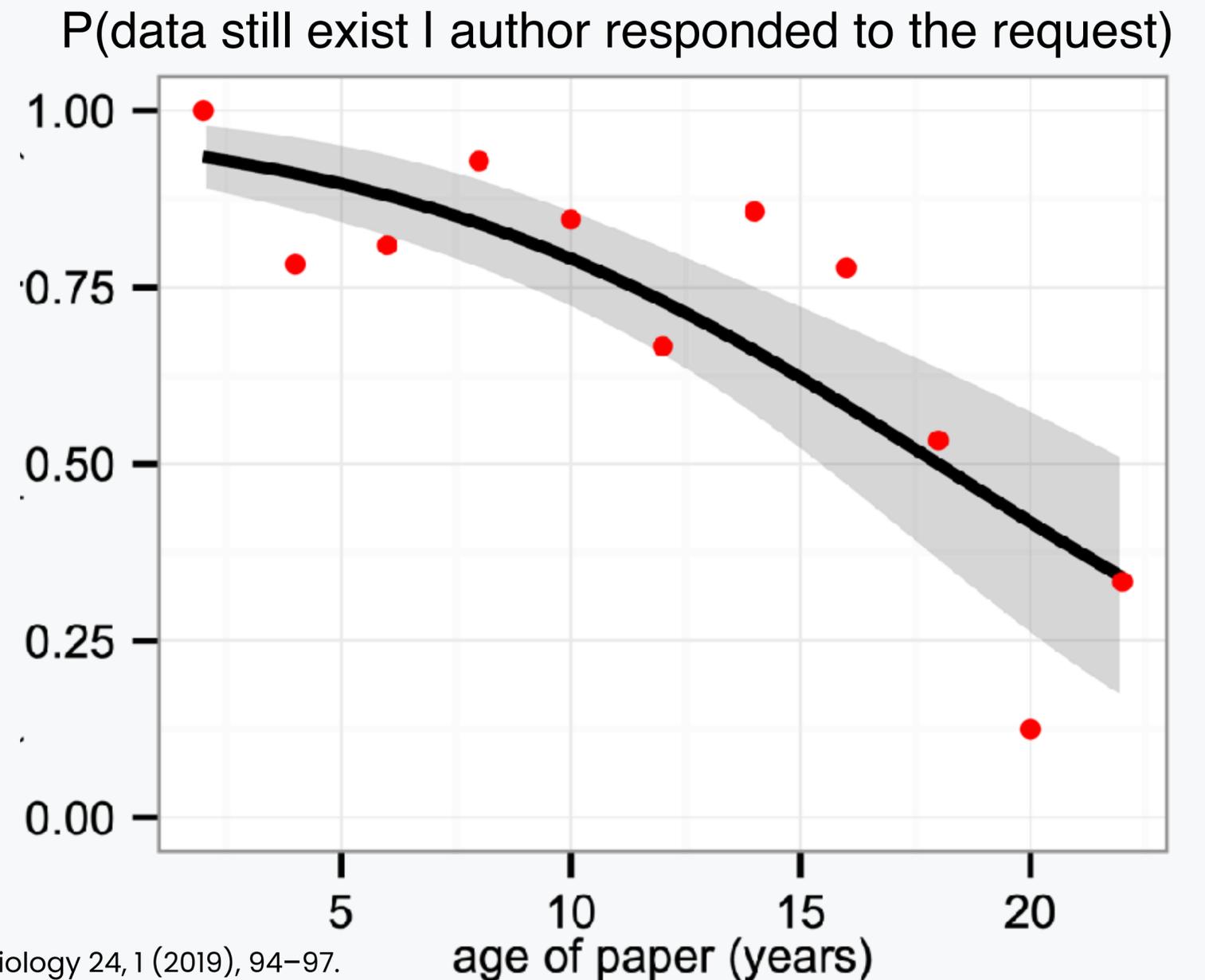
27% success rate



“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

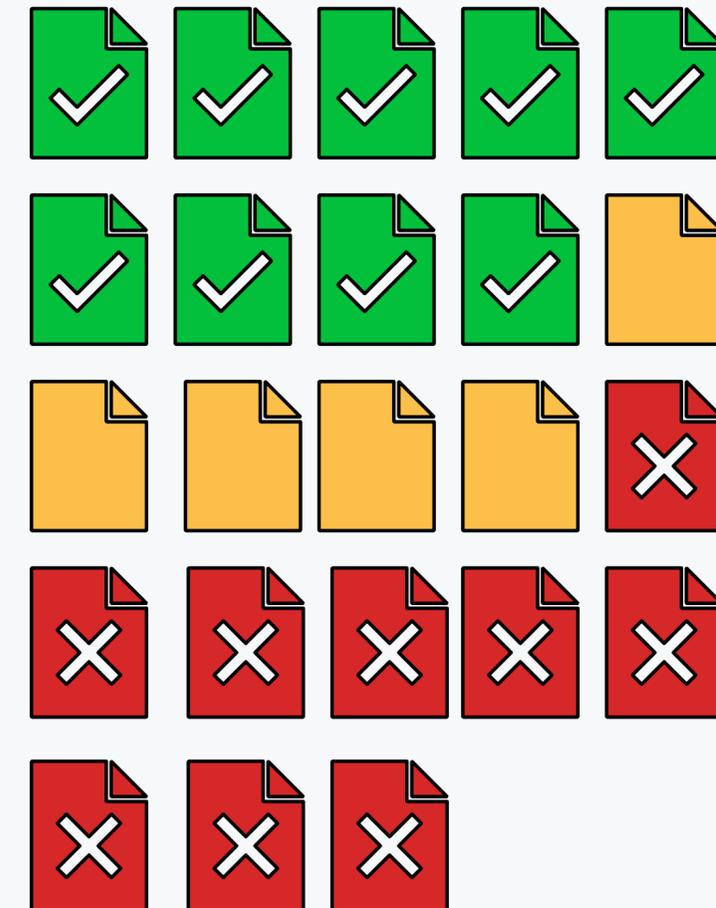


“If researchers want to use data or code from my paper, they can contact me”

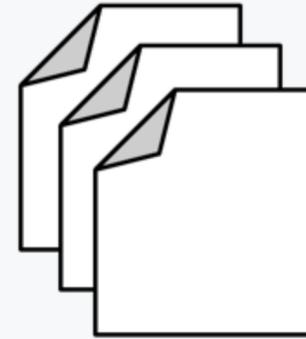
A meta-analysis of highly-cited HCI text-entry experiments

23 papers published in 2008–2020

- 9 complete information for meta analysis
- 5 provided information upon email request
- 9 no/partial responses



HCI research materials



Paper

Study
plan

Stimuli

Interview
scripts

Data

Prototypes

Software
code

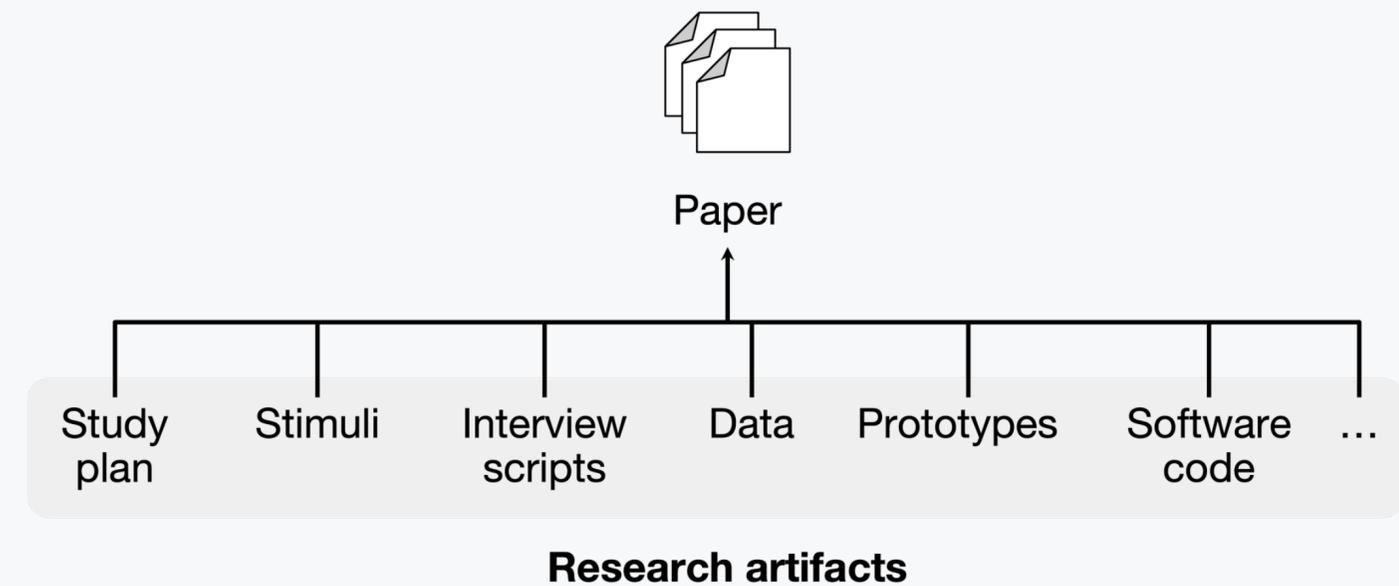
...

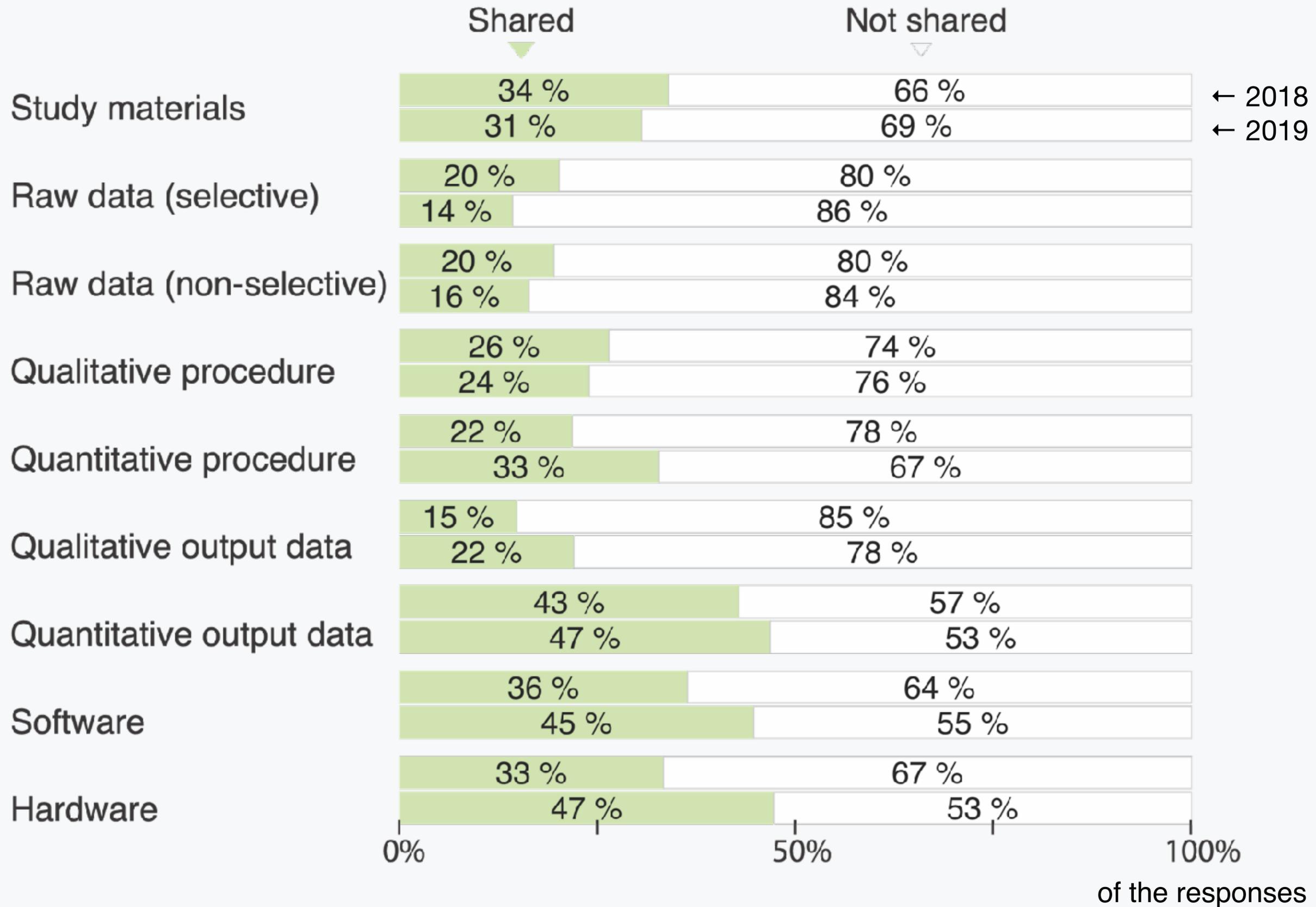
Research artifacts

We asked authors of CHI 2018–19

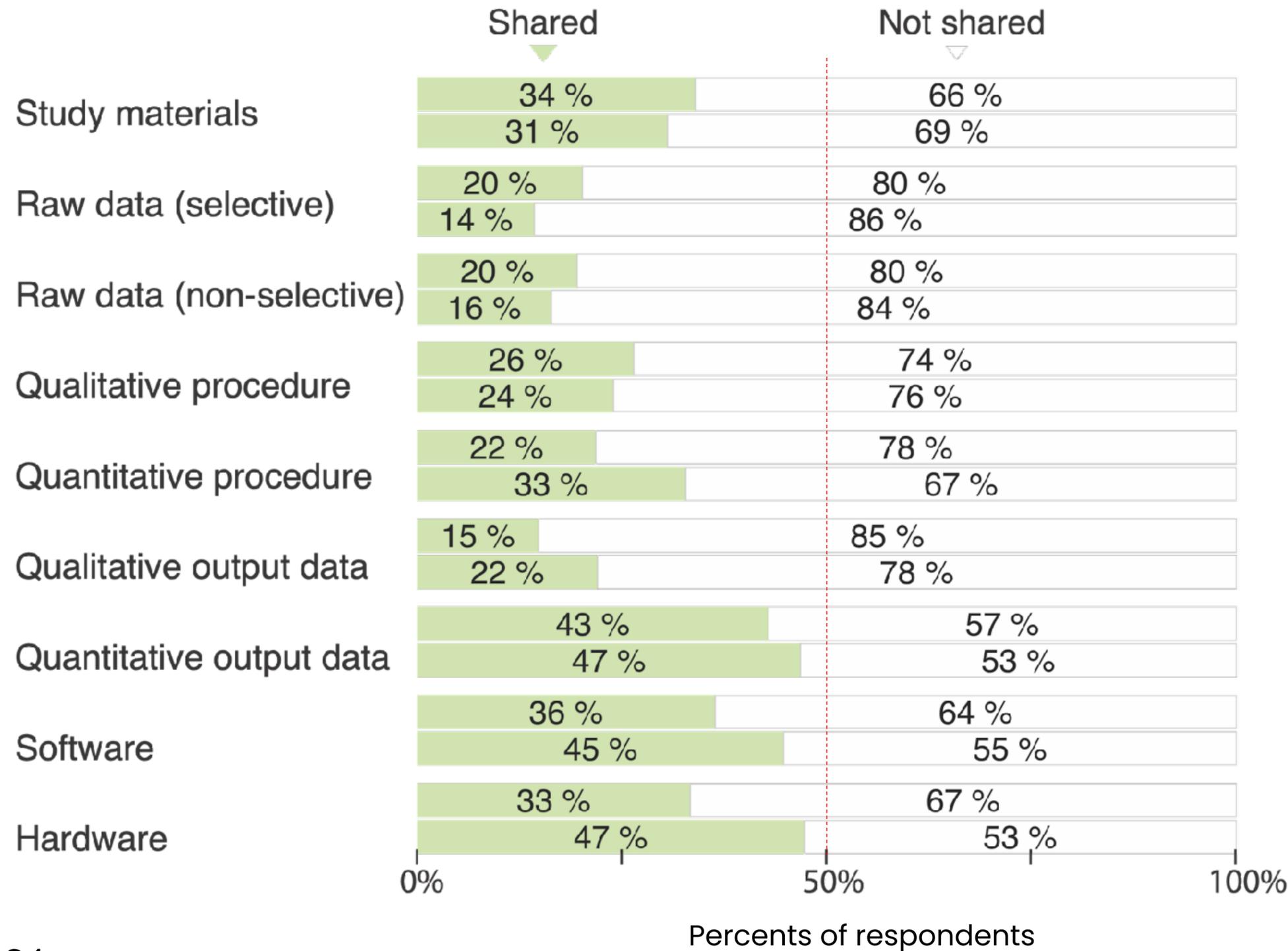
- What types of research artifacts they generate?
- If they share it, how?
- If not, why?

460 responses out of 1356 invitations
34% response rate! (quite high)



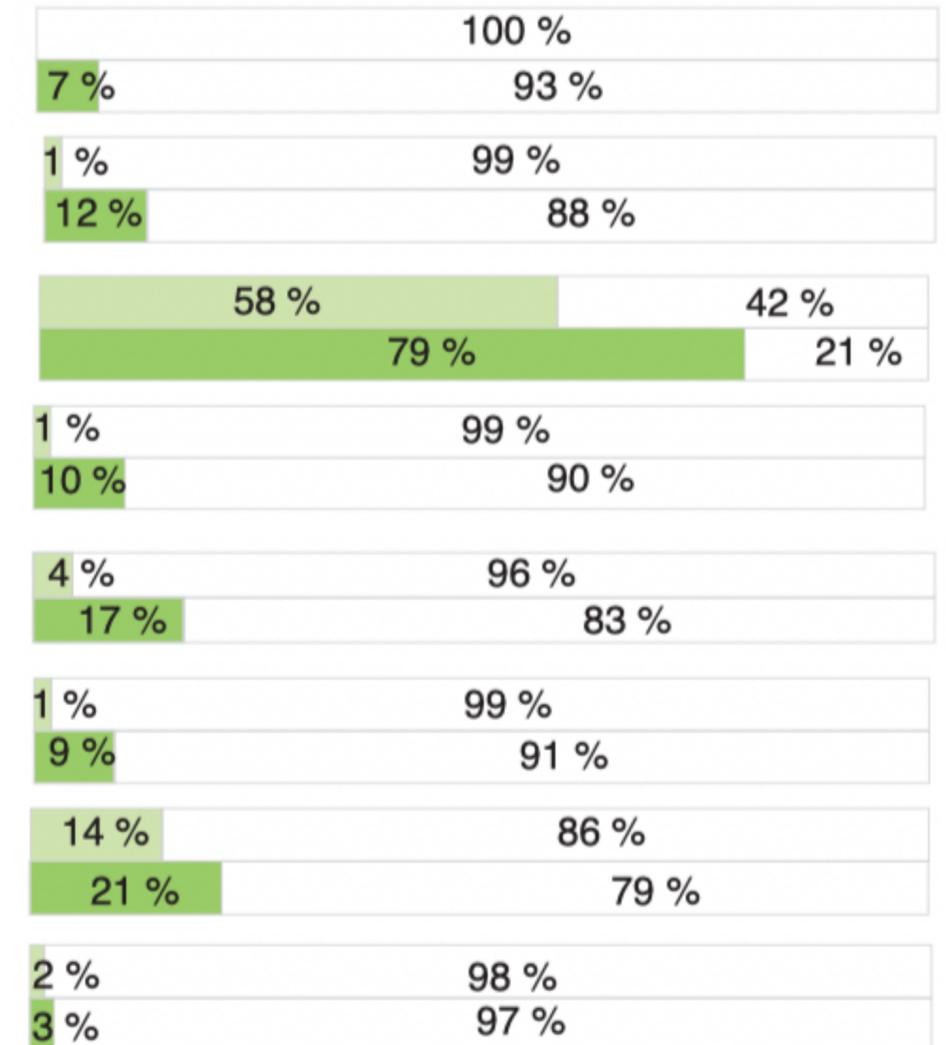


Survey to **authors** of CHI 2018, 2019



Content **analysis of papers** from
CHI 2017, 2022

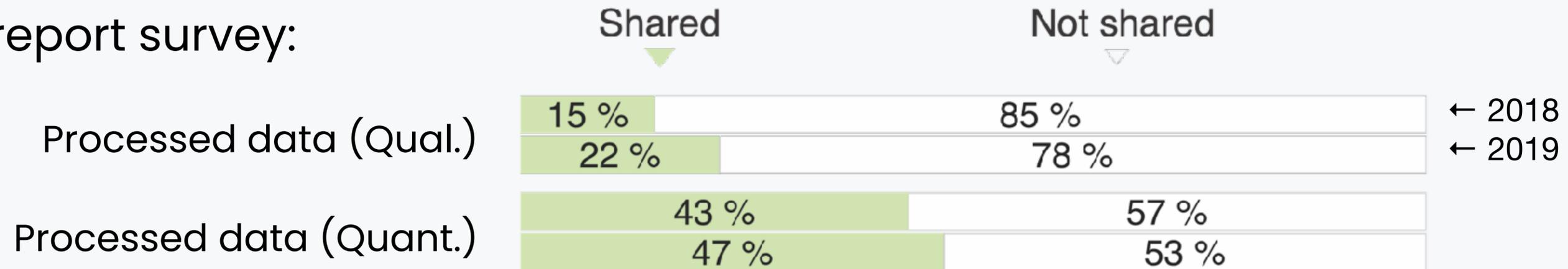
Mapped to equivalent categories on the left



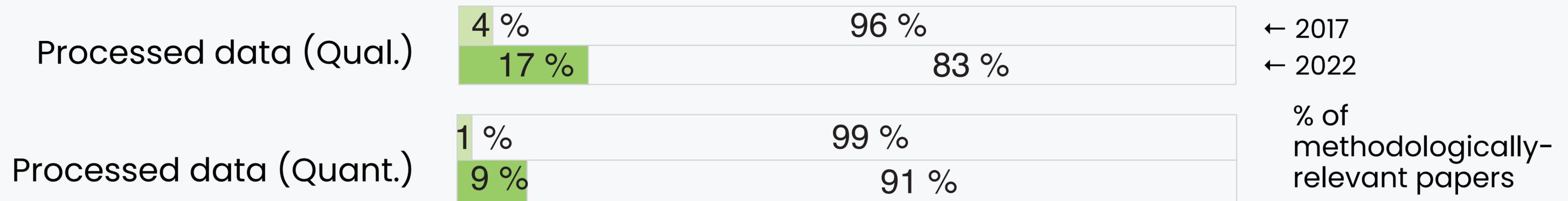
Percents of sampled papers

Have CHI papers improved in transparency?

Self-report survey:

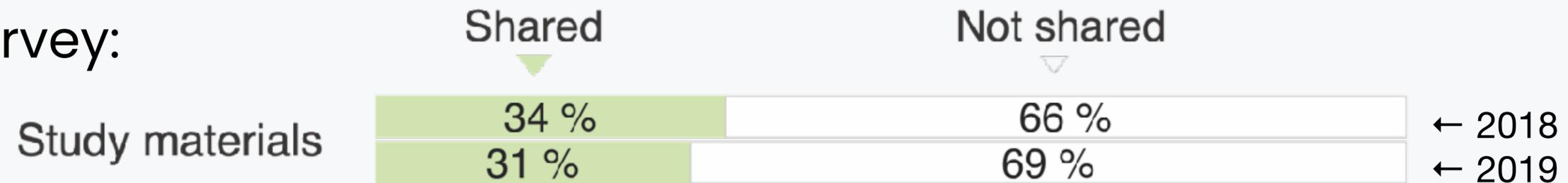


Content of sampled papers:

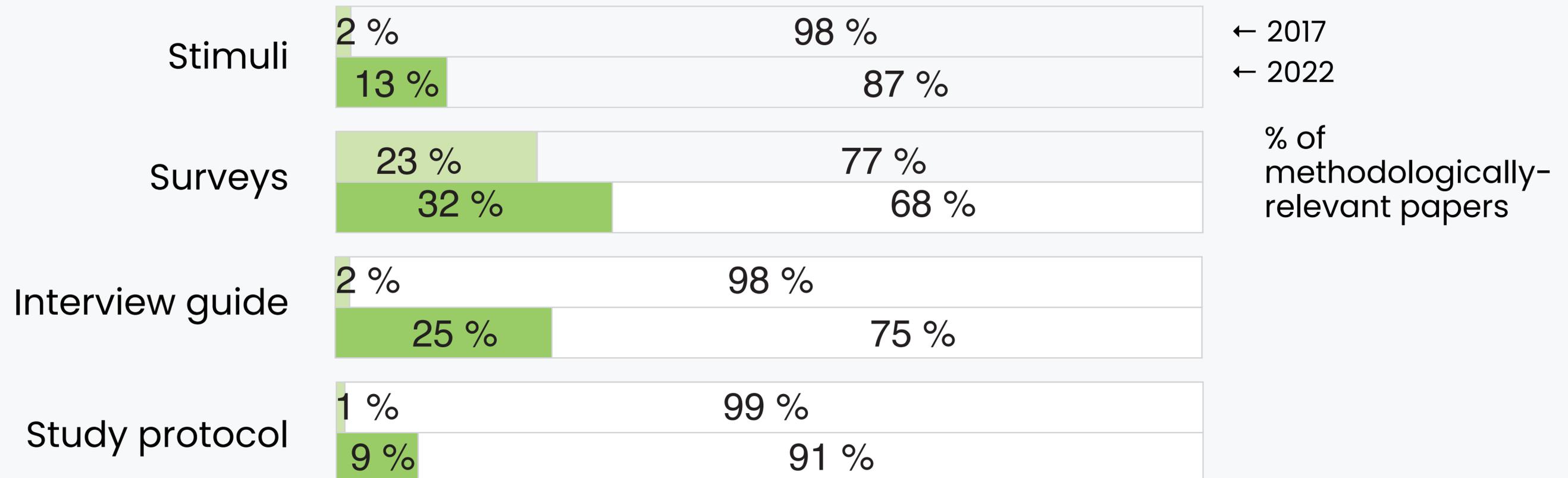


Have CHI papers improved in transparency?

Self-report survey:

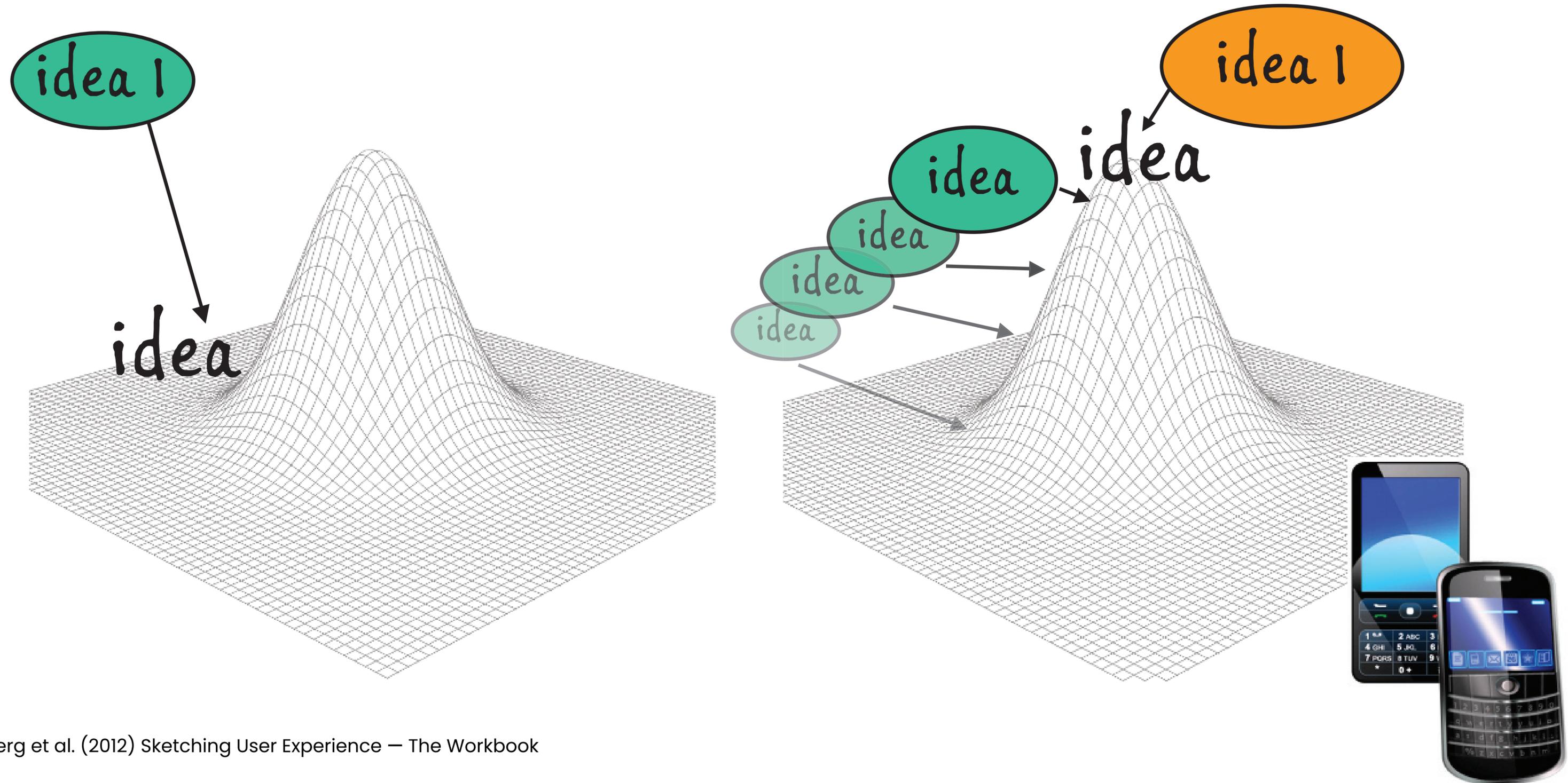


Content of sampled papers:

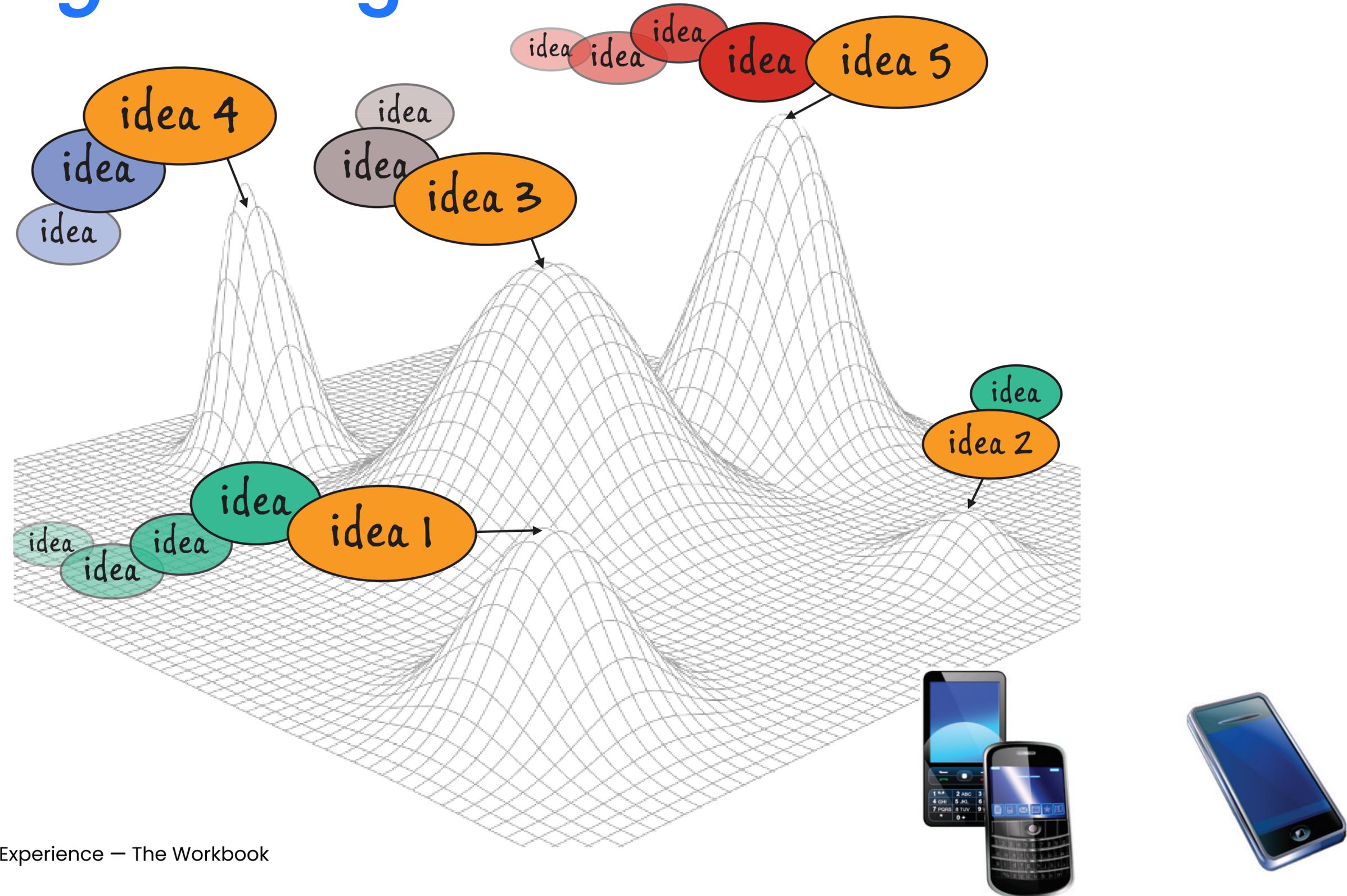


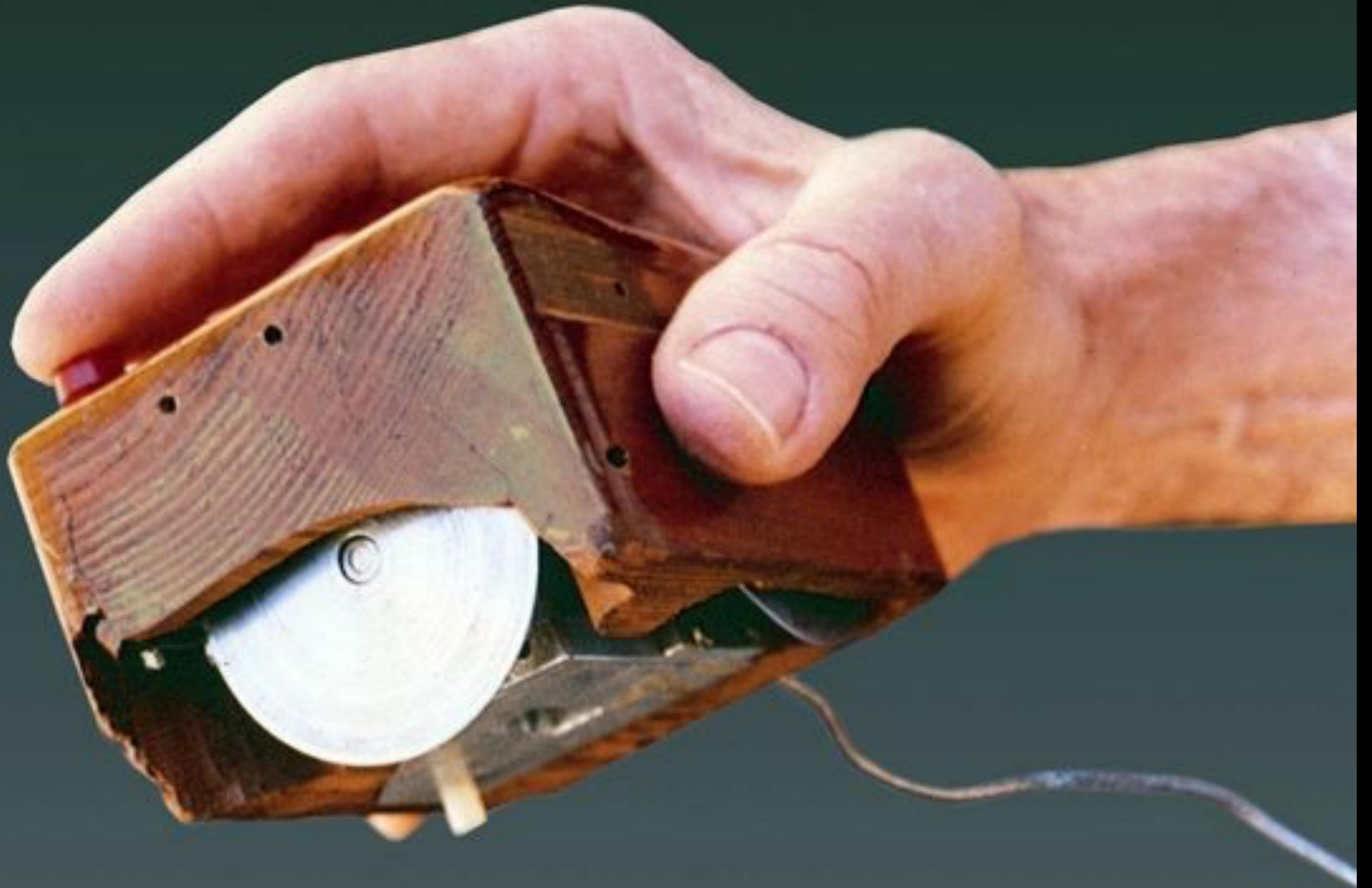
Challenges in motivating Research transparency in Human-Computer Interaction

Getting the design right



Getting the *right* design





Douglas Engelbart's prototype of a mouse

Image: SRI International

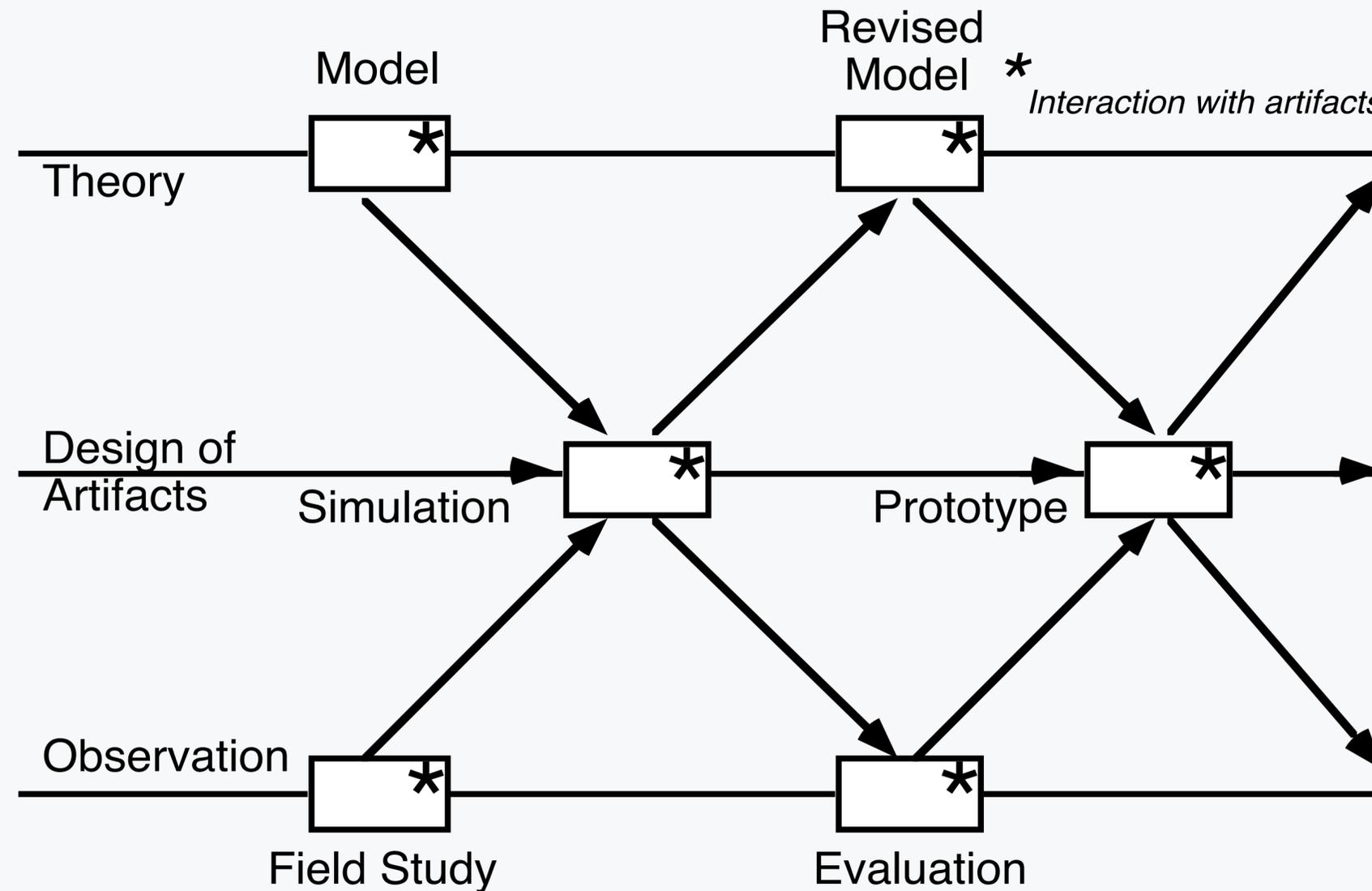


Prototypes of Microsoft Touch Mouse.

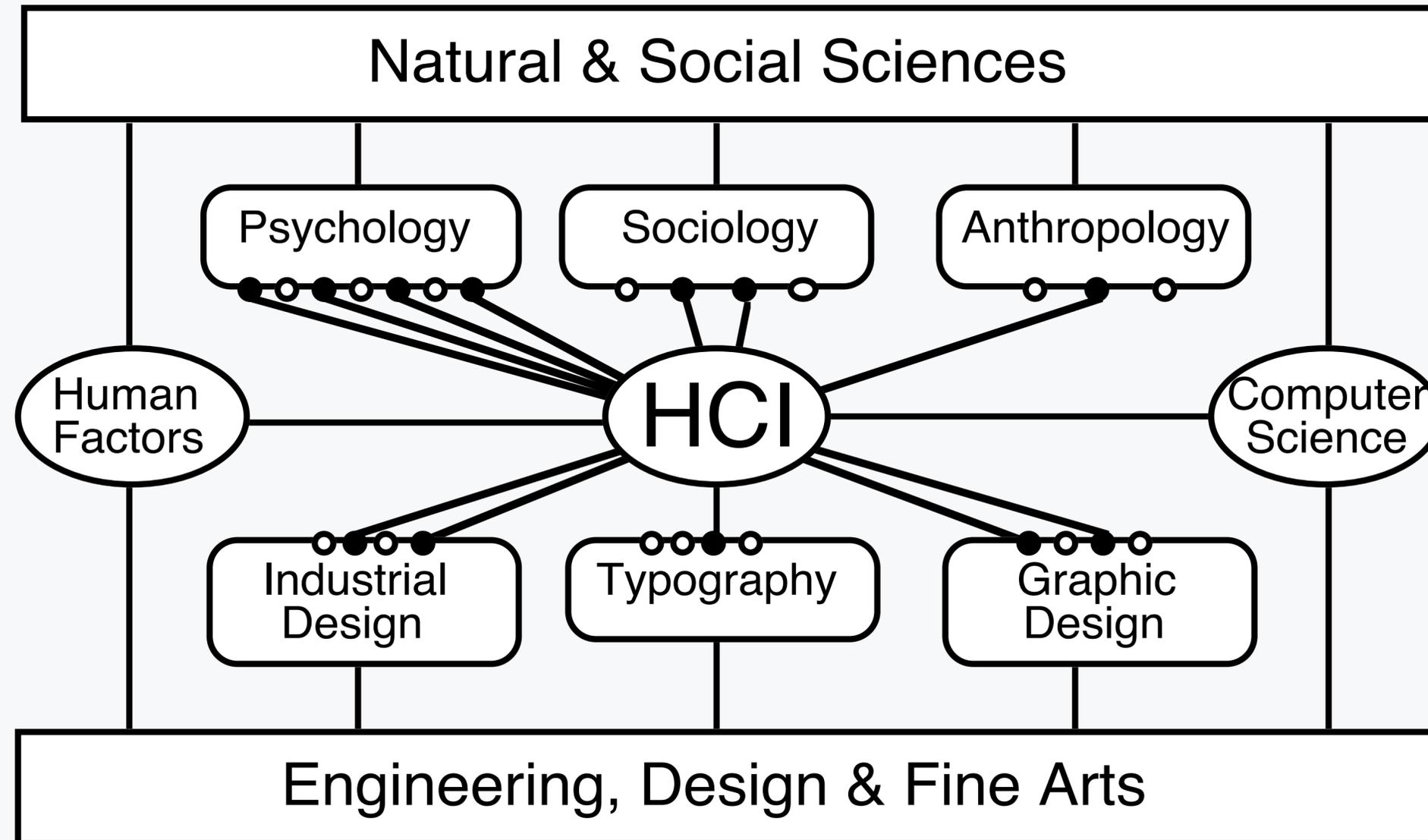


Should research transparency require making all these prototypes available?

HCI integrates science and design to create knowledge

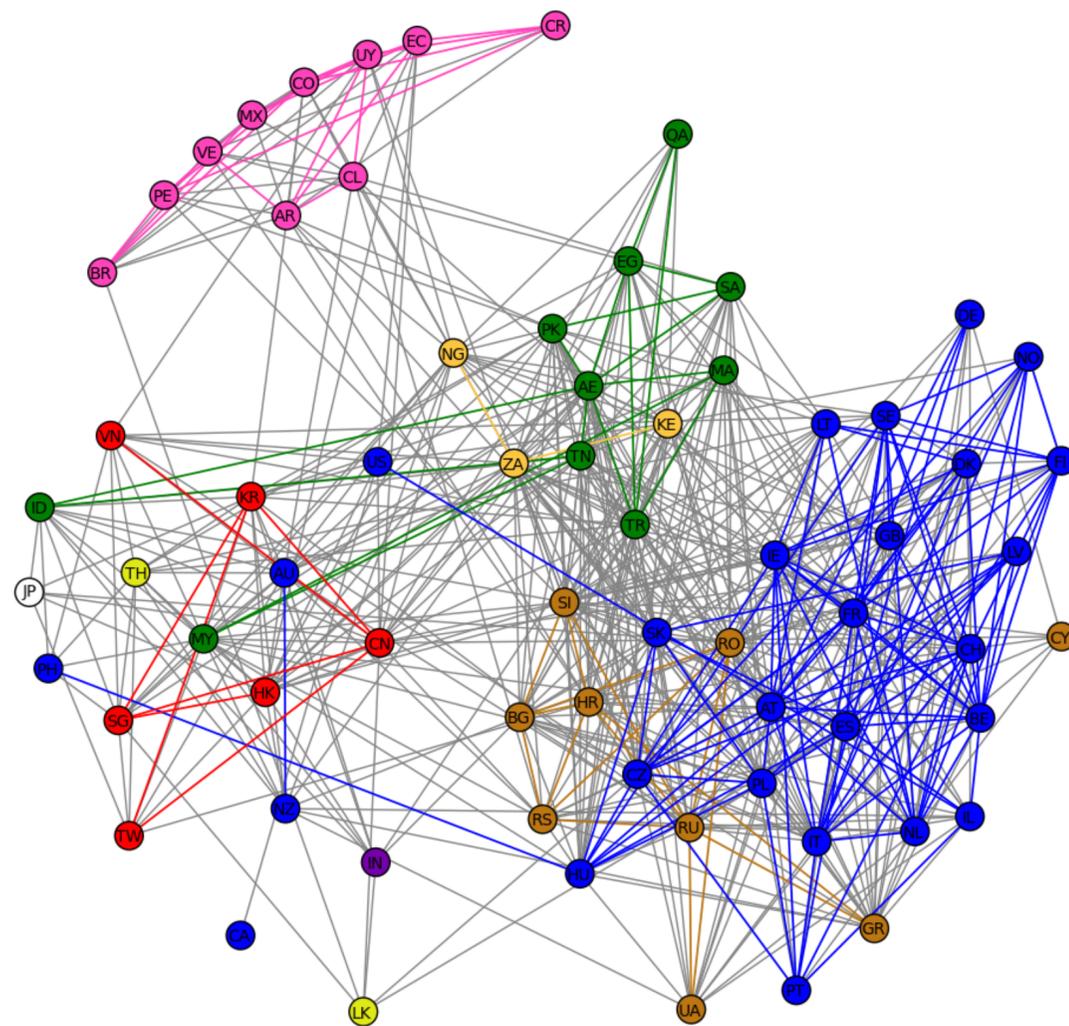


The field of Human-Computer Interaction

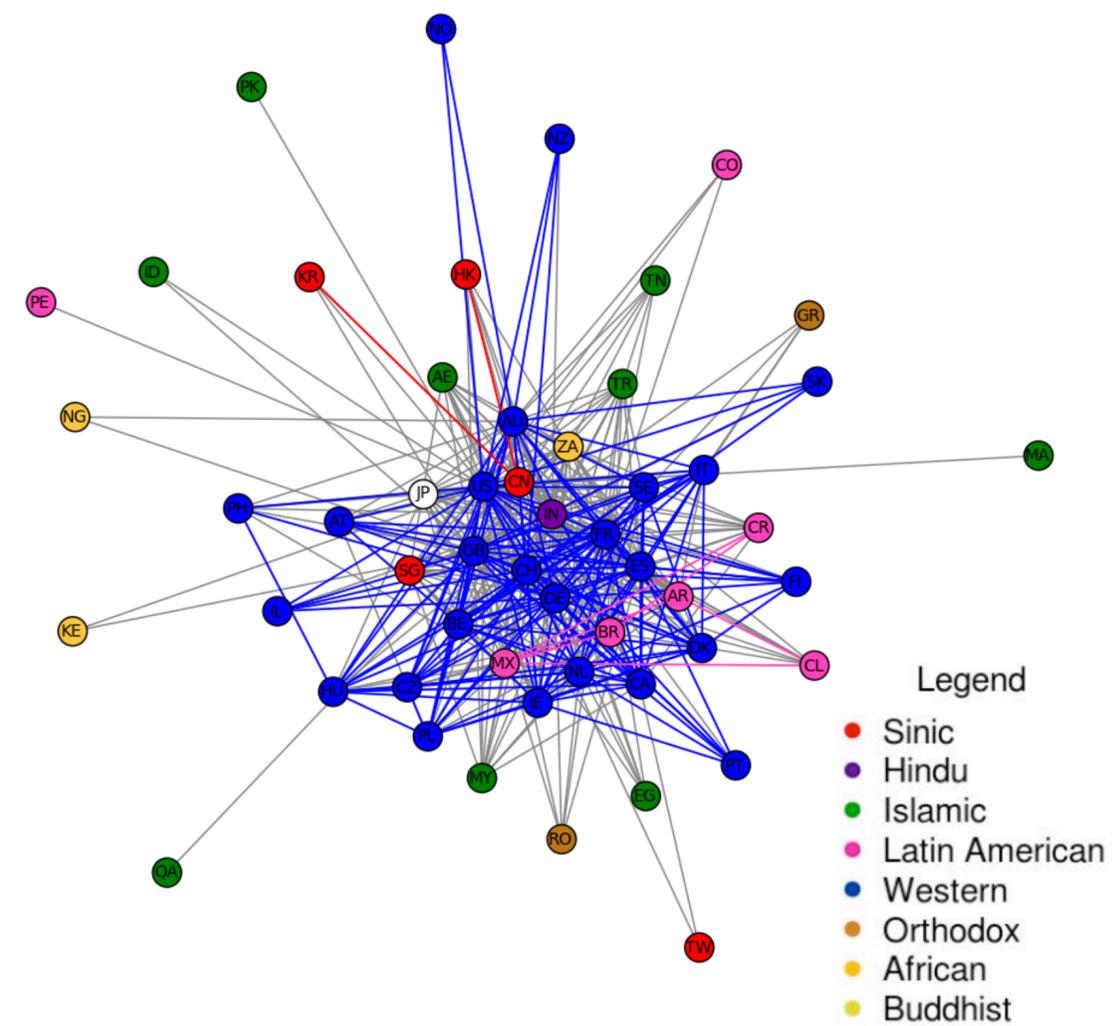


Social network use in a large multi-national company

Company-wide

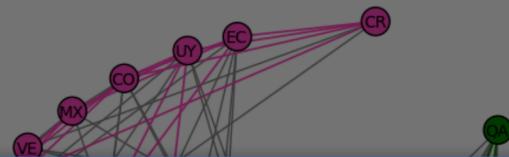


Sales workgroup



Social network use in a large multi-national company

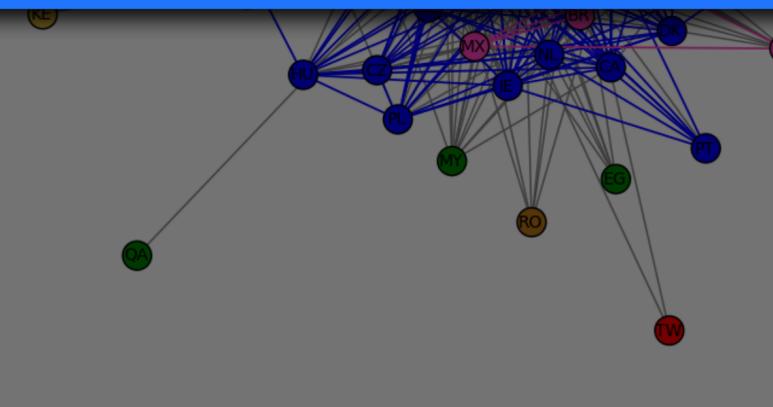
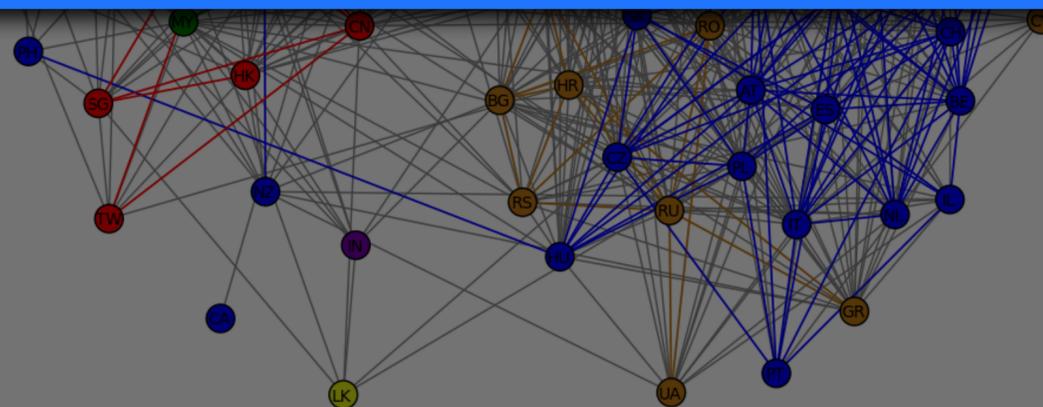
Company-wide



Sales workgroup



If sharing data is mandatory,
would this paper have been published?



- Legend
- Sinic
 - Hindu
 - Islamic
 - Latin American
 - Western
 - Orthodox
 - African
 - Buddhist

Inquiry paradigms

Positivism

Apply methods from natural science (posit, observe, derive logical **truth**)

Post-positivism

All methods are imperfect for apprehending reality; At best, research can identify a **valid belief**

Constructivism et al.

Meaning is in humans' construction of reality or in the human mind. Research could only arrive at **truth claims**

Critical theory et al.

Research & theory should be used to understand power relations and **change situations**

For more in-depth discussion on paradigms, epistemology, ontology

Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation biology*, 28(5), 1167-1177.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.

Identifying opportunities for design for women in patriarchal societies

- Interview and focus group studies; 90 participants in Bangladesh

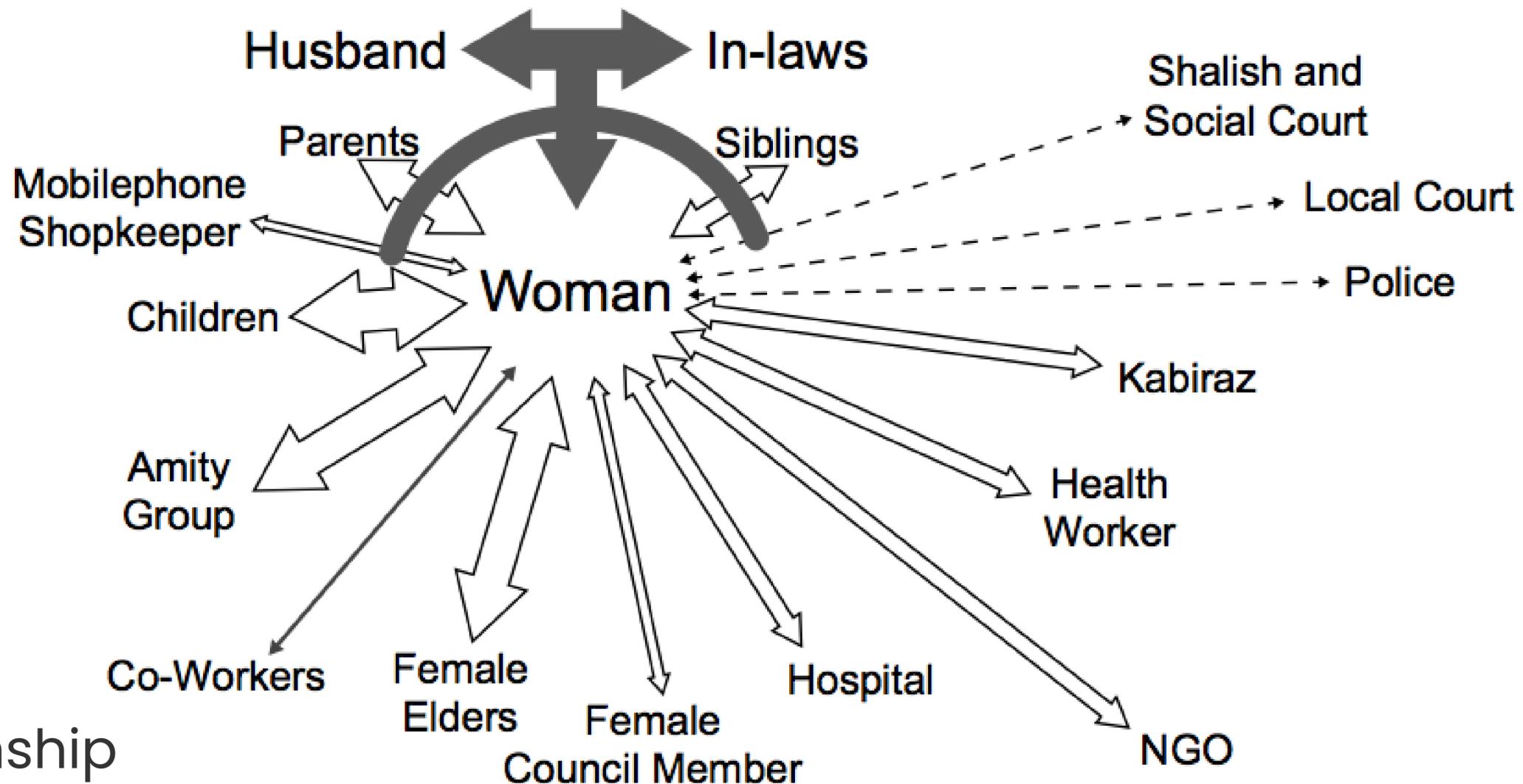
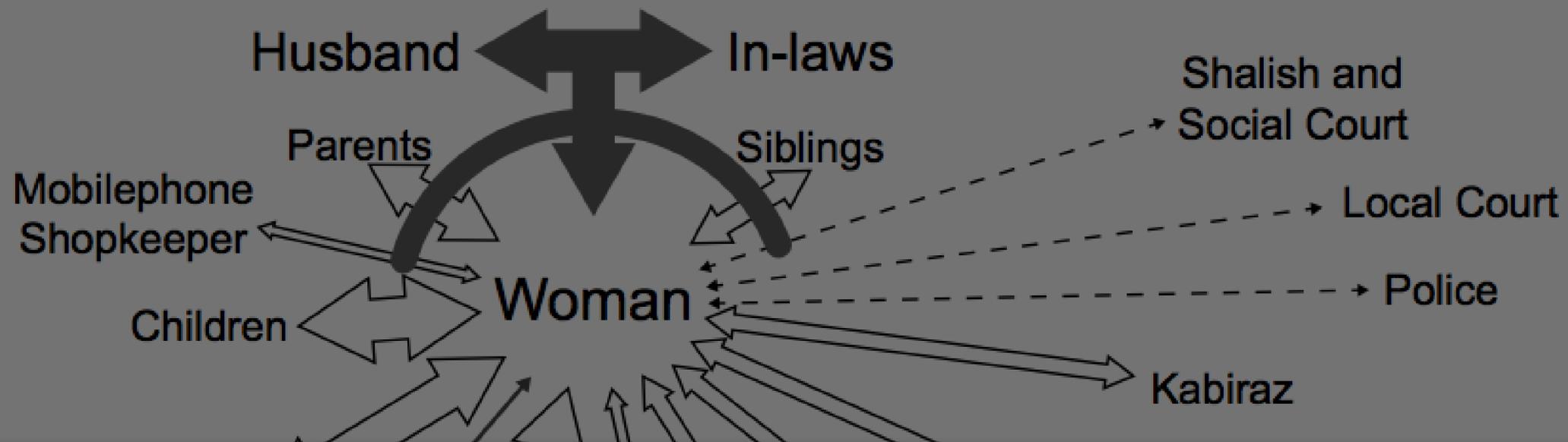


Diagram arrows

- Width: Frequency
- Length: Intimacy
- Grey arc: meditated relationship

Identifying opportunities for design for women in patriarchal societies

- Interview and focus group studies; 90 participants in Bangladesh



If researchers were all Caucasian male, can we trust the research findings?

- Length: Intimacy
- Grey arc: meditated relationship

Identifying opportunities for design for women in patriarchal societies

- Interview and focus group studies; 90 participants in Bangladesh

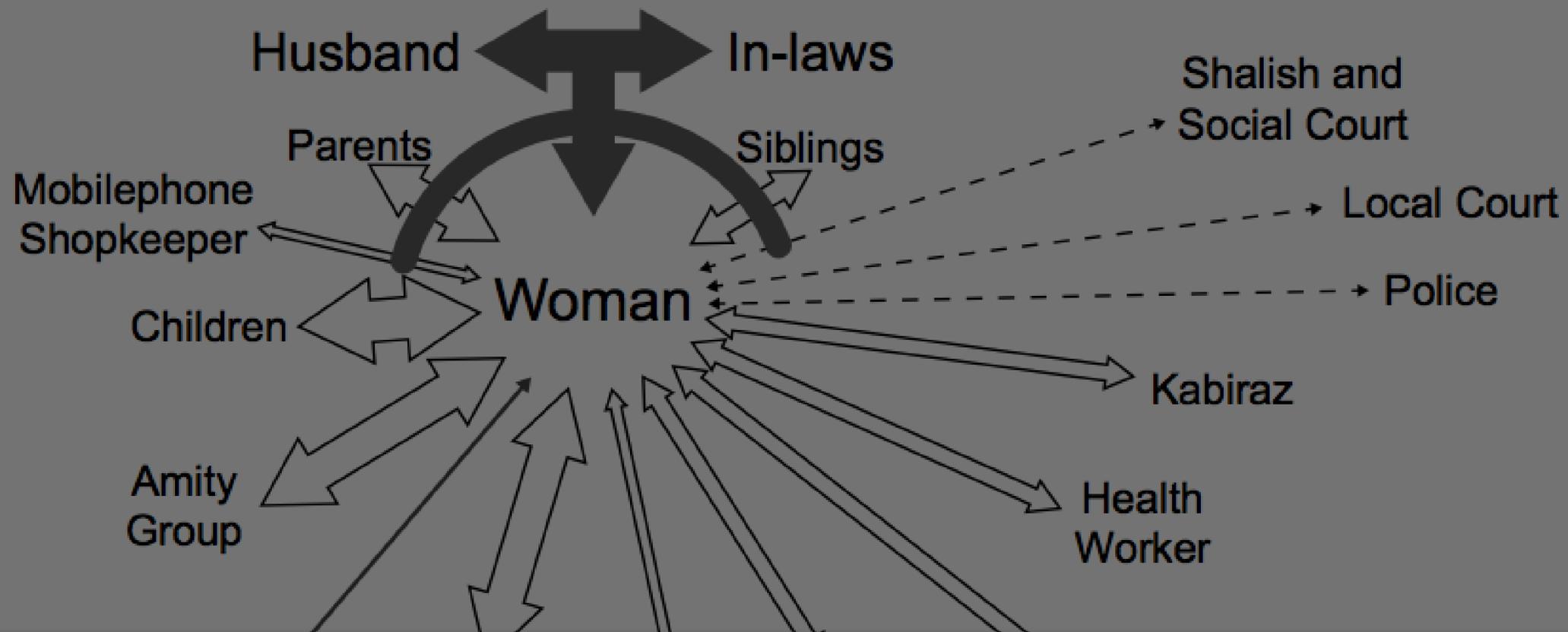


Diagram arrows

- Width: Frequency

If the funding body mandate that all research data has to be public,
is it safe to conduct this work?

Inquiry paradigms

Positivism

Apply methods from natural science (posit, observe, derive logical **truth**)

Post-positivism

All methods are imperfect for apprehending reality; At best, research can identify a **valid belief**

Constructivism et al.

Meaning is in humans' construction of reality or in the human mind. Research could only arrive at **truth claims**

Critical theory et al.

Research & theory should be used to understand power relations and **change situations**

For more in-depth discussion on paradigms, epistemology, ontology

Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation biology*, 28(5), 1167-1177.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.



Teddybear blood radio by Fiona Raby and Anthony Dunne
Image [UK Science Museum Group](#)



What research materials could have been transparent in this project?
How relevant to research process and outcome are those materials?

PSYCHOLOGY

Human factors

Engineering

Information science

Management information system

Computer Science

Office information system

Design

Commercial products

Replicability^[1]

Closely matched method

+

New data

=

Consistent results

[1] Replicability definition is 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.

Replicability^[1]

Closely matched method

+

New data

=

Consistent results



RepliCHI award

- RepliCHI movements in meetings and workshops [2]
- RepliCHI award implemented in CHI Conference 2023 by the conference organizers that year
- Received push-backs
- How about research where replicability is irrelevant?

[1] Replicability definition is according to Clearbout terminology, which differs from ACM's terminology See discussion in Plesser HE (2018) [Reproducibility vs.](#)

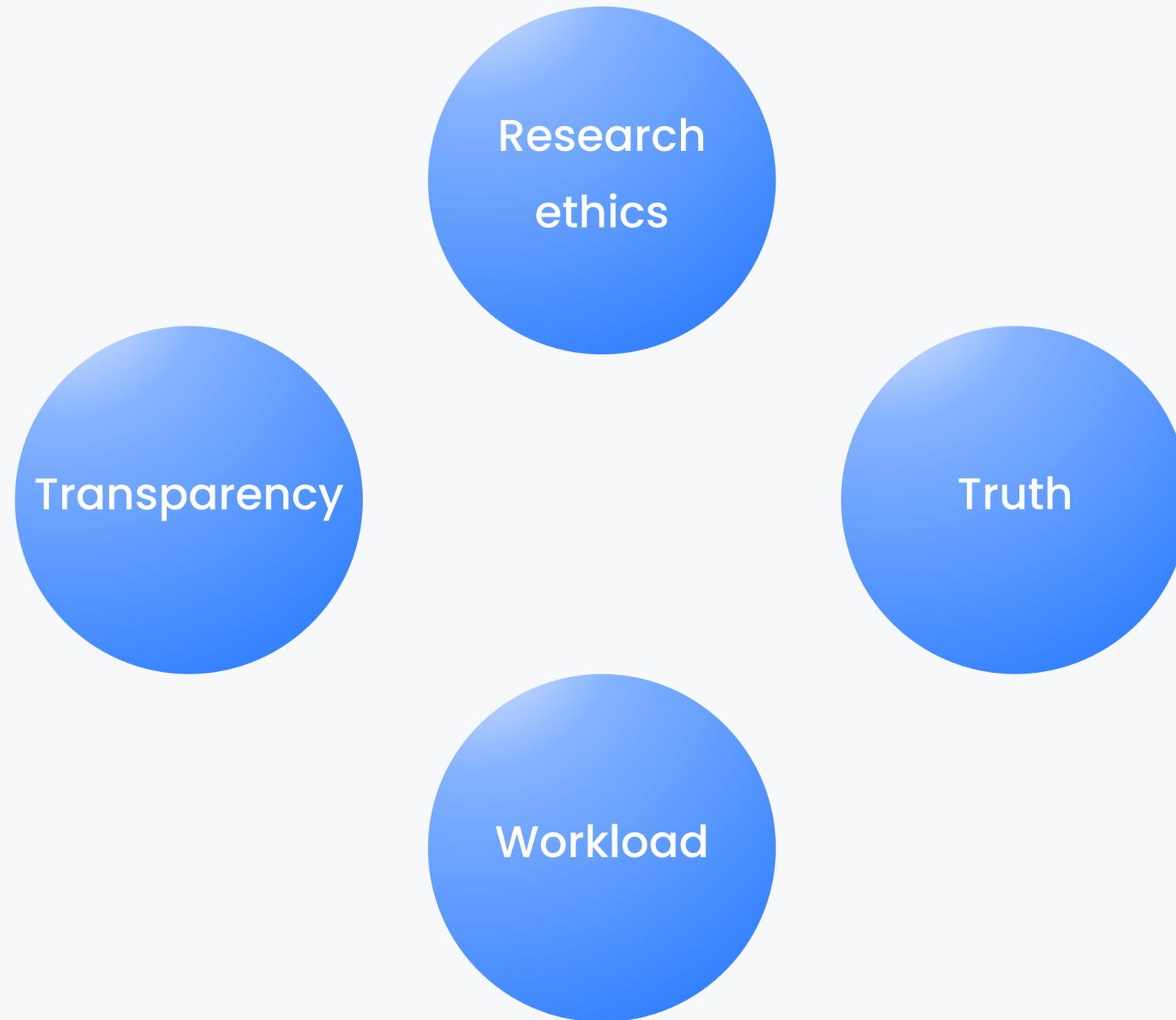
[Replicability: A Brief History of a Confused Terminology](#). Front. Neuroinform. 11:76.

[2] Starting in Wilson et al. (2011) [RepliCHI - CHI should be replicating and validating results more: discuss](#)

Scenarios

- You are a reviewer of a paper. The paper looks fine. But you looked into the statistical analysis code and found an error. What will you do?
- You are an editor who solicited reviews for Paper A from a panel of reviewers. A reviewer pointed out that an error in statistical analysis in the code—whereas it looks fine in the paper. This observation came in the very last minute in the reviewing process. What will you do?
- Continuing the previous point: At the present, there is no expectation of sharing code and data in your field. Paper B is in the same area, using the same statistical analysis; it also looks fine. But Paper B does not share their statistical analysis code. What will you do?

A balancing act?



Why motivating transparency is not straight-forward

Established researchers

Resourceful institutes

Transparency-aware
reviewers

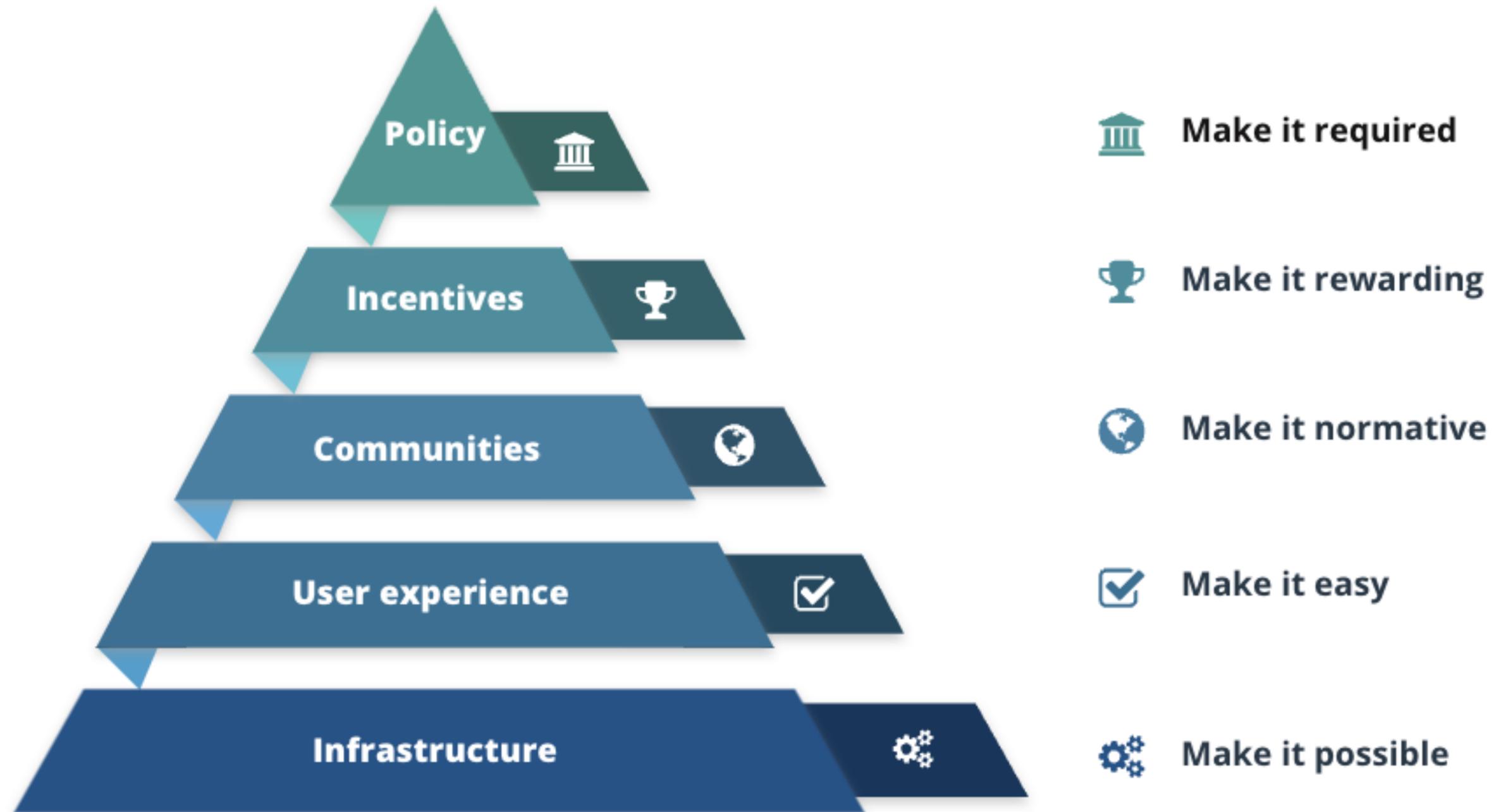


Junior researchers

Institutes in low-
income countries

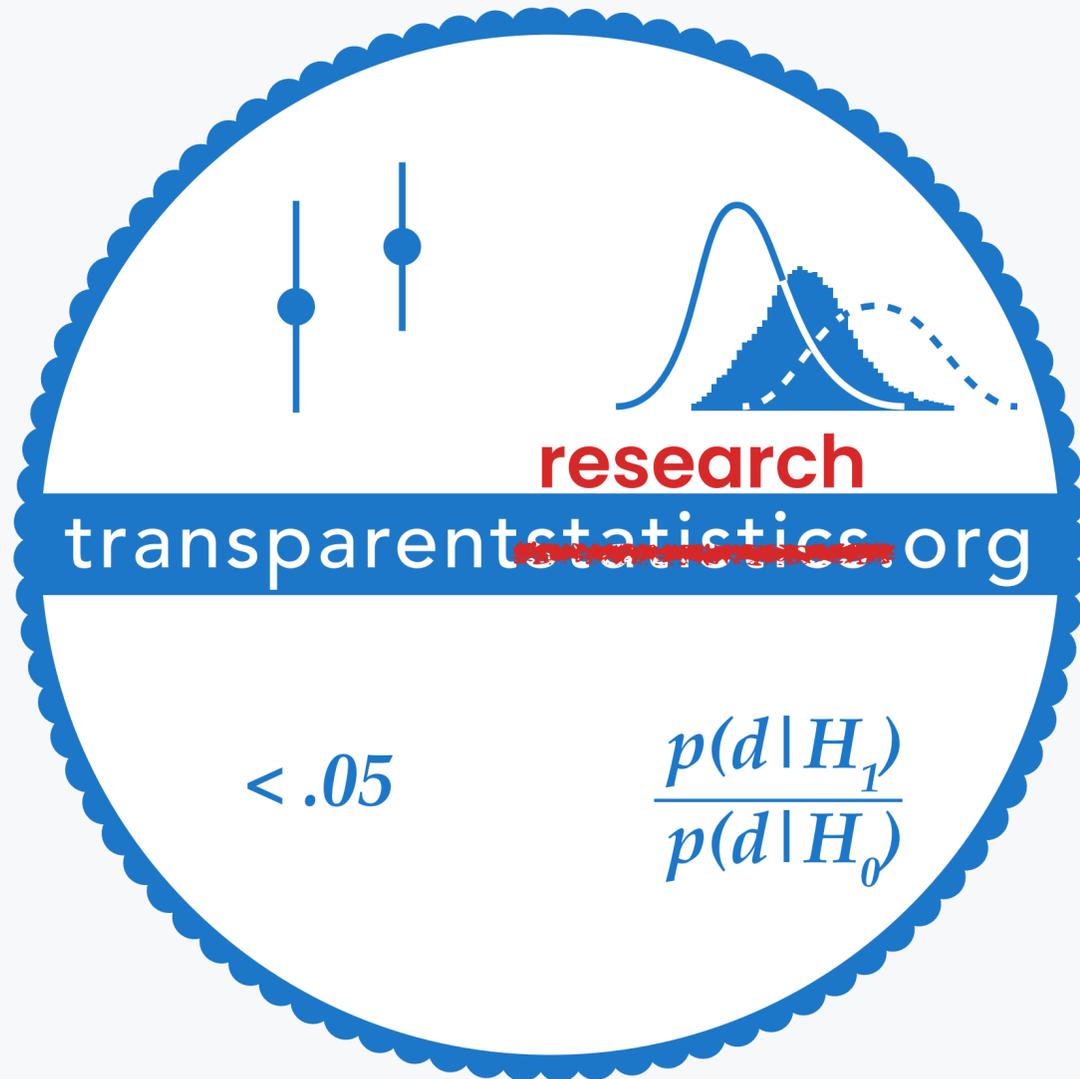
Transparency-unaware
reviewers

Motivating research transparency in HCI



Progressive process of research transparency

- **Become aware** and set realistic expectations
- **Learn** from materials and examples in- and outside of your field
- **Adopt** step-by-step: “be better in your next paper”
- **Educate** your peers and students
- **Influence** your institutions downward and upward



Transparent Statistics Guidelines

Transparent Statistics in HCI Working Group (<http://transparentstatistics.org/>)

2019-06-29

Preface

This document grew out of our community-led effort to improve the [Guide to a Successful Submission](#) on the CHI conference website. The proposal was accepted and integrated into the current guidelines. The recommended changes are available in this version [10.5281/zenodo.105281](https://zenodo.org/record/105281).

Transparency

CHI papers should strive for transparency in their type and methodology. Different types of quantitative studies, and quantitative data analysis, require different levels of transparency.

Contributions that are technical (e.g., new methods, contributions that are quantitative (e.g., new data sets, analyzed results) are expected to be available to rerun the interactive experiments.

Proposal for amending CHI guides for authors and reviewers

The text below was a community-led effort to improve the [Guide to a Successful Submission](#) on the CHI conference website. The proposal was accepted and integrated into the current guidelines. The recommended changes are available in this version [10.5281/zenodo.105281](https://zenodo.org/record/105281).

SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

September 25 – 30, 2022, Dagstuhl Seminar 22392

Transparent Quantitative Research as a User Interface Problem

Organizers

- Kasper Hornbaek (University of Copenhagen, DK)
- Yvonne Jansen (Sorbonne University – Paris, FR)
- Amelia A. McNamara (University of St. Thomas – St. Paul, US)
- Judy Robertson (University of Edinburgh, GB)
- Chat Wacharamanotham (Universität Zürich, CH)

Motivation

Many research fields are currently rethinking their research methods towards more transparent practices. The most rapid progress towards transparency can be observed in fields that were heavily affected by a replication crisis, like psychology, while changes are slower and receive more resistance in interdisciplinary fields, such as human-computer interaction (HCI) and visualization (VIS). In this Dagstuhl Seminar, we want to address the issue of hesitant adoption of transparent research methods by framing it as a user interface problem: the 'interface' of using transparent methods is ill-



and many more...

Tutorials at CHI 2022, 2023, MuC 2023, VIS 2023

- **Planning** research and **sharing** research artifacts
- Introduction to Transparent **Bayesian Data Analysis**
- Introduction to Transparency-oriented **Visualizations**
- **Hands-on exercises** with transparency practices

Materials available for reuse <https://osf.io/27r5z>





Journal of Visualization and Interaction

- Registered reports
- Open access (Diamond)
- Novel article formats
- Open, continuous review



journalovi.org

Organizers: Lonni Besançon, Florian Echtler, Matthew Kay, and Chat Wacharamanotham.

Gatherplot: A Non-Overlapping Scatterplot

</> CODE

Table of contents

AUTHORS

Deokgun Park 

Sung-Hee Kim 

Niklas Elmqvist 

AFFILIATIONS

University of Texas at Arlington

Dong-eui University

University of Maryland, College Park

 **Under Review**

This paper is under review on the experimental track of the [Journal of Visualization and Interaction](#).

Abstract

Introduction

Scatterplots are a common tool for exploring multidimensional datasets, especially in the form of scatterplot matrices (SPLOMs). However, scatterplots suffer from overplotting when categorical variables are mapped to one or two axes, or the same continuous variable is used for both axes. Previous methods such as histograms or violin plots use aggregation, which makes brushing and linking difficult.

Conclusion

We propose gatherplots, an extension of scatterplots to manage the overplotting problem. Gatherplots are a form of *unit visualization*, which avoid aggregation and maintain the identity of individual objects to ease visual perception. In gatherplots, every visual mark that maps to the same position coalesces to form a packed entity, thereby making it easier to see the overview of data groupings. The size and aspect ratio of marks can also be changed dynamically to make it easier to compare the composition of different groups. In the case of a categorical variable vs. a categorical variable, we propose a heuristic to decide bin sizes for optimal space usage. Results from a crowdsourced user study show that gatherplots enable people to assess data distribution more quickly and more correctly than when using jittered scatterplots.

Materials

Source code for Gatherplots can be found at <https://github.com/intuinno/gatherplot>. Research materials associated with the crowdsourced user study can be found on OSF at <https://osf.io/bk9cx/>.

Data Collection

We conducted a crowdsourced user study on [Amazon Mechanical Turk](#) involving participants drawn from the general population. We collected completion time, accuracy, and confidence for five different retrieval, ranking, and comparison tasks under four conditions: scatterplots with jittering, gatherplots with absolute mode, gatherplots with normalized mode, and gatherplots with a toggle to switch between absolute and normalized mode.

Data Analysis

1 Introduction

2 Background

3 The Gather Transformation

4 Gatherplots: A 2D Gathering Representation

5 Evaluation

6 GatherLens: A Gathering Magic Lens

7 Conclusion and Future Work

References

Research Material Statements

Authorship

License

Conflict of Interest

is:issue is:open

Labels 11

Milestones 1

New issue

9 Open 3 Closed

Author Label Projects Milestones Assignee Sort

[REVIEW] Gatherplot Review

#16 opened 4 days ago by jov-anonymous-reviewer-AAAA 3 of 8 tasks

[REVIEW] Gatherplot Review 2

#15 opened on Aug 20 by facet-fan 3 of 8 tasks

Opportunity to capitalize on interactivity to showcase GatherLens?

#13 opened on Aug 12 by mjskay

[REVIEW] [Accessibility] Gatherplots review

#10 opened on Aug 4 by domoritz 17 of 31 tasks

3

[REVIEW] Gatherplot review

#7 opened on Jul 12 by joviewer-xyz 2 of 8 tasks

Writing Your Review

JoVI reviews should be *thorough but kind*: aim for completeness and clarity of critique, but also think about how the authors will read your review (write the kind of review you would appreciate receiving). JoVI does not have a set rejection rate, so any work meeting our key criteria (above) should be a candidate for acceptance.

Modular Reviews

To facilitate understanding by the editor and responses by the authors, reviewers are encouraged to organize concerns into numbered sections. Each section should include:

1. A description of the concern
2. Where the issue occurs
3. Severity: Is this a threat to validity? Or is it a suggestion?
4. How to fix the issue, or how an editor would know if the issue has been fixed

Please put minor issues and typos into one "minor issues" section at the end of your review.

Intentional Citations

Each citation should have a clear purpose: to offload evidence or to build on existing progress.

- References should be used to explain the rationale, methods, and claims.
- Avoid telling authors to add references without stating what evidence or prior use that reference would support. A laundry list of references that "should be cited" is not helpful. If the reviewer wishes to suggest additional references, they must make a clear argument why.
- References can also be used to contrast the article with other very similar articles. However, this style of citation should be kept to a minimum and only used to clarify differences.
- A designated "Related Work" section is not required.
- Missing citations that could be easily remedied in revision is generally not grounds for rejection.

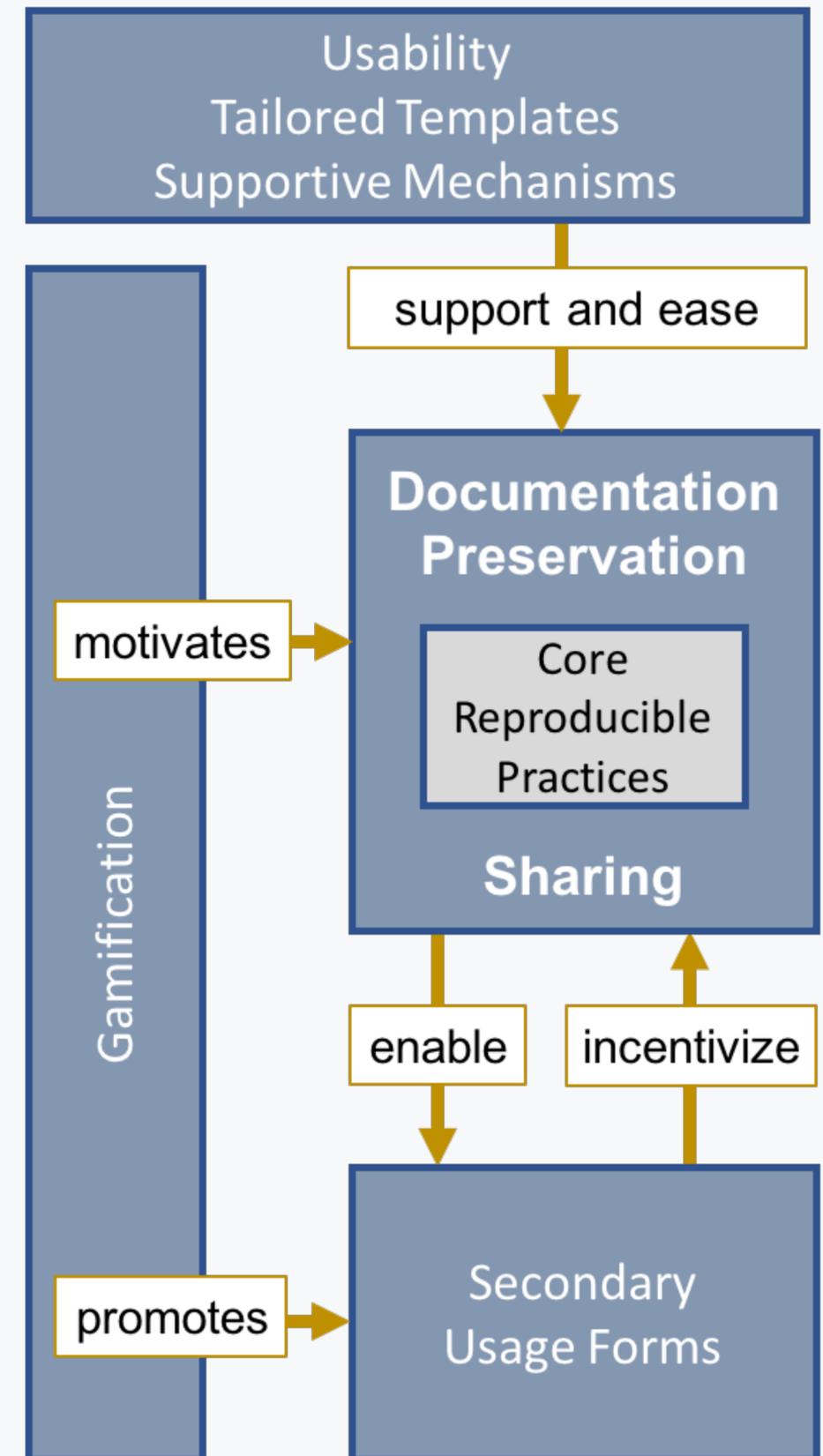
On this page

- [Before Agreeing to Review](#)
- [First Impression](#)
- [Key Assessment Criteria](#)
- [Writing Your Review](#)
- [Modular Reviews](#)
- [Intentional Citations](#)
- [Do's and Don'ts](#)
- [Registered Reports](#)
- [Possible Review Outcomes](#)

- [Edit this page](#)
- [Report an issue](#)

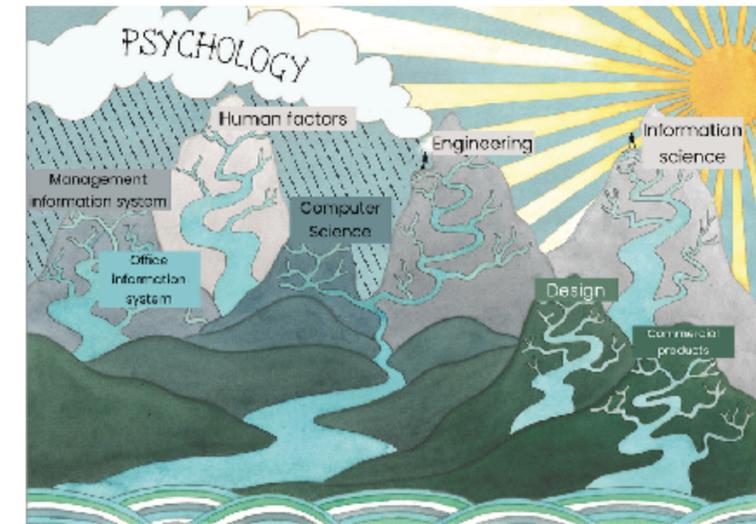
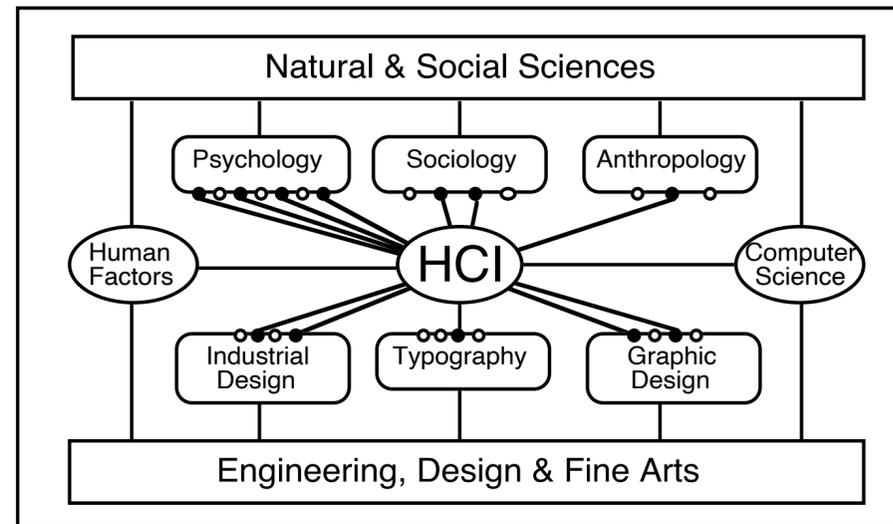
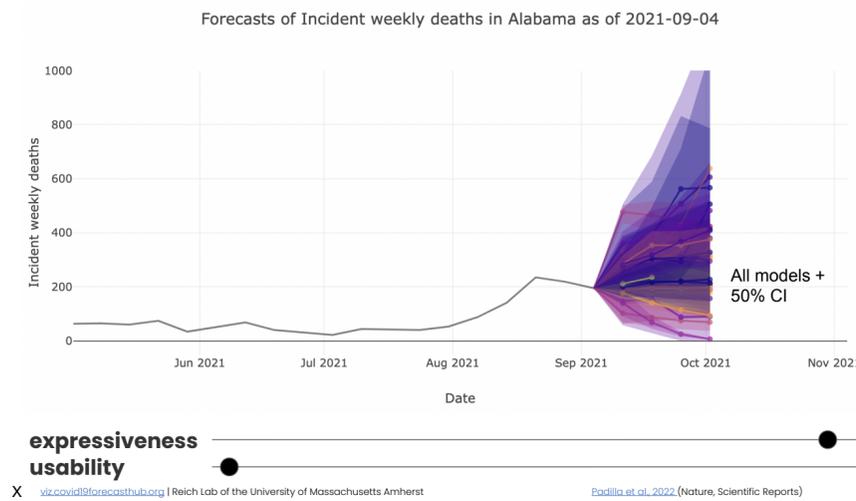
How HCI can contribute to Open Science

- **Mapping** and understanding current research practices
- **Lowering efforts** for Open Science practices
- **Ensuring usability** of services
- **Making connections** between researchers and technology developers



Opportunities for MPI policymakers

- Recognize Open Science as a group of phenomena that can be studied
- Recognize research into designing and developing infrastructure
 - Cross-cutting
 - Specific field
- Fund these research!



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

How could you be more transparent in your next paper?

How could you contribute to encourage your field to be more transparent in the long run?