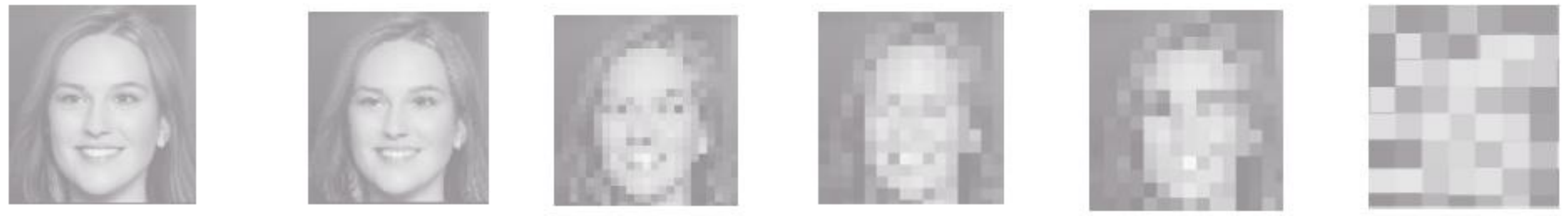
Exploring User-Suitable Metaphors for Differentially Private Data Analyses

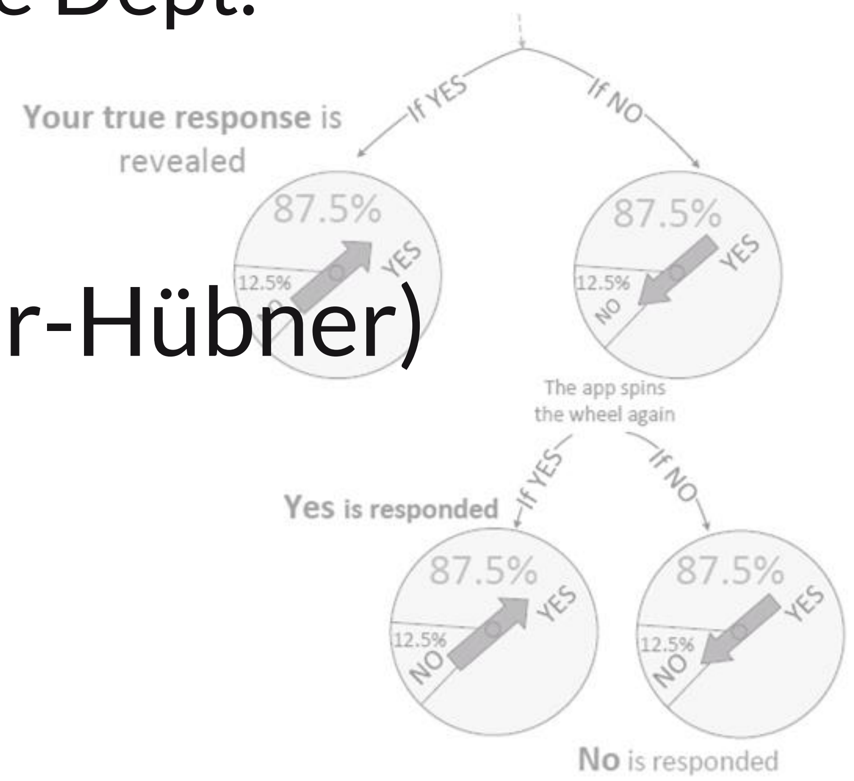
SOUPS– August 7-9, 2022

Presenter: Farzaneh Karegar

PriSec Group, Mathematics and Computer Science Dept.
Karlstad University

(*A joint work with Ala Sarah Alaqra*, Simone Fischer-Hübner)





Funded by the H2020 Framework of the European Commission under Grant Agreement No. 786767 (PAPAYA project) and by the Swedish Knowledge Foundation (TRUEdig project).

Objective

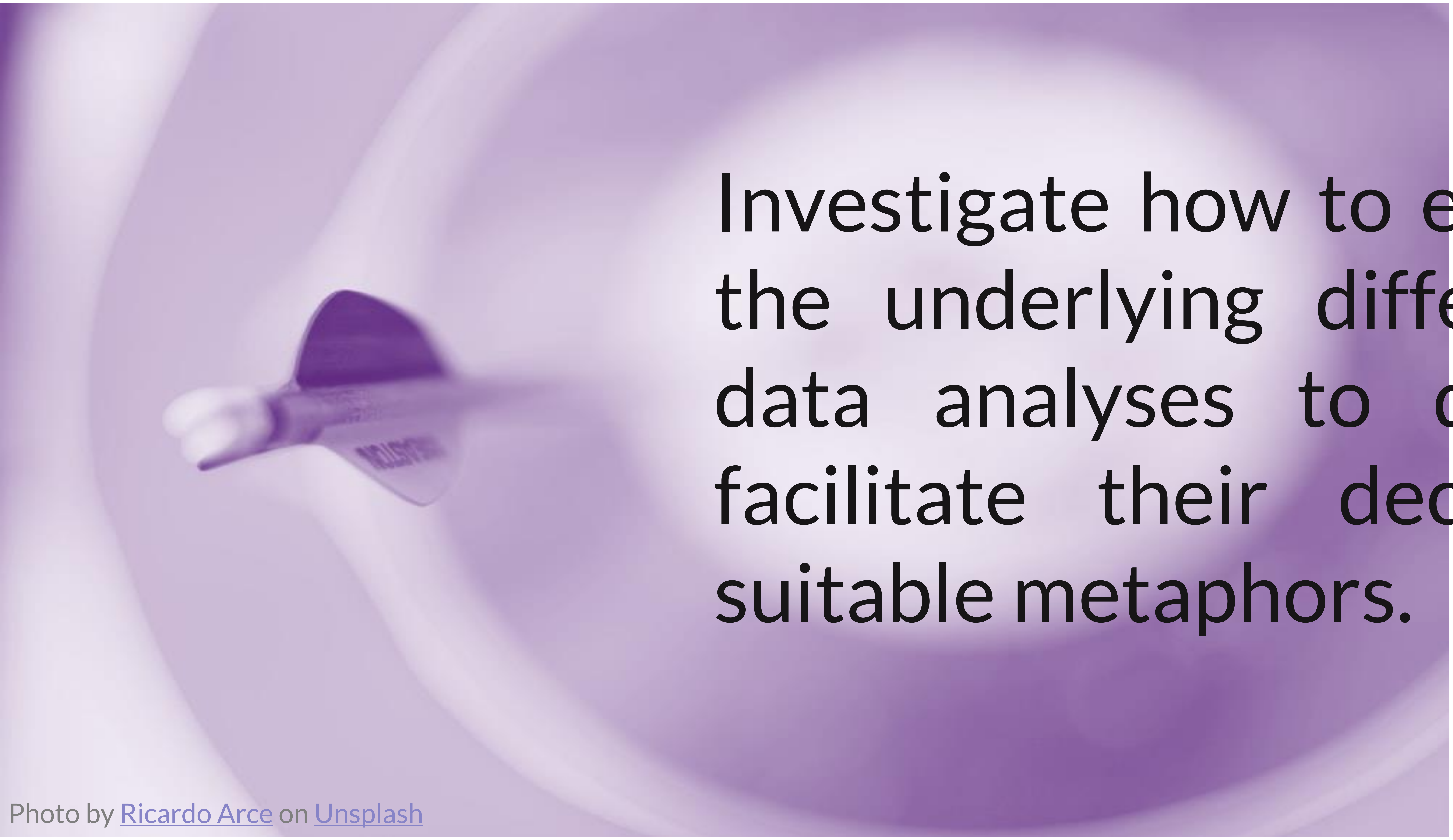
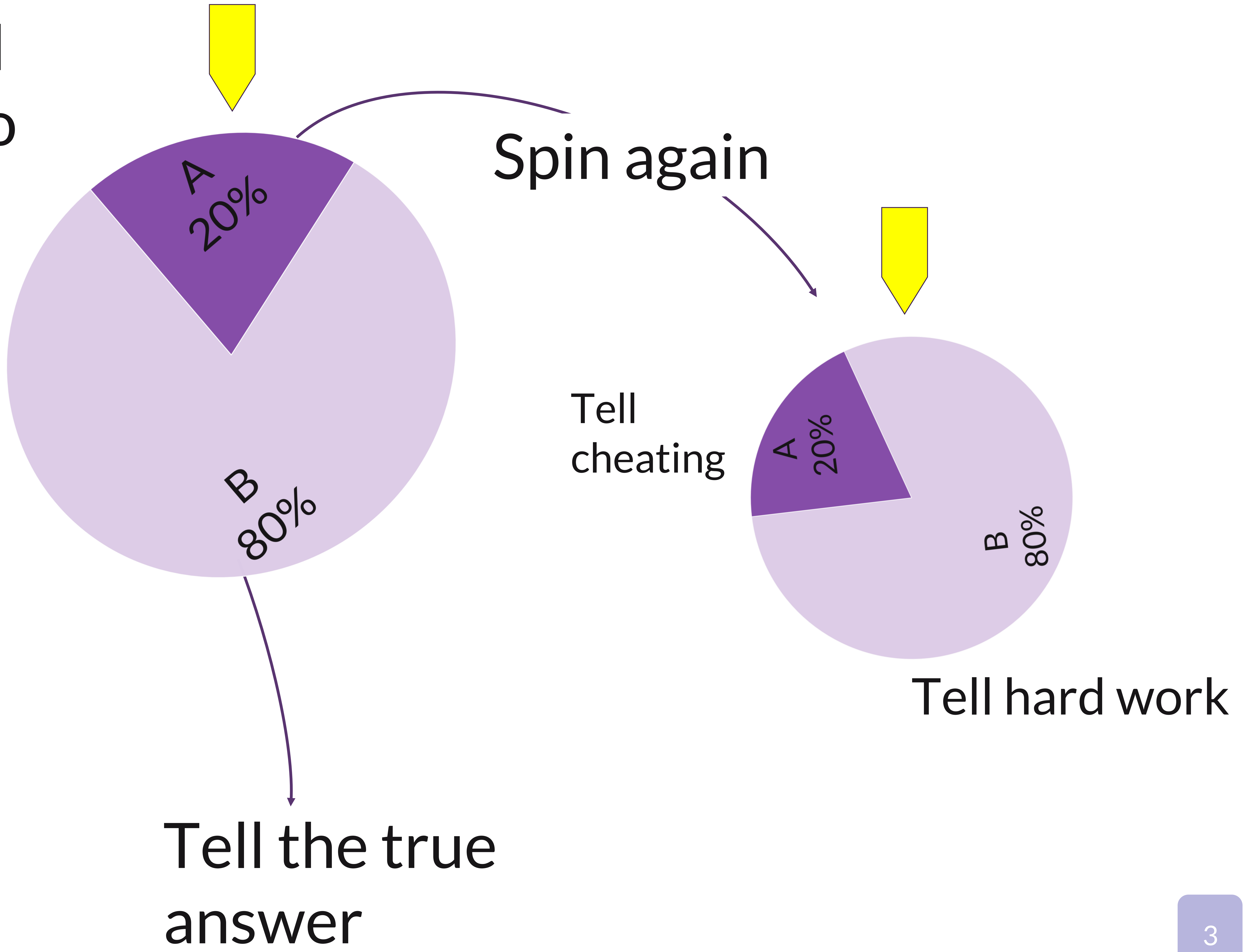


Photo by Ricardo Arce on Unsplash

Investigate how to effectively explain the underlying differentially private data analyses to data subjects to facilitate their decisions by using suitable metaphors.

Do you prefer hard work or cheating to succeed?



Spin again

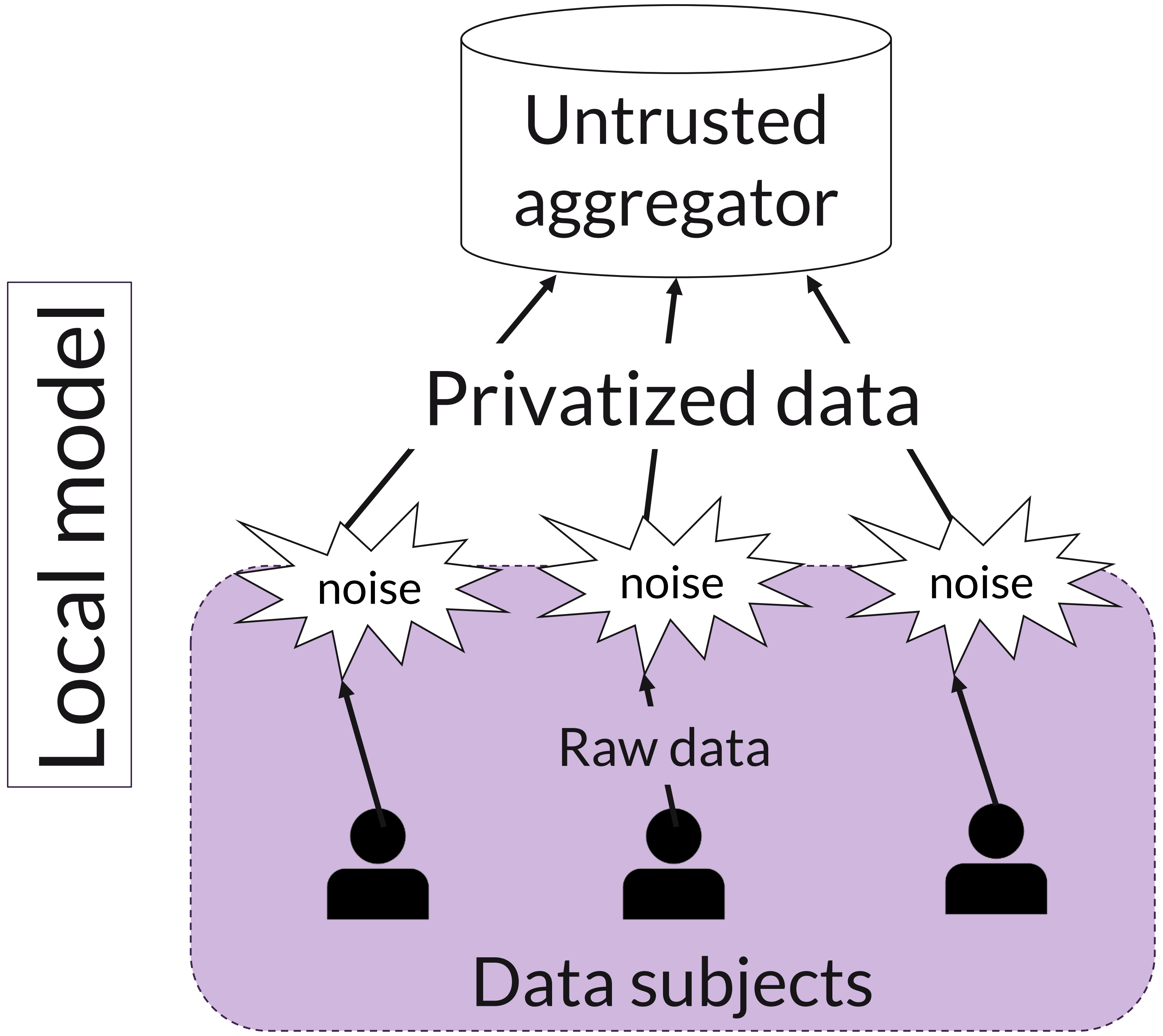
Tell cheating

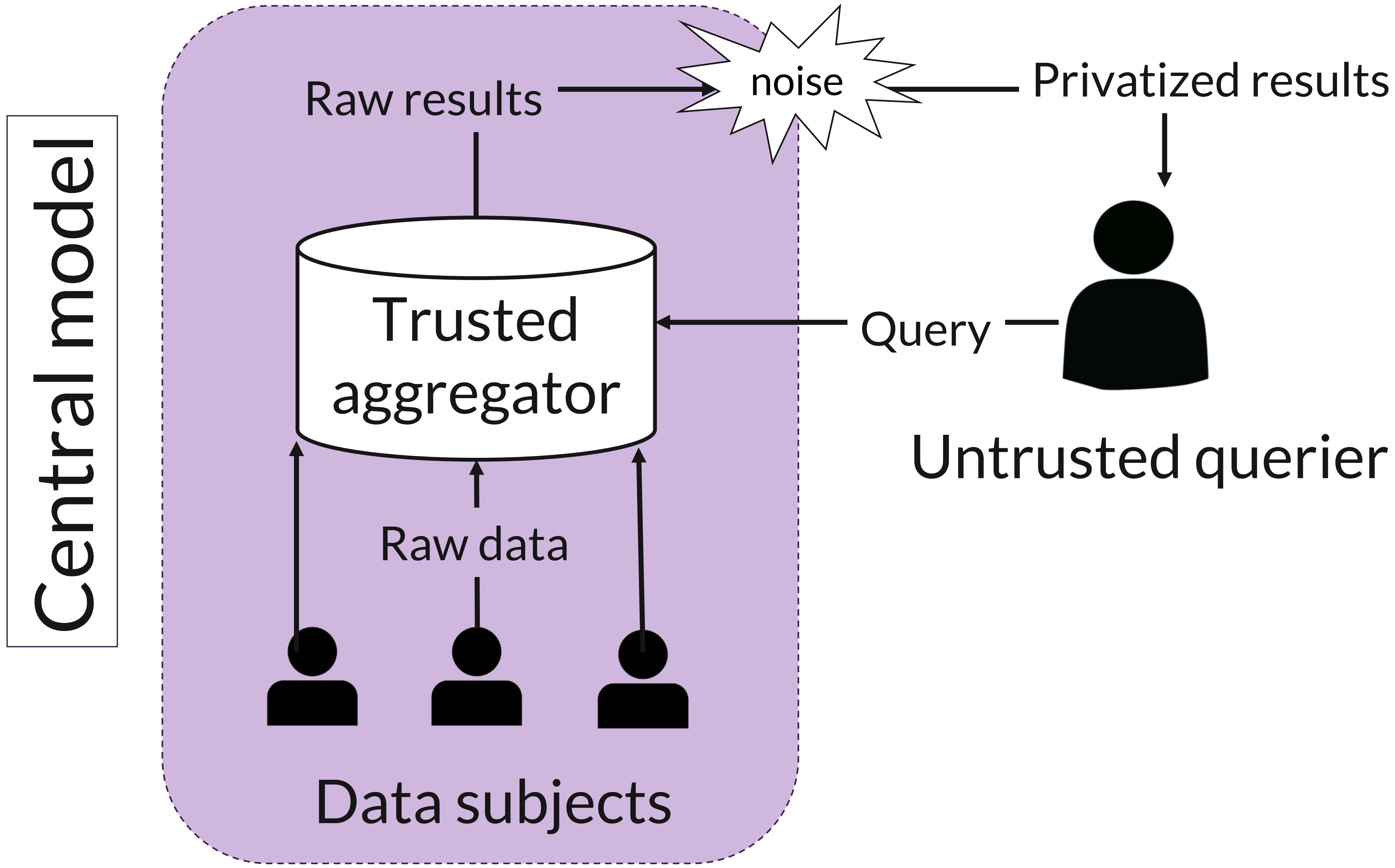
Tell hard work

Tell the true answer

Differential privacy - models

- *Local DP (individual level) – untrusted aggregator*
- *Central DP (aggregated-level) – untrusted querier*



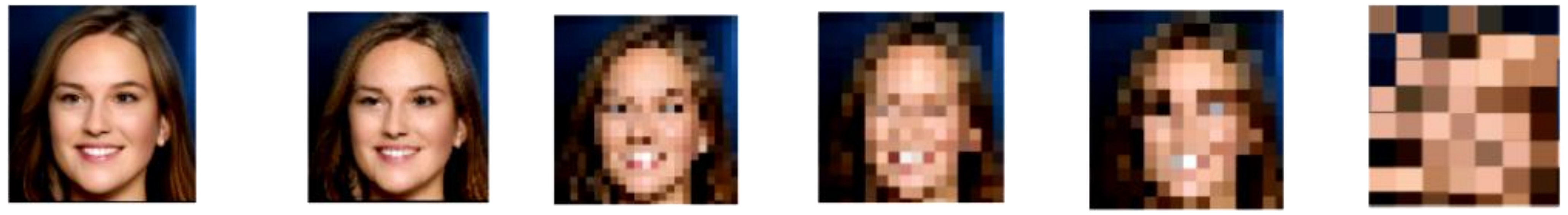


DP descriptions in industry & media outlets do not distinguish different models*.

** Rachel Cummings, Gabriel Kaptchuk, and Elissa M. Redmiles. 2021. "I need a better description": An Investigation Into User Expectations For Differential Privacy. In Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS '21). ACM, 3037–3052.*

Metaphors for local DP – Scenario 1

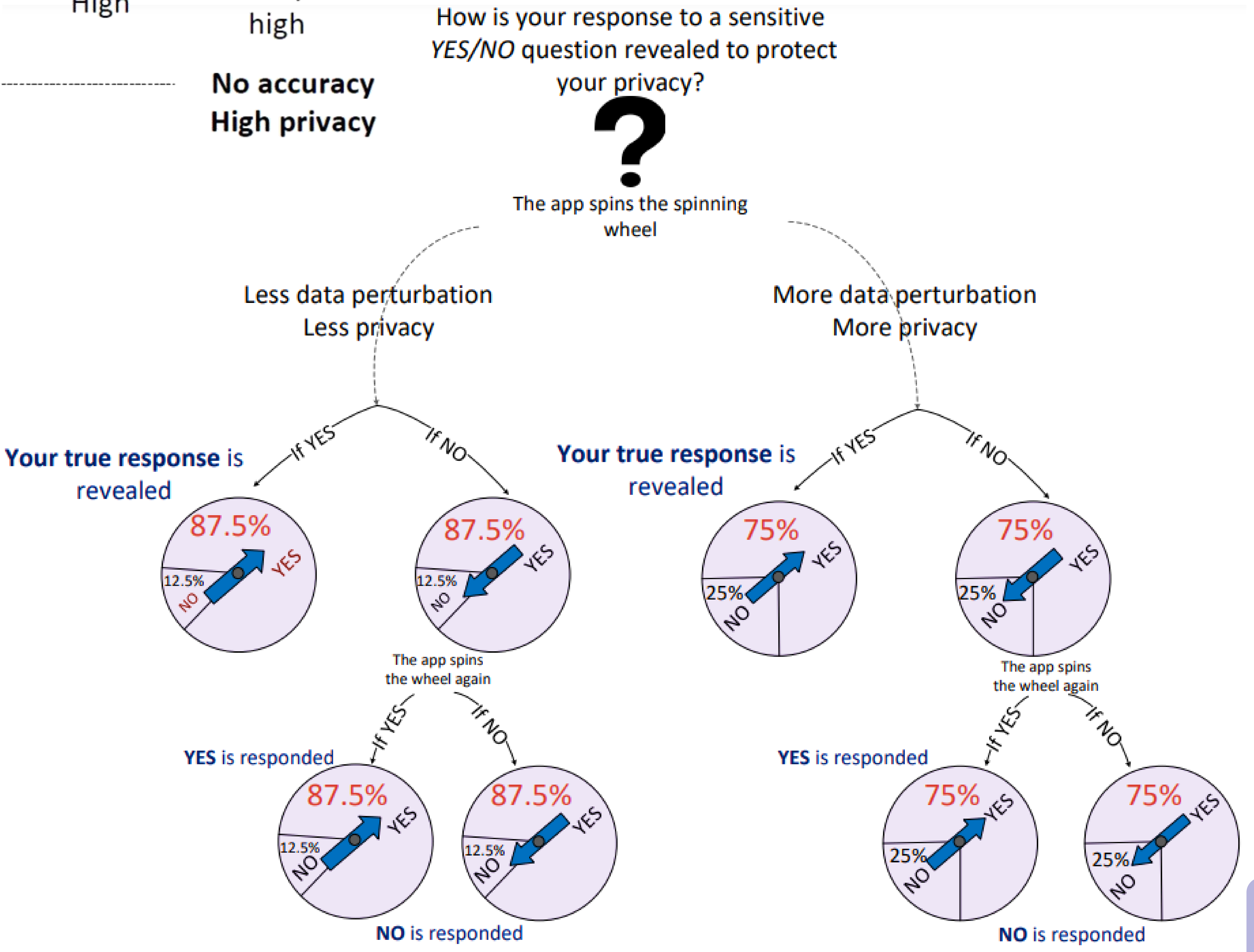
Original data



The amount of **added noise**: No added noise | Very low | Low | Medium | High | Very high

Accuracy of outcome: **Highest accuracy** **No privacy** ----------Decreasing---------- **No accuracy** **High privacy**

Noisy picture (portrait) metaphor



How is your response to a sensitive *YES/NO* question revealed to protect your privacy?

The app spins the spinning wheel

Less data perturbation
Less privacy

More data perturbation
More privacy

Your true response is revealed

Your true response is revealed

The app spins the wheel again

The app spins the wheel again

YES is responded

YES is responded

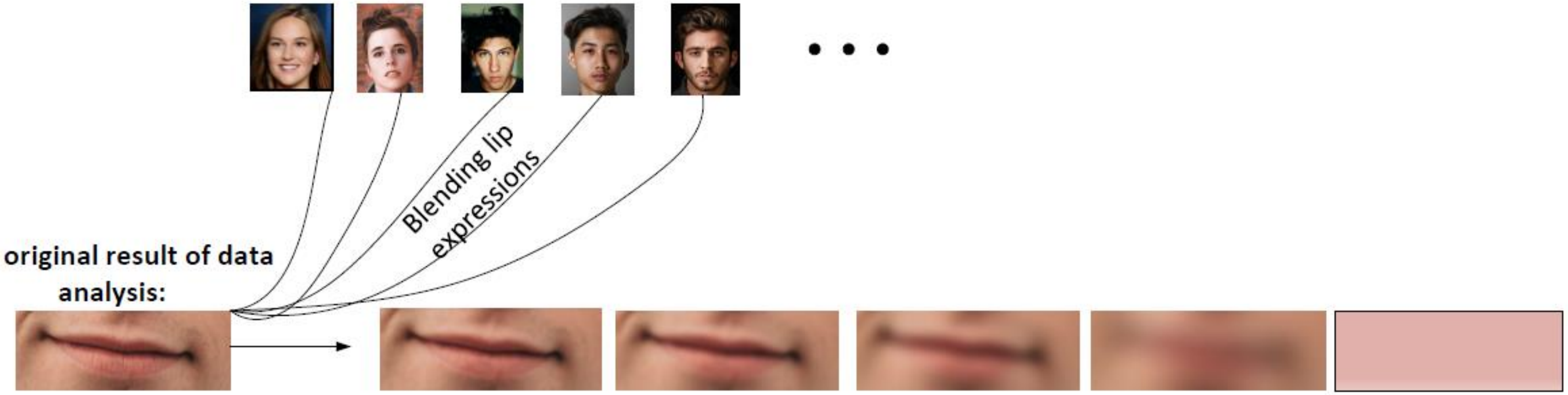
NO is responded

NO is responded

Spinner metaphor

Metaphor for central DP – Scenario 2

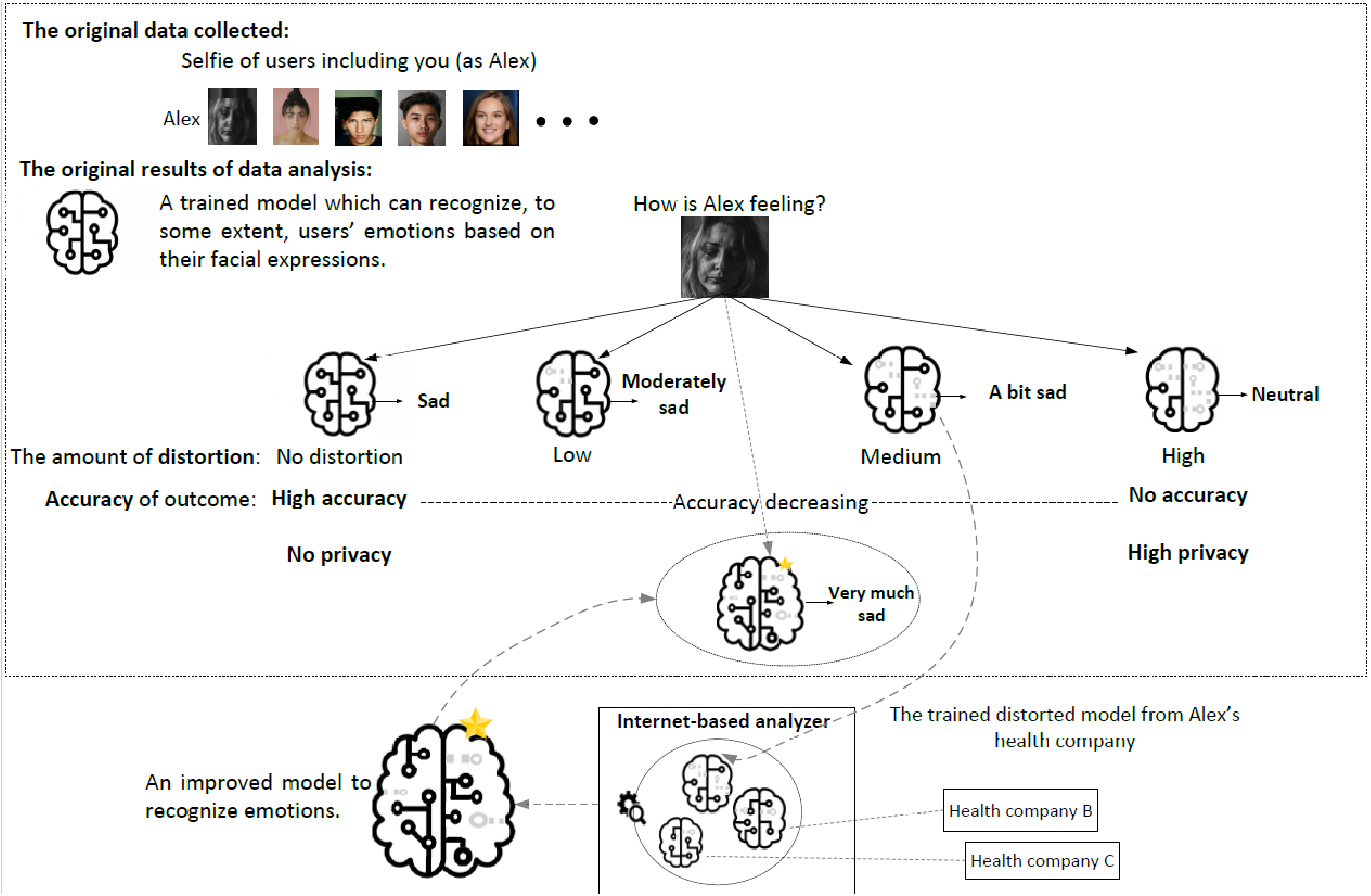
Original data collected:
Selfie of users



The original result of data analysis:

The amount of added noise:	No noise	Very low	Low	Medium	High	Very high
Accuracy of outcome:	**High accuracy** **No privacy**	-----Accuracy decreasing-----				**No accuracy** **High privacy**

Metaphor for central DP – Scenario 3



The original data collected:

Selfie of users including you (as Alex)

Alex • • •

The original results of data analysis:

A trained model which can recognize, to some extent, users’ emotions based on their facial expressions.

How is Alex feeling?

Sad

Moderately sad

A bit sad

Neutral

The amount of **distortion**: No distortion — Low — Medium — High

Accuracy of outcome: **High accuracy** — Accuracy decreasing — **No accuracy**

No privacy — **High privacy**

Very much sad

An improved model to recognize emotions.

Internet-based analyzer

The trained distorted model from Alex’s health company

Health company B

Health company C

Our approach

How to reach our objective

Phase 1: Metaphor generation
Phase 2: Analytical evaluation
Phase 3: Empirical evaluation
Extract metaphors
Define analysis scenarios
Adapt and extend metaphors
Define functionality
A
Analyze the metaphors using the evaluation matrix (1st round)
Expert analysis
B
Adapt the metaphors, functionality list, and scenarios
Analyze the metaphors using the evaluation matrix (2st round)
Evaluate the metaphors in online interviews with lay users

General view of our approach, based on the extended and adapted version of Alty et al.’s framework*.

* Alty, James L., Roger P. Knott, Ben Anderson, and Michael Smyth. "A framework for engineering metaphor at the user interface." *Interacting with computers* 13, no. 2 (2000): 301-322.

Research questions

RQ1

What information of the underlying differentially private systems is required by users to decide about using such systems?

RQ2

What are users’ perceptions of data privacy provided by the proposed metaphors?

RQ3

To what extent are our proposed metaphors suitable for conveying the concept of differential privacy to lay users?

Interviews – design and demographics

- 30 (3 X 10) online interviews with participants recruited via Prolific.

- **Interview design:**
  - Main session with two parts:
    a) Scenario introduction.
       (before exposure to metaphors)
    b) Metaphor introduction.

Photo by Kane Reinholdtsen on Unsplash

- **Demographics:**
  - 13 females, 18 males, one did not answer.
  - Relatively young.
  - Diverse academic background.
  - Non-experts in privacy.

Results - themes

RQ1

- T1: Factors affecting sharing of data.
- T2: Expressed needs for more privacy information.
- T3: Expectation of claimed protection (data access).
- T4: Expressed trust factors of DP protecting data.
- T6: Varied impact of DP descriptions on decisions to share.
- T7: Perceptions of info provided/missing.
- T8: Expressed trust factors (post-explanation).

Pre-explanation themes: before exposure to metaphor

RQ2

RQ3

- T5: Perceptions of claimed protection of DP.
- T9: Perceptions of accuracy-privacy trade-off
- T10: Preferences for distortion levels.
- T11: Varied acceptance/ perceptions of remaining risks.
- T12: Users' input/suggestions on DP alternatives.

Post-explanation themes: after exposure to metaphor

Information needed for trust and data sharing – RQ1

- The mere presence of a privacy technique:
  - seemingly enough.

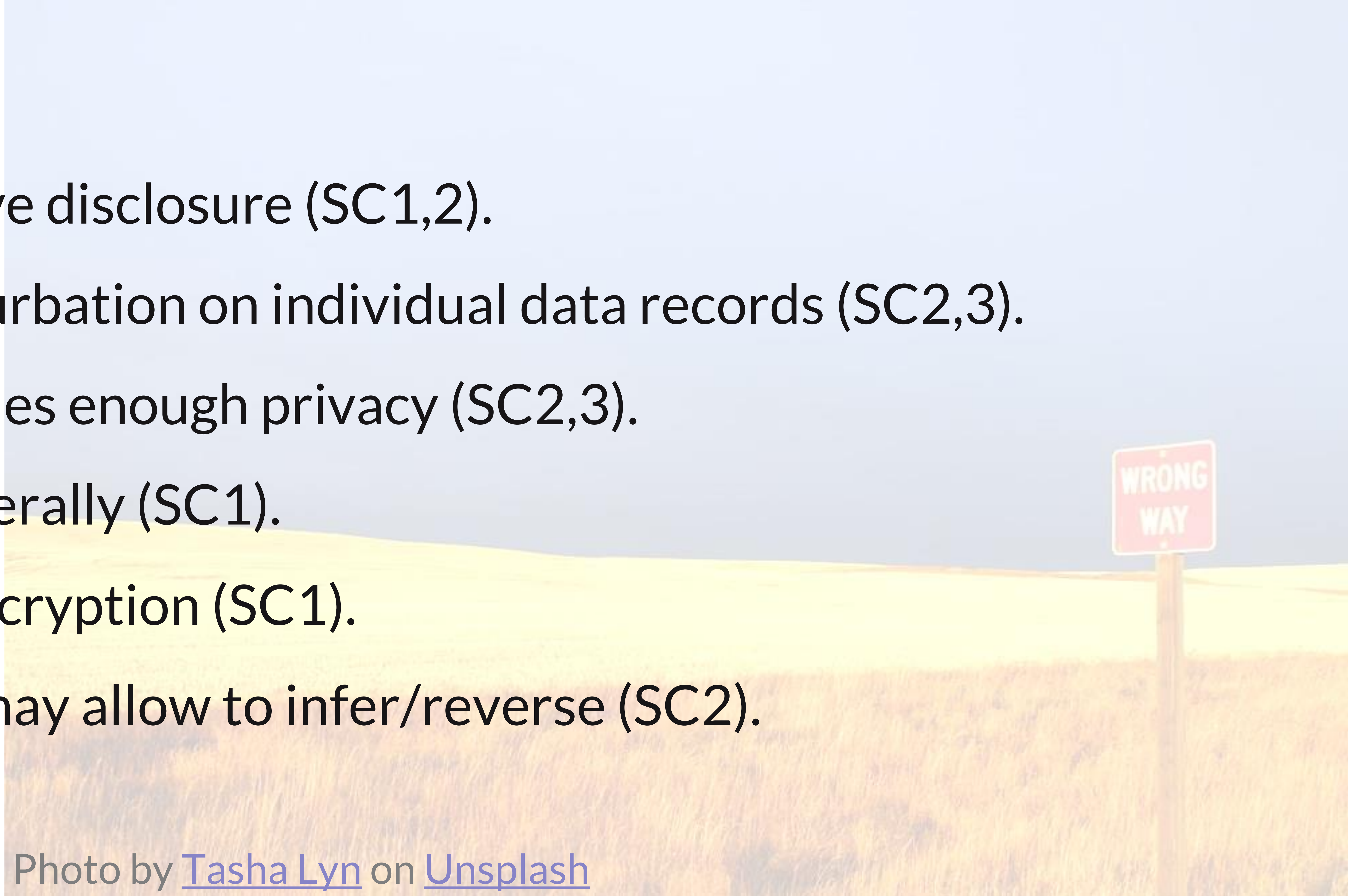
- However:
  - Lack of information on the underlying mechanism/transparency on DP →
    - Varied expectations/interpretations of access to actual data.
    - Different (correct/incorrect) assumptions of DP.
    - Negative impacts on trust and data sharing.
  - (Usable) Transparency of DP is desired by most.

Perceptions of privacy features and the extent of the suitability of metaphors – RQ2/RQ3

- Participants understood (that):
  - Perturbation:
    - leads to privacy.
    - protects against identifiability.
    - provides plausible deniability.
  - The trade-off between accuracy and privacy protection.

- However:
  - Several misconceptions about DP.
  - Varied perceptions and preferences about different aspects.

Misconceptions of DP



- DP is reversible.
- DP enables selective disclosure (SC1,2).
- Perception of perturbation on individual data records (SC2,3).
- Aggregation provides enough privacy (SC2,3).
- Metaphor taken literally (SC1).
- DP perceived as encryption (SC1).
- Knowledge of DP may allow to infer/reverse (SC2).

Photo by Tasha Lyn on Unsplash

Challenges and conclusion

- **Need of emphasising the reduction of identification risks**
  - Guidance needed on adequate risks per context and implications.

- **Misconception triggered by digital-world analogies**
  - Both real-world & **digital-world analogies** need to be considered.

Photo by Samantha Sophia on Unsplash

- **Metaphorical explanations: A quandary**
  - Complement metaphors with suitable additional information.

Thanks!

Any questions?

You can contact me via email:
Farzaneh.Karegar@kau.se