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

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?

Tell the true answer

Spin again

Tell cheating

Tell hard work
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*.

Metaphors for local DP – Scenario 1

Noisy picture (portrait) metaphor

Spinner metaphor
Metaphor for central DP – Scenario 2

Original data collected: Selfie of users

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

Blending lip expressions
Metaphor for central DP – Scenario 3

The original data collected:
Selfie of users including you (as 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?

The amount of distortion:
No distortion
Low
Medium
High

Accuracy of outcome:
High accuracy
No accuracy

Accuracy decreasing

No privacy

High privacy

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

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

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.

• 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).

RQ2
- 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.

RQ3

Pre-explanation themes: before exposure to metaphor

Post-explanation themes: after exposure to metaphor
Information needed for trust and data sharing – RQ1

• The mere presence of a privacy technique:
  o seemingly enough.

• However:
  o 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.
  o (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):
  o Perturbation:
    o leads to privacy.
    o protects against identifiability.
    o provides plausible deniability.
  o The trade-off between accuracy and privacy protection.

• However:
  o Several misconceptions about DP.
  o 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
  o Guidance needed on adequate risks per context and implications.

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

• Metaphorical explanations: A quandary
  o Complement metaphors with suitable additional information.
Thanks!

Any questions?

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