

## Paper

This document contains the survey questions for:

Michael Khavkin, Eran Toch. Valuating the Differential Privacy Budget in Data Sharing: A Conjoint Analysis. IFIP International Summer School on Privacy and Identity Management, 2022.

## Questionnaire

### *Introduction*

The aim of this survey is to evaluate the value that people assign to privacy in different data sharing scenarios, where they are given the opportunity to share their personal data with some data collector in return for a payment. Each scenario can have different properties, relating to both the shared data and its processing.

The survey consists of 10 choice-selection tasks and 5 general demographic questions (used for a general statistical background analysis) - 15 questions in total.

We expect a working time of about 5-10 minutes.

All collected answers will be used exclusively for the study's purposes and will not be passed on to third parties. No personal identifiers are used in this survey, and it is processed anonymously. Your MTurk-ID may be used to approve your participation and the HIT (Human Intelligence Task) towards Amazon Mechanical Turk platform but is ignored and deleted during processing.

We recommend using a computer to fill out this form (and not a mobile phone).

**Thank you for your cooperation!**

*Click the "Next" button to start!*

### *Description*

Imagine you have agreed to participate in a program where you can offer specific personal information of yours to a data collector, in return for a payment. You can define the payment you would like to request from the data collector for sharing your data.

The data collector guarantees that your personal data will be protected and used only for the purpose for which it was intended. The data collector uses a strong notion of privacy in its collection and processing application, called ***differential privacy (DP)***.

That is, the application will store your data, but only use the aggregated statistics extracted from your data, modified with a random injection of noise. These modified results reveal limited information specific to each individual user, so that your personal information cannot be inferred from the resulting statistics with a high probability.

In addition, you can control the level of provided privacy by this technique, which will determine the probability that your individual personal information is inferred by some external party. For example, requesting a noisier modification to your data will result in a higher privacy level, since the data collector will be exposed to less accurate data.

In this questionnaire, we ask for your personal evaluation of different scenarios, characterized by five properties:

- **Sensitive disclosed data** (*Type of collected data*)
- **Sharing purpose** (*The purpose for which personal data is collected*)
- **Retention period** (*Retention period of collected personal data by the data processor*)
- **Differential privacy level: 1 (lowest) - 5 (highest)** (*Expected differential privacy level to be satisfied by the data processor*)
- **Data sharing compensation price** (*Lowest price that individuals are willing to accept for sharing their personal information*)

In the next section, please choose the scenario you prefer the most, according to your personal valuation of privacy with respect to properties of the scenario.

**Question 1**

Please carefully review the choice options detailed below, then choose your most preferred scenario.

Which of these choices do you prefer?

	Choice 1	Choice 2	Choice 3
<b>Retention period</b> (Retention period of collected personal data by the data processor)	Indefinitely	Temporary	Temporary
<b>Differential privacy level: 1 (lowest) - 5 (highest)</b> (Expected differential privacy level to be satisfied by the data processor)	1	3	4
<b>Data sharing compensation price</b> (Lowest price that individuals are willing to accept for sharing their personal information)	1\$	50\$	10\$
<b>Sharing purpose</b> (The purpose for which personal data is collected)	Research	Research	Research
<b>Sensitive disclosed data</b> (Type of collected data)	(Medical) Existence of Diabetes	(Financial) Income	(Behavioral) Browsing History

Choice 1

Choice 2

Choice 3

## *Demographics*

The next section includes several general demographic questions.

These answers are used only for statistical analysis to supports our findings in this study.

What is your gender?

- Male
- Female
- Prefer not to say

What is your age group?

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65 and above
- Prefer not to say

Which of the following options best represents your highest education level?

- Less than high school degree
- High school graduate (high school diploma or equivalent including GED)
- Other post-high school vocational training
- Some college but no degree
- Bachelor's degree
- Bachelor's degree student
- Master's degree
- Master's degree student
- Doctoral degree
- Doctoral degree student
- Professional degree (JD, MD)
- Prefer not to say

Which of the following statements best describes your employment status?

- Full-time
- Part-time
- Own business / Self-employed / Freelance
- Student

- Unemployed
- Prefer not to say

How would you rate your level of computer skills?

- Basic
- Intermediate
- Advanced