Communicating Differential Privacy Guarantees to Data Subjects

Priyanka Nanayakkara

PhD Candidate, Computer Science & Communication

Northwestern University

Based on joint work with Rachel Cummings, Gabriel Kaptchuk, Elissa M. Redmiles, Mary Anne Smart

USENIX PEPR | September 11, 2023 | Santa Clara, California


Photo: Unsplash

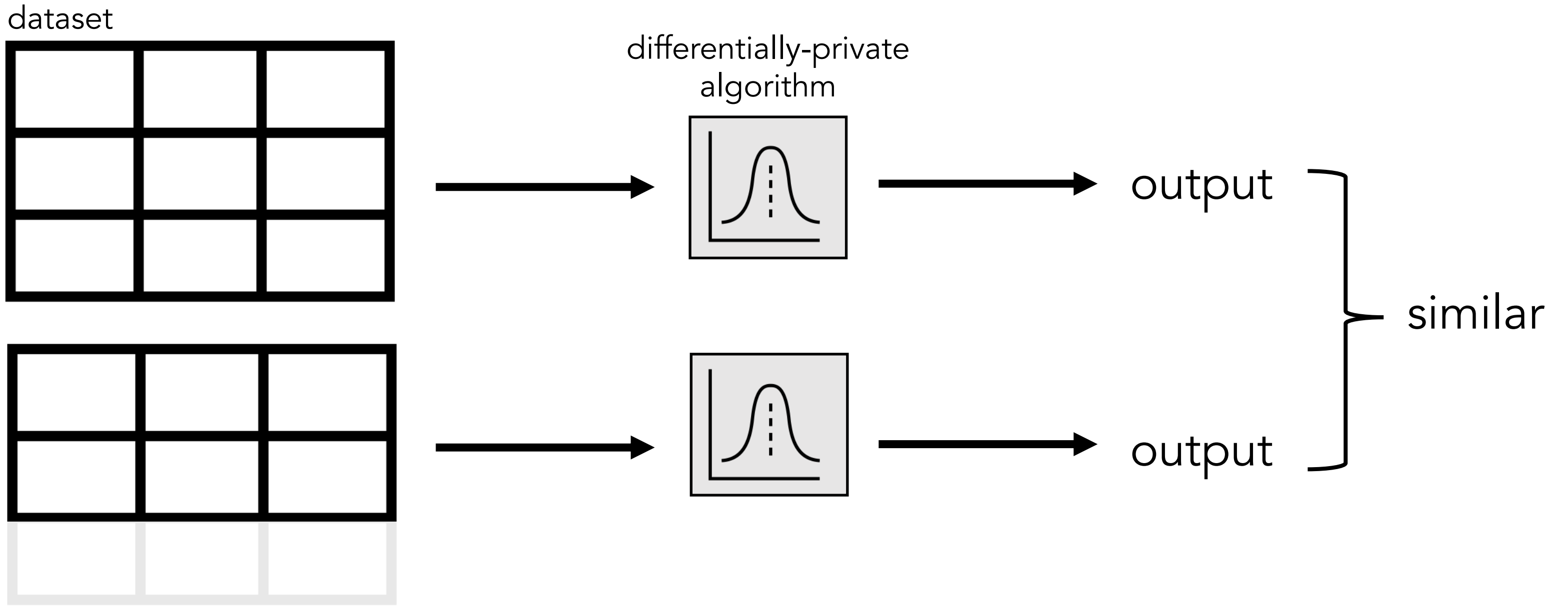
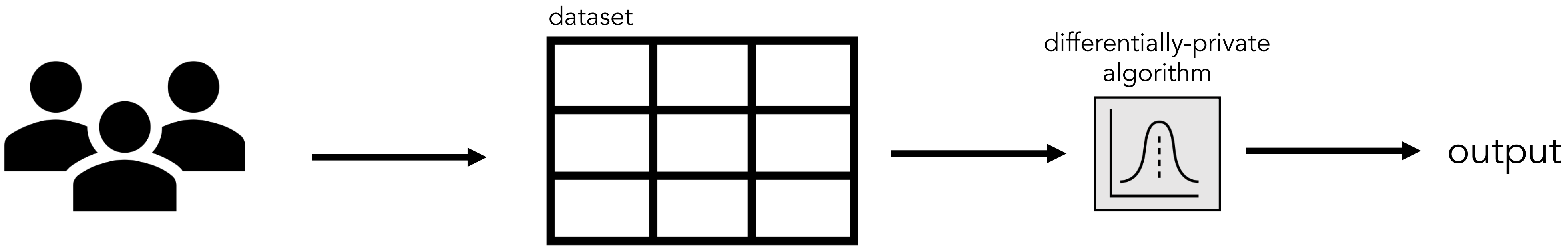


Diagram adapted from
Wood, Altman, Bembenek et al. (2020). Differential Privacy: A Primer for a Non-Technical Audience
Near, Darais, Boeckl (2020). Differential Privacy for Privacy-Preserving Data Analysis: An Introduction to our Blog Series

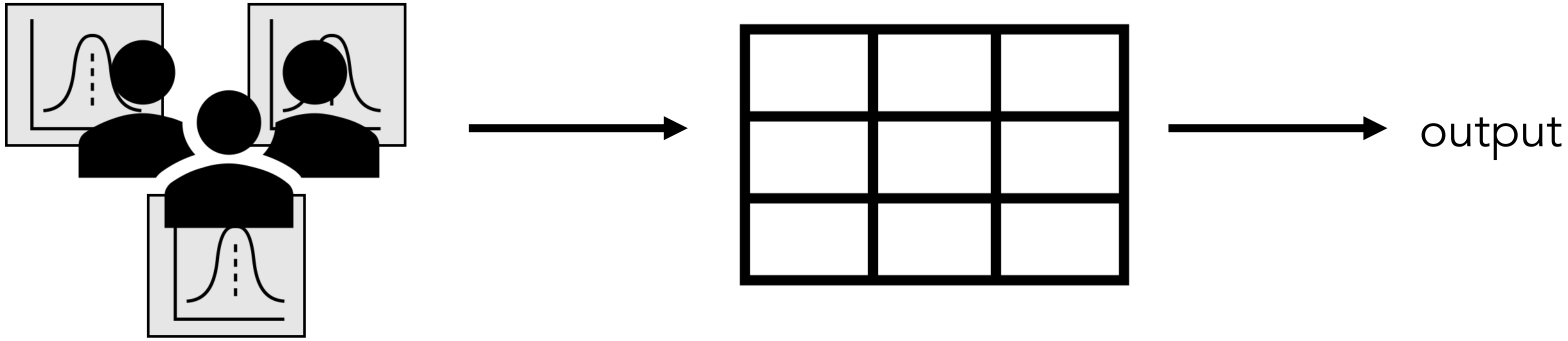
Dwork, McSherry, Nissim, Smith (2006). Calibrating noise to sensitivity in private data analysis

1 Deployment model



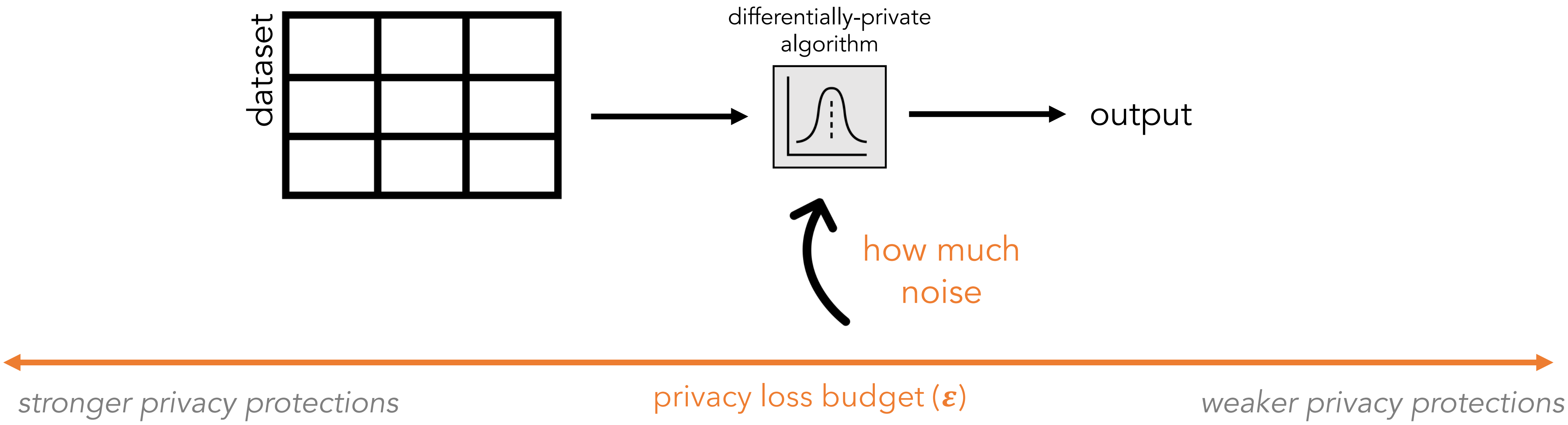
central

local



Kasiviswanathan, Lee, Nissim, Raskhodnikova, Smith. (2011). What can we learn privately?

2 "Privacy loss budget" ε



1. **Deployment model**
2. **"Privacy loss budget" ε**

"differential privacy," the new **gold standard in data privacy** protection.

In ideal implementations, this **risk remains close to zero**, guaranteeing… virtually no adverse effect on them from an informational standpoint.

When a differential privacy algorithm is applied to a data set, those **links get blurred**, and bits of data can no longer be traced to their source.

Differential privacy works by algorithmically **scrambling individual user data** so that it cannot be traced back

TechCrunch

In short, differential privacy **allows general statistical analysis** without revealing information about a particular individual in the data

"differential privacy," which alters the numbers but **does not change core findings** to protect the identities of individual respondents.

Slide courtesy of Gabriel Kaptchuk

Cummings, Kaptchuk, Redmiles (2021). "I need a better description": An Investigation Into User Expectations For Differential Privacy

Design effective explanations that expose information about:

1. **Deployment model**
2. **"Privacy loss budget" ε**

Metaphors

Diagrams

Privacy Labels

Improved comprehension of ***information flows***

Smart, Nanayakkara, Cummings, Kaptchuk, Redmiles (2023). Models Matter: Setting Accurate Privacy Expectations for Local and Central Differential Privacy

1 Deployment model

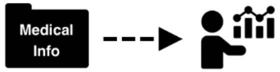
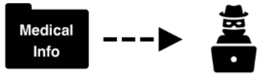
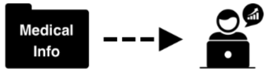
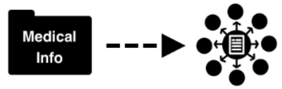
Who Can See Your Data	Without Privacy Protection	With Privacy Protection

Kelley, Bresee, Cranor, Reeder (2009). A “nutrition label” for privacy
Smart, Nanayakkara, Cummings, Kaptchuk, Redmiles (2023). Models Matter: Setting Accurate Privacy Expectations for Local and Central Differential Privacy

1 Deployment model

Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…		
Hackers—like criminals or foreign governments— who successfully attack the non-profit…		
Law enforcement with a court order requesting your information from the non-profit…		
Employees of the non-profit, such as data analysts, who work with the non-profit's data…		
Organizations collaborating with the non-profit that are given access to the non-profit's data…		

1 Deployment model

Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…	 …might be able to see your information.	
Hackers—like criminals or foreign governments— who successfully attack the non-profit…	 …might be able to see your information.	
Law enforcement with a court order requesting your information from the non-profit…	 …might be able to see your information.	
Employees of the non-profit, such as data analysts, who work with the non-profit's data…	 …might be able to see your information.	
Organizations collaborating with the non-profit that are given access to the non-profit's data…	 …might be able to see your information.	

1 Deployment model

Local DP

Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…	…might be able to see your information.	…will not be able to see your information.
Hackers—like criminals or foreign governments— who successfully attack the non-profit…	…might be able to see your information.	…will not be able to see your information.
Law enforcement with a court order requesting your information from the non-profit…	…might be able to see your information.	…will not be able to see your information.
Employees of the non-profit, such as data analysts, who work with the non-profit's data…	…might be able to see your information.	…will not be able to see your information.
Organizations collaborating with the non-profit that are given access to the non-profit's data…	…might be able to see your information.	…will not be able to see your information.

Smart, Nanayakkara, Cummings, Kaptchuk, Redmiles (2023). Models Matter: Setting Accurate Privacy Expectations for Local and Central Differential Privacy

1 Deployment model

Central DP

Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…	…might be able to see your information.	…will not be able to see your information.
Hackers—like criminals or foreign governments— who successfully attack the non-profit…	…might be able to see your information.	…might be able to see your information.
Law enforcement with a court order requesting your information from the non-profit…	…might be able to see your information.	…might be able to see your information.
Employees of the non-profit, such as data analysts, who work with the non-profit's data…	…might be able to see your information.	…might be able to see your information.
Organizations collaborating with the non-profit that are given access to the non-profit's data…	…might be able to see your information.	…might be able to see your information.

1 Deployment model

Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…	…might be able to see your information.	…will not be able to see your information.
Hackers—like criminals or foreign governments— who successfully attack the non-profit…	…might be able to see your information.	…will not be able to see your information.
Law enforcement with a court order requesting your information from the non-profit…	…might be able to see your information.	…will not be able to see your information.
Employees of the non-profit, such as data analysts, who work with the non-profit's data…	…might be able to see your information.	…will not be able to see your information.
Organizations collaborating with the non-profit that are given access to the non-profit's data…	…might be able to see your information.	…will not be able to see your information.

+

To protect your information, your data will be randomly modified before it is sent to the organization. Only the modified version will be stored, so that your exact data is never collected by the organization.

=

Improved trust

2 "Privacy loss budget" ε

stronger privacy protections ← **ε** → *weaker privacy protections*

Increased ***willingness to share data***

- **Example-Based**
- **Odds-Based Text**
- **Odds-Based Visual**

Improved ***risk comprehension*** *&* ***self-efficacy*** *(enough info)*

2 "Privacy loss budget" ε

If you **do not participate**,
x out of 100 potential *DP outputs* will lead adversary A to believe you responded d_{true}.

If you **participate**,
y out of 100 potential *DP outputs* will lead adversary A to believe you responded d_{true}.

2 "Privacy loss budget" ε

Probabilities reflect immediate decisions

If you **do not participate**,
x out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.

If you **participate**,
y out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.

Framing probabilities as frequencies vs. percentages supports statistical reasoning & has been applied in privacy contexts

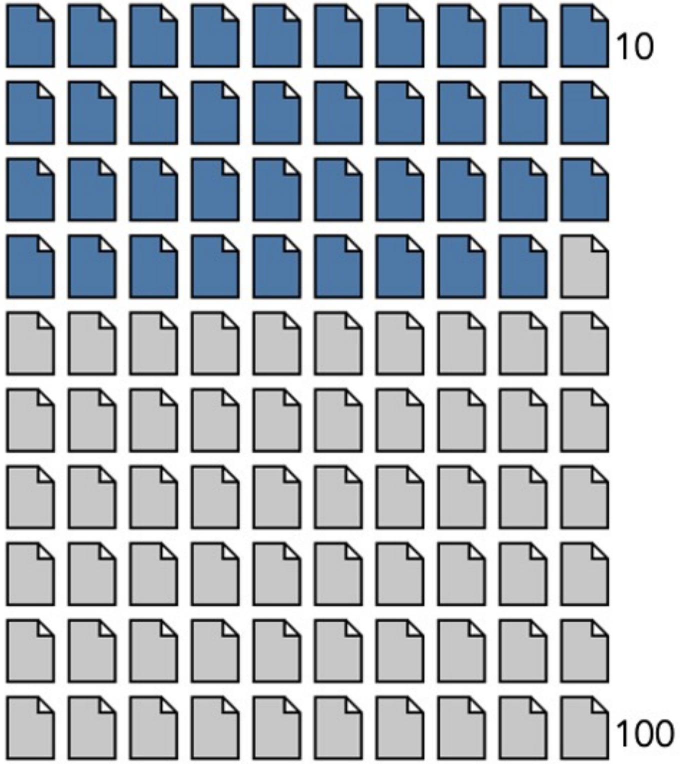
If you **do not participate**,
x out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.

If you **participate**,
y out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.

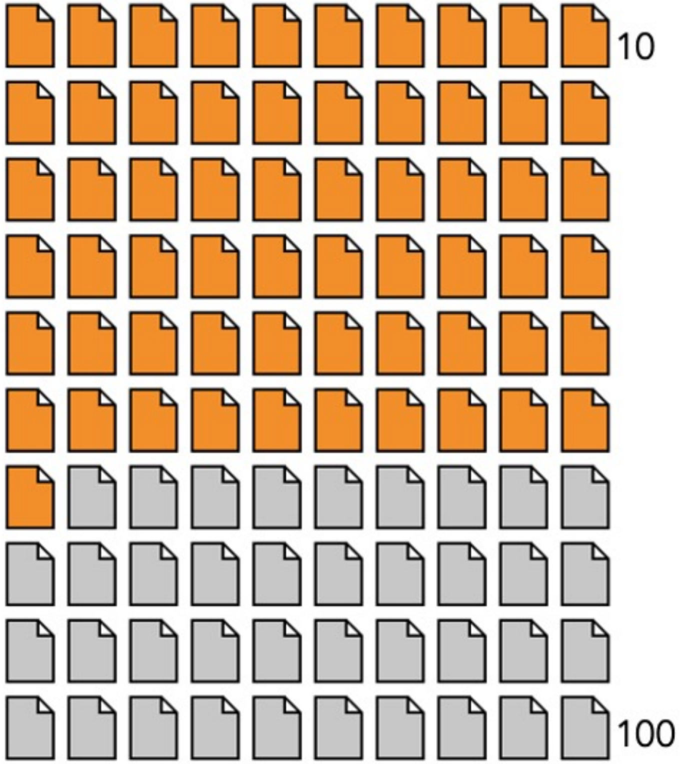
Nanayakkara, Smart, Cummings, Kaptchuk, Redmiles (2023). What Are the Chances? Explaining the Epsilon Parameter in Differential Privacy
Gigerenzer and Hoffrage (1995). How to improve Bayesian reasoning without instruction: Frequency formats
Hoffrage and Gigerenzer (1998). Using natural frequencies to improve diagnostic inferences
Slovic (2000). The perception of risk
Kaptchuk, Goldstein, Hargittai, Hofman, and Redmiles (2020). How good is good enough for COVID19 apps? ...
Franzen, Nuñez von Voigt, Sörries, Tschorsch, Müller-Birn (2022). "Am I private and if so, how many?" ...

2 "Privacy loss budget" ε

If you **do not participate**,
x out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.



If you **participate**,
y out of 100 potential *DP outputs* will lead adversary *A* to believe you responded d_{true}.

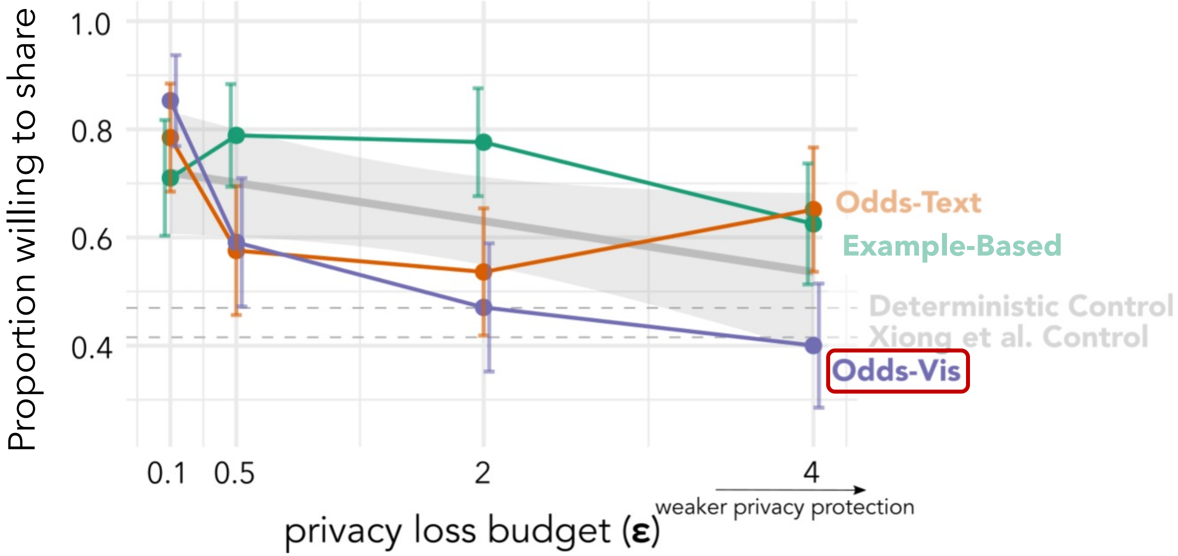


Icon arrays assume x = 39 and y = 61 for illustration purposes.

Nanayakkara, Smart, Cummings, Kaptchuk, Redmiles (2023). What Are the Chances? Explaining the Epsilon Parameter in Differential Privacy
Galesic, Garcia-Retamero, Gigerenzer (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy

2 **"Privacy loss budget" ε**

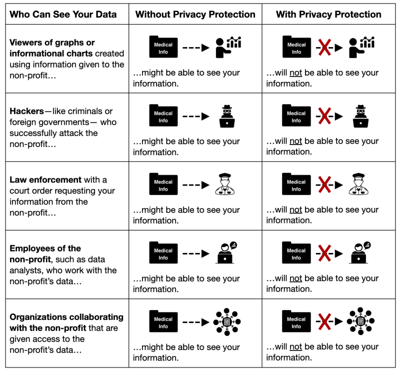
Increased **willingness to share** with increased **privacy strength**



Nanayakkara, Smart, Cummings, Kaptchuk, Redmiles (2023). What Are the Chances? Explaining the Epsilon Parameter in Differential Privacy

Takeaways | How Organizations Can Explain DP

1 Deployment model



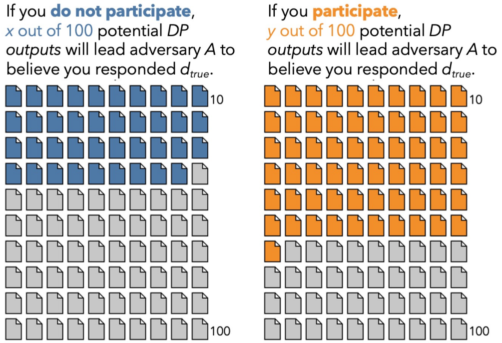
Who Can See Your Data	Without Privacy Protection	With Privacy Protection
Viewers of graphs or informational charts created using information given to the non-profit…	Medical Info …might be able to see your information.	Medical Info …will not be able to see your information.
Hackers—like criminals or foreign governments— who successfully attack the non-profit…	Medical Info …might be able to see your information.	Medical Info …will not be able to see your information.
Law enforcement with a court order requesting your information from the non-profit…	Medical Info …might be able to see your information.	Medical Info …will not be able to see your information.
Employees of the non-profit, such as data analysts, who work with the non-profit's data…	Medical Info …might be able to see your information.	Medical Info …will not be able to see your information.
Organizations collaborating with the non-profit that are given access to the non-profit's data…	Medical Info …might be able to see your information.	Medical Info …will not be able to see your information.

\+ *Process text*

Privacy labels improve comprehension of information flows

Adding process text improves trust

2 "Privacy loss budget" ε



Improves risk comprehension, self efficacy (enough info)

People are sensitive to changes in ε

Smart, Nanayakkara, Cummings, Kaptchuk, Redmiles (2023). Models Matter: Setting Accurate Privacy Expectations for Local and Central Differential Privacy
Nanayakkara, Smart, Cummings, Kaptchuk, Redmiles (2023). What Are the Chances? Explaining the Epsilon Parameter in Differential Privacy

Takeaways | Lessons for Explaining PETs Beyond DP

Expose key decision-making information, even if it's complicated.
Make complexity interpretable.

Describe implications + process to increase comprehension & trust.

Explain utility as well as privacy.

Thank you!
Priyanka Nanayakkara (priyankan@u.northwestern.edu | @priyakalot | @priyakalot@hci.social)
Coauthors: Rachel Cummings, Gabriel Kaptchuk, Elissa M. Redmiles, Mary Anne Smart