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
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

Kasiviswanathan, Lee, Nissim, Raskhodnikova, Smith. (2011). What can we learn privately?
“Privacy loss budget” $\varepsilon$

- **Dataset** → **Differentially-private algorithm** → **Output**

How much noise:

- **Stronger privacy protections**
- **Privacy loss budget ($\varepsilon$)**
- **Weaker privacy protections**
1 Deployment model

2 “Privacy loss budget” $\varepsilon$
“differential privacy,” the new gold standard in data privacy protection.

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.

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

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

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

“differential privacy,” which alters the numbers but does not change core findings to protect the identities of individual respondents.
Design effective explanations that expose information about:

1. Deployment model
2. “Privacy loss budget” $\varepsilon$
Deployment model

1. Metaphors
2. Diagrams
3. Privacy Labels

Improved comprehension of information flows
Deployment model

<table>
<thead>
<tr>
<th>Who Can See Your Data</th>
<th>Without Privacy Protection</th>
<th>With Privacy Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Deployment Model

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

<table>
<thead>
<tr>
<th>Who Can See Your Data</th>
<th>Without Privacy Protection</th>
<th>With Privacy Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewers of graphs or informational charts created using information given to the non-profit...</td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
<tr>
<td>Hackers—who like criminals or foreign governments—who successfully attack the non-profit...</td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
<tr>
<td>Law enforcement with a court order requesting your information from the non-profit...</td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td>Employees of the non-profit, such as data analysts, who work with the non-profit's data...</td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
<tr>
<td>Organizations collaborating with the non-profit that are given access to the non-profit's data...</td>
<td><img src="image9" alt="Diagram" /></td>
<td><img src="image10" alt="Diagram" /></td>
</tr>
</tbody>
</table>
### Deployment model

<table>
<thead>
<tr>
<th>Who Can See Your Data</th>
<th>Without Privacy Protection</th>
<th>With Privacy Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewers of graphs or informational charts created using information given to the non-profit…</td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
<tr>
<td>Hackers—like criminals or foreign governments—who successfully attack the non-profit…</td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
<tr>
<td>Law enforcement with a court order requesting your information from the non-profit…</td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td>Employees of the non-profit, such as data analysts, who work with the non-profit's data…</td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
<tr>
<td>Organizations collaborating with the non-profit that are given access to the non-profit's data…</td>
<td><img src="image9" alt="Diagram" /></td>
<td><img src="image10" alt="Diagram" /></td>
</tr>
</tbody>
</table>
## Deployment model

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

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.

Smart, Nanayakkara, Cummings, Kaptchuk, Redmiles (2023). Models Matter: Setting Accurate Privacy Expectations for Local and Central Differential Privacy
“Privacy loss budget” $\varepsilon$

\begin{align*}
\text{Example-Based} \\
\text{Odds-Based Text} \\
\text{Odds-Based Visual}
\end{align*}

Increased willingness to share data

Improved risk comprehension & self-efficacy (enough info)

Nanayakkara, Smart, Cummings, Kaptchuk, Redmiles (2023). What Are the Chances? Explaining the Epsilon Parameter in Differential Privacy
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 $d_{true}$.
If you **do not participate**, \(x\) out of 100 potential DP outputs will lead adversary A to believe you responded \(d_{\text{true}}\).

If you **participate**, \(y\) out of 100 potential DP outputs will lead adversary A to believe you responded \(d_{\text{true}}\).
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
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?” …
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.
Increased **willingness to share** with increased **privacy strength**
Takeaways | How Organizations Can Explain DP

1. Deployment model

- Privacy labels improve comprehension of information flows
- Adding process text improves trust

2. "Privacy loss budget" $\varepsilon$

- Improves risk comprehension, self efficacy (enough info)
- People are sensitive to changes in $\varepsilon$

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