Whose Data Is It Anyway? Towards a Formal Treatment of Differential Privacy for Surveys

James Bailie* & Jörg Drechsler†

*Harvard University, †Institute for Employment Research, Germany

May 16, 2024

NBER, Washington DC

Motivation

- The U.S. Census Bureau has committed to adopting *formal privacy* for all their data products (US Census Bureau 2022).
- Most of their collections are *surveys.*
- Yet the *"science ... does not yet exist"* for a formally private solution to the American Community Survey (for example).
- In implementing differential privacy (DP), surveys come with their own set of *unique challenges and opportunities.*

Survey Data Pipeline

Register for sampling

The collected data

$\cdots \longrightarrow$ **Population** $\mathfrak{p}$ $\longrightarrow$ **Frame** $\mathfrak{f}$ $\longrightarrow$ **Sample** $\mathfrak{s}$ $\longrightarrow$ **Responding sample** $\mathfrak{r}$ $\longrightarrow \cdots \longrightarrow$ **Published statistics** t

What we want to study

Random set of units

Survey Data Pipeline

Register for sampling

The collected data

$\cdots \longrightarrow$ **Population** $\mathfrak{p}$ $\longrightarrow$ **Frame** $\mathfrak{f}$ $\longrightarrow$ **Sample** $\mathfrak{s}$ $\longrightarrow$ **Responding sample** $\mathfrak{r}$ $\longrightarrow \cdots \longrightarrow$ **Published statistics** t

What we want to study

Random set of units

Survey Data Pipeline

Register for sampling

The collected data

$\cdots \longrightarrow$ **Population** $\mathfrak{p}$ $\longrightarrow$ **Frame** $\mathfrak{f}$ $\longrightarrow$ **Sample** $\mathfrak{s}$ $\longrightarrow$ **Responding sample** $\mathfrak{r}$ $\longrightarrow \cdots \longrightarrow$ **Published statistics** t

What we want to study

Random set of units

Survey Data Pipeline

$\cdots \longrightarrow$ **Population** $\mathfrak{p}$ $\longrightarrow$ **Frame** $\mathfrak{f}$ $\longrightarrow$ **Sample** $\mathfrak{s}$ $\longrightarrow$ **Responding sample** $\mathfrak{r}$ $\longrightarrow \cdots \longrightarrow$ **Published statistics** t

What we want to study

Register for sampling

Random set of units

The collected data

Survey Data Pipeline

$\cdots \longrightarrow$ **Population** $\mathfrak{p}$ $\longrightarrow$ **Frame** $\mathfrak{f}$ $\longrightarrow$ **Sample** $\mathfrak{s}$ $\longrightarrow$ **Responding sample** $\mathfrak{r}$ $\longrightarrow \cdots \longrightarrow$ **Published statistics** t

Population: What we want to study

Frame: Register for sampling

Sample: Random set of units

Responding sample: The collected data

DP Settings for Surveys

$$\cdots \longrightarrow \mathfrak{p} \longrightarrow \mathfrak{f} \longrightarrow \mathfrak{s} \longrightarrow \mathfrak{r} \longrightarrow \cdots \longrightarrow t$$

Two considerations

- Where does the DP mechanism *start* in the data pipeline?
- Which of the previous steps in the pipeline are kept *invariant*?

For example,

1)

$$\begin{tikzcd}
 & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
 & \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

2)

$$\begin{tikzcd}
 & & & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[ur] \arrow[dr] & & & & \\
 & & & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

Ten Possible Settings

$$\cdots \longrightarrow \mathfrak{p} \longrightarrow \mathfrak{f} \longrightarrow \mathfrak{s} \longrightarrow \mathfrak{r} \longrightarrow \cdots \longrightarrow t$$

---

1)
$$\begin{tikzcd}[row sep=small]
 & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
 & \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

2)
$$\begin{tikzcd}[row sep=small]
 & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
 & \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

3)
$$\begin{tikzcd}[row sep=small]
 & & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[ur] \arrow[dr] & & & & & \\
 & & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

4)
$$\begin{tikzcd}[row sep=small]
 & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
 & \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

5)
$$\begin{tikzcd}[row sep=small]
 & & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[ur] \arrow[dr] & & & & & \\
 & & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

6)
$$\begin{tikzcd}[row sep=small]
 & & & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[ur] \arrow[dr] & & & & \\
 & & & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

7)
$$\begin{tikzcd}[row sep=small]
 & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
 & \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

8)
$$\begin{tikzcd}[row sep=small]
 & & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[ur] \arrow[dr] & & & & & \\
 & & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

9)
$$\begin{tikzcd}[row sep=small]
 & & & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[ur] \arrow[dr] & & & & \\
 & & & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

10)
$$\begin{tikzcd}[row sep=small]
 & & & & \mathfrak{r} \arrow[r] & \cdots \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[ur] \arrow[dr] & & & \\
 & & & & \mathfrak{r}' \arrow[r] & \cdots \arrow[r] & t'
\end{tikzcd}$$

Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the max, min are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

because changing a record can change the weights of other records.

- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Privacy Considerations (I)

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $[e^{-\varepsilon}, e^{\varepsilon}]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Considerations (I)

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $[e^{-\varepsilon}, e^{\varepsilon}]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Considerations (I)

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $\left[e^{-\varepsilon}, e^{\varepsilon}\right]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Considerations (I)

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

For these attackers, the posterior-to-posterior ratio of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Considerations (I)

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

For these attackers, the posterior-to-posterior ratio of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Considerations (I)

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

For these attackers, the posterior-to-posterior ratio of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Considerations (II)

- The composition theorem does not hold when there is dependency between the sample designs.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- This will complicate global privacy loss calculations.

Privacy Considerations (II)

- The composition theorem does not hold when there is dependency between the sample designs.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- This will complicate global privacy loss calculations.

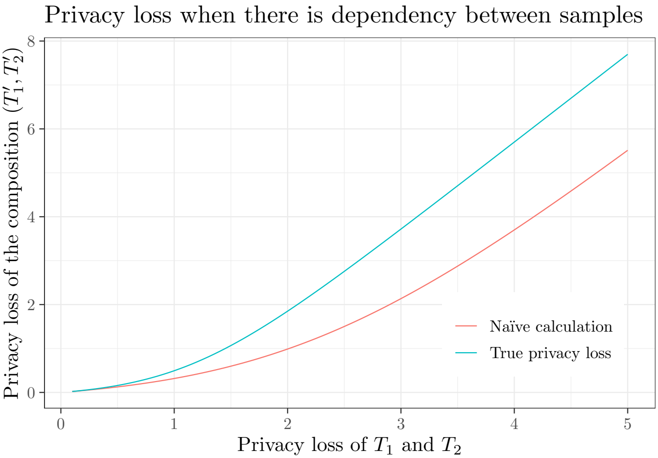
Privacy Considerations (II)

- The composition theorem does not hold when there is dependency between the sample designs.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- This will complicate global privacy loss calculations.

Privacy Considerations (II)

- The composition theorem does not hold when there is dependency between the sample designs.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- This will complicate global privacy loss calculations.

Privacy Considerations (II)



References I

Abowd, John M, Matthew J Schneider, and Lars Vilhuber (2013). "Differential privacy applications to Bayesian and linear mixed model estimation". In: *Journal of Privacy and Confidentiality* 5.1.

Ashmead, Robert, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, and William Sexton (2019). *EFFECTIVE PRIVACY AFTER ADJUSTING FOR INVARIANTS WITH APPLICATIONS TO THE 2020 Census*. Tech. rep.

Asi, Hilal, John C. Duchi, and O. Javidbakht (2022). "Element Level Differential Privacy: The Right Granularity of Privacy". In: *AAAI Workshop on Privacy-Preserving Artificial Intelligence*. Association for the Advancement of Artificial Intelligence.

Balle, Borja, Gilles Barthe, and Marco Gaboardi (Jan. 2020). "Privacy Profiles and Amplification by Subsampling". In: *Journal of Privacy and Confidentiality* 10.1. ISSN: 2575-8527. DOI: 10.29012/jpc.726.

References II

Barber, Rina Foygel and John C. Duchi (Dec. 2014). *Privacy and Statistical Risk: Formalisms and Minimax Bounds.* http://arxiv.org/abs/1412.4451. DOI: 10.48550/arXiv.1412.4451. arXiv: 1412.4451 [cs, math, stat].

Barthe, Gilles and Federico Olmedo (2013). "Beyond Differential Privacy: Composition Theorems and Relational Logic for f-Divergences between Probabilistic Programs". In: *Automata, Languages, and Programming.* Ed. by Fedor V. Fomin, Rūsiņš Freivalds, Marta Kwiatkowska, and David Peleg. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 49–60. ISBN: 978-3-642-39212-2. DOI: 10.1007/978-3-642-39212-2_8.

References III

Beimel, Amos, Shiva Prasad Kasiviswanathan, and Kobbi Nissim (Feb. 2010). "Bounds on the Sample Complexity for Private Learning and Private Data Release". In: *Proceedings of the 7th Theory of Cryptography Conference, TCC 2010, Zurich, Switzerland*. Ed. by Daniele Micciancio. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 437–454. DOI: 10.1007/978-3-642-11799-2_26.

Bhaskar, Raghav, Abhishek Bhowmick, Vipul Goyal, Srivatsan Laxman, and Abhradeep Thakurta (2011). "Noiseless Database Privacy". In: *Advances in Cryptology – ASIACRYPT 2011*. Ed. by Dong Hoon Lee and Xiaoyun Wang. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 215–232. ISBN: 978-3-642-25385-0. DOI: 10.1007/978-3-642-25385-0_12.

References IV

Bun, Mark, Jörg Drechsler, Marco Gaboardi, Audra McMillan, and Jayshree Sarathy (June 2022). "Controlling Privacy Loss in Sampling Schemes: An Analysis of Stratified and Cluster Sampling". In: *Foundations of Responsible Computing (FORC 2022)*, p. 24.

Bun, Mark and Thomas Steinke (2016). "Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds". In: *Theory of Cryptography*. Ed. by Martin Hirt and Adam Smith. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 635–658. ISBN: 978-3-662-53641-4. DOI: 10.1007/978-3-662-53641-4_24.

Charest, Anne-Sophie and Yiwei Hou (2016). "On the meaning and limits of empirical differential privacy". In: *Journal of Privacy and Confidentiality* 7.3, pp. 53–66.

References V

Chatzikokolakis, Konstantinos, Miguel E. Andrés, Nicolás Emilio Bordenabe, and Catuscia Palamidessi (2013). "Broadening the Scope of Differential Privacy Using Metrics". In: *Privacy Enhancing Technologies*. Ed. by Emiliano De Cristofaro and Matthew Wright. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 82–102. DOI: 10.1007/978-3-642-39077-7_5.

Dharangutte, Prathamesh, Jie Gao, Ruobin Gong, and Fang-Yi Yu (2023). "Integer Subspace Differential Privacy". In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-23)*.

Dong, Jinshuo, Aaron Roth, and Weijie J. Su (2022). "Gaussian Differential Privacy". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 84.1, pp. 3–37. ISSN: 1467-9868. DOI: 10.1111/rssb.12454.

References VI

Dwork, Cynthia, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor (2006). "Our Data, Ourselves: Privacy Via Distributed Noise Generation". In: *Advances in Cryptology - EUROCRYPT 2006*. Ed. by Serge Vaudenay. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 486–503. ISBN: 978-3-540-34547-3. DOI: 10.1007/11761679_29.

Dwork, Cynthia, Moni Naor, Toniann Pitassi, and Guy N. Rothblum (June 2010). "Differential Privacy under Continual Observation". In: *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*. STOC '10. https://dl.acm.org/doi/10.1145/1806689.1806787. New York, NY, USA: Association for Computing Machinery, pp. 715–724. ISBN: 978-1-4503-0050-6. DOI: 10.1145/1806689.1806787.

References VII

Ebadi, Hamid, David Sands, and Gerardo Schneider (Jan. 2015). "Differential Privacy: Now It's Getting Personal". In: *ACM SIGPLAN Notices* 50.1, pp. 69–81. ISSN: 0362-1340. DOI: 10.1145/2775051.2677005.

Feldman, Vitaly and Tijana Zrnic (Jan. 2022). *Individual Privacy Accounting via a Rényi Filter.* http://arxiv.org/abs/2008.11193. arXiv: 2008.11193 [cs, stat].

Gao, Jie, Ruobin Gong, and Fang-Yi Yu (June 2022). "Subspace Differential Privacy". In: *Proceedings of the AAAI Conference on Artificial Intelligence.* Vol. 36. 4, pp. 3986–3995. DOI: 10.1609/aaai.v36i4.20315.

Gong, Ruobin and Xiao-Li Meng (2020). "Congenial differential privacy under mandated disclosure". In: *Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference.* FODS '20, pp. 59–70.

References VIII

Hay, Michael, Chao Li, Gerome Miklau, and David Jensen (Dec. 2009). “Accurate Estimation of the Degree Distribution of Private Networks”. In: *2009 Ninth IEEE International Conference on Data Mining*, pp. 169–178. DOI: 10.1109/ICDM.2009.11.

He, Xi, Ashwin Machanavajjhala, and Bolin Ding (2014). “Blowfish privacy: Tuning privacy-utility trade-offs using policies”. In: *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pp. 1447–1458.

Jorgensen, Zach, Ting Yu, and Graham Cormode (Apr. 2015). “Conservative or Liberal? Personalized Differential Privacy”. In: *2015 IEEE 31st International Conference on Data Engineering*. https://ieeexplore.ieee.org/document/7113353, pp. 1023–1034. DOI: 10.1109/ICDE.2015.7113353. (Visited on 09/30/2023).

References IX

Kifer, Daniel, John M. Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev (Sept. 2022). *Bayesian and Frequentist Semantics for Common Variations of Differential Privacy: Applications to the 2020 Census*. Tech. rep. arXiv:2209.03310. DOI: 10.48550/arXiv.2209.03310. eprint: 2209.03310 (cs, stat). (Visited on 10/23/2022).

Kifer, Daniel and Ashwin Machanavajjhala (2011). "No Free Lunch in Data Privacy". In: *Proceedings of the 2011 International Conference on Management of Data - SIGMOD '11*. Athens, Greece: ACM Press, pp. 193–204. ISBN: 978-1-4503-0661-4. DOI: 10.1145/1989323.1989345.

— (2014). "Pufferfish: A framework for mathematical privacy definitions". In: *ACM Transactions on Database Systems (TODS)* 39.1, pp. 1–36.

References X

- McSherry, Frank and Ratul Mahajan (Aug. 2010). “Differentially-Private Network Trace Analysis”. In: *Proceedings of the ACM SIGCOMM 2010 Conference*. SIGCOMM ’10. New York, NY, USA: Association for Computing Machinery, pp. 123–134. ISBN: 978-1-4503-0201-2. DOI: 10.1145/1851182.1851199.
- Mironov, Ilya (Aug. 2017). “Rényi Differential Privacy”. In: *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275. DOI: 10.1109/CSF.2017.11. eprint: 1702.07476. (Visited on 01/14/2020).
- O’Keefe, Christine M and Anne-Sophie Charest (2019). “Bootstrap differential privacy”. In: *Transactions on Data Privacy* 12, pp. 1–28.
- Redberg, Rachel and Yu-Xiang Wang (2021). “Privately Publishable Per-Instance Privacy”. In: *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., pp. 17335–17346. (Visited on 03/29/2023).

References XI

- Seeman, Jeremy, Matthew Reimherr, and Aleksandra Slavkovic (May 2022). *Formal Privacy for Partially Private Data*. http://arxiv.org/abs/2204.01102. arXiv: 2204.01102 [cs, stat].
- Seeman, Jeremy, William Sexton, David Pujol, and Ashwin Machanavajjhala (2023+). "Per-Record Differential Privacy: Modeling Dependence between Individual Privacy Loss and Confidential Records". In.
- Soria-Comas, Jordi, Josep Domingo-Ferrer, David Sánchez, and David Megías (June 2017). "Individual Differential Privacy: A Utility-Preserving Formulation of Differential Privacy Guarantees". In: *IEEE Transactions on Information Forensics and Security* 12.6, pp. 1418–1429. ISSN: 1556-6013, 1556-6021. DOI: 10.1109/TIFS.2017.2663337. (Visited on 03/29/2023).

References XII

US Census Bureau (Dec. 2022). *Disclosure Avoidance Protections for the American Community Survey*. https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html. (Visited on 12/17/2023).

Wang, Yu-Xiang (Nov. 2018). *Per-Instance Differential Privacy*. http://arxiv.org/abs/1707.07708. arXiv: 1707.07708 [cs, stat].

Zhou, Shuheng, Katrina Ligett, and Larry Wasserman (June 2009). “Differential Privacy with Compression”. In: *Proceedings of the 2009 IEEE International Conference on Symposium on Information Theory - Volume 4*. ISIT’09. Coex, Seoul, Korea: IEEE Press, pp. 2718–2722. ISBN: 978-1-4244-4312-3.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

Intuition: DP is a bound on the *derivative* of a data-release mechanism $\frac{d}{d\mathfrak{d}}\mathrm{P}_{\mathfrak{d}}(T \in \cdot)$ at every dataset $\mathfrak{d}$ in every data universe $\mathcal{D}$.

Derivatives measure change in output per change in input. How do we measure change?

1. Data space $\mathcal{D}_0$ (the set of all theoretically-possible datasets).
3. Divergence $d_{\mathcal{D}_0}$ on $\mathcal{D}_0$.
4. Divergence d_{Pr} on the space of (probability distributions over) the output.
2. Allow for multiple data universes $\mathcal{D} \subset \mathcal{D}_0$ from a data multiverse $\mathscr{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

Intuition: DP is a bound on the *derivative* of a data-release mechanism $\frac{d}{d\mathfrak{d}}\mathrm{P}_{\mathfrak{d}}(T \in \cdot)$ at every dataset $\mathfrak{d}$ in every data universe $\mathcal{D}$.

Derivatives measure change in output per change in input. How do we measure change?

1. Data space $\mathcal{D}_0$ (the set of all theoretically-possible datasets).
3. Divergence $d_{\mathcal{D}_0}$ on $\mathcal{D}_0$.
4. Divergence d_{Pr} on the space of (probability distributions over) the output.
2. Allow for multiple data universes $\mathcal{D} \subset \mathcal{D}_0$ from a data multiverse $\mathscr{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr})$

Intuition: DP is a bound on the *derivative* of a data-release mechanism $\frac{d}{d\mathfrak{d}}\mathrm{P}_{\mathfrak{d}}(T \in \cdot)$ at every dataset $\mathfrak{d}$ in every data universe $\mathcal{D}$.

Derivatives measure change in output per change in input. How do we measure change?

1. Data space $\mathcal{D}_0$ (the set of all theoretically-possible datasets).
3. Divergence $d_{\mathcal{D}_0}$ on $\mathcal{D}_0$.
4. Divergence $d_{\Pr}$ on the space of (probability distributions over) the output.
2. Allow for multiple data universes $\mathcal{D} \subset \mathcal{D}_0$ from a data multiverse $\mathscr{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

Intuition: DP is a bound on the *derivative* of a data-release mechanism $\frac{d}{d\mathfrak{d}}\mathrm{P}_{\mathfrak{d}}(T \in \cdot)$ at every dataset $\mathfrak{d}$ in every data universe $\mathcal{D}$.

Derivatives measure change in output per change in input. How do we measure change?

1. Data space $\mathcal{D}_0$ (the set of all theoretically-possible datasets).
3. Divergence $d_{\mathcal{D}_0}$ on $\mathcal{D}_0$.
4. Divergence d_{Pr} on the space of (probability distributions over) the output.
2. Allow for multiple data universes $\mathcal{D} \subset \mathcal{D}_0$ from a data multiverse $\mathscr{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

Intuition: DP is a bound on the *derivative* of a data-release mechanism $\frac{d}{d\mathfrak{d}}\mathrm{P}_{\mathfrak{d}}(T \in \cdot)$ at every dataset $\mathfrak{d}$ in every data universe $\mathcal{D}$.

Derivatives measure change in output per change in input. How do we measure change?

1. Data space $\mathcal{D}_0$ (the set of all theoretically-possible datasets).
3. Divergence $d_{\mathcal{D}_0}$ on $\mathcal{D}_0$.
4. Divergence d_{Pr} on the space of (probability distributions over) the output.
2. Allow for multiple data universes $\mathcal{D} \subset \mathcal{D}_0$ from a data multiverse $\mathscr{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr})$

Definition

A differential privacy flavour is a tuple $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr})$.

A data release mechanism T *satisfies* $\mathrm{DP}(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr})$ with budget ϵ if

$$d_{\Pr}\left(\mathsf{P}_{\mathfrak{d}}(T \in \cdot), \mathsf{P}_{\mathfrak{d}'}(T \in \cdot)\right) \leq \epsilon d_{\mathcal{D}_0}(\mathfrak{d}, \mathfrak{d}'),$$

for all data universes $\mathcal{D} \in \mathscr{D}$ and all datasets $\mathfrak{d}, \mathfrak{d}' \in \mathcal{D}$.

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

4. d_{Pr}: (ϵ, δ)-approximate DP (Dwork et al. 2006) Rényi DP (Mironov 2017) concentrated DP (Bun and Steinke 2016) f-divergence privacy (Barber and Duchi 2014; Barthe and Olmedo 2013) f-DP (including Gaussian DP) (Dong et al. 2022).

3. $d_{\mathcal{D}_0}$: $(\mathcal{R}, \epsilon)$-generic DP (Kifer and Machanavajjhala 2011) edge vs node privacy (Hay et al. 2009; McSherry and Mahajan 2010) d-metric DP (Chatzikokolakis et al. 2013) Blowfish privacy (He et al. 2014) element level DP (Asi et al. 2022) distributional privacy (Zhou et al. 2009) event-level vs user-level DP (Dwork et al. 2010).

2. $\mathscr{D}$: privacy under invariants (Ashmead et al. 2019; Gong and Meng 2020; Gao et al. 2022; Dharangutte et al. 2023) conditioned or empirical DP (Abowd et al. 2013; Charest and Hou 2016) personalized DP (Ebadi et al. 2015; Jorgensen et al. 2015) individual DP (Soria-Comas et al. 2017; Feldman and Zrnic 2022) bootstrap DP (O'Keefe and Charest 2019) stratified DP (Bun et al. 2022) per-record DP (Seeman et al. 2023+) per-instance DP (Wang 2018; Redberg and Wang 2021).

1. $\mathcal{D}_0$: Pufferfish DP (Kifer and Machanavajjhala 2014) noiseless privacy (Bhaskar et al. 2011) privacy under partial knowledge (Seeman et al. 2022) privacy amplification (Beimel et al. 2010; Balle et al. 2020; Bun et al. 2022).

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

4. d_{Pr}: (ϵ, δ)-approximate DP (Dwork et al. 2006) Rényi DP (Mironov 2017) concentrated DP (Bun and Steinke 2016) f-divergence privacy (Barber and Duchi 2014; Barthe and Olmedo 2013) f-DP (including Gaussian DP) (Dong et al. 2022).

3. $d_{\mathcal{D}_0}$: $(\mathcal{R}, \epsilon)$-generic DP (Kifer and Machanavajjhala 2011) edge vs node privacy (Hay et al. 2009; McSherry and Mahajan 2010) d-metric DP (Chatzikokolakis et al. 2013) Blowfish privacy (He et al. 2014) element level DP (Asi et al. 2022) distributional privacy (Zhou et al. 2009) event-level vs user-level DP (Dwork et al. 2010).

2. $\mathscr{D}$: privacy under invariants (Ashmead et al. 2019; Gong and Meng 2020; Gao et al. 2022; Dharangutte et al. 2023) conditioned or empirical DP (Abowd et al. 2013; Charest and Hou 2016) personalized DP (Ebadi et al. 2015; Jorgensen et al. 2015) individual DP (Soria-Comas et al. 2017; Feldman and Zrnic 2022) bootstrap DP (O'Keefe and Charest 2019) stratified DP (Bun et al. 2022) per-record DP (Seeman et al. 2023+) per-instance DP (Wang 2018; Redberg and Wang 2021).

1. $\mathcal{D}_0$: Pufferfish DP (Kifer and Machanavajjhala 2014) noiseless privacy (Bhaskar et al. 2011) privacy under partial knowledge (Seeman et al. 2022) privacy amplification (Beimel et al. 2010; Balle et al. 2020; Bun et al. 2022).

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

4. d_{Pr}: (ϵ, δ)-approximate DP (Dwork et al. 2006) Rényi DP (Mironov 2017) concentrated DP (Bun and Steinke 2016) f-divergence privacy (Barber and Duchi 2014; Barthe and Olmedo 2013) f-DP (including Gaussian DP) (Dong et al. 2022).

3. $d_{\mathcal{D}_0}$: $(\mathcal{R}, \epsilon)$-generic DP (Kifer and Machanavajjhala 2011) edge vs node privacy (Hay et al. 2009; McSherry and Mahajan 2010) d-metric DP (Chatzikokolakis et al. 2013) Blowfish privacy (He et al. 2014) element level DP (Asi et al. 2022) distributional privacy (Zhou et al. 2009) event-level vs user-level DP (Dwork et al. 2010).

2. $\mathscr{D}$: privacy under invariants (Ashmead et al. 2019; Gong and Meng 2020; Gao et al. 2022; Dharangutte et al. 2023) conditioned or empirical DP (Abowd et al. 2013; Charest and Hou 2016) personalized DP (Ebadi et al. 2015; Jorgensen et al. 2015) individual DP (Soria-Comas et al. 2017; Feldman and Zrnic 2022) bootstrap DP (O'Keefe and Charest 2019) stratified DP (Bun et al. 2022) per-record DP (Seeman et al. 2023+) per-instance DP (Wang 2018; Redberg and Wang 2021).

1. $\mathcal{D}_0$: Pufferfish DP (Kifer and Machanavajjhala 2014) noiseless privacy (Bhaskar et al. 2011) privacy under partial knowledge (Seeman et al. 2022) privacy amplification (Beimel et al. 2010; Balle et al. 2020; Bun et al. 2022).

Four Components of a DP Flavour $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}})$

4. d_{Pr}: (ϵ, δ)-approximate DP (Dwork et al. 2006) Rényi DP (Mironov 2017) concentrated DP (Bun and Steinke 2016) f-divergence privacy (Barber and Duchi 2014; Barthe and Olmedo 2013) f-DP (including Gaussian DP) (Dong et al. 2022).
3. $d_{\mathcal{D}_0}$: $(\mathcal{R}, \epsilon)$-generic DP (Kifer and Machanavajjhala 2011) edge vs node privacy (Hay et al. 2009; McSherry and Mahajan 2010) d-metric DP (Chatzikokolakis et al. 2013) Blowfish privacy (He et al. 2014) element level DP (Asi et al. 2022) distributional privacy (Zhou et al. 2009) event-level vs user-level DP (Dwork et al. 2010).
2. $\mathscr{D}$: privacy under invariants (Ashmead et al. 2019; Gong and Meng 2020; Gao et al. 2022; Dharangutte et al. 2023) conditioned or empirical DP (Abowd et al. 2013; Charest and Hou 2016) personalized DP (Ebadi et al. 2015; Jorgensen et al. 2015) individual DP (Soria-Comas et al. 2017; Feldman and Zrnic 2022) bootstrap DP (O'Keefe and Charest 2019) stratified DP (Bun et al. 2022) per-record DP (Seeman et al. 2023+) per-instance DP (Wang 2018; Redberg and Wang 2021).
1. $\mathcal{D}_0$: Pufferfish DP (Kifer and Machanavajjhala 2014) noiseless privacy (Bhaskar et al. 2011) privacy under partial knowledge (Seeman et al. 2022) privacy amplification (Beimel et al. 2010; Balle et al. 2020; Bun et al. 2022).

Five Building Blocks of DP $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr}, \epsilon_{\mathcal{D}})$

1. The protection domain (*what* can be protected?): as defined by the dataset space $\mathcal{D}_0$;
2. The scope of protection (*to where* does the protection extend?): as instantiated by the data multiverse $\mathscr{D}$, which is a collection of data universes $\mathcal{D} \subset \mathcal{X}$;
3. The protection unit (*who* are the units for data perturbation?): as conceptualized by the divergence $d_{\mathcal{X}}$ on the dataset space $\mathcal{X}$;
4. The standard of protection (*how* to measure the output variations?): as captured by the divergence $d_{\Pr}$ on the output probability distributions; and
5. The intensity of protection (*how much* protection is afforded?): as quantified by the privacy-loss budget $\epsilon_{\mathcal{D}}$.

Five Building Blocks of DP $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr}, \epsilon_{\mathcal{D}})$

1. The protection domain (*what* can be protected?): as defined by the dataset space $\mathcal{D}_0$;
2. The scope of protection (*to where* does the protection extend?): as instantiated by the data multiverse $\mathscr{D}$, which is a collection of data universes $\mathcal{D} \subset \mathcal{X}$;
3. The protection unit (*who* are the units for data perturbation?): as conceptualized by the divergence $d_{\mathcal{X}}$ on the dataset space $\mathcal{X}$;
4. The standard of protection (*how* to measure the output variations?): as captured by the divergence $d_{\Pr}$ on the output probability distributions; and
5. The intensity of protection (*how much* protection is afforded?): as quantified by the privacy-loss budget $\epsilon_{\mathcal{D}}$.

Five Building Blocks of DP $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}}, \epsilon_{\mathcal{D}})$

1. The protection domain (*what* can be protected?): as defined by the dataset space $\mathcal{D}_0$;
2. The scope of protection (*to where* does the protection extend?): as instantiated by the data multiverse $\mathscr{D}$, which is a collection of data universes $\mathcal{D} \subset \mathcal{X}$;
3. The protection unit (*who* are the units for data perturbation?): as conceptualized by the divergence $d_{\mathcal{X}}$ on the dataset space $\mathcal{X}$;
4. The standard of protection (*how* to measure the output variations?): as captured by the divergence d_{Pr} on the output probability distributions; and
5. The intensity of protection (*how much* protection is afforded?): as quantified by the privacy-loss budget $\epsilon_{\mathcal{D}}$.

Five Building Blocks of DP $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\mathrm{Pr}}, \epsilon_{\mathcal{D}})$

1. The protection domain (*what* can be protected?): as defined by the dataset space $\mathcal{D}_0$;

2. The scope of protection (*to where* does the protection extend?): as instantiated by the data multiverse $\mathscr{D}$, which is a collection of data universes $\mathcal{D} \subset \mathcal{X}$;

3. The protection unit (*who* are the units for data perturbation?): as conceptualized by the divergence $d_{\mathcal{X}}$ on the dataset space $\mathcal{X}$;

4. The standard of protection (*how* to measure the output variations?): as captured by the divergence d_{Pr} on the output probability distributions; and

5. The intensity of protection (*how much* protection is afforded?): as quantified by the privacy-loss budget $\epsilon_{\mathcal{D}}$.

Five Building Blocks of DP $(\mathcal{D}_0, \mathscr{D}, d_{\mathcal{D}_0}, d_{\Pr}, \epsilon_{\mathcal{D}})$

1. The protection domain (*what* can be protected?): as defined by the dataset space $\mathcal{D}_0$;
2. The scope of protection (*to where* does the protection extend?): as instantiated by the data multiverse $\mathscr{D}$, which is a collection of data universes $\mathcal{D} \subset \mathcal{X}$;
3. The protection unit (*who* are the units for data perturbation?): as conceptualized by the divergence $d_{\mathcal{X}}$ on the dataset space $\mathcal{X}$;
4. The standard of protection (*how* to measure the output variations?): as captured by the divergence $d_{\Pr}$ on the output probability distributions; and
5. The intensity of protection (*how much* protection is afforded?): as quantified by the privacy-loss budget $\epsilon_{\mathcal{D}}$.